

# Tartalomjegyzék

## I. A lineáris algebra forrásai 15

### 1 Vektorok 19

#### Vektorok a 2- és 3-dimenziós térben 19

Írányított szakasz, kötött és szabad vektor 19 • Vektor magadása egy írányított szakasszal 20 • Vektor megadása hossz és írány segítségével 21 • Vektorműveletek a 2- és 3-dimenziós térben 21 • A lineáris kombináció definíciója 23 • Lineáris függetlenség 25 • Speciális lineáris kombinációk\* 26

#### Távolság, szög, orientáció 29

Skaláris szorzás 29 • Hosszúság és szög 30 • Pithagorász-tétel 31 • Két fontos egyenlőtlenség 31 • Egységvektorral való szorzás és a merőleges vetítés 32 • Merőlegesség és orientáció 33 • Vektori szorzás 34 • Parallelepipedon térfogata és előjeles térfogata 37 • Vegyes szorzat 38

#### Vektorok koordinátás alakban 41

Descartes-féle koordinátarendszer 41 • Műveletek koordinátás alakban megadott vektorokkal 42 • A derékszögű koordinátarendszer 44 • Az  $\mathbb{R}^n$  halmaz 46 • Vektorok összeadása és skalárral szorzása  $\mathbb{R}^n$ -ben 47 • Lineáris kombináció, lineáris függetlenség, lineáris összefüggőség 48 • Skaláris szorzás  $\mathbb{R}^n$ -ben 50 • Távolság és szög  $\mathbb{R}^n$ -ben 51 • Korrelációs együttható\* 54 • Bitvektorok, kódvektorok\* 55 • Vektorműveletek  $\mathbb{Z}_m^n$ -ben\* 56

#### Megoldások 60

### 2 Lineáris egyenletrendszerek és megoldásuk 63

#### Egyenes és sík egyenletei 63

Alakzatok és egyenletek 63 • Síkbeli egyenes egyenletei 65  
 • Síkbeli pont egyenletei 68 • A 3-dimenziós tér síkjainak  
 egyenletei 69 • Térbeli egyenes egyenletei 71 • Térbeli pont  
 egyenletei 74 • Egyenletek  $\mathbb{R}^n$ -ben 74

### *A lineáris egyenletrendszer és két modellje* 77

Lineáris egyenlet és egyenletrendszer 77 • Ekvivalens lineáris  
 egyenletrendszerek 79 • Mátrixok 80 • Egyenletrendszer  
 mátrixa és bővített mátrixa 81 • Sormodell: hipersíkok  
 metszete 83 • Oszlopmodell: vektor előállítása lineáris  
 kombinációként 85

### *Megoldás kiküszöböléssel* 88

Elemi sorműveletek 88 • Lépcsős alak 88 • Gauss-módszer 89  
 • Redukált lépcsős alak 93 • Gauss–Jordan-módszer 94 • A  
 redukált lépcsős alak egyértelműsége 96 • Szimultán  
 egyenletrendszerek 97 • Kiküszöbölés  $\mathbb{Z}_p$ -ben\* 99

### *Megoldás a gyakorlatban* 103

A kiküszöbölés műveletigénye 103 • Numerikusan instabil  
 egyenletrendszerek 103 • Részleges főelemkiválasztás 105  
 • Skálázás 107 • Iteratív módszerek 108 • Jacobi-iteráció 109  
 • Gauss–Seidel-iteráció 110 • Az iterációk konvergenciája 111

### *Megoldások* 115

## 3 *Megoldhatóság és a megoldások tere* 119

### *Homogén és inhomogén egyenletrendszerek megoldásai* 119

Kötött változók száma, mátrix rangja 119 • Egyenletrendszer  
 megoldhatóságának feltétele 121 • Homogén lineáris  
 egyenletrendszer megoldásai 123 • Vektortér 124 • Alterek  
 tulajdonságai és szemléltetésük 124 • Kifeszített altér 126 • Az  
 inhomogén lineáris egyenletrendszer megoldásai 127 • Lineáris  
 függetlenség és összefüggőség 129

### *Alterek tulajdonságai és az egyenletrendszerek* 132

Sor- és oszloptér 132 • Bázis 133 • Vektor egy bázisra  
 vonatkozó koordináták alakja 135 • Dimenzió és rang 136 • A  
 sortér és a nulltér merőlegessége 139 • A lineáris algebra  
 alaptétele 141 • A lineáris egyenletrendszer megoldásainak  
 jellemzése 142 • Elemi bázistranszformáció\* 144

### *Megoldások* 148

## *II. Mátrixok algebrája és geometriája* 151

### 4 *Mátrixműveletek definíciói* 155

## Műveletek táblázatokkal – műveletek mátrixokkal 155

Táblázatok összeadása 155 • Táblázat szorzása számmal 156  
 • Táblázatok szorzása 156 • Lineáris helyettesítés 157  
 • Mátrixok 158 • Elemenkénti mátrixműveletek 159  
 • Mátrixok lineáris kombinációi 160 • Mátrixszorzás 161  
 • Műveletek blokkmátrixokkal 162 • Kronecker-szorzat és a  
 vec-függvény\* 164 • Műveletek hipermátrixokkal\* 165

## A mátrixszorzás használata 170

Skaláris szorzat és diadikus szorzat mátrixszorzatos alakja 170  
 • Lineáris egyenletrendszer mátrixszorzatos alakja 171  
 • Lineáris helyettesítés mátrixszorzatos alakja 172 • Szorzás  
 vektorral 172 • Szorzás standard egységvektorral 173 • A  
 báziscsere mátrixszorzatos alakja 173 • Bázisfelbontás\* 174  
 • Egységmátrix, elemi mátrixok 175 • Vektorokra partícionált  
 mátrixok 177 • Altér felírása mátrixszorzattal 180 • \* 182

## Megoldások 183

## 5 Mátrixműveletek algebrája 187

### Az alaplóműveletek tulajdonságai 187

Az összeadás és a skalárral való szorzás tulajdonságai 187 • A  
 szorzás tulajdonságai 187 • Mátrix hatványozása 190 • A  
 transzponálás tulajdonságai 192 • Mátrixszorzás inverze –  
 mátrixok osztása\* 193 • Mátrix inverze 193 • Elemi mátrixok  
 inverze 195 • Az inverz kiszámítása 195 • Az inverz  
 tulajdonságai 198 • Az invertálhatóság és az egyenletrendszerek  
 megoldhatósága 199 • Invertálhatóság és bázis 201  
 • Báziscsere 202

### Műveletek speciális mátrixokkal 206

Diagonális mátrixok 206 • Permutáló mátrixok és kigyók 206  
 • Háromszögmátrixok 208 • Szimmetrikus és ferdén  
 szimmetrikus mátrixok 209 • Mátrix és diád összegének  
 inverze\* 210 • Gyorsszorítás\* 212

### Mátrixfelbontások 215

Az LU-felbontás 215 • Egyenletrendszer megoldása  
 LU-felbontással 218 • Mátrix invertálása LU-felbontással 219  
 • Az LU-felbontás a gyakorlatban 220 • PLU-felbontás\* 221

## Megoldások 227

## 6 Determináns 231

Parallelogramma előjeles területe 231 • Parallelepipedon előjeles  
 térfogata 232

### A determináns mint sorvektorainak függvénye 233

A determináns definíciója 233 • Mikor 0 a determináns értéke 234 • A determináns értékének kiszámítása 236 • Elemi mátrixok determinánsa 237 • Permutáló mátrix determinánsa\* 238 • Mátrixműveletek és determináns 239  
 • Determinánsok soronkénti additivitása 241

### *A determináns mint elemeinek függvénye* 247

Kígyók determinánsa 247 • Előjeles aldetermináns 249  
 • Determináns kifejtése 252 • Vandermonde-determináns 253  
 • Cramer-szabály és a mátrix inverze 255 • Blokkmátrixok determinánsa\* 259

### *Megoldások* 264

## 7 *Mátrixleképezések és geometriájuk* 271

### *Mátrixleképezés, lineáris leképezés* 271

A mátrixleképezés fogalma 271 • Műveletek mátrixleképezések között 272 • Mátrixleképezések tulajdonságai 273 • Lineáris leképezés 274 •  $\mathbb{R}^n$ -ből  $\mathbb{R}^m$ -be képző lineáris leképezések 275  
 • A mátrixleképezés hatásának szemléltetései 277 • Lineáris leképezés mátrixa különböző bázisokban 279 • Hasonlóság 280

### *Alkalmazás: differenciálhatóság\** 283

Vektor-vektor függvények differenciálhatósága 283  
 • Jacobi-mátrix 285 • Jacobi-determináns és az integrál transzformációja 288 • Függvények kompozíciójának deriváltja 291

### *2- és 3-dimenziós geometriai transzformációk mátrixa* 293

Forgatás 293 • Merőleges vetítés 296 • Tükrözés 298  
 • Vetítés 298 • Eltolás 299

### *Merőleges vetítés és a legjobb közelítés* 302

Alterek összege és direkt összege 302 • Merőleges vetítés  $\mathbb{R}^n$  egy alterére 305 • Melyik mátrix merőleges vetítés mátrixa? 306  
 • Altértől való távolság 307 • Egyenletrendszer optimális megoldása 309 • Lineáris és polinomiális regresszió 310  
 • Vetítés 312

### *Pszéudoinverz\** 316

A pszéudoinverz fogalma 316 • A pszéudoinverz tulajdonságai 319 • A pszéudoinverz és a minimális abszolút értékű optimális megoldás 321

### *Ortonormált bázis – ortogonális mátrix* 324

Ortogonalis és ortonormált bázis 324 • Ortogonalis mátrixok 326 • Ortogonalis mátrixok geometriája 328 • A 2- és 3-dimenziós tér ortogonalis transzformációi 329  
 • Givens-forgatás, Householder-tükrözés\* 331  
 • Gram–Schmidt-ortogonalizáció\* 333 • A QR-felbontás\* 335

- QR-felbontás primitív ortogonális transzformációkkal\* 337
- Egyenletrendszer optimális megoldása QR-felbontással\* 339

### *Komplex és véges test feletti terek\** 343

- Komplex vektorok skaláris szorzata 343
- $\mathbb{C}^n$  kitüntetett alterei 345
- Önadjungált mátrixok 346
- Távolság és a merőleges vetítés komplex terekben 346
- Unitér mátrixok 346
- Fourier-mátrixok 347
- Diszkrét Fourier-transzformáció 350
- Periodikus összetevők szűrése 352
- Gyors Fourier-transzformáció 354
- Vektorok konvolúciója 357

### *Megoldások* 357

## **III. Mátrixok sajátosságai** 363

### **8 Sajátérték, diagonalizálás** 367

#### *Sajátérték, sajátvektor, sajátaltér* 367

- A sajátérték és a sajátvektor fogalma 367
- Karakterisztikus polinom 369
- A valós  $2 \times 2$ -es mátrixok sajátaltereinek jellemzése 371
- Mátrix összes sajátértékének és sajátvektorának meghatározása 372
- A karakterisztikus egyenlet komplex gyökei 375
- A karakterisztikus egyenlet többszörös gyökei: az algebrai és a geometriai multiplicitás 376
- Sajátértékek és a mátrix hatványai 377
- Speciális mátrixok sajátértékei 378

#### *Hasonlóság, diagonalizálhatóság* 381

- Lineáris transzformációk sajátértékei 381
- Hasonló mátrixok sajátértékei 382
- Mátrixok diagonalizálása és sajátfelbontása 383
- Bal sajátvektorok és a sajátfelbontás diadikus alakja\* 385
- Diagonalizálható mátrixok polinomjai és a Cayley–Hamilton-tétel\* 386
- Különböző sajátértékek sajátalterei 388
- Sajátértékek multiplicitása és a diagonalizálhatóság\* 390
- Diagonalizálható mátrixok spektrálfelbontása\* 393
- Sajátalterek direkt összege 395

#### *A sajátérték kiszámítása* 396

- Gersgorin-körök 396
- Hatványmódszer 398

### *Megoldások* 403

### **9 Diagonalizálás ortonormált bázisban** 405

#### *Ortogonalis és unitér diagonalizálás* 405

- Valós mátrixok ortogonális diagonalizálása, valós spektráltétel 405
- Schur-felbontás\* 409
- Mátrixok unitér diagonalizálása\* 411

#### *Kvadratikus formák* 414

Homogén másodfokú polinomok mátrixszorzatos alakja 414  
 • Főtengelytétel 415 • Kvadratikus formák és mátrixok  
 definitisége 417 • Definitiség és főminorok 418 • Szélsőérték 419

*Megoldások* 419

## 10 *Szinguláris érték* 421

*Szinguláris érték, szinguláris vektor, SVD* 421

Szinguláris érték és szinguláris vektorok 421 • Szinguláris  
 felbontás 422 • A szinguláris felbontás meghatározása 425  
 • Szinguláris érték szerinti felbontás létezése 427 • Szinguláris  
 felbontás geometriai interpretációja 428 • Polárfelbontás 430  
 • Pszeudoinverz 431 • Információtömörítés 432

*Vektor- és mátrixnorma* 435

Vektor abszolút értéke – az euklideszi norma 435 • A  
 $p$ -norma 435 • A norma általános fogalma 436  
 • Vektornormák ekvivalenciája 438 • Vektornormák  
 mátrixokon 439 • A mátrixnorma általános fogalma 441  
 • Indukált norma 441 • Az 1-, 2- és  $\infty$ -norma mátrixokra 443

*Megoldások* 446

## 11 *Jordan-féle normálalak* 447

*Normálalak és invariáns altér* 447

Invariáns alterek 447 • Blokkdiagonális mátrixok 448  
 • Általánosított sajátvektorok és a Jordan-blokk 449  
 • Jordan-normálalak 453 • A Jordan-alak egyértelműsége 456  
 • Minimálpolinom\* 459 • A Jordan-bázis konstrukciója\* 461

*Mátrixfüggvények* 467

Diagonalizálható mátrixok függvényei 467 • Mátrixfüggvény  
 kiszámítása a Jordan-alakból 468 • Mátrixfüggvény kiszámítása  
 polinominterpolációval 470

*Megoldások* 473

## 12 *Nemnegatív mátrixok* 475

*A Perron–Frobenius-elmélet* 475

Mátrixok összehasonlítása 475 • Pozitív mátrixok 476  
 • Nemnegatív mátrixok 478 • Irreducibilis mátrixok 481  
 • Primitív és imprimitív mátrixok 483

*Sztochasztikus mátrixok* 487

Markov-láncok, sztochasztikus mátrixok 487 • Duplán  
 sztochasztikus mátrixok 487 • A Leontief-modell 489

*Megoldások* 491

## **IV. Kitekintő** 493

### **13 Terek** 497

*Testek, gyűrűk* 497

Test 497 • Gyűrű 498 • Egyműveletes struktúrák\* 499

*Vektortér* 501

*Modulus* 502

### **14 Mátrixegyenletek és -egyenlőtlenségek** 503

*Lineáris mátrixegyenletek* 503

*Irodalomjegyzék* 507

*Tárgymutató* 509





# Listák

## Tételek, állítások, következmények

|                                                                         |    |                                                                      |     |
|-------------------------------------------------------------------------|----|----------------------------------------------------------------------|-----|
| 1.2. Parallelogramma-módszer . . . . .                                  | 22 | 1.57. Ortogonális vektorrendszer lineáris független-sége . . . . .   | 53  |
| 1.5. A vektorműveletek tulajdonságai . . . . .                          | 23 | 2.5. Síkbeli egyenes explicit vektoregyenlete . . . . .              | 65  |
| 1.7. Vektorral párhuzamos vektorok . . . . .                            | 24 | 2.6. Síkbeli egyenes implicit vektoregyenlete . . . . .              | 66  |
| 1.8. Két vektorral egy síkba eső vektorok . . . . .                     | 24 | 2.7. Síkbeli egyenes explicit egyenletrendszere . . . . .            | 66  |
| 1.9. Térbeli vektorok . . . . .                                         | 25 | 2.8. Síkbeli egyenes (implicit) egyenlete . . . . .                  | 66  |
| 1.11. Síkbeli vektor felbontása . . . . .                               | 26 | 2.10. Sík explicit vektoregyenlete . . . . .                         | 69  |
| 1.12. Térbeli vektor felbontása . . . . .                               | 26 | 2.11. Sík implicit vektoregyenlete . . . . .                         | 69  |
| 1.13. Két ponton átmenő egyenes jellemzése . . . . .                    | 26 | 2.12. Sík explicit egyenletrendszere . . . . .                       | 70  |
| 1.14. Intervallum pontjainak jellemzése . . . . .                       | 27 | 2.13. Sík implicit egyenlete . . . . .                               | 70  |
| 1.17. Mikor 0 a skaláris szorzat? . . . . .                             | 29 | 2.15. Térbeli egyenes explicit vektoregyenlete . . . . .             | 71  |
| 1.18. A skaláris szorzás műveleti tulajdonságai . . . . .               | 30 | 2.16. Térbeli egyenes explicit egyenletrendszere . . . . .           | 72  |
| 1.19. Pithagorász-tétel . . . . .                                       | 31 | 2.17. Térbeli egyenes implicit egyenletrendszere . . . . .           | 72  |
| 1.21. Cauchy–Bunyakovszkij–Schwarz-egyenlőtlenség                       | 31 | 2.24. Ekvivalens átalakítások . . . . .                              | 79  |
| 1.22. Háromszög-egyenlőtlenség . . . . .                                | 32 | 2.28. Sormodell . . . . .                                            | 85  |
| 1.23. Egységvektorral való szorzás geometriai jelentése                 | 32 | 2.30. Oszlopmodell . . . . .                                         | 86  |
| 1.24. Vektor felbontása merőleges összetevőkre . . . . .                | 33 | 2.36. Lépcsős alakra hozás . . . . .                                 | 91  |
| 1.29. Mikor 0 a vektori szorzat? . . . . .                              | 36 | 2.44. A redukált lépcsős alak egyértelmű . . . . .                   | 96  |
| 1.30. Vektori szorzat abszolút értékének geometriai jelentése . . . . . | 36 | 2.50. A kiküszöbölés műveletigénye . . . . .                         | 103 |
| 1.31. Vektori szorzás műveleti tulajdonságai . . . . .                  | 37 | 2.59. Elégséges feltétel az iterációk konvergenciájára .             | 112 |
| 1.35. Ekvivalenciareláció . . . . .                                     | 40 | 3.1. Főelemek oszlopai . . . . .                                     | 119 |
| 1.38. Vektorműveletek koordinátás alakja . . . . .                      | 43 | 3.4. Köötött és szabad változók száma . . . . .                      | 120 |
| 1.40. Skaláris szorzat ortonormált koordinátarendszerben . . . . .      | 44 | 3.5. A megoldhatóság mátrixrangos feltétele . . . . .                | 121 |
| 1.41. Vektori szorzat ortonormált koordinátarendszerben . . . . .       | 45 | 3.6. Homogén lineáris egyenletrendszer megoldhatósága . . . . .      | 122 |
| 1.46. Az összeadás és skalárral szorzás tulajdonságai                   | 47 | 3.8. Megoldások lineáris kombinációja . . . . .                      | 123 |
| 1.48. Lineáris függetlenség . . . . .                                   | 49 | 3.12. Megoldások altere . . . . .                                    | 126 |
| 1.49. Lineáris összefüggőség . . . . .                                  | 50 | 3.15. A kifeszített altér altér . . . . .                            | 126 |
| 1.51. A skaláris szorzás tulajdonságai . . . . .                        | 51 | 3.17. Homogén és inhomogén egyenletrendszer megoldásai . . . . .     | 127 |
| 1.54. Cauchy–Bunyakovszkij–Schwarz-egyenlőtlenség                       | 52 | 3.19. Inhomogén egyenletrendszer megoldhatósága .                    | 129 |
| 1.55. Háromszög-egyenlőtlenség $\mathbb{R}^n$ -ben . . . . .            | 53 | 3.21. Lineáris függetlenség eldöntése . . . . .                      | 130 |
| 1.56. Skaláris szorzat és abszolút érték $\mathbb{R}^n$ -ben . . . . .  | 53 | 3.23. Elemi sorműveletek hatása a sor- és oszlopvektorokra . . . . . | 132 |
|                                                                         |    | 3.24. Mátrix lépcsős alakjának vektori . . . . .                     | 133 |
|                                                                         |    | 3.28. Bázis ekvivalens definíciói . . . . .                          | 135 |
|                                                                         |    | 3.30. Bázis-tétel . . . . .                                          | 136 |
|                                                                         |    | 3.33. Dimenzió = rang . . . . .                                      | 138 |
|                                                                         |    | 3.36. Dimenziótétel . . . . .                                        | 139 |

|                                                                                |     |                                                                                                  |     |
|--------------------------------------------------------------------------------|-----|--------------------------------------------------------------------------------------------------|-----|
| 3.38. A sortér és a nulltér merőlegessége . . . . .                            | 140 | 6.19. Determináns rendjének csökkentése . . . . .                                                | 250 |
| 3.39. A lineáris algebra alaptétele . . . . .                                  | 141 | 6.21. Determinánsok kifejtési tétele . . . . .                                                   | 252 |
| 3.40. A négy kitüntetett altér . . . . .                                       | 142 | 6.25. Vandermonde-determináns értéke . . . . .                                                   | 255 |
| 3.41. Lineáris egyenletrendszer megoldásai . . . . .                           | 142 | 6.26. Cramer-szabály . . . . .                                                                   | 256 |
| 3.43. Elemi bázistranszformáció . . . . .                                      | 145 | 6.28. Mátrix inverzének elemei . . . . .                                                         | 257 |
| 4.11. Műveletek blokkmátrixokkal . . . . .                                     | 163 | 6.30. Determinánsok szorzata blokkmátrixban . . . . .                                            | 259 |
| 4.13. A Kronecker-szorzat tulajdonságai . . . . .                              | 164 | 6.31. $2 \times 2$ -es blokkmátrix determinánsa . . . . .                                        | 260 |
| 4.14. A Kronecker-szorzat és a vec-függvény tulajdonságai . . . . .            | 165 | 7.2. Mátrixleképezések alapműveletei . . . . .                                                   | 272 |
| 4.21. Mátrixszorzás és lineáris kombináció . . . . .                           | 172 | 7.3. Inverz mátrixleképezések . . . . .                                                          | 273 |
| 4.22. Mátrix elemeinek, sor- és oszlopvektorainak előállítás . . . . .         | 173 | 7.4. Mátrixleképezések alaptulajdonságai . . . . .                                               | 273 |
| 4.25. Koordináták változása a bázis cseréjénél . . . . .                       | 174 | 7.7. Síkbeli forgatás, tükrözés, vetítés . . . . .                                               | 275 |
| 4.26. Bázisfelbontás . . . . .                                                 | 174 | 7.8. Lineáris leképezés ekvivalens definíciói . . . . .                                          | 275 |
| 4.31. Elemi sorműveletek mátrixszorzással . . . . .                            | 177 | 7.9. Az $\mathbb{R}^n \rightarrow \mathbb{R}^m$ lineáris leképezések mátrixleképezések . . . . . | 275 |
| 4.32. A szorzat oszlopai és sorai . . . . .                                    | 179 | 7.12. Lineáris leképezés mátrixai közti kapcsolat . . . . .                                      | 279 |
| 4.33. Szorzat rangja . . . . .                                                 | 179 | 7.15. Hasonló mátrixok hatása . . . . .                                                          | 281 |
| 5.1. Mire vigyázzunk a mátrixszorzásnál? . . . . .                             | 188 | 7.16. Hasonlóságra invariáns tulajdonságok . . . . .                                             | 281 |
| 5.2. Mátrixszorzás algebrai tulajdonságai . . . . .                            | 188 | 7.18. Jacobi-mátrix . . . . .                                                                    | 285 |
| 5.3. Hatványozás azonosságai . . . . .                                         | 190 | 7.21. Láncszabály . . . . .                                                                      | 291 |
| 5.6. Transzponálás tulajdonságai . . . . .                                     | 192 | 7.23. A forgatás mátrixa . . . . .                                                               | 293 |
| 5.9. Sorművelet inverzének mátrixa . . . . .                                   | 195 | 7.26. Tengely körüli forgatás – Rodrigues-formula . . . . .                                      | 294 |
| 5.10. Az inverz létezéséhez elég egy feltétel . . . . .                        | 195 | 7.29. Egyenesre való merőleges vetítés mátrixa . . . . .                                         | 296 |
| 5.11. Inverz kiszámítása elemi sorműveletekkel . . . . .                       | 196 | 7.30. Síkra való merőleges vetítés mátrixa . . . . .                                             | 297 |
| 5.13. $2 \times 2$ -es mátrix inverze . . . . .                                | 197 | 7.32. Síkbeli tükrözés mátrixa . . . . .                                                         | 298 |
| 5.14. Az inverz alaptulajdonságai . . . . .                                    | 198 | 7.33. Síkra való tükrözés mátrixa . . . . .                                                      | 298 |
| 5.15. Az invertálhatóság és az egyenletrendszerek . . . . .                    | 199 | 7.35. Alterek összege . . . . .                                                                  | 302 |
| 5.19. Invertálhatóság és bázis . . . . .                                       | 201 | 7.36. Kiegészítő alterek tulajdonságai . . . . .                                                 | 302 |
| 5.20. Szinguláris mátrixok . . . . .                                           | 202 | 7.38. A merőleges kiegészítő altér tulajdonságai . . . . .                                       | 304 |
| 5.21. Az áttérés mátrixának inverze . . . . .                                  | 202 | 7.39. Altérre való vetítés mátrixa . . . . .                                                     | 305 |
| 5.24. Műveletek diagonális mátrixokkal . . . . .                               | 206 | 7.41. Merőleges vetítés mátrixai . . . . .                                                       | 306 |
| 5.28. Műveletek permutáló mátrixokkal . . . . .                                | 207 | 7.43. Legjobb közelítés tétele . . . . .                                                         | 307 |
| 5.30. Műveletek háromszögmátrixokkal . . . . .                                 | 209 | 7.44. Vektor felbontása összetevőkre . . . . .                                                   | 308 |
| 5.33. Műveletek (ferdén) szimmetrikus mátrixokkal . . . . .                    | 209 | 7.46. Egyenletrendszer optimális megoldása . . . . .                                             | 309 |
| 5.34. Felbontás szimmetrikus és ferdén szimmetrikus mátrix összegére . . . . . | 209 | 7.47. Lineáris regresszió . . . . .                                                              | 311 |
| 5.35. $A^T A$ és $AA^T$ szimmetrikus . . . . .                                 | 210 | 7.48. Linearizálható regressziós modellek . . . . .                                              | 311 |
| 5.36. Sherman – Morrison-formula . . . . .                                     | 210 | 7.51. A projekció tulajdonságai . . . . .                                                        | 313 |
| 5.42. Az LU-felbontás létezése és egyértelműsége . . . . .                     | 217 | 7.53. A vetítés ekvivalens definíciója . . . . .                                                 | 314 |
| 6.2. Ránézésre 0 determinánsok . . . . .                                       | 234 | 7.54. Mikor merőleges egy vetítés? . . . . .                                                     | 314 |
| 6.3. Zérus értékű determináns . . . . .                                        | 234 | 7.57. A pszeudoinverz mátrixa . . . . .                                                          | 317 |
| 6.5. Egyenletrendszer megoldhatósága és a determináns . . . . .                | 235 | 7.59. Moore–Penrose-tétel . . . . .                                                              | 320 |
| 6.6. Háromszögmátrix determinánsa . . . . .                                    | 236 | 7.60. $A^+ A$ és $AA^+$ merőleges vetítés . . . . .                                              | 320 |
| 6.8. Elemi mátrixok determinánsa . . . . .                                     | 238 | 7.61. Optimális megoldás pszeudoinverzrel . . . . .                                              | 321 |
| 6.9. Permutáló mátrix determinánsa . . . . .                                   | 238 | 7.64. Ortogonális vektorok függetlensége . . . . .                                               | 324 |
| 6.10. Determinánsok szorzásszabálya . . . . .                                  | 239 | 7.65. Legjobb közelítés ONB esetén . . . . .                                                     | 325 |
| 6.12. Transzponált determinánsa . . . . .                                      | 240 | 7.69. Szemiortogonális mátrixok ekvivalens definíciói . . . . .                                  | 327 |
| 6.14. Soronkénti additivitás . . . . .                                         | 241 | 7.70. Ortogonális mátrixok ekvivalens definíciói . . . . .                                       | 327 |
| 6.15. Felbontás kigyók determinánsainak összegére . . . . .                    | 248 | 7.72. Ortogonális mátrixhoz tartozó mátrixleképezés . . . . .                                    | 328 |
| 6.16. Determinánsfüggvény létezése . . . . .                                   | 249 | 7.73. Ortogonális mátrixok tulajdonságai . . . . .                                               | 329 |
|                                                                                |     | 7.74. . . . .                                                                                    | 330 |
|                                                                                |     | 7.76. Egy vektor tükrözése egy másikba . . . . .                                                 | 332 |
|                                                                                |     | 7.78. Gram–Schmidt-ortogonalizáció . . . . .                                                     | 333 |

|                   |                                                               |     |
|-------------------|---------------------------------------------------------------|-----|
| 7.82.             | QR-felbontás létezése és egyértelműsége . . .                 | 337 |
| 7.85.             | Legkisebb négyzetek QR-felbontással . . .                     | 340 |
| 7.89.             | Az adjungált tulajdonságai . . .                              | 344 |
| 7.90.             | A komplex skaláris szorzás tulajdonságai . . .                | 344 |
| 7.91.             | Komplex mátrix kitüntetett alterei . . .                      | 345 |
| 7.94.             | Fourier-összeg helyettesítési értékei . . .                   | 347 |
| 7.95.             | A Fourier-mátrixok tulajdonságai . . .                        | 349 |
| 7.97.             | A DFT tulajdonságai . . .                                     | 351 |
| 7.100.            | Gyors Fourier-transzformáció . . .                            | 355 |
| 8.4.              | A sajátvektorok alterei . . .                                 | 368 |
| 8.8.              | Háromszögmátrixok sajátértékei . . .                          | 370 |
| 8.9.              | Determináns, nyom és a sajátértékek . . .                     | 371 |
| 8.11.             | $2 \times 2$ -es szimmetrikus mátrixok sajátalterei . . .     | 372 |
| 8.16.             | Mátrix invertálhatósága és a 0 sajátérték . . .               | 377 |
| 8.17.             | Mátrix hatványainak sajátértékei és sajátvektorai . . .       | 377 |
| 8.18.             | Mátrix hatványinak hatása . . .                               | 378 |
| 8.19.             | Speciális mátrixok sajátértéke . . .                          | 379 |
| 8.20.             | Speciális komplex mátrixok sajátértékei . . .                 | 379 |
| 8.23.             | Sajátértekhez kapcsolódó invarianták . . .                    | 382 |
| 8.25.             | Diagonalizálhatóság szükséges és elégséges fel-tétele . . .   | 383 |
| 8.28.             | Diagonalizálható mátrix polinomja . . .                       | 386 |
| 8.29.             | Cayley–Hamilton-tétel . . .                                   | 387 |
| 8.30.             | Különböző sajátértékek sajátvektorai . . .                    | 388 |
| 8.31.             | Különböző sajátértékek és a diagonalizálhatóság . . .         | 389 |
| 8.33.             | Algebrai és geometriai multiplicitás kapcsolata . . .         | 390 |
| 8.34.             | Diagonalizálhatóság és a geometriai multiplicitás . . .       | 391 |
| 8.36.             | Diagonalizálható mátrixok spektrálfelbontása . . .            | 393 |
| 8.39.             | A direkt összeg tulajdonságai . . .                           | 395 |
| 8.40.             | Diagonalizálható mátrixok sajátalterei . . .                  | 395 |
| 8.42.             | Gersgorin-körök tulajdonságai . . .                           | 397 |
| 8.44.             | Domináns főátlójú mátrix invertálhatósága . . .               | 398 |
| 8.46.             | Hatványmódszer . . .                                          | 401 |
| 9.2.              | Szimmetrikus mátrix sajátalterei . . .                        | 406 |
| 9.3.              | Valós spektráltétel . . .                                     | 406 |
| 9.5.              | Schur-felbontás . . .                                         | 409 |
| 9.9.              | Unitér diaginalizálhatóság . . .                              | 411 |
| 9.12.             | Főtengelytétel . . .                                          | 415 |
| 9.16.             | Definitég meghatározása a sajátértékekből . . .               | 418 |
| 9.17.             | Definit mátrixok főminorai és vezető főminorai . . .          | 418 |
| 10.5.             | Az SVD létezése és $\Sigma$ egyértelműsége . . .              | 427 |
| 10.6.             | Egységgömb képe . . .                                         | 429 |
| 10.7.             | Polárfelbontás . . .                                          | 430 |
| 10.9.             | A pszeudoinverz kiszámítása . . .                             | 431 |
| 10.11.            | Kis rangú approximáció tétele – Eckart–Young-tétel . . .      | 432 |
| 10.16.            | Minden vektornorma ekvivalens . . .                           | 439 |
| 10.18.            | Frobenius-norma ekvivalens alakjai . . .                      | 440 |
| 10.19.            | . . .                                                         | 440 |
| 10.23.            | Indukált norma tulajdonságai . . .                            | 442 |
| 10.24.            | $l_1$ , $l_2$ és $\infty$ -norma kiszámítása . . .            | 443 |
| 11.2.             | Invariáns al tér bázisa . . .                                 | 448 |
| 11.4.             | Blokkdiagonális mátrixok és az invariáns alterek . . .        | 448 |
| 11.8.             | Jordan-normálalak . . .                                       | 453 |
| 11.10A            | Jordan-alak egyértelműsége . . .                              | 456 |
| 11.14A            | minimálpolinom tulajdonságai . . .                            | 460 |
| 11.15J            | Jordan normálalak és minimálpolinom . . .                     | 461 |
| 11.23A            | definíciók ekvivalenciája . . .                               | 473 |
| 12.2.             | Perron-tétel: pozitív sajátérték és sajátvektor . . .         | 476 |
| 12.3.             | Perron-tétel: egyszeres és domináns sajátérték . . .          | 478 |
| 12.4.             | Perron–Frobenius-tétel – gyenge változat . . .                | 479 |
| 12.5.             | Collatz–Wielandt-tétel . . .                                  | 479 |
| 12.6.             | Nemnegatív mátrixok spektrálsugarának bec-s-lése . . .        | 480 |
| 12.7.             | Reducibilis és irreducibilis mátrixok . . .                   | 481 |
| 12.9.             | Perron–Frobenius-tétel – erős változat . . .                  | 483 |
| 12.10F            | eltétel mátrix primitivitására . . .                          | 483 |
| 12.12P            | rerron–Frobenius-tétel – sajátértékek a spek-trál-körön . . . | 485 |
| 12.13S            | tochasticikus mátrix sajátértékei . . .                       | 487 |
| 12.14F            | robenius–König-tétel . . .                                    | 488 |
| 12.15P            | ozitív kígyó . . .                                            | 488 |
| 12.16B            | irkhoff-tétel . . .                                           | 488 |
| 13.3.             | Vektortér alap tulajdonságai . . .                            | 501 |
| 14.2.             | A lineáris mátrixegyenlet megoldása . . .                     | 503 |
| 14.4.             | A Sylvester-egyenlet egyértelmű megoldhatósága . . .          | 504 |
| <i>Definíciók</i> |                                                               |     |
| .                 | Irányított szakasz, kötött vektor . . .                       | 19  |
| .                 | Vektor . . .                                                  | 20  |
| .                 | Zérusvektor . . .                                             | 20  |
| .                 | Vektor hossza . . .                                           | 21  |
| .                 | Vektorok szöge . . .                                          | 21  |
| 1.1.              | Két vektor összege – háromszög módszer . . .                  | 21  |
| 1.3.              | Vektorok különbsége . . .                                     | 22  |
| 1.4.              | Vektor szorozása skálárral . . .                              | 23  |
| 1.6.              | Lineáris kombináció . . .                                     | 23  |
| 1.10.             | Vektorok függetlensége . . .                                  | 25  |
| 1.15.             | Két vektor skalárszorzata . . .                               | 29  |
| .                 | Egységvektor . . .                                            | 32  |
| 1.26.             | Vektori szorzás . . .                                         | 35  |
| 1.33.             | Vegyesszorzat . . .                                           | 38  |
| .                 | Vektor koordinátái 2D-ben . . .                               | 41  |
| .                 | Vektor koordinátái 3D-ben . . .                               | 41  |
| 1.44.             | . . .                                                         | 46  |
| 1.45.             | Vektorműveletek $\mathbb{R}^n$ -ben . . .                     | 47  |
| 1.50.             | Skalárszorzás $\mathbb{R}^n$ -ben . . .                       | 51  |
| 1.52.             | Abszolút érték, szög, merőlegesség, távolság . . .            | 51  |
| .                 | Korrrelációs együttható . . .                                 | 54  |
| 1.58.             | Kód . . .                                                     | 56  |
| 2.3.              | Alakzat (implicit) egyenletrendszere . . .                    | 64  |

|                                                              |     |                                                                           |     |
|--------------------------------------------------------------|-----|---------------------------------------------------------------------------|-----|
| 2.4. Alakzat (explicit) egyenletrendszere . . . . .          | 65  | 7.17. Differenciálhatóság . . . . .                                       | 284 |
| . Hipersík . . . . .                                         | 75  | . Kiegészítő altér . . . . .                                              | 302 |
| 2.20. Lineáris egyenlet . . . . .                            | 77  | 7.37. Direkt összeg . . . . .                                             | 304 |
| 2.21. Lineáris egyenletrendszer . . . . .                    | 78  | . Altérre való merőleges vetület . . . . .                                | 305 |
| 2.22. Lineáris egyenletrendszer megoldása . . . . .          | 79  | . Teljes oszloprangú vagy sorrangú mátrix . . . . .                       | 305 |
| 2.23. Ekvivalens egyenletrendszerek . . . . .                | 79  | . Optimális megoldás . . . . .                                            | 309 |
| 2.31. Elemi sorműveletek . . . . .                           | 88  | . Normálegyenlet-rendszer . . . . .                                       | 309 |
| 2.32. Lépcsős alak . . . . .                                 | 88  | . Regressziós egyenes . . . . .                                           | 311 |
| 2.39. Redukált lépcsős alak . . . . .                        | 93  | 7.50. Vetítés altérre . . . . .                                           | 312 |
| . rref függvény . . . . .                                    | 97  | 7.55. A Moore–Penrose-féle pszeudinverz . . . . .                         | 316 |
| 2.45. Szimultán egyenletrendszerek . . . . .                 | 97  | . Ortogonális és ortonormált bázis . . . . .                              | 324 |
| 2.58. Soronként domináns főátlójú mátrix . . . . .           | 112 | 7.67. Ortogonális és szemiorotogonális mátrix . . . . .                   | 326 |
| 3.2. Mátrix rangja . . . . .                                 | 120 | . Givens-forgatás . . . . .                                               | 331 |
| 3.9. Vektortér . . . . .                                     | 124 | . Householder-tükrözés . . . . .                                          | 332 |
| 3.10. Altér . . . . .                                        | 124 | 7.80. QR-felbontás . . . . .                                              | 335 |
| 3.13. Nulltér . . . . .                                      | 126 | 7.87. Komplex mátrix adjungáltja . . . . .                                | 343 |
| 3.14. Kifeszített altér . . . . .                            | 126 | 7.88. Komplex vektorok skaláris szorzata . . . . .                        | 344 |
| 3.18. Sortér, oszloptér . . . . .                            | 129 | . . . . .                                                                 | 346 |
| 3.25. Bázis . . . . .                                        | 133 | . Komplex vektorok hossza, távolsága, szöge, me-<br>rőlegessége . . . . . | 346 |
| 3.31. Dimenzió . . . . .                                     | 137 | 7.93. Unitér mátrix . . . . .                                             | 346 |
| 3.34. Vektorrendszer rangja . . . . .                        | 138 | . . . . .                                                                 | 347 |
| . Merőleges altér és merőleges kiegészítő altér . . . . .    | 140 | 7.96. Diszkrét Fourier-transzformáció (DFT) . . . . .                     | 350 |
| 4.1. Lineáris helyettesítés . . . . .                        | 157 | . . . . .                                                                 | 354 |
| 4.3. Adott típusú mátrixok tere . . . . .                    | 159 | 8.2. Sajátérték, sajátvektor . . . . .                                    | 368 |
| . Mátrixok egyenlősége . . . . .                             | 159 | 8.5. Sajátaltér . . . . .                                                 | 368 |
| 4.4. Mátrixok összege, különbsége . . . . .                  | 159 | . . . . .                                                                 | 369 |
| 4.6. Zérusmátrix . . . . .                                   | 160 | 8.21. Lineáris transzformáció sajátértéke, sajátvektora . . . . .         | 381 |
| 4.7. Mátrix szorzása skalárral . . . . .                     | 160 | 8.24. Diagonalizálhatóság . . . . .                                       | 383 |
| 4.9. Mátrixok szorzása . . . . .                             | 161 | . Bal sajátvektor . . . . .                                               | 385 |
| 4.15. Hipermátrix . . . . .                                  | 165 | 8.38. . . . .                                                             | 395 |
| 4.16. Hipermátrix transzponáltja . . . . .                   | 166 | . Gersgorin-körök . . . . .                                               | 396 |
| 4.17. Hipermátrixok külső szorzata . . . . .                 | 167 | . Szigorúan domináns saját pár . . . . .                                  | 398 |
| 4.18. Diadikus szorzat . . . . .                             | 170 | 9.1. Ortogonális diagonalizálhatóság . . . . .                            | 405 |
| 4.24. Áttérés mátrixa . . . . .                              | 174 | 9.7. Unitér diagonalizálhatóság . . . . .                                 | 411 |
| 4.28. Egységmátrix . . . . .                                 | 176 | 9.8. Normális mátrix . . . . .                                            | 411 |
| 4.29. Elemi mátrixok . . . . .                               | 176 | . . . . .                                                                 | 414 |
| 5.7. Mátrix inverze . . . . .                                | 194 | 9.11. Kvadratikus forma . . . . .                                         | 415 |
| . . . . .                                                    | 194 | 9.14. Kvadratikus formák és mátrixok definitsége . . . . .                | 417 |
| 5.26. Permutáló mátrix, kígyó . . . . .                      | 207 | . Főminor, vezető főminor . . . . .                                       | 418 |
| 5.29. Háromszögmátrix . . . . .                              | 208 | 10.1. Szinguláris érték . . . . .                                         | 422 |
| 5.31. Szimmetrikus és ferdén szimmetrikus mátrixok . . . . . | 209 | 10.2. Szinguláris felbontás . . . . .                                     | 424 |
| 5.39. LU-felbontás . . . . .                                 | 215 | . Polárfelbontás . . . . .                                                | 430 |
| 5.45. PLU-felbontás . . . . .                                | 221 | 10.12 Euklideszi norma . . . . .                                          | 435 |
| . . . . .                                                    | 231 | 10.13 $p$ -norma . . . . .                                                | 436 |
| 6.1. Determináns . . . . .                                   | 233 | 10.14 Norma . . . . .                                                     | 437 |
| 6.17. Előjeles aldetermináns . . . . .                       | 249 | 10.15 Normák ekvivalenciája . . . . .                                     | 439 |
| 6.24. Vandermonde-determináns . . . . .                      | 254 | 10.17 Frobenius-norma . . . . .                                           | 440 |
| . . . . .                                                    | 271 | 10.20 Mátrixnorma . . . . .                                               | 441 |
| 7.5. Lineáris leképezés . . . . .                            | 274 | 10.21. . . . .                                                            | 441 |
| 7.14. Hasonlóság . . . . .                                   | 280 | 10.22 Indukált norma . . . . .                                            | 441 |
| . Lineáris leképezés rangja . . . . .                        | 281 |                                                                           |     |

|                                                                |     |                                                               |     |
|----------------------------------------------------------------|-----|---------------------------------------------------------------|-----|
| 11.1. Invariáns altér . . . . .                                | 447 | 2.42. Gauss – Jordan-módszer, egy megoldás . . . . .          | 94  |
| 11.5. Általánosított sajátvektor . . . . .                     | 450 | 2.43. Gauss – Jordan-módszer, végtelen sok megoldás . . . . . | 95  |
| 11.7. Jordan-blokk . . . . .                                   | 453 | 2.46. Szimultán egyenletrendszer megoldása . . . . .          | 98  |
| 11.13. Minimálpolinom . . . . .                                | 459 | 2.47. Szimultán egyenletrendszer bővített mátrixa . . . . .   | 98  |
| 11.18. Spektrumon definiált függvény . . . . .                 | 469 | 2.48. Egyenletrendszer $\mathbb{Z}_2$ fölött . . . . .        | 99  |
| 11.19. Mátrixfüggvény a Jordan-alakból . . . . .               | 469 | 2.49. Egyenletrendszer $\mathbb{Z}_5$ fölött . . . . .        | 100 |
| 11.21. Mátrixfüggvény interpolációs polinommal . . . . .       | 471 | 2.51. Instabil egyenletrendszer . . . . .                     | 104 |
| 12.1. Primitív, ireducibilis és reducibilis mátrixok . . . . . | 475 | 2.52. Gauss-módszer lebegőpontos számokkal . . . . .          | 105 |
| . . . . .                                                      | 477 | 2.53. Részleges főelemkiválasztás . . . . .                   | 106 |
| . Sztochasztikus vektorok és mátrixok . . . . .                | 487 | 2.54. Sor szorzása . . . . .                                  | 107 |
| . Duplán sztochasztikus mátrixok . . . . .                     | 487 | 2.55. Jacobi-iteráció . . . . .                               | 109 |
| 13.1. Test . . . . .                                           | 497 | 2.56. Gauss – Seidel-iteráció . . . . .                       | 110 |
| 13.2. . . . .                                                  | 501 | 2.57. Divergens iteráció . . . . .                            | 111 |
| 14.1. Lineáris mátrixegyenlet . . . . .                        | 503 | 3.3. Mátrix rangjának kiszámítása . . . . .                   | 120 |

### *Kidolgozott példák*

|                                                                             |    |                                                                                |     |
|-----------------------------------------------------------------------------|----|--------------------------------------------------------------------------------|-----|
| 1.16. Skaláris szorzat kiszámítása a definíció alapján . . . . .            | 29 | 3.20. Kifeszített altér vektorai . . . . .                                     | 129 |
| 1.20. Skaláris szorzat kiszámítása . . . . .                                | 31 | 3.22. Vektorok lineáris függetlenségének eldöntése . . . . .                   | 130 |
| 1.25. Merőleges összetevőkre bontás . . . . .                               | 33 | 3.26. Altér bázisának meghatározása . . . . .                                  | 134 |
| 1.27. Vektori szorzat meghatározása . . . . .                               | 35 | 3.27. Vektor felírása a bázisvektorok lineáris kombi-<br>nációjaként . . . . . | 134 |
| 1.28. $\mathbf{i}$ , $\mathbf{j}$ , $\mathbf{k}$ vektori szorzata . . . . . | 36 | 3.29. Vektor koordinátás alakja a $\mathcal{B}$ bázisban . . . . .             | 136 |
| 1.32. Parallelepipedon térfogata . . . . .                                  | 37 | 3.32. Mátrix transzponáltja . . . . .                                          | 138 |
| 1.34. Vegyes szorzat . . . . .                                              | 38 | 3.35. Dimenzió kiszámítása . . . . .                                           | 138 |
| 1.36. Vektorok koordinátái . . . . .                                        | 41 | 3.37. Vektorokra merőleges altér . . . . .                                     | 140 |
| 1.37. Pontok koordinátái . . . . .                                          | 42 | 3.42. Lineáris egyenletrendszer sortérbe eső megol-<br>dása . . . . .          | 143 |
| 1.39. Skaláris szorzás koordinátarendszerben . . . . .                      | 43 | 3.44. Egyenletrendszer megoldása elemi bázistransz-<br>formációval . . . . .   | 146 |
| 1.42. Parallelogramma területe . . . . .                                    | 45 | 4.2. Lineáris helyettesítések kompozíciója . . . . .                           | 158 |
| 1.43. Paralelepipedon térfogata . . . . .                                   | 46 | 4.5. Mátrixok összege, különbsége . . . . .                                    | 160 |
| 1.47. . . . .                                                               | 48 | 4.8. Mátrixok lineáris kombinációja . . . . .                                  | 160 |
| 1.53. Vektorok szöge és távolsága . . . . .                                 | 52 | 4.10. Mátrixok szorzása . . . . .                                              | 162 |
| 1.59. Lineáris kombináció $\mathbb{Z}_m^n$ -ben . . . . .                   | 56 | 4.12. Műveletek blokkmátrixokkal . . . . .                                     | 163 |
| 1.60. One time pad – a tökéletes titkosítás . . . . .                       | 57 | 4.19. Skaláris és diadikus szorzat . . . . .                                   | 170 |
| 1.61. Paritásbit . . . . .                                                  | 58 | 4.20. Szimultán egyenletrendszer mátrixszorzatos<br>alakja . . . . .           | 171 |
| 2.1. Az $x + y = 1$ egyenlet . . . . .                                      | 63 | 4.23. Áttérés standard bázisra . . . . .                                       | 173 |
| 2.2. Az $x^2 + y^2 = 1$ egyenlet . . . . .                                  | 63 | 4.27. Bázisfelbontás . . . . .                                                 | 175 |
| 2.9. Síkbeli egyenes egyenletei . . . . .                                   | 67 | 4.30. Mátrix balról szorzása elemi mátrixszal . . . . .                        | 177 |
| 2.14. Sík egyenletei . . . . .                                              | 70 | 4.34. Nulltér . . . . .                                                        | 180 |
| 2.18. Térbeli egyenes egyenletrendszerei . . . . .                          | 73 | 5.4. Mátrix hatványozása . . . . .                                             | 190 |
| 2.19. Egyenes és sík explicit vektoregyenlete . . . . .                     | 75 | 5.5. Polinom helyettesítési értéke . . . . .                                   | 191 |
| 2.25. Mátrixok és elemeik . . . . .                                         | 81 | 5.8. $\mathbf{I} - \mathbf{A}$ inverze nilpotens $\mathbf{A}$ esetén . . . . . | 194 |
| 2.26. Mátrix használata a megoldáshoz . . . . .                             | 82 | 5.12. Az inverz kiszámítása . . . . .                                          | 196 |
| 2.27. Sormodell két kétismeretlenes egyenlettel . . . . .                   | 83 | 5.16. Egyenletrendszer megoldása mátrixinvertálással . . . . .                 | 200 |
| 2.29. Oszlopmodell . . . . .                                                | 86 | 5.17. Mátrixegyenlet megoldása mátrixinvertálással . . . . .                   | 200 |
| 2.33. Lépcsős alak . . . . .                                                | 89 | 5.18. Mátrix elemi mátrixok szorzatára bontása . . . . .                       | 200 |
| 2.34. Gauss-módszer, egy megoldás . . . . .                                 | 89 | 5.22. Az áttérés mátrixának inverze . . . . .                                  | 202 |
| 2.35. Gauss-módszer, végtelen sok megoldás . . . . .                        | 90 | 5.23. Műveletek diagonális mátrixokkal . . . . .                               | 206 |
| 2.37. Homogén lineáris egyenletrendszer megoldása . . . . .                 | 92 |                                                                                |     |
| 2.38. Síkok metszésvonalának meghatározása . . . . .                        | 92 |                                                                                |     |
| 2.40. Redukált lépcsős alak . . . . .                                       | 93 |                                                                                |     |
| 2.41. Redukált lépcsős alakra hozás . . . . .                               | 94 |                                                                                |     |



|                                                                         |     |                                                                       |     |
|-------------------------------------------------------------------------|-----|-----------------------------------------------------------------------|-----|
| 5.25. Sorok permutációja mátrixszorzással . . . . .                     | 206 | 7.81. QR-felbontás kiszámítása . . . . .                              | 336 |
| 5.27. Kígyók . . . . .                                                  | 207 | 7.83. QR-felbontás Givens-forgatásokkal . . . . .                     | 337 |
| 5.32. Szimmetrikus és ferdén szimmetrikus mátrixok . . . . .            | 209 | 7.84. QR-felbontás Householder-tükrözéssel . . . . .                  | 339 |
| 5.37. Inverz változása . . . . .                                        | 211 | 7.86. Egyenletrendszer optimális megoldása . . . . .                  | 340 |
| 5.38. Inverz változása számpéldán . . . . .                             | 211 | 7.92. Önadjungált mátrixok . . . . .                                  | 346 |
| 5.40. Az LU-felbontás kiszámítása . . . . .                             | 216 | 7.98. DFT kiszámítása . . . . .                                       | 352 |
| 5.43. Egyenletrendszer megoldása LU-felbontással . . . . .              | 218 | 7.99. Magas frekvenciájú összetevők szűrése . . . . .                 | 353 |
| 5.44. Mátrix invertálása LU-felbontással . . . . .                      | 219 | 8.1. Jó bázis tükrözéshez . . . . .                                   | 367 |
| 5.46. PLU-felbontás . . . . .                                           | 222 | 8.3. Sajátérték, sajátvektor . . . . .                                | 368 |
| 5.47. . . . .                                                           | 223 | 8.6. Sajátaltér bázisának meghatározása . . . . .                     | 368 |
| 6.4. Zérus értékű determinánsok . . . . .                               | 235 | 8.7. Karakterisztikus polinom felírása . . . . .                      | 370 |
| 6.7. Determináns kiszámítása háromszög alakra hozással . . . . .        | 237 | 8.10. $2 \times 2$ -es mátrixok sajátvektorainak ábrázolása . . . . . | 371 |
| 6.11. Determináns kiszámolása PLU-felbontásból . . . . .                | 240 | 8.12. Az összes sajátérték és sajátvektor meghatározása . . . . .     | 373 |
| 6.13. Determináns kiszámítása elemi oszlopműveletekkel . . . . .        | 241 | 8.13. Magasabbfokú karakterisztikus egyenlet . . . . .                | 374 |
| 6.18. Előjeles aldetermináns . . . . .                                  | 250 | 8.14. Komplex sajátértékek és komplex elemű sajátvektorok . . . . .   | 375 |
| 6.20. Determináns rendjének csökkentése . . . . .                       | 251 | 8.15. Sajátérték algebrai és geometriai multiplicitása . . . . .      | 376 |
| 6.22. Kifejtési tétel . . . . .                                         | 253 | 8.22. Lineáris transzformáció sajátértéke, sajátaltére . . . . .      | 381 |
| 6.23. Interpoláció másodfokú polinomokra . . . . .                      | 253 | 8.26. Mátrix diagonalizálása . . . . .                                | 384 |
| 6.27. Cramer-szabály . . . . .                                          | 256 | 8.27. Sajátfelbontás diadikus alakja és a bal sajátvektorok . . . . . | 385 |
| 6.29. Mátrix inverze . . . . .                                          | 258 | 8.32. Diagonalizálhatóság megállapítása . . . . .                     | 389 |
| 7.1. Vektori szorzással definiált mátrixleképezés . . . . .             | 272 | 8.35. Lineáris transzformáció diagonalizálása . . . . .               | 391 |
| 7.6. A deriválás és az integrálás lineáris leképezés . . . . .          | 274 | 8.37. Spektrálfelbontás . . . . .                                     | 394 |
| 7.10. . . . .                                                           | 276 | 8.41. Gersgorin-körök . . . . .                                       | 397 |
| 7.11. Mátrixleképezés ábrázolása az egységnyezet-rács képével . . . . . | 277 | 8.43. Gersgorin-körök használata . . . . .                            | 398 |
| 7.13. Lineáris leképezés mátrixa másik bázisban . . . . .               | 280 | 8.45. . . . .                                                         | 399 |
| 7.19. Jacobi-mátrix kiszámítása . . . . .                               | 286 | 9.4. . . . .                                                          | 407 |
| 7.20. Függvényérték becslése Jacobi-mátrixszal . . . . .                | 287 | 9.6. Schur-felbontás . . . . .                                        | 410 |
| 7.22. Láncszabály . . . . .                                             | 291 | 9.10. Másodfokú polinom mátrixszorzatos alakja . . . . .              | 414 |
| 7.24. Forgatás egy tetszőleges pont körül . . . . .                     | 293 | 9.13. Főtengely-transzformáció . . . . .                              | 416 |
| 7.25. Koordinátatengely körüli forgatás a térben . . . . .              | 294 | 9.15. Definitiség meghatározása a sajátértékekből . . . . .           | 417 |
| 7.27. Forgatás mátrixa . . . . .                                        | 295 | 10.3. Szinguláris értékek meghatározása . . . . .                     | 425 |
| 7.28. A forgatás mátrixának inverze . . . . .                           | 296 | 10.4. Szinguláris felbontás . . . . .                                 | 427 |
| 7.31. Síkra eső merőleges vetület kiszámítása . . . . .                 | 297 | 10.8. Polárfelbontás kiszámítása . . . . .                            | 431 |
| 7.34. Vetítés síkra . . . . .                                           | 298 | 10.10A pszeudoinverz kiszámítása SVD-ből . . . . .                    | 432 |
| 7.40. Merőleges vetület kiszámítása . . . . .                           | 305 | 11.3. Invariáns altér . . . . .                                       | 448 |
| 7.42. . . . .                                                           | 307 | 11.6. Jordan-lánc és Jordan-bázis konstrukciója . . . . .             | 451 |
| 7.45. . . . .                                                           | 308 | 11.9. Normálalakok . . . . .                                          | 455 |
| 7.49. . . . .                                                           | 312 | 11.11 Jordan-blokkok mérete . . . . .                                 | 458 |
| 7.52. Projekció mátrixa . . . . .                                       | 313 | 11.12 Jordan-blokkok mérete . . . . .                                 | 458 |
| 7.56. Néhány pszeudoinverz . . . . .                                    | 316 | 11.16 Jordan-bázis előállítás . . . . .                               | 463 |
| 7.58. A pszeudoinverz kiszámítása . . . . .                             | 318 | 11.17 Mátrixok hatványai . . . . .                                    | 467 |
| 7.62. Egyenletrendszer optimális megoldása . . . . .                    | 321 | 11.20 Mátrix exponenciális függvénye . . . . .                        | 470 |
| 7.63. Egyenletrendszer optimális megoldása . . . . .                    | 322 | 11.22 Exponenciális függvény Hermite-polinommal . . . . .             | 472 |
| 7.66. Egy pont síkra való merőleges vetülete . . . . .                  | 326 | 11.24 Fourier-mátrix függvényei . . . . .                             | 473 |
| 7.68. Ortogonális mátrixok . . . . .                                    | 326 | 12.8. . . . .                                                         | 481 |
| 7.71. Ortogonális mátrixok inverze . . . . .                            | 328 | 12.11 Primitív mátrixok . . . . .                                     | 484 |
| 7.75. Forgatás tengelye és szöge . . . . .                              | 330 | 12.17 Leontief zárt modell . . . . .                                  | 489 |
| 7.77. Householder-tükrözés . . . . .                                    | 333 | 12.18 Leontief nyílt modell . . . . .                                 | 490 |
| 7.79. Gram-Schmidt-ortogonalizáció . . . . .                            | 334 | 14.3. Lineáris mátrixegyenlet megoldása . . . . .                     | 504 |
|                                                                         |     | 14.5. Sylvester-egyenlet egyértelmű megoldhatósága . . . . .          | 505 |

**I. rész**

# **A lineáris algebra forrásai**





A lineáris algebra két fő forrásának egyike a geometria, másika az algebra vidékéről ered. Mindkét forrás jól jellemezhető egy-egy elemi fogalommal: az egyik a vektor, a másik a lineáris egyenletrendszer. E könyv első része e két fogalmat vizsgálja egészen elemi, középiskolai szintről indulva. A lineáris algebra mélyebb fogalmai már itt fölbukkannak, de csak nagyon egyszerű és a legkevésbé absztrakt formájukban. Az első rész végére látni fogjuk, hogy e két forrás már ezen a bevezető szinten szétválaszthatatlanul egyetlen folyamattá válik.



Hang gliding @ Pule (CC) on flickr by purplematffish



# 1

## Vektorok

Általánosan elterjedt nézet szerint a természeti jelenségek leírásakor sok összefüggést számszerű adatokkal, ún. *skalárokkal* vagy *skalármennyiségekkel* fejezünk ki, míg mások leírásához a számadat mellett egy irány megadása is szükséges; ez utóbbiakat nevezzük *vektoroknak*. A valóság ennél sokkal színesebb: a téridő 4-dimenziós vektoraitól, a bitvektorokon, a gazdasági számításokban használt többszázezer-dimenziós, vagy az internetkeresők által kezelt sokmillió-dimenziós vektorokon át a matematika különböző területein gyümölcsöző absztrakt vektorfogalomig széles a skála.

### Vektorok a 2- és 3-dimenziós térben

---

*E szakaszban a vektor szemléletes, geometriai fogalmával ismerkedünk. A vektorok összeadásán és skalárral való szorzásán keresztül a lineáris kombináció és a lineáris függetlenség fogalmáig jutunk. E szakasz kulcsfogalma: egy vektor lineárisan független vektorok lineáris kombinációjaként való előállítása.*

---

**Irányított szakasz, kötött és szabad vektor** Tekintsünk egy sárkányrepülőt repülés közben. Számtalan skalár- és vektormennyiség írja le állapotát. A földtől való távolság, a légnyomás, a légellenállási együttható vagy az emelkedés szöge skalármennyiségek, míg vektormennyiségek a sebesség- és gyorsulásvektor, a szárnyra ható felhajtóerő, a gravitációs erő, a szél ereje vagy az elmozdulást leíró vektor.

A vektor fogalma kapcsolatban van az irányított szakasz fogalmával. Irányított szakaszon olyan szakaszt értünk, melynek végpontjain megadunk egy sorrendet, azaz kijelöljük, hogy melyik a *kezdő*- és melyik a *végpontja*. Más szóhasználatban az irányított szakaszt szokás *kötött vektornak* is nevezni. Az  $A$  kezdőpontú és  $B$  végpontú irányított szakaszt  $\overrightarrow{AB}$  jelöli.

*Skalár, skaláris: a lépcső, létra jelentésű latin scalae (scālae) szóból ered. E szó származéka a skála szó is, mely jól őrizi az eredeti jelentést. A skalár vagy skaláris szót a matematikában szám vagy számszerű értelemben használjuk, például olyankor, amikor egy mennyiségről azt akarjuk hangsúlyozni, hogy irány nélküli, azaz nem vektor jellegű.*

Több jelenség leírására a kötött vektor alkalmas. Természetes példa az elmozdulásvektor, mely megadja, hogy egy tárgy a tér mely pontjából melyik pontjába jutott. Másik példa kötött vektorra a rugalmas testen alakváltozást okozó erőt leíró vektor (1.1. ábra).

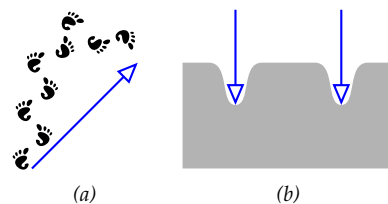
Alkalmazásokban gyakran előfordul, hogy egy jelenség különböző irányított szakaszokkal is ugyanúgy leírható. Például ha egy tárgy mozgását egy olyan irányított szakasszal jellemezzük, melynek hossza az időegység alatt megtett út hosszával egyenlő, iránya pedig a mozgás irányát jelzi, akkor mindegy hogy a tér melyik pontjából indítjuk e szakaszt, a mozgást ugyanúgy leírja (1.2. ábra). Ekkor tehát nem a két pont, hanem azok viszonya a kérdés, azaz hogy az egyik pont a másiktól milyen *távolságra*, és milyen *irányban* van. Az, hogy a két pont pontosan hol van, nem lényeges. Ekkor bármely két irányított szakasz, mely párhuzamosan egymásba tolható, ugyanazt a viszonyt fejezi ki. Az így kapott fogalmat a fizikában *szabad vektornak* nevezik. Ez a lineáris algebra vektor-fogalmának egyik forrása: a *vektor* a geometriában irányított szakasszal reprezentálható azt hozzávéve, hogy két irányított szakasz pontosan akkor reprezentálja ugyanazt a vektort, ha párhuzamosan egymásba tolhatók (ld. 1.3 ábra).

Vektorok jelölésére félkövér kisbetűket használunk, pl.  $\mathbf{x}$ ,  $\mathbf{u}$ ,  $\mathbf{v}$ , stb. A műszaki és fizikai szakirodalomban a félkövér nagy betű is előfordul, pl. az  $\mathbf{F}$  erő, a  $\mathbf{B}$  indukció is vektormennyiségek.

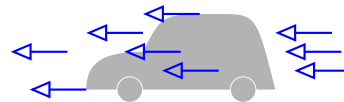
**Vektor megadása egy irányított szakasszal** Egy vektor megadható egy irányított szakasszal, azaz két pont és a köztük lévő sorrend kijelölésével. Valójában ennyi adat felesleges, hisz egy irányított szakasz önmagával párhuzamosan eltolva ugyanazt a vektort adja meg, ezért például kiköthető, hogy a kezdőpont a sík (tér) egy előre kijelölt rögzített pontja legyen. Ezt a közös kezdőpontot nevezzük *origónak*. Egy origóból induló irányított szakaszt egyértelműen definiál a végpontja, így a vektorok megadásához elég egyetlen pont, a végpont megadása. Ezzel a sík vagy tér pontjai és vektorai közt kölcsönösen egyértelmű megfeleltetést létesíthetünk (1.4. ábra). Az origóból egy  $P$  pontba húzott irányított  $\overrightarrow{OP}$  szakaszt a ponthoz tartozó *helyvektornak* is szokás nevezni. Világos, hogy minden vektor reprezentánsai közt pontosan egy helyvektor van.

A későbbiekben gyakran fogunk egy ponthalmazt az origóból a ponthalmaz pontjaiba mutató vektorokkal jellemezni. Amikor vektorok végpontjairól beszélünk, mindig a vektorokat megadó, az origóból indított irányított szakaszok végpontjaira gondolunk.

Az olyan vektort, melynek kezdő és végpontja egybeesik, *zérusvektornak* vagy *nullvektornak* nevezzük. A zérusvektort általában félkövér zérussal, azaz  $\mathbf{0}$ -val jelöljük. A pontok és vektorok közti megfeleltetésben a zérusvektornak az origó felel meg.

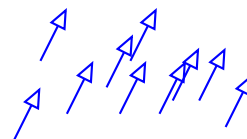


1.1. ábra: Kötött vektorok: (a) elmozdulásvektor (lábnyomokkal), (b) rugalmas testen alakváltozást okozó erő vektora



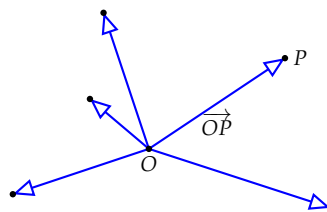
1.2. ábra: Példa szabad vektorra

**VEKTOR:** a *hordozó*, *vivő*, *utazó* jelentésű latin *vector* szóból származik. A tudomány más területein hordozó anyag, az élettanban vírus-hordozó értelemben használják.



1.3. ábra: Ugyanazt a vektort reprezentáló irányított szakaszok

**VEKTOROK JELÖLÉSE:** Műszaki, fizikai szövegek szedésének tipográfiai szabályait az ISO 31-11 szabvány írja le. Eszerint a vektorok félkövér betűkkel szedendők. Kézírásban aláhúzással, vagy fölé írt nyíllal szokás jelezni a vektort (pl.  $\underline{x}$ ,  $\underline{u}$ ,  $\underline{v}$ , ...), de körültekintő jelölésrendszer és jegyzetelés esetén elhagyhatók a jelzések. Felsőbb matematikai művek nem használják e szabványt, mondván, kiderül a szövegből, hogy vektort jelölnek-e a betűk ( $x$ ,  $u$ ,  $v$ , ...).



1.4. ábra: A sík pontjai és vektorai közötti kölcsönösen egyértelmű megfeleltetés: egy  $P$  pontnak az  $\overrightarrow{OP}$  vektor felel meg, az origónak a nullvektor.

**Vektor megadása hossz és irány segítségével** Ha tudunk távolságot mérni és irányt meghatározni, akkor a vektor megadható hosszával és irányával is. A vektor *hosszát*, azaz két végpontjának távolságát, a vektor *abszolút értékének* is nevezzük. Az  $\mathbf{a}$  vektor abszolút értékét  $|\mathbf{a}|$  jelöli. Vektor abszolút értékét a vektor *euklideszi normájának* is nevezik, ugyanis speciális esete egy később részletezendő fogalomnak, a normának. Az  $\mathbf{a}$  vektor (euklideszi) normájának jelölése az abszolút értékre emlékeztet:  $\|\mathbf{a}\|$ .

Az irány fogalmát az 1.38. feladatban definiáljuk. Itt megelégszünk annyival, hogy két nemzérus vektort *azonos irányúnak* vagy *egyirányúnak* nevezünk, ha a kezdőpontjukból induló, és a végpontjukon áthaladó félegyenesek párhuzamos eltolással fedésbe hozhatók (1.5 (a) ábra). Két nemzérus vektort *kollineárisnak* vagy *párhuzamosnak* nevezünk, ha az őket tartalmazó egyenesek párhuzamosak. Két vektort, amely párhuzamos, de nem egyirányú, *ellenkező irányúnak* nevezünk (1.5 (b) ábra). A zérusvektor irányát tetszőlegesnek tekintjük, így az bármely vektorral egyirányú. Belátható, hogy a vektort egyértelműen meghatározza hossza és iránya.

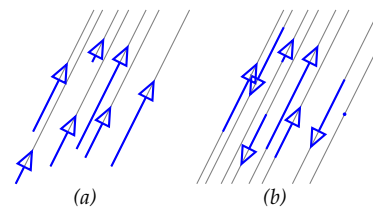
Vektor irányának meghatározásakor gyakran hívjuk segítségül a szög fogalmát. Két vektor szögén azt a szöget értjük, melyet a sík vagy tér egy tetszőleges pontjából kiinduló és az adott vektorokkal egyirányú félegyenesek zárnak be (1.6. ábra). Az  $\mathbf{a}$  és  $\mathbf{b}$  vektorok szögét  $(\mathbf{a}, \mathbf{b})_\angle$  jelöli. Két vektor szöge tehát mindig  $0^\circ$  és  $180^\circ$  – radiánban mérve 0 és  $\pi$  – közé esik, beleértve a határokat is. Egyirányú vektorok szöge 0, ellenkező irányúaké  $\pi$ .

**Vektorműveletek a 2- és 3-dimenziós térben** A vektorműveletek – az összeadás és a számmal való szorzás – definíciója természetes módon adódik, ha a vektorok tipikus alkalmazásaira gondolunk. Pl. magától értetődő, hogy két elmozdulás összegén az elmozgatások egymás után való elvégzését, egy eltolás kétszeresén egy azonos irányú, de kétszer olyan hosszú eltolást értsünk.

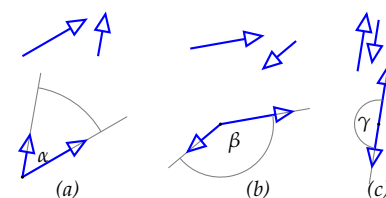
**1.1. DEFINÍCIÓ (KÉT VEKTOR ÖSSZEGE – HÁROMSZÖGMÓDSZER).** Legyen adva két vektor,  $\mathbf{a}$  és  $\mathbf{b}$ . Vegyünk föl egy tetszőleges  $O$  pontot. Indítsunk belőle egy  $\mathbf{a}$ -val egyenlő  $\vec{OP}$  vektort, ennek végpontjából pedig egy  $\mathbf{b}$ -vel egyenlő  $\vec{PQ}$  vektort. Az  $\vec{OQ}$  vektort az  $\mathbf{a}$  és  $\mathbf{b}$  vektorok összegének nevezzük és  $\mathbf{a} + \mathbf{b}$ -vel jelöljük (ld. 1.7. ábra).

Könnyen belátható, hogy az eredmény független az  $O$  pont megválasztásától, tehát vektorok összeadásának művelete definiálható e módszerrel (a bizonyítás leolvasható az 1.8. ábráról).

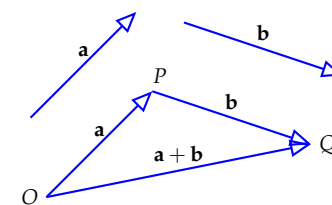
Egy másik módszert is ismertetünk két nem kollineáris vektor összegének megszerkesztésére:



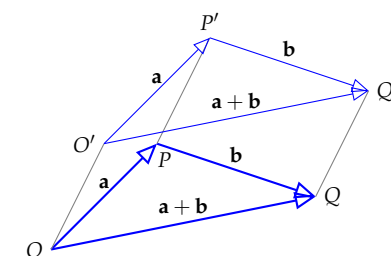
1.5. ábra: (a) egyirányú vektorok, (b) kollineáris (párhuzamos) vektorok, vannak közöttük egyirányúak és ellenkező irányúak



1.6. ábra: Két vektor szöge ( $0 \leq \alpha, \beta, \gamma \leq \pi$ ). Az ábra felső felén a két adott vektor, alatta szögük meghatározásának módja szerepel.



1.7. ábra: Az  $\mathbf{a}$  és  $\mathbf{b}$  vektor összege

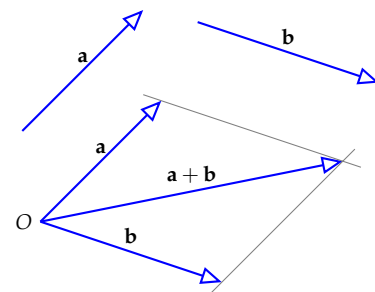
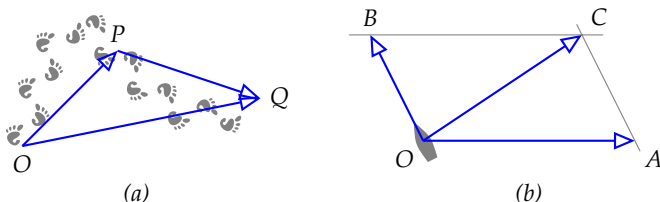


1.8. ábra: Az összeg független az  $O$  pont megválasztásától, ugyanis  $\vec{OQ} = \vec{O'Q'}$ .

**1.2. ÁLLÍTÁS (PARALLELOGRAMMA-MÓDSZER).** A közös kezdőpontból indított  $\mathbf{a}$  és  $\mathbf{b}$  vektorok összege megkapható abból a paralelogrammából, melynek két szomszédos oldala  $\mathbf{a}$  és  $\mathbf{b}$ , ekkor az összeg a közös kezdőpontból indított és a paralelogramma szemközi csúcsába futó vektor.

► Ha  $\mathbf{a}$  és  $\mathbf{b}$  nem kollineárisak, akkor összegük pl. megkapható úgy, hogy  $\mathbf{a}$  végpontján át egy  $\mathbf{b}$  egyenesével,  $\mathbf{b}$  végpontján át egy  $\mathbf{a}$  egyenesével párhuzamos egyenest húzunk. A közös kezdőpontból e két egyenes metszéspontjába futó vektor lesz az összeg (ld. 1.9. ábra).

Az alkalmazásokban hol a háromszög-, hol a paralelogramma-módszer tűnik kézenfekvőbbnek (ld. 1.10).



1.9. ábra: Parallelogramma-módszer

1.10. ábra: Az (a) ábrán a lábnyomok  $O$ -ból  $P$ -be, majd onnan  $Q$ -ba vezetnek. Az  $\overrightarrow{OP}$  és a  $\overrightarrow{PQ}$  elmozdulásvektorok összege  $\overrightarrow{OQ}$  (háromszögmódszer). A (b) ábrán a csónak az  $\overrightarrow{OB}$  irányba evez, de a folyó  $\overrightarrow{OA}$  irányba folyik. A két sebesség eredője, azaz összege  $\overrightarrow{OC}$  (parallelogramma módszer).

Ha  $\mathbf{a}$  és  $\mathbf{b}$  két térbeli vektor, akkor a háromszögmódszerben és a paralelogramma-módszerben is az  $\mathbf{a}$ ,  $\mathbf{b}$  és  $\mathbf{a} + \mathbf{b}$  vektorokat reprezentáló irányított szakaszok egy síkba esnek. Általában azt mondjuk, hogy néhány térbeli vektor egy síkba esik, más szóval *komplanáris*, ha van olyan sík, hogy mindegyik vektort reprezentáló irányított szakasz párhuzamosan betolható e síkba. Eszerint tehát az  $\mathbf{a}$ ,  $\mathbf{b}$  és  $\mathbf{a} + \mathbf{b}$  vektorok mindig komplanárisak.

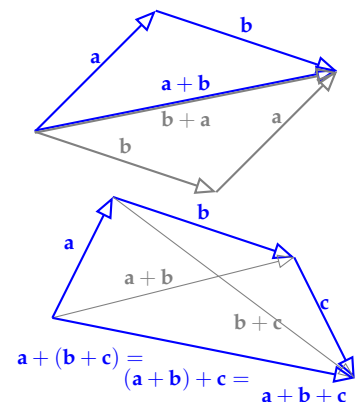
A vektorösszeadás két fontos tulajdonsága a kommutativitás ( $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ ) és az asszociativitás ( $\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}$ ) könnyen leolvasható az 1.11. ábráról. Az asszociativitás következtében több tag összeadásánál elhagyható a zárójel, például az ábrabeli három vektor összegére  $\mathbf{a} + \mathbf{b} + \mathbf{c}$  írható.

Az  $\mathbf{a}$  és  $\mathbf{b}$  vektorokat közös kezdőpontból indítva – a háromszögmódszerrel – azonnal látható, hogy csak egyetlen olyan  $\mathbf{x}$  vektor létezik, melyre  $\mathbf{a} = \mathbf{b} + \mathbf{x}$  (ld. 1.12 (a) ábra). Ennek felhasználásával definiálható vektorok különbsége.

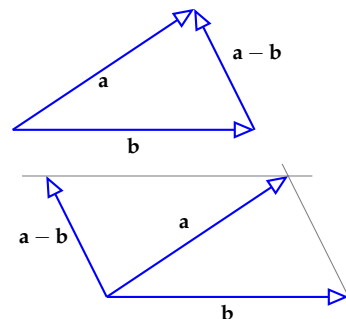
**1.3. DEFINÍCIÓ (VEKTOROK KÜLÖNBSÉGE).** Adva van az  $\mathbf{a}$  és  $\mathbf{b}$  vektor. Azt az egyértelműen létező  $\mathbf{x}$  vektort, melyre  $\mathbf{a} = \mathbf{b} + \mathbf{x}$ , az  $\mathbf{a}$  és  $\mathbf{b}$  különbségének nevezzük és  $\mathbf{a} - \mathbf{b}$ -vel jelöljük.

Könnyen fejben tartható a különbségvektor megszerkesztése akár a háromszög-, akár a paralelogrammamódszerrel (ld. 1.12. ábra), ha a definícióra gondolunk, azaz arra, hogy  $\mathbf{a} - \mathbf{b}$  az  $\mathbf{a}$  vektor, melyet  $\mathbf{b}$ -hez adva  $\mathbf{a}$ -t kapunk, azaz

$$\mathbf{a} = \mathbf{b} + (\mathbf{a} - \mathbf{b}).$$



1.11. ábra: A vektorösszeadás kommutativitása és asszociativitása.



1.12. ábra: A különbségvektor meghatározása háromszög- és paralelogrammamódszerrel.

Az 1.13. ábráról az is leolvasható, hogy ha a  $\mathbf{b}$  vektorral egyenlő hosszúságú, de ellenkező irányú vektort  $-\mathbf{b}$  jelöli, akkor fennáll az  $\mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b})$  összefüggés, és így az is igaz, hogy  $\mathbf{b} + (-\mathbf{b}) = \mathbf{0}$ .

Érdekes megjegyezni, hogy ha  $P$  és  $Q$  két tetszőleges pont, akkor az  $\overrightarrow{OQ} - \overrightarrow{OP}$  vektort akkor is ismerjük, ha az  $O$  pontot nem, hisz az a  $\overrightarrow{PQ}$  vektor. Sok hasonló jelenség vezetett a *torzor* fogalmához, melyet egy rövid széljegyzetben ismertetünk.

**1.4. DEFINÍCIÓ (VEKTOR SZORZÁSA SKALÁRRAL).** Legyen  $k$  valós szám. Az  $\mathbf{a}$  vektor  $k$ -szorosán azt a vektort értjük, melynek hossza az  $\mathbf{a}$  hosszának  $|k|$ -szorosa, iránya

- tetszőleges, ha  $k = 0$  vagy  $\mathbf{a} = \mathbf{0}$ ,
- megegyezik  $\mathbf{a}$  irányával, ha  $k > 0$ , és
- ellentétes, ha  $k < 0$  (ld. 1.14. ábra).

A skalárral való szorzás definíciójából azonnal látszik, hogy minden  $\mathbf{a}$  vektorra  $1\mathbf{a} = \mathbf{a}$ ,  $0\mathbf{a} = \mathbf{0}$  és  $(-1)\mathbf{a} = -\mathbf{a}$ .

E paragrafus végén összefoglaljuk a vektorműveletek legfontosabb tulajdonságait, melyek segítségével később általánosítani fogjuk a vektor fogalmát. Az eddig nem bizonyított tulajdonságok igazolását az olvasóra hagyjuk.

**1.5. TÉTEL (A VEKTORMŰVELETEK TULAJDONSÁGAI).** Ha  $\mathbf{a}$ ,  $\mathbf{b}$  és  $\mathbf{c}$  a 2- vagy 3-dimenziós tér tetszőleges vektorai,  $\mathbf{0}$  a zérusvektor és  $r$ ,  $s$  két tetszőleges valós szám, akkor fennállnak az alábbi azonosságok:

- |                                                                                      |                                                             |
|--------------------------------------------------------------------------------------|-------------------------------------------------------------|
| a) $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$                               | e) $r(s\mathbf{a}) = (rs)\mathbf{a}$                        |
| b) $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$ | f) $r(\mathbf{a} + \mathbf{b}) = r\mathbf{a} + r\mathbf{b}$ |
| c) $\mathbf{a} + \mathbf{0} = \mathbf{a}$                                            | g) $(r + s)\mathbf{a} = r\mathbf{a} + s\mathbf{a}$          |
| d) $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$                                         | h) $1\mathbf{a} = \mathbf{a}$ és $0\mathbf{a} = \mathbf{0}$ |

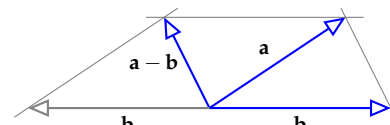
**A lineáris kombináció definíciója** Ha vektorokra a skalárral való szorzás és az összeadás műveletét alkalmazzuk, akkor e vektorok egy lineáris kombinációját kapjuk. Pontosabban:

**1.6. DEFINÍCIÓ (LINEÁRIS KOMBINÁCIÓ).** Az  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k$  vektorok lineáris kombinációján egy

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \dots + c_k\mathbf{a}_k$$

alakú vektort értünk, ahol  $c_1, c_2, \dots, c_k$  valós számok. Azt mondjuk, hogy a  $\mathbf{v}$  vektor előáll az  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k$  vektorok lineáris kombinációjaként, ha vannak olyan  $c_1, c_2, \dots, c_k$  valós számok, hogy  $\mathbf{v} = c_1\mathbf{a}_1 + \dots + c_k\mathbf{a}_k$ .

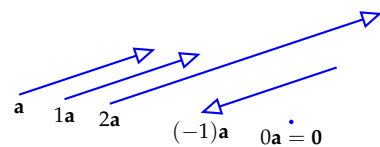
Ha egy vektort egy skalárral beszorzunk, az előző definíció szerint egy lineáris kombinációját kapjuk, mely vele párhuzamos, azaz kollineáris. Így egy nemzérus vektor összes lineáris kombinációja csupa vele párhuzamos vektor (ld. 1.15. ábrát). Ennél több is igaz:



1.13. ábra: Az  $\mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b})$  szemléltetése.

**TORZOR:** a modern matematika fogalma. Néhány példa, mielőtt definiálnánk: (1) Az energiát a newtoni fizikában nem tudjuk mérni, csak az energiakülönbséget. Ha viszont megállapodunk abban, hogy egy adott rendszernek melyik állapota tartozik a 0 energiaszinthez, beszélhetünk a rendszer energiájáról is. (2) A pontba mutató vektor fogalmának nincs értelme, amíg nincs kijelölve az origó, viszont két pontba mutató vektor különbségét az origótól függetlenül is meg tudjuk határozni. (3) Egy  $f$  függvény  $I$  intervallumon vett határozatlan integrálja  $F + C$  alakú, ahol  $C$  konstans. Nincs értelme megkérdezni, hogy  $f$  egy konkrét primitív függvényében mennyi a  $C$  értéke, de két primitív függvény különbsége mindig egy konstans. (4) Egy hasonló jelenség a zenéből: bármely két hang közti távolság meghatározható, de azt nem mondhatjuk egy hangra, hogy az a „fá”, amíg nem rögzítjük, melyik a „dó”.

A torzort egy *kommutatív csoport* nevű algebrai struktúrával definiálhatjuk, mely egy kommutatív, asszociatív, null-elemes, invertálható művelettel ellátott és e műveletre zárt halmaz. Kommutatív csoport például a valósok az összeadásra nézve, a vektorok az összeadásra nézve, vagy  $\mathbb{Z}_{12}$  az összeadásra nézve (ld. a ?? példát és a ?? definíciót). Legyen  $G$  egy kommutatív csoport, és  $X$  egy nem üres halmaz, melyen definiálva van bármely két elem különbsége, ami  $G$ -beli. Ekkor  $X$ -et  *$G$ -torzornak* nevezzük, ha bármely  $x_0, x_1, x_2 \in X$  elem esetén, ha  $x_1 - x_0 = g_1$  és  $x_2 - x_0 = g_2$ , akkor  $x_1 - x_2 = g_1 - g_2$ . Más-ként fogalmazva,  $X$  örzi  $G$  struktúráját a zéruselem nélkül úgy, hogy bármely elemét zéruselemnek választva azonnal megkapjuk  $G$ -t.



1.14. ábra: Vektor skalárszorosai



**1.7. TÉTEL (VEKTORRAL PÁRHUZAMOS VEKTOROK).** Ha  $\mathbf{a}$  nem zérusvektor, akkor bármely vele párhuzamos  $\mathbf{v}$  vektor az  $\mathbf{a}$  skalárszorosa, azaz van olyan  $c$  valós szám, hogy  $\mathbf{v} = c\mathbf{a}$ , más szóval  $\mathbf{v}$  előáll az  $\mathbf{a}$  valamely lineáris kombinációjaként. Ez az előállítás egyértelmű.

**BIZONYÍTÁS.** Ha a két vektor egyirányú, az előállításban szereplő  $c$  konstans egyszerűen a  $\mathbf{v}$  és  $\mathbf{a}$  vektorok abszolút értékének hányadosa, ha ellenkező irányúak, e hányados  $(-1)$ -szerese.  $\square$

E tétel következménye, hogy ha  $\mathbf{a}$  nem zérusvektor, akkor az  $\mathbf{a}$  összes lineáris kombinációjának halmaza és az  $\mathbf{a}$ -val párhuzamos vektorok halmaza megegyezik. Másként fogalmazva: egy nemzérus vektor összes lineáris kombinációjának végpontja egy origón átmenő egyenest ad.

A háromszögmódszerből jól látszik, hogy tetszőleges két vektor bármely lineáris kombinációja velük komplanáris vektor lesz. Az állítás megfordítása is igaz:

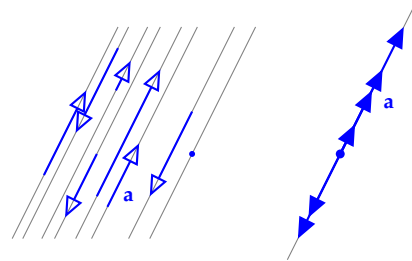
**1.8. TÉTEL (KÉT VEKTORRAL EGY SÍKBA ESŐ VEKTOROK).** Ha  $\mathbf{a}_1$  és  $\mathbf{a}_2$  nem párhuzamos vektorok, akkor bármely velük egy síkba eső  $\mathbf{v}$  vektor előáll az  $\mathbf{a}_1$  és  $\mathbf{a}_2$  valamely lineáris kombinációjaként, azaz van olyan  $v_1$  és  $v_2$  konstans, hogy  $\mathbf{v} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2$ . Ez az előállítás egyértelmű.

**BIZONYÍTÁS.** A bizonyításnak a felbontás létezését biztosító része könnyen leolvasható az 1.16. ábráról. A  $\mathbf{v}$  végpontjából húzzunk az  $\mathbf{a}_1$  és az  $\mathbf{a}_2$  vektorokkal párhuzamos egyeneseket. Az így létrejött – esetleg elfajuló – paralelogramma két oldala az előző tétel szerint  $\mathbf{a}_1$ , illetve  $\mathbf{a}_2$  konstansszorosa, melyek összege a paralelogramma szabály szerint épp  $\mathbf{v}$ . Előállítottuk tehát  $\mathbf{v}$ -t  $\mathbf{a}_1$  és  $\mathbf{a}_2$  lineáris kombinációjaként. Meg kell még mutatnunk, hogy ez az előállítás egyértelmű. Legyen

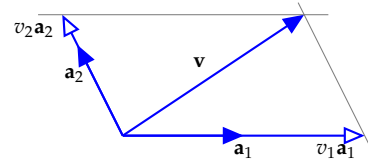
$$\mathbf{v} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2 = w_1\mathbf{a}_1 + w_2\mathbf{a}_2.$$

a  $\mathbf{v}$  vektor két előállítása. Ekkor átrendezés után  $(v_1 - w_1)\mathbf{a}_1 = (w_2 - v_2)\mathbf{a}_2$ . Mivel  $\mathbf{a}_1$  és  $\mathbf{a}_2$  nem kollineárisak, konstansszorosaik csak akkor egyezhetnek meg, ha mindkettő a zérusvektor. Ugyanakkor  $\mathbf{a}_1 \neq \mathbf{0}$  és  $\mathbf{a}_2 \neq \mathbf{0}$ , ezért az előző egyenlőség csak akkor áll fenn, ha  $(v_1 - w_1) = (w_2 - v_2) = 0$ , azaz ha  $v_1 = w_1$  és  $v_2 = w_2$ . Tehát a felbontás egyértelmű.  $\square$

Látható tehát, hogy két nem párhuzamos vektor összes lineáris kombinációjának halmaza megegyezik a két vektorral komplanáris vektorok halmazával, egyszerűbben fogalmazva: két nem párhuzamos vektor összes lineáris kombinációjának végpontja egy origón átmenő síkot ad.



1.15. ábra: Egy nemzérus  $\mathbf{a}$  vektor, és néhány lineáris kombinációja kétféle reprezentációban.



1.16. ábra: A  $\mathbf{v}$  egyértelműen előáll  $\mathbf{v} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2$  alakban, ha  $\mathbf{a}_1$  és  $\mathbf{a}_2$  nem párhuzamos.



Abban nincs semmi meglepő, hogy a tér három nem egy síkba eső vektorának bármely lineáris kombinációja térbeli vektor, az állítás megfordítása viszont igen fontos:

**1.9. TÉTEL (TÉRBELI VEKTOROK).** Ha  $\mathbf{a}_1$ ,  $\mathbf{a}_2$  és  $\mathbf{a}_3$  nem egy síkba eső vektorok, akkor a tér bármely  $\mathbf{v}$  vektora előáll az  $\mathbf{a}_1$ ,  $\mathbf{a}_2$  és  $\mathbf{a}_3$  valamely lineáris kombinációjaként, azaz van olyan  $v_1$ ,  $v_2$  és  $v_3$  konstans, hogy

$$\mathbf{v} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3. \quad (1.1)$$

Ez az előállítás egyértelmű.

**BIZONYÍTÁS.** A  $\mathbf{v}$  vektor  $V$  végpontján át párhuzamos egyenest húzunk az  $\mathbf{a}_3$  vektorral, mely az  $\mathbf{a}_1$  és  $\mathbf{a}_2$  vektorok síkját egy  $C$  pontban metszi (1.17. (a) ábra). Az  $\overrightarrow{OC}$  vektor az előző tétel szerint egyértelműen előáll  $\mathbf{a}_1$  és  $\mathbf{a}_2$  lineáris kombinációjaként, azaz  $\overrightarrow{OC} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2$  (ld. 1.17. (b) ábra). Másrészt  $\mathbf{v} = \overrightarrow{OV} = \overrightarrow{OC} + \overrightarrow{CV}$ , ahol  $\overrightarrow{CV} \parallel \mathbf{a}_3$ , így  $\overrightarrow{CV} = v_3\mathbf{a}_3$  valamely  $v_3$  valósra. Tehát  $\mathbf{v} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3$ .

Be kell még látnunk az előállítás egyértelműségét! Tegyük fel, hogy

$$\mathbf{v} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3 = w_1\mathbf{a}_1 + w_2\mathbf{a}_2 + w_3\mathbf{a}_3$$

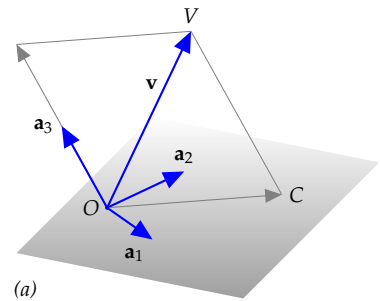
a  $\mathbf{v}$  két felbontása. Ekkor  $(v_1 - w_1)\mathbf{a}_1 + (v_2 - w_2)\mathbf{a}_2 + (v_3 - w_3)\mathbf{a}_3 = \mathbf{0}$ . Így ha  $v_1 \neq w_1$ , akkor  $\mathbf{a}_1$  kifejezhető  $\mathbf{a}_2$  és  $\mathbf{a}_3$  lineáris kombinációjaként:

$$\mathbf{a}_1 = -\frac{v_2 - w_2}{v_1 - w_1}\mathbf{a}_2 - \frac{v_3 - w_3}{v_1 - w_1}\mathbf{a}_3.$$

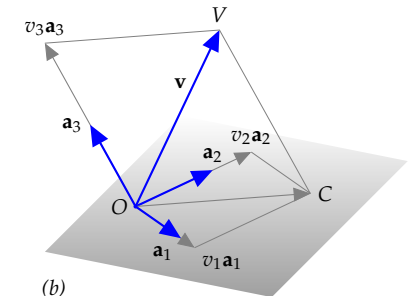
Ez ellentmond annak, hogy  $\mathbf{a}_1$ ,  $\mathbf{a}_2$  és  $\mathbf{a}_3$  nem esnek egy síkba. Így tehát  $v_1 = w_1$ . Hasonlóan kapjuk, hogy  $v_2 = w_2$  és  $v_3 = w_3$ , azaz az (1.1) előállítás egyértelmű.  $\square$

**Lineáris függetlenség** Az előző két tételből világos, hogy a tér három vektora vagy egy síkba esik, ekkor valamelyikük a másik kettő lineáris kombinációja, vagy nem esik egy síkba, és akkor egyikük sem áll elő a másik kettő lineáris kombinációjaként. Ekkor viszont a tér minden vektora előáll az ő lineáris kombinációjuként. Látjuk, alapvető, hogy egy vektor kifejezhető-e más vektorok lineáris kombinációjaként.

**1.10. DEFINÍCIÓ (VEKTOROK FÜGGETLENSÉGE).** Azt mondjuk, hogy egy  $\mathbf{v}$  vektor lineárisan független az  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$  ( $n \geq 1$ ) vektoroktól, ha  $\mathbf{v}$  nem fejezhető ki e vektorok lineáris kombinációjaként. Azt mondjuk, hogy az  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$  ( $n \geq 2$ ) vektorok lineárisan függetlenek ha e vektorok egyike sem fejezhető ki a többi lineáris kombinációjaként. Ha legalább egyikük kifejezhető a többi lineáris kombinációjaként, azaz legalább egyikük lineárisan függ a többitől, akkor e vektorokat lineárisan összefüggőknek nevezzük. Az egyetlen vektorból álló vektorrendszert lineárisan függetlennek tekintjük, ha a vektor nem a zérusvektor.



(a)



(b)

1.17. ábra: A térbeli  $\mathbf{v}$  vektor előállítása három nem egy síkba eső vektor lineáris kombinációjaként.

Például egy térbeli vektor, mely nem esik egy adott síkba, független a síkba eső vektorok bármely rendszerétől (1.18. ábra).

Egy kocka egy csúcsból kiinduló élvektorai lineárisan függetlenek (1.19. ábra).

Általában: bármely két nem kollineáris vektor lineárisan független, hasonlóképp, a tér bármely három nem komplanáris, azaz nem egy síkba eső vektora lineárisan független.

Az 1.8. tétel tehát a következőképp fogalmazható át:

**1.11. TÉTEL (SÍKBELI VEKTOR FELBONTÁSA).** Ha  $\mathbf{a}_1$  és  $\mathbf{a}_2$  egy sík két lineárisan független vektora, akkor a sík minden  $\mathbf{v}$  vektora egyértelműen előáll  $\mathbf{v}$  vektorok lineáris kombinációjaként, azaz egyértelműen léteznek olyan  $v_1$  és  $v_2$  valós számok, hogy

$$\mathbf{v} = v_1 \mathbf{a}_1 + v_2 \mathbf{a}_2.$$

Hasonlóképp az 1.9. tétel így fogalmazható át:

**1.12. TÉTEL (TÉRBELI VEKTOR FELBONTÁSA).** Ha  $\mathbf{a}_1$ ,  $\mathbf{a}_2$  és  $\mathbf{a}_3$  három lineárisan független térbeli vektor, akkor a tér minden  $\mathbf{v}$  vektora egyértelműen előáll  $\mathbf{v}$  vektorok lineáris kombinációjaként, azaz egyértelműen léteznek olyan  $v_1$ ,  $v_2$  és  $v_3$  valós számok, hogy

$$\mathbf{v} = v_1 \mathbf{a}_1 + v_2 \mathbf{a}_2 + v_3 \mathbf{a}_3.$$

A koordinátákról szóló szakaszban e két tétel lesz alapja a koordinátarendszer bevezetésének.

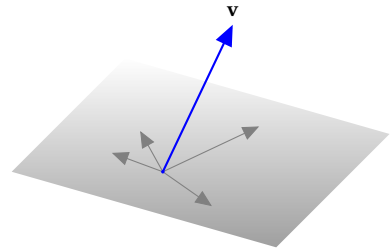
**Speciális lineáris kombinációk\*** A sík és a tér bizonyos konfigurációi jól jellemezhetők lineáris kombinációkkal, ha a kombinációs együtthatókra bizonyos feltételeket kötünk ki.

**1.13. ÁLLÍTÁS (KÉT PONTON ÁTMENŐ EGYENES JELLEMZÉSE).** Legyen  $O$ ,  $A$  és  $B$  a sík vagy a tér három pontja. Az  $r\vec{OA} + s\vec{OB}$  alakú lineáris kombináció végpontja pontosan akkor mutat az  $A$  és  $B$  ponton átmenő egyenes egy pontjába, ha  $r + s = 1$ .

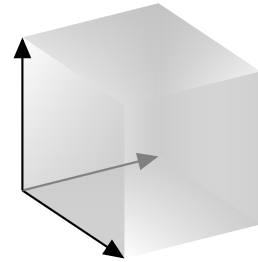
**BIZONYÍTÁS.** Legyen  $\mathbf{a} = \vec{OA}$ ,  $\mathbf{b} = \vec{OB}$ , és  $\mathbf{x}$  mutasson az  $AB$  egyenes valamely  $X$  pontjára, azaz legyen  $\mathbf{x} = \vec{OB} + r\vec{BA}$  valamilyen  $r$  valós számra, tehát

$$\mathbf{x} = \mathbf{b} + r(\mathbf{a} - \mathbf{b}), \text{ azaz } \mathbf{x} = r\mathbf{a} + (1 - r)\mathbf{b}.$$

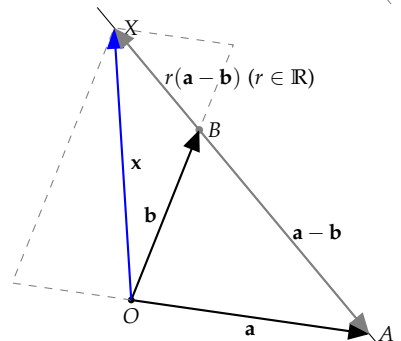
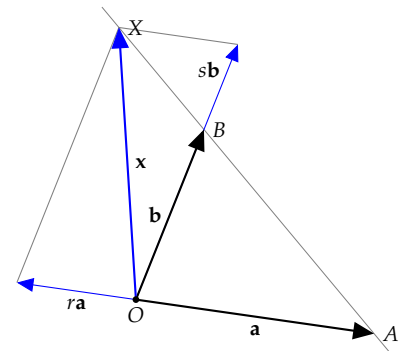
A fenti gondolatmenet lépésein visszafelé haladva látható, hogy minden valós  $r$  számra az  $r\mathbf{a} + (1 - r)\mathbf{b}$  vektor végpontja az  $AB$  egyenesen van. Fogalmazhatunk úgy is, hogy az  $\mathbf{a}$  és  $\mathbf{b}$  vektorok végpontján átmenő egyenes összes pontját pontosan azok az  $r\mathbf{a} + s\mathbf{b}$  alakú lineáris kombinációk adják, amelyeknél  $r + s = 1$  (ld. ?? ábra).  $\square$



1.18. ábra: A síkba nem eső  $\mathbf{v}$  vektor nem áll elő a síkbeli vektorok lineáris kombinációjaként.



1.19. ábra: Egy kocka három egy csúcsból induló élvektora lineárisan független.

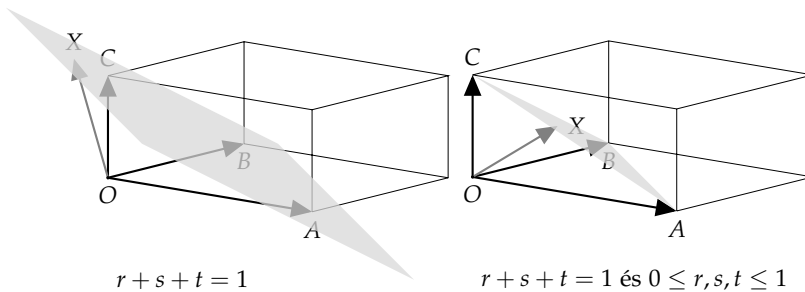


1.20. ábra: Az  $X$  pont pontosan akkor van az  $AB$  egyenesen, ha azon  $r$  és  $s$  valósokra, melyekre  $\vec{OX} = r\vec{OA} + s\vec{OB}$ ,  $r + s = 1$  teljesül. Ezen az ábrán  $r = -0.5$ ,  $s = 1.5$ .

**1.14. ÁLLÍTÁS (INTERVALLUM PONTJAINAK JELLEMZÉSE).** Legyen  $O$ ,  $A$  és  $B$  a sík vagy a tér három pontja. Az  $r\vec{OA} + s\vec{OB}$  vektor pontosan akkor mutat az  $A$  és  $B$  pontot összekötő szakasz valamely pontjába, ha  $r + s = 1$  és  $0 \leq r, s \leq 1$ .

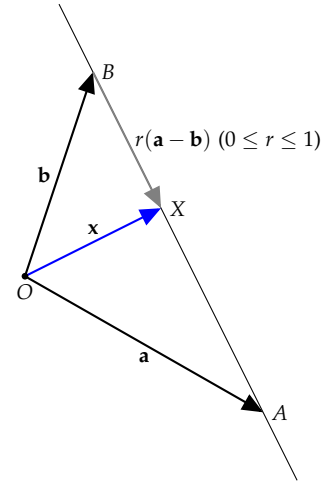
**BIZONYÍTÁS.** Megismételjük az előző feladat megoldását a különbséggel, hogy itt a  $\vec{BX} = r\vec{BA}$  összefüggés csak 0 és 1 közé eső  $r$  értékekre igaz. Tehát  $\mathbf{x} = r\mathbf{a} + (1 - r)\mathbf{b}$ , ahol  $0 \leq r \leq 1$ . Másként fogalmazva az  $\mathbf{a}$  és  $\mathbf{b}$  vektorok végpontjait összekötő szakasz összes pontját pontosan azok az  $r\mathbf{a} + s\mathbf{b}$  alakú lineáris kombinációk adják, amelyekben  $r + s = 1$  és  $0 \leq r, s \leq 1$  (ld. 1.21 ábra).  $\square$

Hasonló összefüggés igaz három vektor esetén is, azaz megmutatható, hogy a nem kollineáris  $\mathbf{a}$ ,  $\mathbf{b}$  és  $\mathbf{c}$  vektorok végpontjaira fektetett sík pontjaiba pontosan azok a vektorok mutatnak, melyeket  $r\mathbf{a} + s\mathbf{b} + t\mathbf{c}$  alakba írva  $r + s + t = 1$ . Ha még azt is kikötjük e három számról, hogy legyen  $0 \leq r, s, t \leq 1$ , akkor az  $r\mathbf{a} + s\mathbf{b} + t\mathbf{c}$  alakú vektorok a három vektor végpontja által meghatározott háromszög pontjaiba mutatnak (ld. az 1.22. ábrát és a ?? feladatot).



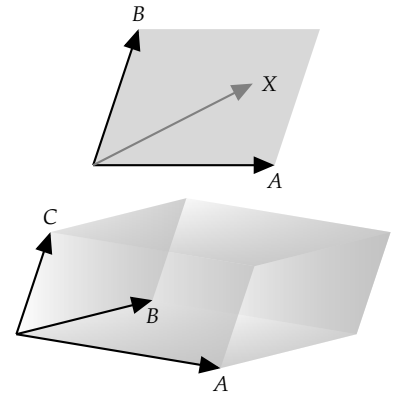
Szemléletesen világos, például a mellékelt 1.23. ábráról leolvasható, de nem bizonyítjuk, hogy két tetszőleges nem kollineáris vektor összes olyan lineáris kombinációja, amelyben az együtthatók 0 és 1 közé esnek, egy parallelogrammát ad. Pontosabban fogalmazva egy  $r\mathbf{a} + s\mathbf{b}$  alakú vektor végpontja pontosan akkor tartozik az  $\mathbf{a}$  és  $\mathbf{b}$  által meghatározott (kifeszített) parallelogrammához, ha  $0 \leq r, s \leq 1$ .

Hasonló mondható három, nem egy síkba eső vektorról: egy  $r\mathbf{a} + s\mathbf{b} + t\mathbf{c}$  alakú vektor végpontja pontosan akkor tartozik az  $\mathbf{a}$ ,  $\mathbf{b}$  és  $\mathbf{c}$  által kifeszített paralelepipedonhoz, ha  $0 \leq r, s, t \leq 1$  (1.23. ábra).



1.21. ábra: Az  $X$  pont pontosan akkor van az  $\overline{AB}$  intervallumban, ha valamely 0 és 1 közé eső  $r$  és  $s$  valósokra  $\vec{OX} = r\vec{OA} + s\vec{OB}$ , és  $r + s = 1$ .

1.22. ábra: Az  $X$  pont pontosan akkor esik az  $A$ ,  $B$  és  $C$  pontokon átmenő síkba, ha  $\vec{OX} = r\vec{OA} + s\vec{OB} + t\vec{OC}$  és  $r + s + t = 1$ . Az  $X$  az  $ABC$  háromszögbe pedig pontosan akkor esik, ha ezen kívül még  $0 \leq r, s, t \leq 1$  is fennáll.

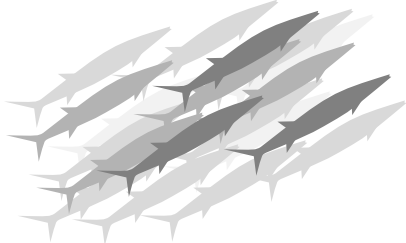


1.23. ábra: A parallelogramma és a paralelepipedon olyan lineáris kombinációkkal állítható elő, ahol az együtthatók 0 és 1 közé esnek.

## Feladatok

### Vektor

1.1. Egy matematikán kívüli szemléltetés a vektor fogalmához: hogyan fejeznénk be az alábbi hasonlatot? „Ha az irányított szakasz a hal, akkor a vektor a...”



1.2. **VEKTOROK: IGAZ – HAMIS** Melyek igazak, melyek hamisak az alábbi állítások közül?

- Ha az  $\mathbf{a}$  és  $\mathbf{b}$  vektorok hajlásszöge  $\alpha$ , akkor  $\mathbf{a}$  és  $-\mathbf{b}$  hajlásszöge  $\pi - \alpha$ .
- Ha  $A$  és  $B$  két adott pont, akkor az  $\overrightarrow{OA} + \overrightarrow{OB}$  vektor független az  $O$  megválasztásától.
- Ha  $A$  és  $B$  két adott pont, akkor az  $\overrightarrow{OA} - \overrightarrow{OB}$  vektor független az  $O$  megválasztásától.
- Ha két vektor egyirányú, akkor egyikük a másik skalárszorosa.
- Ha két vektor egyike a másik skalárszorosa, akkor egyirányúak.
- Ha két vektor egyike a másik skalárszorosa, akkor párhuzamosak.

### Vektorműveletek a 2- és 3-dimenziós térben

1.3. Adva van a síkban két tetszőleges vektor,  $\mathbf{a}$  és  $\mathbf{b}$ . Szerkesszük meg a következő vektorokat: a)  $\mathbf{c} = 2\mathbf{a} + \mathbf{b}$ , b)  $\mathbf{d} = 2\mathbf{a} - \mathbf{b}$ , c)  $\mathbf{e} = \frac{2}{3}\mathbf{a} + \frac{1}{3}\mathbf{b}$ , d)  $\mathbf{f} = \frac{2}{5}\mathbf{a} + \frac{3}{5}\mathbf{b}$ .

1.4. Legyen  $\mathbf{u} = \mathbf{a} + \mathbf{b}$ ,  $\mathbf{v} = \mathbf{a} - \mathbf{b}$ . Fejezzük ki az  $\mathbf{a}$  és  $\mathbf{b}$  vektor segítségével a következő vektorokat: a)  $2\mathbf{u} + 2\mathbf{v}$ , b)  $3\mathbf{u} - 3\mathbf{v}$ , c)  $3\mathbf{u} - \mathbf{v}$ , d)  $2\mathbf{u} - \frac{1}{2}\mathbf{v}$ .

1.5. Tekintsük az  $ABCD$  négyzetet. Határozzuk meg a következő összegeket! a)  $\overrightarrow{AB} + \overrightarrow{CD}$ , b)  $\overrightarrow{AB} + \overrightarrow{BC} + \overrightarrow{CD}$ , c)  $\overrightarrow{AB} - \overrightarrow{AC}$ , d)  $\overrightarrow{AB} - \overrightarrow{CB}$ , e)  $\overrightarrow{AC} + \overrightarrow{DB}$ , f)  $\overrightarrow{AC} - \overrightarrow{DB}$ , g)  $2\overrightarrow{AB} + \overrightarrow{BD}$

1.6. Tekintsük az  $ABCD$  négyzetet. Jelölje a  $BC$  oldal felezőpontját  $E$ , a  $CD$  oldal felezőpontját  $F$ , a négyzet középpontját  $O$ . Fejezzük ki az egymásra merőleges  $\mathbf{b} = \overrightarrow{AB}$  és  $\mathbf{d} = \overrightarrow{AD}$  vektorok segítségével az  $\overrightarrow{AE}$ ,  $\overrightarrow{AF}$ ,  $\overrightarrow{AO}$ ,  $\overrightarrow{EF}$ ,  $\overrightarrow{OF}$  vektorokat!

1.7. Tekintsük az  $ABCD$  tetraédert! Határozzuk meg az

- $\overrightarrow{AB} + \overrightarrow{BC} + \overrightarrow{CD} + \overrightarrow{DA}$ ,
- $\overrightarrow{AB} - \overrightarrow{CB} + \overrightarrow{CD} - \overrightarrow{AD}$ ,
- $\overrightarrow{AD} - \overrightarrow{AC} - \overrightarrow{BD}$

vektorokat.

1.8. Tekintsük a szabályos  $ABCDEF$  hatszöget, melynek geometriai középpontját jelölje  $O$ . Fejezzük ki az  $\mathbf{a} = \overrightarrow{OA}$  és  $\mathbf{b} = \overrightarrow{OB}$  vektorok segítségével az a)  $\overrightarrow{OC}$ , b)  $\overrightarrow{OE}$ , c)  $\overrightarrow{OF}$ , d)  $\overrightarrow{AC}$ , e)  $\overrightarrow{BD}$ , f)  $\overrightarrow{BF}$ , g)  $\overrightarrow{AB} + \overrightarrow{CD} + \overrightarrow{EF}$  vektorokat!

1.9. Adva van  $n$  tetszőleges, nem feltétlenül különböző  $P_1, P_2, \dots, P_n$  pont a térben. Mivel egyenlő a

$$\overrightarrow{P_1P_2} + \overrightarrow{P_2P_3} + \overrightarrow{P_3P_4} + \dots + \overrightarrow{P_{n-1}P_n}$$

és a

$$\overrightarrow{P_1P_2} + \overrightarrow{P_2P_3} + \overrightarrow{P_3P_4} + \dots + \overrightarrow{P_{n-1}P_n} + \overrightarrow{P_nP_1}$$

összeg?

1.10. Mutassuk meg, hogy az  $\mathbf{a}$ ,  $\mathbf{b}$  és  $\mathbf{c}$  vektorok pontosan akkor lehetnek (egy esetleg szakasszá vagy ponttá elfajuló) háromszög oldalvektorai, ha az

$$\mathbf{a} + \mathbf{b} + \mathbf{c}, \quad \mathbf{a} + \mathbf{b} - \mathbf{c}, \quad \mathbf{a} - \mathbf{b} + \mathbf{c}, \quad \mathbf{a} - \mathbf{b} - \mathbf{c}$$

vektorok legalább egyike zérus. Másként fogalmazva: ha a három vektor összege  $\mathbf{0}$ , vagy valamelyik vektor egyenlő a másik kettő összegével.

1.11. Legyen  $\mathbf{a}$  és  $\mathbf{b}$  két tetszőleges vektor. Mutassuk meg, hogy van olyan (esetleg elfajuló) háromszög, melynek oldalvektorai  $2\mathbf{a} - \mathbf{b}$ ,  $\mathbf{a} + 2\mathbf{b}$  és  $3\mathbf{a} + \mathbf{b}$ .

### Lineáris kombináció, lineáris függetlenség

1.12. **LINEÁRIS ÖSSZEFÜGGŐSÉG: IGAZ – HAMIS** Melyek igazak, melyek hamisak az alábbi állítások közül?

- Ha három vektor a térben lineárisan összefüggő, akkor bármelyikük a másik kettő lineáris kombinációja.
- Megadható a térben három vektor, hogy egyikük sem lineárisan független a többitől.
- Megadható a térben három vektor,  $\mathbf{a}$ ,  $\mathbf{b}$  és  $\mathbf{c}$ , hogy  $\mathbf{a}$  független a  $\mathbf{b}$  és  $\mathbf{c}$  vektoroktól, de  $\mathbf{b}$  nem független az  $\mathbf{a}$  és  $\mathbf{c}$  vektoroktól.
- Ha három térbeli vektor egy síkba esik, akkor mindegyik kifejezhető a másik kettő lineáris kombinációjaként!

### Speciális lineáris kombinációk

1.13. **SZAKASZT  $m : n$  ARÁNYBAN OSZTÓ PONT** Ha az  $\overline{AB}$  szakaszt a  $P$  pont úgy bontja ketté, hogy  $|\overline{AP}| : |\overline{PB}| = m : n$ , akkor bármely  $O$  pontra igaz, hogy

$$\overrightarrow{OP} = \frac{n}{m+n}\overrightarrow{OA} + \frac{m}{m+n}\overrightarrow{OB}.$$

Speciálisan, az  $\overline{AB}$  szakasz felezőpontjába az

$$\frac{\overrightarrow{OA} + \overrightarrow{OB}}{2}$$

vektor mutat.

## Távolság, szög, orientáció

A címben jelzett három alapfogalomhoz három – vektorok közti – szorzás-művelet visz közelebb. Mindegyik művelet igen szokatlan tulajdonságokkal rendelkezik: az egyik eredményül nem vektort, hanem skalárt ad, a másik nem felcserélhető, és kétváltozós (bináris) műveletként csak a 3-dimenziós térben definiálható, a harmadik pedig nem két- hanem háromváltozós művelet.

**Skaláris szorzás** A fizikában az erő által végzett munka az út hosszának és az erő elmozdulás irányába eső merőleges vetülete hosszának szorzata. Vagyis két vektorjellegű mennyiségből egy skalármennyiséget kapunk eredményül. Ha  $\mathbf{F}$  jelöli az erővektort,  $\mathbf{s}$  az elmozdulásvektort,  $\mathbf{F}_s$  az erőnek az elmozdulás irányába eső merőleges vetületi vektorát és  $\gamma$  az  $\mathbf{F}$  és  $\mathbf{s}$  vektorok hajlásszögét, akkor a munka értéke  $|\mathbf{F}_s||\mathbf{s}| = |\mathbf{F}||\mathbf{s}| \cos \gamma$ . Ez vezet a következő definícióhoz:

**1.15. DEFINÍCIÓ (KÉT VEKTOR SKALÁRIS SZORZATA).** Két vektor skaláris szorzatán a vektorok abszolút értékének és az általuk bezárt szög koszinuszának szorzatát értjük. Az  $\mathbf{a}$  és  $\mathbf{b}$  vektorok skaláris szorzatát  $\mathbf{a} \cdot \mathbf{b}$  jelöli, tehát

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos(\mathbf{a}, \mathbf{b})_{\angle},$$

ahol a két vektor által bezárt szög  $(\mathbf{a}, \mathbf{b})_{\angle}$ .

Ha  $\mathbf{a}$  és  $\mathbf{b}$  valamelyike zérusvektor, akkor a két vektor szöge, s így annak koszinusza sem határozható meg egyértelműen, a skaláris szorzat viszont ekkor is egyértelmű, és pedig 0, hisz a zérusvektor abszolút értéke 0, és 0 bármivel vett szorzata 0.

Szokás  $\mathbf{a}$  és  $\mathbf{b}$  skaláris szorzatát  $\mathbf{ab}$ -vel is jelölni, de ezt egy később bevezetendő művelettel (a mátrixszorzással) való összekeverés elkerülése érdekében e könyvben nem fogjuk használni.

**1.16. PÉLDA (SKALÁRIS SZORZAT KISZÁMÍTÁSA A DEFINÍCIÓ ALAPJÁN).** Mennyi a skaláris szorzata egy 1 és egy 2 egység hosszú és egymással  $60^\circ$ -os szöget bezáró két vektornak?

**MEGOLDÁS.** A szorzat  $1 \cdot 2 \cdot \cos 60^\circ = 1 \cdot 2 \cdot \frac{1}{2} = 1$ . □

**1.17. TÉTEL (MIKOR 0 A SKALÁRIS SZORZAT?).** Két vektor skaláris szorzata pontosan akkor 0, ha a két vektor merőleges egymásra.

**BIZONYÍTÁS.** ( $\Leftarrow$ ) Ha  $\mathbf{a} \perp \mathbf{b}$ , akkor  $(\mathbf{a}, \mathbf{b})_{\angle} = \pi/2$ , azaz  $\cos(\mathbf{a}, \mathbf{b})_{\angle} = 0$ , tehát  $\mathbf{a} \cdot \mathbf{b} = 0$ .

( $\Rightarrow$ ) Ha  $\mathbf{a} \cdot \mathbf{b} = 0$ , azaz  $|\mathbf{a}||\mathbf{b}| \cos(\mathbf{a}, \mathbf{b})_{\angle} = 0$ , akkor  $|\mathbf{a}| = 0$ ,  $|\mathbf{b}| = 0$  vagy  $\cos(\mathbf{a}, \mathbf{b})_{\angle} = 0$ . Ha valamelyik vektor zérusvektor, akkor iránya bármely vektorára merőlegesnek tekinthető. Ha viszont  $\mathbf{a} \neq \mathbf{0}$  és  $\mathbf{b} \neq$

$\mathbf{0}$ , akkor  $\cos(\mathbf{a}, \mathbf{b})_{\angle} = 0$ , a  $\cos$  függvénynek a  $[0, \pi]$  intervallumban csak  $\pi/2$ -ben van zérushelye, tehát a két vektor merőleges egymásra.  $\square$

► A tétel bizonyításából látszik, a zérusvektorra úgy tekintünk, mint ami bármely vektorra merőleges. Korábban – a skalárral való szorzásnál – a zérusvektorra úgy tekintettünk, mint ami bármely vektorral párhuzamos, hisz skalárszorosa. A zérusvektorra tehát úgy tekintünk, mint ami egy adott vektorral akkora szöget zár be, mint amekkorára épp szükségünk van. Ezt megtehetjük, hisz a zérusvektor iránya tetszőleges. Ez megóv minket attól, hogy minden tételbe a zérusvektor esetét külön, mint valami rendhagyó esetet bele kelljen fogalmaznunk.

**1.18. TÉTEL (A SKALÁRIS SZORZÁS MŰVELETI TULAJDONSÁGAI).** Ha  $\mathbf{a}$ ,  $\mathbf{b}$  és  $\mathbf{c}$  tetszőleges térbeli (síkbeli) vektorok és  $r$  tetszőleges valós szám, akkor igazak az alábbi összefüggések:

- a)  $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$  (kommutativitás)
- b)  $(\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c}$  (disztributivitás)
- c)  $r(\mathbf{a} \cdot \mathbf{b}) = (r\mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (r\mathbf{b})$
- d)  $\mathbf{a} \cdot \mathbf{a} > 0$ , ha  $\mathbf{a} \neq \mathbf{0}$ , és  $\mathbf{a} \cdot \mathbf{a} = 0$ , ha  $\mathbf{a} = \mathbf{0}$ .

A bizonyítást az olvasóra hagyjuk.

Mivel két vektor skaláris szorzata skalár, ezért az asszociativitás (csoportosíthatóság) kérdése föl sem vethető, hisz az  $(\mathbf{a} \cdot \mathbf{b})\mathbf{c}$  szorzatban két különböző szorzásművelet szerepel. Mindezzel együtt  $(\mathbf{a} \cdot \mathbf{b})\mathbf{c} \neq \mathbf{a}(\mathbf{b} \cdot \mathbf{c})$  (ld. az 1.18. feladatot).

**Hosszúság és szög** Egy vektor hossza, és ezzel két pont távolsága, valamint két vektor hajlásszöge kifejezhető a skaláris szorzat segítségével.

Egy tetszőleges  $\mathbf{a}$  vektorra  $\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}||\mathbf{a}| \cos 0 = |\mathbf{a}||\mathbf{a}|$ , tehát

$$|\mathbf{a}|^2 = \mathbf{a} \cdot \mathbf{a}, \text{ azaz } |\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}}.$$

E képlet szerint tehát egy vektor hossza megegyezik az önmagával vett skaláris szorzatának gyökével. Ebből az is adódik, hogy két pont távolsága megegyezik az őket összekötő vektor önmagával vett skaláris szorzatának négyzetgyökével.

Két pontot összekötő vektor egyenlő az oda mutató helyvektorok különbségével, így ha a két pontba mutató helyvektor  $\mathbf{a}$  és  $\mathbf{b}$ , akkor a pontok távolsága – és ezt fogjuk a vektorok távolságának is tekinteni

$$d(\mathbf{a}, \mathbf{b}) = |\mathbf{a} - \mathbf{b}|.$$

Két vektor skaláris szorzatának és a vektorok hosszának ismeretében a szögük meghatározható:

$$\cos(\mathbf{a}, \mathbf{b})_{\angle} = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|}, \quad (1.2)$$

mivel a  $[0, \pi]$  intervallumon a koszinusz függvény kölcsönösen egyértelmű.

**Pithagorász-tétel** A távolságot vagy hosszúságot skaláris szorzattal is ki tudjuk fejezni, így segítségével a rá vonatkozó összefüggések is vizsgálhatók.

1.19. TÉTEL (PITHAGORÁSZ-TÉTEL). Az  $\mathbf{a}$  és  $\mathbf{b}$  vektorokra pontosan akkor teljesül az

$$|\mathbf{a} + \mathbf{b}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2$$

összefüggés, ha  $\mathbf{a}$  és  $\mathbf{b}$  merőlegesek egymásra.

**BIZONYÍTÁS.**

$$\begin{aligned} |\mathbf{a} + \mathbf{b}|^2 &= (\mathbf{a} + \mathbf{b}) \cdot (\mathbf{a} + \mathbf{b}) \\ &= \mathbf{a} \cdot \mathbf{a} + \mathbf{a} \cdot \mathbf{b} + \mathbf{b} \cdot \mathbf{a} + \mathbf{b} \cdot \mathbf{b} && \text{disztributivitás} \\ &= \mathbf{a} \cdot \mathbf{a} + 2(\mathbf{a} \cdot \mathbf{b}) + \mathbf{b} \cdot \mathbf{b} && \text{kommutativitás} \\ &\stackrel{?}{=} \mathbf{a} \cdot \mathbf{a} + \mathbf{b} \cdot \mathbf{b} && ? \\ &= |\mathbf{a}|^2 + |\mathbf{b}|^2, \end{aligned}$$

Világos, hogy a ?-lel megjelölt egyenlőség pontosan akkor teljesül, ha  $\mathbf{a} \cdot \mathbf{b} = 0$ , azaz ha  $\mathbf{a}$  és  $\mathbf{b}$  merőlegesek egymásra.  $\square$

1.20. PÉLDA (SKALÁRIS SZORZAT KISZÁMÍTÁSA). Számítsuk ki az 1.24 ábrán látható két vektor skaláris szorzatát (a szomszédos rácsvonalak távolsága 1 egység).

**MEGOLDÁS.** Az  $\mathbf{a}$  vektor hossza  $\sqrt{2^2 + 2^2} = 2\sqrt{2}$ , a  $\mathbf{b}$  vektor hossza  $\sqrt{4^2 + 3^2} = 5$ , az  $\mathbf{a}$  vektornak a vízszintes rácsvonalakkal bezárt szöge  $\pi/4$ , a  $\mathbf{b}$  vektornál e szög szögfüggvényei  $\cos \gamma = \frac{4}{5}$ ,  $\sin \gamma = \frac{3}{5}$ . Így

$$\cos\left(\gamma + \frac{\pi}{4}\right) = \cos \gamma \cos \frac{\pi}{4} - \sin \gamma \sin \frac{\pi}{4} = \frac{4}{5} \frac{\sqrt{2}}{2} - \frac{3}{5} \frac{\sqrt{2}}{2} = \frac{1}{5} \frac{\sqrt{2}}{2},$$

tehát a skaláris szorzat  $\mathbf{a} \cdot \mathbf{b} = 2\sqrt{2} \cdot 5 \cdot \frac{1}{5} \frac{\sqrt{2}}{2} = 2$ .  $\square$

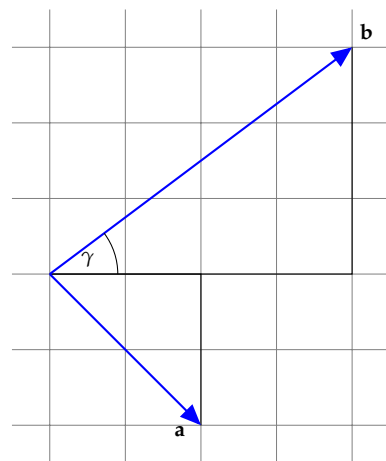
**Két fontos egyenlőtlenség** Mivel a koszinusz függvény értéke abszolút értékben sosem nagyobb 1-nél, ezért a skaláris szorzat definíciójából azonnal látszik, hogy

$$|\mathbf{a} \cdot \mathbf{b}| = |\mathbf{a}| |\mathbf{b}| |\cos(\mathbf{a}, \mathbf{b})| \leq |\mathbf{a}| |\mathbf{b}|.$$

Ezzel bizonyítottuk a következő tételt:

1.21. TÉTEL (CAUCHY–BUNYAKOVSKIJ–SCHWARZ-EGYENLŐTLENSÉG). Két vektor skaláris szorzatának abszolút értéke sosem nagyobb abszolút értékeik szorzatánál, azaz

$$|\mathbf{a} \cdot \mathbf{b}| \leq |\mathbf{a}| |\mathbf{b}|.$$



1.24. ábra: Két vektor skaláris szorzata



A Cauchy–Bunyakovszkij–Schwarz-egyenlőtlenség segítségével bizonyítjuk a geometriából jól ismert háromszög-egyenlőtlenséget. Ennek az az értelme, hogy e bizonyítás változtatás nélkül működni fog sokkal általánosabb körülmények között is.

**1.22. TÉTEL (HÁROMSZÖG-EGYENLŐTLENSÉG).** *Bármely két  $\mathbf{a}$  és  $\mathbf{b}$  vektorra*

$$|\mathbf{a} + \mathbf{b}| \leq |\mathbf{a}| + |\mathbf{b}|.$$

**BIZONYÍTÁS.** Mivel az egyenlőtlenség mindkét oldalán nemnegatív szám áll, ezért vele ekvivalens egyenlőtlenséghez jutunk, ha mindkét oldalt négyzetre emeljük.

$$\begin{aligned} |\mathbf{a} + \mathbf{b}|^2 &= (\mathbf{a} + \mathbf{b}) \cdot (\mathbf{a} + \mathbf{b}) \\ &= |\mathbf{a}|^2 + 2(\mathbf{a} \cdot \mathbf{b}) + |\mathbf{b}|^2 \quad \text{ld. az 1.19. tétel bizonyítását} \\ &\leq |\mathbf{a}|^2 + 2|\mathbf{a} \cdot \mathbf{b}| + |\mathbf{b}|^2 \\ &\leq |\mathbf{a}|^2 + 2|\mathbf{a}||\mathbf{b}| + |\mathbf{b}|^2 \\ &= (|\mathbf{a}| + |\mathbf{b}|)^2. \end{aligned}$$

És ezt akartuk bizonyítani.  $\square$

**Egységvektorral való szorzás és a merőleges vetítés** Minden olyan vektort, melynek abszolút értéke 1, *egységvektornak* nevezünk.

Ha  $\mathbf{a}$  egy tetszőleges nemzérus vektor, akkor  $\mathbf{a}/|\mathbf{a}|$  egységvektor, ugyanis abszolút értéke 1:

$$\left| \frac{\mathbf{a}}{|\mathbf{a}|} \right| = \frac{1}{|\mathbf{a}|} |\mathbf{a}| = 1.$$

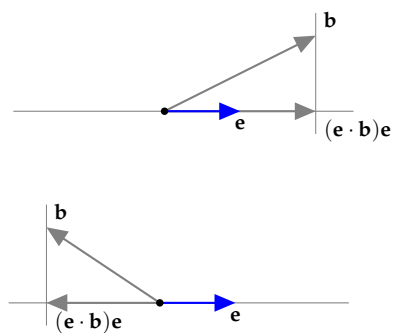
**1.23. TÉTEL (EGYSÉGVEKTORRAL VALÓ SZORZÁS GEOMETRIAI JELENTÉSE).** *Ha  $\mathbf{e}$  egységvektor, akkor a  $\hat{\mathbf{b}} = (\mathbf{e} \cdot \mathbf{b})\mathbf{e}$  vektor a  $\mathbf{b}$  vektornak az  $\mathbf{e}$  egyenesére való merőleges vetülete. Az  $\mathbf{e} \cdot \mathbf{b}$  szorzat  $\mathbf{e}$  vetület előjeles hossza, mely pozitív, ha  $\hat{\mathbf{b}}$  és  $\mathbf{e}$  egyirányúak, és negatív, ha ellenkező irányúak.*

**BIZONYÍTÁS.** Ha  $\mathbf{e}$  egységvektor, azaz abszolút értéke 1, akkor  $\mathbf{e} \cdot \mathbf{b} = |\mathbf{b}| \cos(\mathbf{e}, \mathbf{b})$ , ez pedig a koszinusz függvény definíciója szerint  $\mathbf{b}$  merőleges vetületének előjeles hosszát jelenti. E szám  $\mathbf{e}$ -szerese pedig egy  $\mathbf{e}$  irányú, és ilyen hosszú vektort ad, mely épp  $\mathbf{b}$  vetületi vektora.  $\square$

Jelölje a  $\mathbf{b}$  vektornak az  $\mathbf{a}$  egyenesére eső merőleges vetületi vektorát  $\text{proj}_{\mathbf{a}} \mathbf{b}$ . Eszerint ha  $\mathbf{e}$  egységvektor, akkor

$$\text{proj}_{\mathbf{e}} \mathbf{b} = (\mathbf{e} \cdot \mathbf{b})\mathbf{e}.$$

Alapvető feladat egy vektornak egy másikkal párhuzamos és rá merőleges vektorok összegére való felbontása, amit másként *merőleges összetevőkre bontásnak* nevezünk.



1.25. ábra: Az  $\mathbf{b}$  vektor és az  $\mathbf{e}$  egységvektor egyenesére eső vetülete. A felső ábrán  $\mathbf{e} \cdot \mathbf{b} > 0$ , az alsón  $\mathbf{e} \cdot \mathbf{b} < 0$ .



**1.24. TÉTEL** (VEKTOR FELBONTÁSA MERŐLEGES ÖSSZETEVŐKRE). Ha  $\mathbf{a}$  és  $\mathbf{b}$  a sík vagy a tér két vektora, és  $\mathbf{a} \neq \mathbf{0}$ , akkor  $\mathbf{b}$ -nek az  $\mathbf{a}$  egyenesére eső merőleges vetülete

$$\text{proj}_{\mathbf{a}} \mathbf{b} = \frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{a} \cdot \mathbf{a}} \mathbf{a}.$$

A  $\mathbf{b}$ -nek az  $\mathbf{a}$  egyenesére merőleges összetevője

$$\mathbf{b} - \text{proj}_{\mathbf{a}} \mathbf{b} = \mathbf{b} - \frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{a} \cdot \mathbf{a}} \mathbf{a}.$$

**BIZONYÍTÁS.** Az első képlet az egységvektorral szorzás geometriai jelentéséről szóló 1.23. tételből következik. Legyen  $\mathbf{e} = \frac{\mathbf{a}}{|\mathbf{a}|}$  az  $\mathbf{a}$ -irányú egységvektor. Ekkor

$$\text{proj}_{\mathbf{e}} \mathbf{b} = (\mathbf{e} \cdot \mathbf{b}) \mathbf{e} = \left( \frac{\mathbf{a}}{|\mathbf{a}|} \cdot \mathbf{b} \right) \frac{\mathbf{a}}{|\mathbf{a}|} = \frac{1}{|\mathbf{a}|^2} (\mathbf{a} \cdot \mathbf{b}) \mathbf{a} = \frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{a} \cdot \mathbf{a}} \mathbf{a}.$$

(Az utolsó egyenlőségnél kihasználtuk, hogy  $|\mathbf{a}|^2 = \mathbf{a} \cdot \mathbf{a}$ .) Mivel  $\mathbf{e}$  és  $\mathbf{a}$  párhuzamosak, ezért  $\text{proj}_{\mathbf{a}} \mathbf{b} = \text{proj}_{\mathbf{e}} \mathbf{b}$ , ami bizonyítja első állításunkat. Az állítás második fele abból adódik, hogy a két összetevő összege  $\mathbf{b}$ .  $\square$

**1.25. PÉLDA** (MERŐLEGES ÖSSZETEVŐKRE BONTÁS). Az 1.24 ábrabeli  $\mathbf{b}$  vektort bontsuk fel az  $\mathbf{a}$ -val párhuzamos és rá merőleges összetevőkre.

**MEGOLDÁS.** Bár a megoldás az 1.27 ábráról is leolvasható, kövessük végig a számítást: az 1.20. példa szerint  $\mathbf{a} \cdot \mathbf{b} = 2$ ,  $|\mathbf{a}| = 2\sqrt{2}$ , ezért

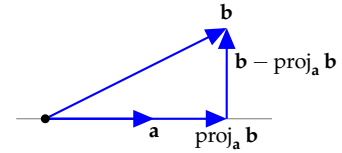
$$\text{proj}_{\mathbf{a}} \mathbf{b} = \frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{a} \cdot \mathbf{a}} \mathbf{a} = \frac{2}{8} \mathbf{a} = \frac{1}{4} \mathbf{a},$$

míg az  $\mathbf{a}$ -ra merőleges összetevő  $\mathbf{b} - \frac{1}{4} \mathbf{a}$ .  $\square$

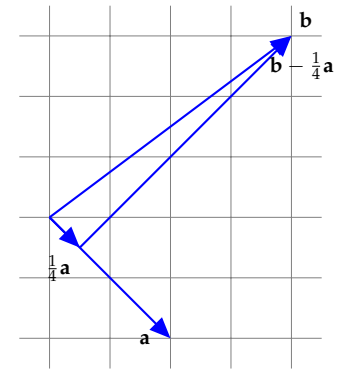
**Merőlegesség és orientáció** Legyenek  $\mathbf{a}$  és  $\mathbf{b}$  egymásra merőleges nem-zérus vektorok a síkban. Ekkor  $\mathbf{a}$  és  $-\mathbf{b}$  is merőlegesek, azaz  $(\mathbf{a}, \mathbf{b})_{\perp} = (\mathbf{a}, -\mathbf{b})_{\perp} = \pi/2$ . Csak az  $\mathbf{a}$  ismeretében meg tudjuk-e különböztetni a  $\mathbf{b}$  és  $-\mathbf{b}$  vektorokat? Hasonló kérdés a térben is fölmerül: ha  $\mathbf{c}$  merőleges a nem kollineáris  $\mathbf{a}$  és  $\mathbf{b}$  vektorok mindegyikére, akkor  $-\mathbf{c}$  is. Vajon  $\mathbf{c}$  és  $-\mathbf{c}$  megkülönböztethető-e egymástól csak  $\mathbf{a}$ -hoz és  $\mathbf{b}$ -hez való viszonyukat tekintve?

A válaszhoz az *irányítás*, más szóval *orientáció* fogalma vezet. E fogalmat precízen később definiáljuk (ld. ??? szakasz), az alapgondolat viszont egyszerű. A síkban a két független vektorból álló párokat két osztályba sorolhatjuk aszerint, hogy a tenyérrel fölfelé fordított jobb vagy bal kezünk első két ujjával szemléltethetők (1.28. ábra) (hüvelyk az első, mutató a második vektor).

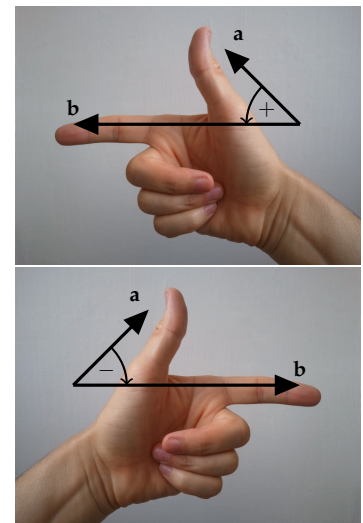
Hasonlóképp: a térben a független vektorokból álló hármasokat két osztályba sorolhatjuk aszerint, hogy jobb vagy bal kezünk első három



1.26. ábra: Az  $\mathbf{b}$  vektor felbontása az  $\mathbf{a}$  vektorral párhuzamos és rá merőleges vektorok összegére.

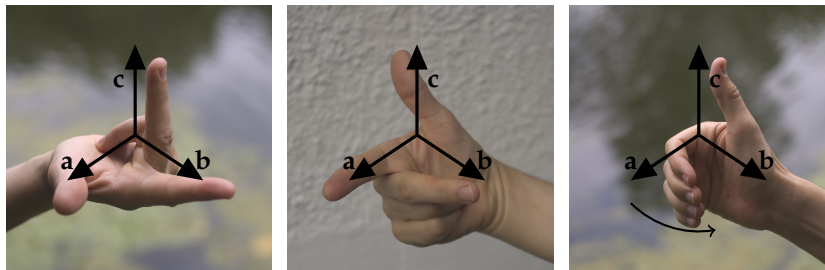


1.27. ábra: Vektor felbontása merőleges összetevőkre

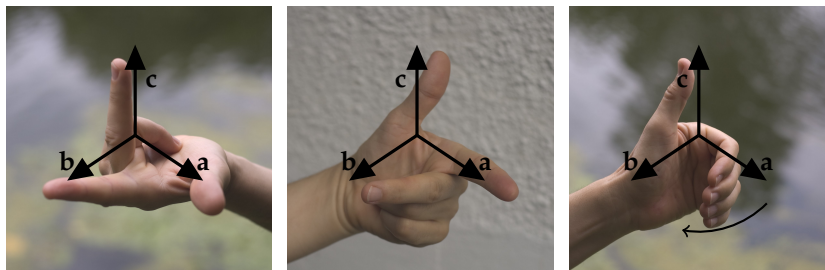


1.28. ábra: Két vektor egymáshoz való viszonya jobbrendszer (felső ábra) vagy balrendszer (alsó ábra) alkot. A közöttük zárt irányított szög az előbbi esetben pozitív, utóbbiban negatív.

ujjával szemléltethetőek (1.29. és 1.30. ábra). Kézenfekvőnek tűnik az első vektornak a hüvelyk, a másodiknak a mutató, a harmadiknak a középső ujjunkat megfeleltetni, de azokban a kultúrákban, ahol a mutató és középső ujjal mutatják a kettőt, ott a mutató-középső-hüvelyk a sorrend. Aszerint, hogy egy vektorpár a síkban, illetve egy vektorhármast a térben melyik osztályba esik, azt mondjuk, hogy *jobbrendszert*, illetve *balrendszert* alkot. Az 1.29. ábra harmadik képén látható mód (az ökölbe szoruló jobb kéz mozgása) azt is megmutatja, hogy milyen egy egyenes körül való pozitív forgás iránya. A síkban ezt azzal is ki tudjuk fejezni, hogy két független vektor szögét előjellel látjuk el, nevezetesen pozitívvál, ha jobbrendszert, és negatívvál, ha balrendszert alkotnak. Az így kapott szöget a két vektor *irányított szögének* nevezzük. Az  $\mathbf{a}$  és  $\mathbf{b}$  irányított szögét  $(\mathbf{a}, \mathbf{b})_{\triangleleft}$  jelöli. Tehát míg  $(\mathbf{a}, \mathbf{b})_{\angle} = (\mathbf{b}, \mathbf{a})_{\angle}$ , addig  $(\mathbf{a}, \mathbf{b})_{\triangleleft} = -(\mathbf{b}, \mathbf{a})_{\triangleleft}$ , és ha  $(\mathbf{a}, \mathbf{b})_{\triangleleft} = \pi/2$ , akkor  $(\mathbf{a}, -\mathbf{b})_{\triangleleft} = -\pi/2$ . Ez tehát a válasz a paragrafus elején feltett kérdésre.



1.29. ábra: Az  $\mathbf{a}$ ,  $\mathbf{b}$  és  $\mathbf{c}$  vektorok ebben a sorrendben jobbrendszert alkotnak, ha irányuk a jobb kezünkkel mutatható a mellékelt három ábra bármelyike szerint: (1) hüvelyk-mutató-középső ujj, (2) mutató-középső-hüvelykujj, (3) a hüvelyk mutatja a  $\mathbf{c}$  vektort, ökölbe szoruló kezünk ujjai pedig az  $\mathbf{a}$  felől a  $\mathbf{b}$  felé haladnak.



1.30. ábra: Az  $\mathbf{a}$ ,  $\mathbf{b}$  és  $\mathbf{c}$  vektorok ebben a sorrendben balrendszert alkotnak, ha irányuk a bal kezünkkel mutatható a következők bármelyike szerint: (1) hüvelyk-mutató-középső ujj, (2) mutató-középső-hüvelykujj, (3) a hüvelyk mutatja a  $\mathbf{c}$  vektort, ökölbe szoruló kezünk ujjai pedig az  $\mathbf{a}$  felől a  $\mathbf{b}$  felé haladnak.

**Vektori szorzás** A fizikában több olyan jelenség is van, melyben két térbeli vektorhoz keresünk egy mindkettőre merőleges harmadikat. Legismertebb példa a *forogatónyomaték*.

Hasson egy  $\mathbf{F}$  erő egy test  $\mathbf{P}$  pontjában, és legyen a test rögzítve az  $O$  pontjában. A  $P$  ponton átmenő,  $\mathbf{F}$  irányú egyenesnek az  $O$ -tól való távolságát az erő karjának nevezzük. Az  $\mathbf{F}$  hatására a test  $O$  körül elfordul. Ennek jellemzésére tudnunk kell a forgás tengelyét, a forgás „nagyságát”, és azt, hogy a tengely körüli két forgásirány közül melyikről van szó. Erre alkalmas lehet egy vektor – ezt nevezzük

*forgatónyomatéknak* –, melynek iránya a forgástengellyel párhuzamos, hossza a forgás nagyságát írja le, és a forgástengellyel párhuzamos két vektorirány a két forgásirányhoz tartozik. Hogyan definiálható a forgatónyomaték-vektor, ha tudjuk, hogy abszolút értéke az erőkar hosszának és az erő abszolút értékének szorzata?

Az erő karja  $|\vec{OP}| \sin(\vec{OP}, \mathbf{F})_{\angle}$ , így az  $\mathbf{M}$  forgatónyomaték abszolút értéke:

$$|\mathbf{M}| = |\mathbf{F}| |\vec{OP}| \sin(\vec{OP}, \mathbf{F})_{\angle}.$$

A forgás tengelye nyilván merőleges  $\mathbf{F}$ -re és  $\vec{OP}$ -re is, csak abban kell megegyezni, hogy az  $\vec{OP}$ ,  $\mathbf{F}$  és  $\mathbf{M}$  vektorok jobb- vagy balrendszert alkossanak. A fizikusok a jobbrendszert választották.

A forgatónyomaték és több hasonló fizikai fogalom a következő definícióhoz vezet:

**1.26. DEFINÍCIÓ (VEKTORI SZORZÁS).** A 3-dimenziós tér két vektorának vektori szorzatán azt a vektort értjük, melynek

- abszolút értéke a két vektor abszolút értékének és közbezárt szöge szinuszának szorzata,
- iránya merőleges mindkét vektor irányára és – ha a szorzat nem a nullvektor, akkor – az első tényező, a második tényező és a szorzat ebben a sorrendben jobbrendszert alkot.

► A vektor abszolút értéke tényleg nem negatív, mert a szinusz függvény a  $[0, \pi]$  intervallumon nem negatív.

► E definíció bármely két 3-dimenziós vektor vektori szorzatát egyértelműen definiálja, ugyanis minden olyan esetben, amikor nem tudnánk eldönteni, hogy a vektorok jobbrendszert alkotnak-e, a szorzat a nullvektor (gondoljuk meg!).

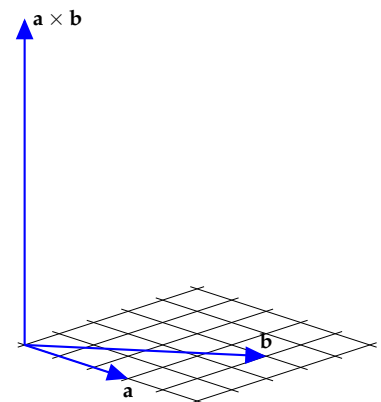
► Az  $\mathbf{a}$  és  $\mathbf{b}$  vektorok vektori szorzatát  $\mathbf{a} \times \mathbf{b}$  jelöli, amit „a kereszt b”-nek olvasunk. Képletekkel megfogalmazva:  $\mathbf{a} \times \mathbf{b}$  egy vektor, melyre

$$|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}| |\mathbf{b}| \sin(\mathbf{a}, \mathbf{b})_{\angle},$$

$\mathbf{a} \times \mathbf{b} \perp \mathbf{a}$ ,  $\mathbf{a} \times \mathbf{b} \perp \mathbf{b}$ , továbbá  $\mathbf{a}$ ,  $\mathbf{b}$  és  $\mathbf{a} \times \mathbf{b}$  ebben a sorrendben jobbrendszert alkot, ha  $|\mathbf{a} \times \mathbf{b}| \neq 0$ .

**1.27. PÉLDA (VEKTORI SZORZAT MEGHATÁROZÁSA).** Tegyük fel, hogy a tér két vektora 3 illetve 5 hosszú, az általuk bezárt szög koszinusza  $\frac{4}{5}$ . Mit tudunk a vektori szorzatról?

**MEGOLDÁS.** Ha  $\cos \gamma = \frac{4}{5}$ , akkor  $\sin \gamma = \sqrt{1 - (\frac{4}{5})^2} = \frac{3}{5}$ , így a vektori szorzat hossza  $|\mathbf{a}| |\mathbf{b}| \sin(\mathbf{a}, \mathbf{b})_{\angle} = 3 \cdot 5 \cdot \frac{3}{5} = 9$ , iránya merőleges mindkét vektorra és  $\mathbf{a}$ ,  $\mathbf{b}$ ,  $\mathbf{a} \times \mathbf{b}$  ebben a sorrendben jobbrendszert alkot (ld. 1.31. ábra). □



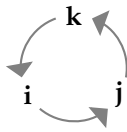
1.31. ábra: Az  $\mathbf{a}$ ,  $\mathbf{b}$  és az  $\mathbf{a} \times \mathbf{b}$  vektorok

**1.28. PÉLDA** ( $\mathbf{i}, \mathbf{j}, \mathbf{k}$  VEKTORI SZORZATA). Legyen  $\mathbf{i}, \mathbf{j}, \mathbf{k}$  három, páronként egymásra merőleges, ebben a sorrendben jobbrendszert alkotó egységvektor. Készítsünk művelettáblát vektori szorzataikról!

**MEGOLDÁS.** Mivel  $(\mathbf{i}, \mathbf{i})_{\angle} = 0$ , ezért  $|\mathbf{i} \times \mathbf{i}| = 0$ , így  $\mathbf{i} \times \mathbf{i} = \mathbf{0}$ . Hasonlóan  $\mathbf{j} \times \mathbf{j} = \mathbf{0}$  és  $\mathbf{k} \times \mathbf{k} = \mathbf{0}$ .

Mivel  $|\mathbf{i}| = |\mathbf{j}| = 1$  és  $(\mathbf{i}, \mathbf{j})_{\angle} = 90^\circ$ , ezért  $|\mathbf{i} \times \mathbf{j}| = 1$ , azaz  $\mathbf{i} \times \mathbf{j}$  is egységvektor. Ráadásul  $\mathbf{i} \times \mathbf{j}$  merőleges  $\mathbf{i}$ -re és  $\mathbf{j}$ -re, és  $\mathbf{i}, \mathbf{j}$  valamint  $\mathbf{i} \times \mathbf{j}$  jobbrendszert alkotnak épp úgy, mint  $\mathbf{i}, \mathbf{j}$  és  $\mathbf{k}$ . Ebből következik, hogy  $\mathbf{i} \times \mathbf{j} = \mathbf{k}$ . Hasonlóképp  $\mathbf{j} \times \mathbf{k} = \mathbf{i}$  és  $\mathbf{k} \times \mathbf{i} = \mathbf{j}$ . Ha  $\mathbf{i}, \mathbf{j}, \mathbf{k}$  jobbrendszert alkot, akkor  $\mathbf{j}, \mathbf{i}$  és  $\mathbf{k}$  balrendszert, így  $\mathbf{j} \times \mathbf{i} = -\mathbf{k}$ . Mindezeket összefoglalva a következő művelettáblát kapjuk.

| $\times$     | $\mathbf{i}$  | $\mathbf{j}$  | $\mathbf{k}$  |
|--------------|---------------|---------------|---------------|
| $\mathbf{i}$ | $\mathbf{0}$  | $\mathbf{k}$  | $-\mathbf{j}$ |
| $\mathbf{j}$ | $-\mathbf{k}$ | $\mathbf{0}$  | $\mathbf{i}$  |
| $\mathbf{k}$ | $\mathbf{j}$  | $-\mathbf{i}$ | $\mathbf{0}$  |



E három vektor közti szorzatok könnyen megjegyezhetők, ha egy szabályos háromszög csúcsaira írjuk őket pozitív körüljárás szerint, mint azt a táblázat melletti ábra mutatja. Ekkor két különböző vektor szorzata a harmadik, ha a két vektor pozitív körüljárás szerint követi egymást. Ha negatív körüljárás szerint követik egymást, a szorzat a harmadik vektor  $-1$ -szerese.  $\square$

**1.29. TÉTEL** (MIKOR  $\mathbf{0}$  A VEKTORI SZORZAT?). Két térbeli vektor vektori szorzata pontosan akkor zérusvektor, ha a két vektor párhuzamos.

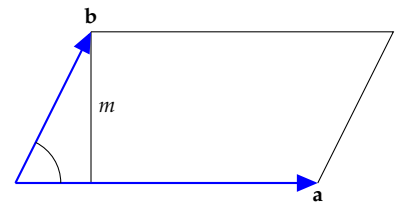
**BIZONYÍTÁS.** Ha  $\mathbf{a}$  vagy  $\mathbf{b}$  valamelyike zérusvektor, akkor egyrészt a két vektor tekinthető párhuzamosnak, másrészt  $\mathbf{a} \times \mathbf{b} = \mathbf{0}$ , az állítás tehát igaz, ezért a továbbiakban feltesszük, hogy a két tényező egyike sem zérusvektor.

( $\Leftarrow$ ) Ha  $\mathbf{a}$  és  $\mathbf{b}$  párhuzamosak, akkor  $(\mathbf{a}, \mathbf{b})_{\angle} = 0$  vagy  $\pi$ , tehát  $\sin(\mathbf{a}, \mathbf{b})_{\angle} = 0$ , így  $|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}||\mathbf{b}|0 = 0$ , azaz  $\mathbf{a} \times \mathbf{b} = \mathbf{0}$ .

( $\Rightarrow$ ) Ha  $\mathbf{a} \times \mathbf{b} = \mathbf{0}$ , azaz  $|\mathbf{a}||\mathbf{b}|\sin(\mathbf{a}, \mathbf{b})_{\angle} = 0$ , akkor  $|\mathbf{a}| \neq 0$  és  $|\mathbf{b}| \neq 0$  miatt  $\sin(\mathbf{a}, \mathbf{b})_{\angle} = 0$ . A szinusz függvénynek a  $[0, \pi]$  intervallumban a  $0$  és a  $\pi$  helyen van zérushelye, tehát a két vektor vagy egyirányú, vagy ellenkező irányú, vagyis párhuzamos.  $\square$

**1.30. TÉTEL** (VEKTORI SZORZAT ABSZOLÚT ÉRTÉKÉNEK GEOMETRIAI JELENTÉSE). Két vektor vektori szorzatának abszolút értéke a két vektor által kifeszített paralelogramma területének mérőszámával egyenlő.

**BIZONYÍTÁS.** Az  $\mathbf{a}$  és  $\mathbf{b}$  vektorok által kifeszített paralelogramma oldalainak hossza  $|\mathbf{a}|$  és  $|\mathbf{b}|$ , az  $\mathbf{a}$  oldalhoz tartozó magassága pedig  $m = |\mathbf{b}|\sin(\mathbf{a}, \mathbf{b})_{\angle}$ . A paralelogramma területe  $|\mathbf{a}|m = |\mathbf{a}||\mathbf{b}|\sin(\mathbf{a}, \mathbf{b})_{\angle} = |\mathbf{a} \times \mathbf{b}|$  (1.32. ábra).  $\square$



1.32. ábra:  $|\mathbf{a} \times \mathbf{b}|$  megegyezik a paralelogramma területével

1.31. TÉTEL (VEKTORI SZORZÁS MŰVELETI TULAJDONSÁGAI). *Tetszőleges  $\mathbf{a}$ ,  $\mathbf{b}$  és  $\mathbf{c}$  vektorokra, valamint tetszőleges  $r$  valós számra igazak az alábbi összefüggések:*

$$a) \quad \mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a} \quad (\text{alternáló tulajdonság})$$

$$b) \quad (\mathbf{a} + \mathbf{b}) \times \mathbf{c} = \mathbf{a} \times \mathbf{c} + \mathbf{b} \times \mathbf{c} \\ \mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c} \quad (\text{disztributivitás})$$

$$c) \quad r(\mathbf{a} \times \mathbf{b}) = (r\mathbf{a}) \times \mathbf{b} = \mathbf{a} \times (r\mathbf{b})$$

$$d) \quad |\mathbf{a} \times \mathbf{b}| = \sqrt{|\mathbf{a}|^2 |\mathbf{b}|^2 - |\mathbf{a} \cdot \mathbf{b}|^2}$$

- E tétel  $a)$  pontja szerint a vektori szorzás *nem kommutatív*!
- A vektori szorzás nem is asszociatív. Az 1.28. példa eredményét használva könnyen látható, hogy

$$(\mathbf{i} \times \mathbf{j}) \times \mathbf{j} \neq \mathbf{i} \times (\mathbf{j} \times \mathbf{j}),$$

ugyanis  $(\mathbf{i} \times \mathbf{j}) \times \mathbf{j} = \mathbf{k} \times \mathbf{j} = -\mathbf{i}$ , másrészt  $\mathbf{i} \times (\mathbf{j} \times \mathbf{j}) = \mathbf{i} \times \mathbf{0} = \mathbf{0}$ .

- A tétel bizonyítását az 1.32. feladatra hagyjuk.

*Parallelepipedon térfogata és előjeles térfogata* Az 1.30. tételben megmutattuk, hogy a vektori szorzat abszolút értéke a két vektor által kifeszített paralelogramma területét adja. Hogy számítható ki a parallelepipedon térfogata?

1.32. PÉLDA (PARALLELEPIPEDON TÉRFOGATA). *Határozzuk meg az  $\mathbf{a}$ ,  $\mathbf{b}$  és  $\mathbf{c}$  vektorok által kifeszített parallelepipedon térfogatát!*

**MEGOLDÁS.** Az  $\mathbf{a}$  és  $\mathbf{b}$  által kifeszített paralelogramma területe  $|\mathbf{a} \times \mathbf{b}|$ , és mivel  $\mathbf{a} \times \mathbf{b}$  merőleges a paralelogramma síkjára, ezért a parallelepipedon magassága  $\mathbf{c}$ -nek az  $\mathbf{a} \times \mathbf{b}$  egyenesére eső merőleges vetületi hosszával egyenlő. Ez az  $\mathbf{a} \times \mathbf{b}$  irányú egységvektorral való skaláris szorzással számolható. Az egységvektor

$$\mathbf{e} = \frac{\mathbf{a} \times \mathbf{b}}{|\mathbf{a} \times \mathbf{b}|},$$

a magasság  $|\mathbf{e} \cdot \mathbf{c}|$ , és így a térfogat (azaz az alapterületszer magasság) értéke

$$|\mathbf{a} \times \mathbf{b}| \left| \frac{\mathbf{a} \times \mathbf{b}}{|\mathbf{a} \times \mathbf{b}|} \cdot \mathbf{c} \right| = |(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}|.$$

Tehát a parallelepipedon térfogata  $|(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}|$ . □

- A  $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$  skalárt az  $\mathbf{a}$ ,  $\mathbf{b}$  és  $\mathbf{c}$  vektorok által kifeszített parallelepipedon *előjeles térfogatának* nevezzük.
- Ez pontosan akkor negatív, ha a  $\mathbf{c}$  vektor  $\mathbf{a} \times \mathbf{b}$  egyenesére eső merőleges vetülete és  $\mathbf{a} \times \mathbf{b}$  ellenkező irányú. Vagyis ha a  $\mathbf{c}$  vektor az  $\mathbf{a}$  és  $\mathbf{b}$  síkjának másik oldalán van, mint az  $\mathbf{a} \times \mathbf{b}$  vektor, azaz ha  $\mathbf{a}$ ,  $\mathbf{b}$  és  $\mathbf{c}$  balrendszert alkot! Tehát e skalár előjele a három vektor orientációját adja.

► A három vektor pontosan akkor esik egy síkba, azaz pontosan akkor lineárisan összefüggők, ha  $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = 0$ .

**Vegyes szorzat** Az előző paragrafusban megmutattuk az  $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$  kifejezés fontosságát. Ez vezet a következő definícióhoz:

**1.33. DEFINÍCIÓ (VEGYES SZORZAT).** A 3-dimenziós tér három tetszőleges  $\mathbf{a}$ ,  $\mathbf{b}$  és  $\mathbf{c}$  vektorából képzett

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$$

kifejezést a három vektor vegyes szorzatának nevezzük.

- A vegyes szorzat eredménye skalár.
- Az  $\mathbf{a}$ ,  $\mathbf{b}$  és  $\mathbf{c}$  vektorok vegyes szorzatának szokás jelölése  $\mathbf{abc}$ , de mi a későbbi fejezetekben nem fogjuk használni.
- Mivel a skaláris szorzás kommutatív, ezért  $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})$ .
- A paralelepipedon térfogatára ugyanazt az értéket kell kapnunk, bármelyik oldallapot is választjuk alapnak, így e három vektorból a vektorok különböző sorrendjeivel képzett vegyes szorzatok csak előjelükben térhetnek el egymástól. Mivel az előjel az orientáció függvénye, ezért – figyelembe véve az előző megjegyzést is – kapjuk, hogy

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = (\mathbf{b} \times \mathbf{c}) \cdot \mathbf{a} = (\mathbf{c} \times \mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}).$$

Az ellenkező előjelű szorzatok:

$$(\mathbf{c} \times \mathbf{b}) \cdot \mathbf{a} = (\mathbf{b} \times \mathbf{a}) \cdot \mathbf{c} = (\mathbf{a} \times \mathbf{c}) \cdot \mathbf{b} = \mathbf{c} \cdot (\mathbf{b} \times \mathbf{a}) = \mathbf{b} \cdot (\mathbf{a} \times \mathbf{c}) = \mathbf{a} \cdot (\mathbf{c} \times \mathbf{b}).$$

Ezeket a vegyes szorzatra használt jelöléssel fölírva:

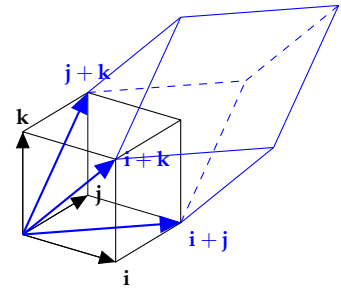
$$\mathbf{abc} = \mathbf{bca} = \mathbf{cab} = -\mathbf{acb} = -\mathbf{cba} = -\mathbf{bac}.$$

**1.34. PÉLDA (VEGYES SZORZAT).** Határozzuk meg egy egységélű kocka egy csúcsból induló három lapátló-vektorának vegyes szorzatát (1.33 ábra)!

**MEGOLDÁS.** Jelölje a kocka egyik csúcsából induló három élvektorát  $\mathbf{i}$ ,  $\mathbf{j}$  és  $\mathbf{k}$ . E három vektor ebben a sorrendben alkotson jobbrendszer. Ekkor az előző megjegyzés szerint  $\mathbf{ijk} = \mathbf{jki} = \mathbf{kij} = 1$ ,  $\mathbf{kji} = \mathbf{jik} = \mathbf{ikj} = -1$ . Mivel a vegyes szorzat egy paralelepipedon térfogatát vagy annak ellentettjét adja, ezért ha egy szorzatban egy vektor többször is szerepel, akkor annak értéke 0. Például  $\mathbf{iji} = (\mathbf{i} \times \mathbf{j}) \cdot \mathbf{i} = \mathbf{k} \cdot \mathbf{i} = 0$ . A három lapátló-vektor:  $\mathbf{i} + \mathbf{j}$ ,  $\mathbf{j} + \mathbf{k}$ ,  $\mathbf{k} + \mathbf{i}$ . Ezek vegyes szorzata

$$\begin{aligned} ((\mathbf{i} + \mathbf{j}) \times (\mathbf{j} + \mathbf{k})) \cdot (\mathbf{k} + \mathbf{i}) &= \mathbf{ijk} + \mathbf{iji} + \mathbf{ikk} + \mathbf{iki} + \mathbf{jjk} + \mathbf{jji} + \mathbf{jkk} + \mathbf{jki} \\ &= 1 + 0 + 0 + 0 + 0 + 0 + 0 + 1 \\ &= 2, \end{aligned}$$

tehát a három lapátló-vektor vegyes szorzata 2. Ez azt is jelenti, hogy e három vektor által kifeszített paralelepipedon térfogata 2.  $\square$



1.33. ábra:  $(\mathbf{i} + \mathbf{j})(\mathbf{j} + \mathbf{k})(\mathbf{k} + \mathbf{i}) = 2$



## Feladatok

### Skaláris szorzás

A következő feladatokban megadott adatok alapján számítsuk ki az  $\mathbf{a} \cdot \mathbf{b}$  skaláris szorzatot! Legyen  $\gamma = (\mathbf{a}, \mathbf{b})_{\angle}$ .

1.14.  $|\mathbf{a}| = 1, |\mathbf{b}| = 2, \gamma = \frac{\pi}{3}$ .

1.15.  $|\mathbf{a}| = \sqrt{2}, |\mathbf{b}| = 2, \gamma = \frac{3\pi}{4}$ .

1.16.  $|\mathbf{a}| = 1, |\mathbf{b}| = 2, \gamma = \pi$ .

1.17.  $|\mathbf{a}| = \sqrt{2}, |\mathbf{b}| = 2, \gamma = \frac{\pi}{2}$ .

1.18. Igazoljuk, hogy általában

$$(\mathbf{a} \cdot \mathbf{b})\mathbf{c} \neq \mathbf{a}(\mathbf{b} \cdot \mathbf{c}).$$

1.19. **SKALÁRIS SZORZÁS: IGAZ – HAMIS** Melyek igazak, melyek hamisak az alábbi állítások közül?

- Két egységvektor skaláris szorzata  $-1$  és  $+1$  közé eső szám.
- Egy  $\mathbf{v}$  vektor szorzata egy egységvektorral megegyezik  $\mathbf{v}$ -nek az egységvektor egyenesére eső merőleges vetületével.
- A skaláris szorzás kommutatív.
- A skaláris szorzás asszociatív (ld. 1.18. feladat).
- A nullvektor bármely vektorra merőleges.
- Két vektor pontosan akkor merőleges, ha skaláris szorzatuk 0.

Egyszerűsítsük az alábbi két kifejezést!

1.20.  $(\mathbf{a} + \mathbf{b}) \cdot (\mathbf{a} - \mathbf{b})$

1.21.  $(\mathbf{a} + 2\mathbf{b}) \cdot \mathbf{a} - 2\mathbf{a} \cdot \mathbf{b}$

1.22. Mennyi az  $\mathbf{a}$  és  $\mathbf{b}$  szöge, ha  $|\mathbf{a}| = 3, |\mathbf{b}| = 4, |\mathbf{a} + \mathbf{b}| = 5$ .

1.23. Igaz-e, hogy  $|\mathbf{a} + \mathbf{b} + \mathbf{c}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2 + |\mathbf{c}|^2$  pontosan akkor áll fenn, ha  $\mathbf{a}, \mathbf{b}$  és  $\mathbf{c}$  három egymásra páronként merőleges vektor?

1.24. Határozzuk meg az  $\mathbf{e}_1\mathbf{e}_2 + \mathbf{e}_1\mathbf{e}_3 + \mathbf{e}_2\mathbf{e}_3$  értékét, ha  $\mathbf{e}_1, \mathbf{e}_2$  és  $\mathbf{e}_3$  egységvektorok és  $\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 = \mathbf{0}$ .

### Merőlegesség és orientáció: vektori szorzás

1.25. Számítsuk ki  $|\mathbf{a} \times \mathbf{b}|$  értékét, ha  $|\mathbf{a}| = 1, |\mathbf{b}| = 2, (\mathbf{a}, \mathbf{b})_{\angle} = \frac{\pi}{6}$ .

1.26. Számítsuk ki  $\mathbf{a} \times \mathbf{b}$  értékét, ha  $|\mathbf{a}| = 1, |\mathbf{b}| = 2, (\mathbf{a}, \mathbf{b})_{\angle} = \pi$ .

Egyszerűsítsük az alábbi két kifejezést!

1.27.  $(\mathbf{a} + \mathbf{b}) \times (\mathbf{a} - \mathbf{b})$

1.28.  $(\mathbf{i} + \mathbf{j} + \mathbf{k}) \times (\mathbf{i} + \mathbf{j})$

1.29. Tekintsünk egy egységélű kockát, melynek egyik csúcsát jelölje  $P$ . Számítsuk ki a  $P$ -ből induló

- valamelyik két lapátló-vektor

- egyik lapátló- és a testátló-vektor

skaláris szorzatát, valamint a  $P$ -ből induló

- valamelyik élvektor és egy vele egy lapon lévő lapátló-vektor,
- valamelyik élvektor és a vele nem egy lapon lévő lapátló-vektor vektori szorzatát.

1.30. Mutassuk meg, hogy ha  $\mathbf{u}$  merőleges a  $\mathbf{v}$  és  $\mathbf{w}$  vektorokra, akkor merőleges minden lineáris kombinációjukra is.

1.31. Három lineárisan független vektornak hány sorrendje van? E sorrendek közül hány alkot jobb- és hány balrendszert?

1.32. Igazoljuk az 1.31. tétel állításait!

1.33. **SZÖGFELEZŐ** Legyenek  $\mathbf{a}$  és  $\mathbf{b}$  nemzérus vektorok. Mutassuk meg, hogy a  $|\mathbf{b}|\mathbf{a} + |\mathbf{a}|\mathbf{b}$  vektor felezi  $\mathbf{a}$  és  $\mathbf{b}$  szögét!

1.34. **HÁROMSZÖG SZÖGFELEZŐJE** Az előző feladat eredményét felhasználva mutassuk meg, hogy a háromszög egyik szögének szögfelezője a szemközti oldalt a két szomszédos oldal hosszának arányában osztja fel.

1.35. **MIT CSERÉL FÖL A TÜKÖR?** Hogy lehet az, hogy a tükrör fölcseréli a jobbat a ballal, de nem cseréli föl a föntet a lenttel?

### Projekt: ekvivalencia reláció

Egy  $X$  halmazon értelmezett (bináris) reláción az  $X$  elem-párjainak egy  $R$  halmazát értjük. Ha egy  $(a, b)$  pár benne van ebben a halmazban, azt mondjuk, hogy  $a$  az  $R$  relációban van  $b$ -vel, és úgy jelöljük, hogy  $a R b$ . Például, ha  $X$  az összes valaha élt ember halmaza, akkor az összes olyan  $(a, b)$  emberpár halmaza, ahol  $a$  anyja  $b$ -nek, egy reláció (anya-gyermek reláció). Ha  $X$  a valósok halmaza, és  $R$  azokból az  $(a, b)$  párokból áll, melyekre  $a$  kisebb vagy egyenlő mint  $b$ , akkor  $R$  egy reláció, melyet a valósok rendezési relációjának nevezünk. E reláció szokásos jele  $\leq$ , így ha  $(a, b) \in R$ , akkor az  $a \leq b$  jelölést használjuk.

Egy halmaz diszjunkt részhalmazok uniójára való felbontását szokás  $s$  halmaz elemei osztályozásának (particionálásának) nevezni. Egy ilyen osztályozáshoz természetes módon hozzárendelhető egy reláció, melyet a halmazon értelmezett ekvivalencia relációnak nevezünk, és amelyben két elem pontosan akkor van relációban (pontosan akkor ekvivalensek), ha azonos osztályba tartoznak. Kérdés, egy relációról hogyan állapítható meg, hogy ekvivalenciareláció-e?

**1.35. TÉTEL (EKVIVALENCIARELÁCIÓ).** Legyen  $R$  egy tetszőleges reláció az  $X$  halmazon.  $R$  pontosan akkor ekvivalenciareláció, ha tetszőleges  $a, b, c \in X$  elemre fennáll az alábbi három tulajdonság:

- a)  $R$  reflexív, azaz  $a R a$ , vagyis minden elem relációban van önmagával,
- b)  $R$  szimmetrikus, azaz ha  $a R b$ , akkor  $b R a$ ,
- c)  $R$  tranzitív, azaz ha  $a R b$  és  $b R c$ , akkor  $a R c$ .

**1.36.** Legyen  $R$  a fenti tétel szerinti reláció, és jelölje  $R_a$  az  $a$ -val relációban lévő elemek halmazát. Mutassuk meg,

hogy bármely két  $a, b \in X$  elemre  $R_a$  és  $R_b$  vagy azonos, vagy diszjunkt. Ezzel bizonyítsuk az előző tételt!

**1.37. SZABAD VEKTOR FOGALMA** Mutassuk meg, hogy a 3-dimenziós tér szabad vektorai definiálhatók egy – az irányított szakaszok halmazán értelmezett – ekvivalenciareláció ekvivalenciaosztályaival. Mi ez a reláció?

**1.38. VEKTOR IRÁNYA** Milyen halmazon értelmezett ekvivalenciareláció segítségével definiálható a vektor irányának és állásának fogalma?



## Vektorok koordinátás alakban

A koordináták bevezetésével egyrészt új algebrai eszközökhöz jutunk a vektorok és a különféle geometriai alakzatok vizsgálatában, másrészt lehetővé válik a vektor fogalmának kiterjesztése. Így jutunk a sokdimenziós terek fogalmához, ami nélkülözhetetlen a közgazdaságtanban, az internetes keresők matematikájában, vagy véges struktúrák fölötti változatában a kódelméletben és a kriptográfiában.

**Descartes-féle koordinátarendszer** Descartes 1637-ben *La Géométrie* című művében egy szép ötlettel összekapcsolta a geometriát az algebraival. Alapgondolata az volt, hogy a geometria alapelemei (pl. pontok) és a valós számok/számpárok/számhármasok közt kölcsönösen egyértelmű megfeleltetés hozható létre, így bizonyos geometriai alakzatok algebrai egyenletekkel leírhatók és vizsgálhatók válnak.

Az 1.11. tétel szerint a sík bármely  $\mathbf{v}$  vektora felírható két adott lineárisan független  $\mathbf{e}_1$  és  $\mathbf{e}_2$  vektor lineáris kombinációjaként, és e felírás egyértelmű. Ha e lineáris kombináció  $\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2$  alakú, akkor a  $\mathbf{v}$  vektorhoz a  $(v_1, v_2)$  számpárt rendeljük, és ezt a  $\mathbf{v}$  vektor *koordinátás alakjának*, a  $v_1$  és  $v_2$  skalárokat pedig a  $\mathbf{v}$  *koordinátáinak* nevezzük. Azt mondjuk, hogy az  $\{\mathbf{e}_1, \mathbf{e}_2\}$  vektorpár e koordinátázási rendszer – egyszerűbben *koordinátarendszer* – *bázisa*, az  $\mathbf{e}_1$  és  $\mathbf{e}_2$  vektorok a *bázisvektorok* vagy *alapvektorok*. Tetszőleges vektor koordinátáinak meghatározásához elég a bázisvektorokat ismerni.

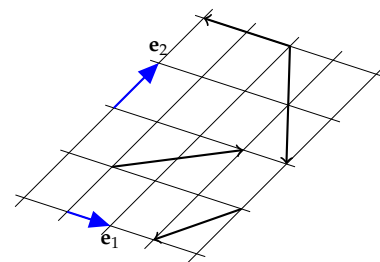
**1.36. PÉLDA (VEKTOROK KOORDINÁTÁI).** Határozzuk meg az 1.34. ábrán megadott vektoroknak – az  $\mathbf{e}_1$  és  $\mathbf{e}_2$  vektorokra, mint bázisra vonatkozó – koordinátáit!

**MEGOLDÁS.** A megoldás könnyen leolvasható az 1.35. ábráról. Áttekinthetőbb, ha az összes vektort egyetlen pontból indítjuk. Ezt mutatja az 1.36. ábra. □

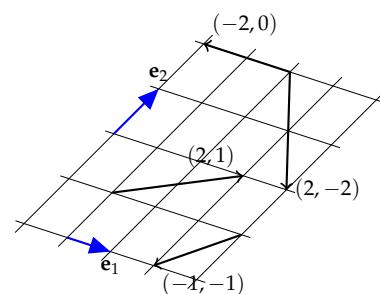
A koordinátarendszer a 3-dimenziós térben is hasonló módon építhető fel. Az 1.12. tétel szerint a tér bármely  $\mathbf{v}$  vektora felírható három adott lineárisan független  $\mathbf{e}_1$ ,  $\mathbf{e}_2$  és  $\mathbf{e}_3$  vektor lineáris kombinációjaként, és e felírás egyértelmű. Ha e lineáris kombináció  $\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3$  alakú, akkor a  $\mathbf{v}$  vektorhoz a  $(v_1, v_2, v_3)$  számhármaszt rendeljük, és ezt a  $\mathbf{v}$  vektor *koordinátás alakjának*, a  $v_1$ ,  $v_2$  és  $v_3$  skalárokat pedig a  $\mathbf{v}$  *koordinátáinak* nevezzük. Bázis az  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$  vektorhármas.

Sőt, a koordinátázás az 1-dimenziós térben is megvalósítható: ha  $\mathbf{e}$  egy nemnulla vektor (tehát lineárisan független vektorrendszer!), akkor bármely vele párhuzamos  $\mathbf{v}$  vektor egyértelműen felírható  $\mathbf{v} = v\mathbf{e}$  alakban. E  $v$  skalár lesz a  $\mathbf{v}$  koordinátás alakja (a zárójel használata

René Descartes (Renatus Cartesius) (1596–1650) francia filozófus és matematikus, a modern filozófia atyja, az analitikus geometria egyik megalkotója. Filozófiáját a pusztá hitre alapozott állításokkal szemben a racionális érvelések útján kívánta fölépíteni (lásd *descartes-i kételkedés* és „gondolkodom, tehát vagyok”). Orvostudományt és jogot tanult, végül hadmérnöki képesítést szerzett. Több háborúban is részt vett. 1619-ben egy Magyarországot is érintő hosszú útján egy Ulm melletti parasztházban három álmot látott, melyek megfejtése „egy csodálatos tudományhoz” vezette, ami filozófiája alapjává vált.



1.34. ábra: Mik a vektorok koordinátái?



1.35. ábra: A megoldás

itt szükségtelen). Így a  $v \leftrightarrow v$  hozzárendelés kölcsönösen egyértelmű megfeleltetést létesít a vektorok és a skalárok közt.

Ha kijelölünk egy pontot az egyenesen/síkban/térben – ez lesz az origó –, akkor az egyenes/sík/tér pontjai és a helyvektorok végpontjai közti kölcsönösen egyértelmű megfeleltetéssel egyúttal a pontok is koordinátát kapnak.

**1.37. PÉLDA (PONTOK KOORDINÁTÁI).** Határozzuk meg az 1.37. ábrán kijelölt pontok – megadott bázisvektorokra és origóra vonatkozó – koordinátáit!

**MEGOLDÁS.** E feladat megoldása lényegében azonos az előzőével, mint hogy a kijelölt pontokba mutató helyvektorok megegyeznek az ott megadott vektorokkal (ld. 1.38. ábra).  $\square$

A helyvektorok és a pontok közti kölcsönösen egyértelmű megfeleltetést a jelölésben is kifejezzük azzal, hogy nem teszünk különbséget a vektor és a pont koordinátás alakja közt, azaz a  $v = (a, b)$  vektorhoz adott origó mellett rendelt pontot is  $(a, b)$  jelöli. Ha a pontnak nevet is adunk, pl. e pont a  $P$  pont, akkor a  $P(a, b)$  jelölés a nevet és a koordinátákat is megadja. Ekkor a helyvektort  $\overrightarrow{OP}$  is jelölheti. Tehát  $v = \overrightarrow{OP}$ . Szokás a vektorok koordinátás alakját úgynevezett *oszlopvektor* alakba is írni, mi e könyvben ekkor kerek helyett szögletes zárójelet használunk, például:

$$\overrightarrow{OP} = \begin{bmatrix} a \\ b \end{bmatrix}.$$

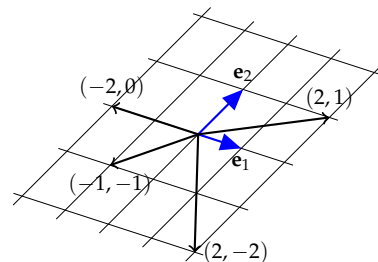
E jelölés előnyeivel hamarosan találkozunk.

Látható, hogy ha egy pont az első tengelyen van, és azon az egyenesen  $x$  az 1-dimenziós koordinátája, akkor síkbeli koordinátás alakja  $(x, 0)$  lesz. Hasonlóképp a második tengely minden pontjának  $(0, y)$  a koordinátás alakja. Az origóé  $(0, 0)$  (ld. 1.39. ábra). Az alapvektorok koordinátás alakja  $e_1 = (1, 0)$  és  $e_2 = (0, 1)$ .

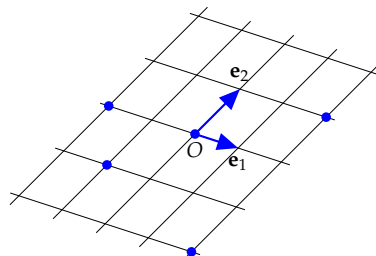
Hasonlóképp a 3-dimenziós esetben a koordinátatengelyekre eső pontok 3-dimenziós koordinátás alakja  $(x, 0, 0)$ ,  $(0, y, 0)$ , illetve  $(0, 0, z)$  attól függően, hogy melyik tengelyről van szó. Az origón átmenő és 2 tengelyt tartalmazó síkokat *koordinátasíkoknak* nevezzük. Ezekből is három van. Könnyen látható, hogy a koordinátasíkok pontjainak alakja  $(x, y, 0)$ ,  $(x, 0, z)$ , illetve  $(0, y, z)$ . Az origóé  $(0, 0, 0)$ , míg az alapvektoroké  $e_1 = (1, 0, 0)$ ,  $e_2 = (0, 1, 0)$ ,  $e_3 = (0, 0, 1)$  (ld. 1.40. ábra).

**Műveletek koordinátás alakban megadott vektorokkal** Legyen adva a térben egy koordinátarendszer és abban két tetszőleges  $u = (u_1, u_2, u_3)$  és  $v = (v_1, v_2, v_3)$  vektor. E paragrafusban megkeressük a vektorműveletek koordinátás alakját. A kérdés tehát az, hogy hogyan kapható meg  $u + v$ ,  $u - v$ ,  $cu$ ,  $u \cdot v$ ,  $u \times v$  koordinátás alakja.

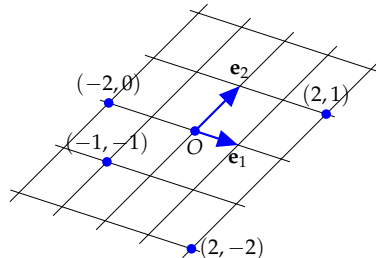
Mindegyik műveletnél felhasználjuk a vektoroknak a bázisvektorok



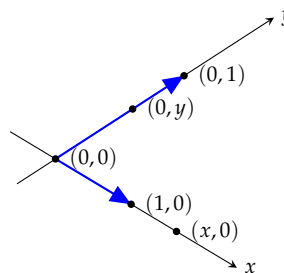
1.36. ábra: A megoldás helyvektorokkal ábrázolva.



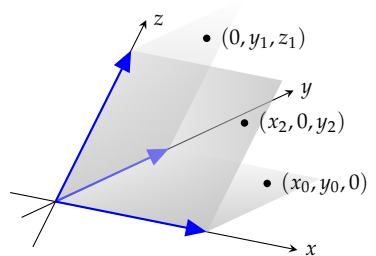
1.37. ábra: Mik a pontok koordinátái?



1.38. ábra: Pontok és koordinátáik



1.39. ábra: Pontok a koordinátarendszer tengelyein.



1.40. ábra: Pontok a koordinátasíkokon

lineáris kombinációjaként való előállítását. Az adott két vektor összege:

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= (u_1, u_2, u_3) + (v_1, v_2, v_3) \\ &= (u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + u_3\mathbf{e}_3) + (v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3) \\ &= (u_1 + v_1)\mathbf{e}_1 + (u_2 + v_2)\mathbf{e}_2 + (u_3 + v_3)\mathbf{e}_3 \\ &= (u_1 + v_1, u_2 + v_2, u_3 + v_3).\end{aligned}$$

Hasonló képlet adódik a különbségre:

$$\mathbf{u} - \mathbf{v} = (u_1, u_2, u_3) - (v_1, v_2, v_3) = (u_1 - v_1, u_2 - v_2, u_3 - v_3).$$

A skalárral való szorzás is a koordinátánként való végrehajtás lehetőségét mutatja:

$$\begin{aligned}c\mathbf{u} &= c(u_1, u_2, u_3) = c(u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + u_3\mathbf{e}_3) \\ &= cu_1\mathbf{e}_1 + cu_2\mathbf{e}_2 + cu_3\mathbf{e}_3 \\ &= (cu_1, cu_2, cu_3).\end{aligned}$$

Összefoglalva tehát a következő állítást kapjuk:

**1.38. ÁLLÍTÁS (VEKTORMŰVELETEK KOORDINÁTÁS ALAKJA).** *Adva van a térben egy koordinátarendszer és abban két tetszőleges  $\mathbf{u} = (u_1, u_2, u_3)$  és  $\mathbf{v} = (v_1, v_2, v_3)$  vektor, valamint egy tetszőleges  $c \in \mathbb{R}$  valós szám. Ekkor a vektorok összegének, különbségének és skalárszorosának koordinátás alakja rendre*

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= (u_1, u_2, u_3) + (v_1, v_2, v_3) = (u_1 + v_1, u_2 + v_2, u_3 + v_3), \\ \mathbf{u} - \mathbf{v} &= (u_1, u_2, u_3) - (v_1, v_2, v_3) = (u_1 - v_1, u_2 - v_2, u_3 - v_3), \\ c\mathbf{u} &= c(u_1, u_2, u_3) = (cu_1, cu_2, cu_3).\end{aligned}$$

Az oszlopvektor jelölést használva

$$\mathbf{u} \pm \mathbf{v} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \pm \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} u_1 \pm v_1 \\ u_2 \pm v_2 \\ u_3 \pm v_3 \end{bmatrix}, \quad c\mathbf{u} = c \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} cu_1 \\ cu_2 \\ cu_3 \end{bmatrix}.$$

A síkbeli vektorokra hasonló állítások igazak, csak két koordinátával. Érdekesebb a helyzet a skaláris szorzással. Látni fogjuk, hogy a skaláris szorzás koordinátás alakja függ a koordinátarendszertől.

**1.39. PÉLDA (SKALÁRIS SZORZÁS KOORDINÁTARENDSZERBEN).** *Tekintsünk egy olyan síkbeli koordinátarendszert, ahol az első alapvektor hossza 1, a másodiké 2, és a kettőjük közti szög  $\pi/3$ . Számítsuk ki az  $\mathbf{a} = (1, 1)$  és  $\mathbf{b} = (-5/2, 1)$  vektorok skaláris szorzatát.*

**MEGOLDÁS.** Az alapvektorok hosszát és szögét ismerve ki tudjuk számítani az alapvektorok skaláris szorzatait:

$$\mathbf{e}_1 \cdot \mathbf{e}_1 = 1, \quad \mathbf{e}_2 \cdot \mathbf{e}_2 = 2^2 = 4, \quad \mathbf{e}_1 \cdot \mathbf{e}_2 = 1 \cdot 2 \cdot \cos \frac{\pi}{3} = 1.$$

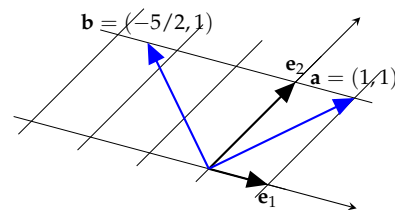
Ezt fölhasználva kapjuk, hogy

$$\begin{aligned} \mathbf{a} \cdot \mathbf{b} &= (\mathbf{e}_1 + \mathbf{e}_2) \cdot \left(-\frac{5}{2}\mathbf{e}_1 + \mathbf{e}_2\right) \\ &= -\frac{5}{2}\mathbf{e}_1 \cdot \mathbf{e}_1 + \left(1 - \frac{5}{2}\right)\mathbf{e}_1 \cdot \mathbf{e}_2 + \mathbf{e}_2 \cdot \mathbf{e}_2 \\ &= -\frac{5}{2} - \frac{3}{2} + 4 \\ &= 0, \end{aligned}$$

tehát a két vektor merőleges egymásra (ld. az 1.41. ábrát).

Érdekességgént meghatározzuk e koordináta-rendszerben a skaláris szorzás általános képletét:

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= (u_1\mathbf{e}_1 + u_2\mathbf{e}_2) \cdot (v_1\mathbf{e}_1 + v_2\mathbf{e}_2) \\ &= u_1v_1\mathbf{e}_1 \cdot \mathbf{e}_1 + (u_1v_2 + u_2v_1)\mathbf{e}_1 \cdot \mathbf{e}_2 + u_2v_2\mathbf{e}_2 \cdot \mathbf{e}_2 \\ &= u_1v_1 + u_1v_2 + u_2v_1 + 4u_2v_2. \end{aligned}$$



1.41. ábra: Két vektor skaláris szorzata

**A derékszögű koordináta-rendszer** A természeti törvények különös fontosságot adnak az egymásra merőleges irányoknak, ezért például igen gyakran érdemes olyan koordináta-rendszert választani, amelyben az alapvektorok merőlegesek, más szóval *ortogonálisak* egymásra. A bázisvektorok szöge mellett azok hosszát is érdemes standardizálni, nevezetesen egységnyi hosszúnak választani, így mindegyik koordináta egyúttal távolságot is jelent. Az egységvektorokból álló ortogonális bázist *ortonormálnak* nevezzük.

Az egységes tárgyalás érdekében a bázisvektorok körüljárását is előírhatjuk: általánosan elterjedt szokás a jobbrendszert választani. Az így konstruált bázis vektorait síkban gyakran **i**, **j**, térben **i**, **j** és **k** jelöli.

A két és háromdimenziós térben a skaláris szorzat egyszerű alakot ölt, ha a koordináta-rendszer alapvektorai ortonormáltak.

**1.40. ÁLLÍTÁS (SKALÁRIS SZORZAT ORTONORMÁLT KOORDINÁTARENDSZERBEN).** A síkbeli  $\mathbf{u} = (u_1, u_2)$  és  $\mathbf{v} = (v_1, v_2)$ , illetve a térbeli  $\mathbf{u} = (u_1, u_2, u_3)$  és  $\mathbf{v} = (v_1, v_2, v_3)$  vektorok skaláris szorzata ortonormált koordináta-rendszerben

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2, \text{ illetve } \mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + u_3v_3.$$

**BIZONYÍTÁS.** A síkbeli esetben kihasználjuk, hogy  $\mathbf{i} \cdot \mathbf{i} = \mathbf{j} \cdot \mathbf{j} = 1$  és  $\mathbf{i} \cdot \mathbf{j} = 0$ :

$$\begin{aligned}\mathbf{u} \cdot \mathbf{v} &= (u_1 \mathbf{i} + u_2 \mathbf{j}) \cdot (v_1 \mathbf{i} + v_2 \mathbf{j}) \\ &= u_1 v_1 \mathbf{i} \cdot \mathbf{i} + (u_1 v_2 + u_2 v_1) \mathbf{i} \cdot \mathbf{j} + u_2 v_2 \mathbf{j} \cdot \mathbf{j} \\ &= u_1 v_1 + u_2 v_2\end{aligned}$$

A térbeli eset hasonlóan bizonyítható. □

**1.41. ÁLLÍTÁS (VEKTORI SZORZAT ORTONORMÁLT KOORDINÁTARENDSZERBEN).** A térbeli  $\mathbf{a} = (a_1, a_2, a_3)$  és  $\mathbf{b} = (b_1, b_2, b_3)$  vektorok vektori szorzata derékszögű koordinátarendszerben

$$\mathbf{a} \times \mathbf{b} = (a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1).$$

► Az  $\mathbf{a} \times \mathbf{b}$  koordinátáinak könnyű memorizálására két sémát mutatunk a széljegyzetben.

**BIZONYÍTÁS.** Az alapvektorok egymással való vektori szorzatait már kiszámoltuk az 1.28. példában. Kihasználva, hogy  $\mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = \mathbf{0}$ ,  $\mathbf{i} \times \mathbf{j} = \mathbf{k}$ ,  $\mathbf{j} \times \mathbf{i} = -\mathbf{k}$ , ..., a következőt kapjuk:

$$\begin{aligned}\mathbf{a} \times \mathbf{b} &= (a_1 \mathbf{i} + a_2 \mathbf{j} + a_3 \mathbf{k}) \times (b_1 \mathbf{i} + b_2 \mathbf{j} + b_3 \mathbf{k}) \\ &= a_2 b_3 \mathbf{j} \times \mathbf{k} + a_3 b_2 \mathbf{k} \times \mathbf{j} + a_3 b_1 \mathbf{k} \times \mathbf{i} + a_1 b_3 \mathbf{i} \times \mathbf{k} + a_1 b_2 \mathbf{i} \times \mathbf{j} + a_2 b_1 \mathbf{j} \times \mathbf{i} \\ &= a_2 b_3 \mathbf{i} - a_3 b_2 \mathbf{i} + a_3 b_1 \mathbf{j} - a_1 b_3 \mathbf{j} + a_1 b_2 \mathbf{k} - a_2 b_1 \mathbf{k} \\ &= (a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1).\end{aligned}$$
 □

**1.42. PÉLDA (PARALLELOGRAMMA TERÜLETE).** Mutassuk meg, hogy az  $(a, b)$  és  $a(c, d)$  vektorok által kifeszített paralelogramma területe

$$|ad - bc|.$$

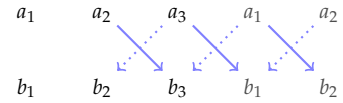
Mi a jelentése az  $ad - bc$  előjelének?

**MEGOLDÁS.** Két 3-dimenziós vektor által kifeszített paralelogramma területe a vektori szorzatuk abszolút értéke. Ággyazzuk be a megadott két vektort a tér egyik koordinátságjába, tekintsük például az  $(a, b, 0)$  és  $(c, d, 0)$  vektorokat. Vektori szorzatuk

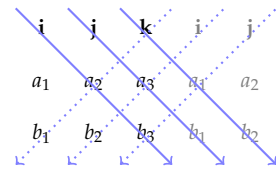
$$(a, b, 0) \times (c, d, 0) = (0, 0, ad - bc),$$

ennek abszolút értéke  $|ad - bc|$ .

Mivel az  $(a, b, 0)$ ,  $(c, d, 0)$  és  $(0, 0, ad - bc)$  vektorok jobbrendszert alkotnak, ezért  $ad - bc$  pontosan akkor pozitív, ha a síkban az  $(a, b)$  és  $(c, d)$  vektorok jobbrendszert alkotnak, és  $ad - bc$  pontosan akkor negatív, ha az  $(a, b)$  és  $(c, d)$  vektorok balrendszert alkotnak (gondoljuk meg!). □



$$a) \quad (a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1)$$



$$b) \quad (a_2 b_3 - a_3 b_2) \mathbf{i} + (a_3 b_1 - a_1 b_3) \mathbf{j} + (a_1 b_2 - a_2 b_1) \mathbf{k}$$

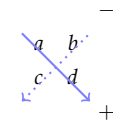
1.42. ábra: A vektori szorzat kiszámítása a két vektor koordinátáiból:

a) Írjuk a két vektort egymás alá, majd az első két koordinátát másoljuk a vektorok végére, végül az X alakba rakott nyíl pároknál a ↘ nyíl végein lévő számok szorzatából vonjuk ki a ↙ szerinti szorzatot. Az eredmény:

$$(a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1).$$

b) Írjuk a két vektor koordinátái fölé az  $\mathbf{i}$ ,  $\mathbf{j}$ ,  $\mathbf{k}$  vektorokat, majd az előzőhöz hasonlóan másoljuk e táblázat után az első két oszlopot, végül a ↘ nyíl menti szorzatokból vonjuk ki a ↙ menti szorzatokat. Megjegyezzük, hogy e séma a később tanulandó determinánsok kiszámítására emlékeztet, formálisan föl is szokás írni a determinánsokat jelölő függőleges zárójelek segítségével:

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}.$$



1.43. ábra: A paralelogramma előjeles területe  $ad - bc$ , melynek memorizálására a fenti séma használatos. Ez megegyezik két 2-dimenziós vektor – később tanulandó – determinánsával, melyet úgy jelöljük, hogy a két vektor koordinátáiból képzett táblázatot függőleges zárójelek közé zárjuk:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

Vigyázzunk, e jel nem az abszolút értéket jelöli, ahhoz egy további zárójelpár szükséges, azaz

$$|ad - bc| = \left| \begin{vmatrix} a & b \\ c & d \end{vmatrix} \right|.$$

1.43. PÉLDA (PARALELEPIPEDON TÉRFOGATA). Mutassuk meg, hogy az  $\mathbf{a} = (a_1, a_2, a_3)$ ,  $\mathbf{b} = (b_1, b_2, b_3)$  és  $\mathbf{c} = (c_1, c_2, c_3)$  vektorok által kifeszített paralelepipedon térfogata

$$|a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_1b_3c_2 - a_2b_1c_3 - a_3b_2c_1|$$

Mi a jelentése az  $a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_1b_3c_2 - a_2b_1c_3 - a_3b_2c_1$  előjelének?

**MEGOLDÁS.** Az 1.32. példában láttuk, hogy a három vektor által kifeszített paralelepipedon térfogata  $|(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}|$ , amit a koordinátás alakból kiszámolhatunk:

$$(a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1) \cdot (c_1, c_2, c_3) = a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_1b_3c_2 - a_2b_1c_3 - a_3b_2c_1. \quad \square$$

**Az  $\mathbb{R}^n$  halmaz** Láttuk, hogy 2-dimenziós, illetve 3-dimenziós vektorjellegű mennyiségek leírhatók egy rendezett számpárral, illetve számhármassal. Izgalmas a fordított helyzet, amikor legalább 4, de akár több millió szorosan összefüggő adatból képzett, rendezett szám- $n$ -essel dolgozunk. Vajon értelmes dolog-e e szám- $n$ -eket egy  $n$ -dimenziós tér vektorainak, vagy pontjainak tekinteni? És van-e értelme a 2- és 3-dimenziós térben használt fogalmak általánosításának  $n$  dimenzióra? A válasz mindegyik kérdésre határozott igen, amit a fizika 4-dimenziós tér-idő fogalma, számtalan gazdasági, vagy internettel kapcsolatos probléma megoldása fényesen bizonyít.

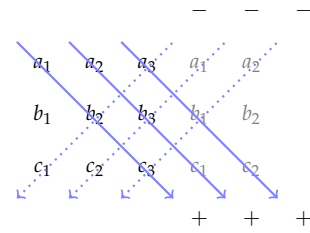
1.44. DEFINÍCIÓ. Egy tetszőleges  $H$  halmaz elemeiből képzett rendezett elem- $n$ -esek halmazát  $H^n$ -nel jelöljük.

Például ha  $H = \{0, 1\}$ , akkor  $H^3$  a  $H$  elemeiből képzett rendezett elemhármassok halmaza, azaz:

$$\{(0, 0, 0), (0, 0, 1), (0, 1, 0), (1, 0, 0), (0, 1, 1), (1, 0, 1), (1, 1, 0), (1, 1, 1)\}.$$

A fenti jelölésnek megfelelően  $\mathbb{R}^n$  a valós számokból képzett rendezett szám- $n$ -esek halmazát jelöli. Eszerint a sík pontjait és vektorait  $\mathbb{R}^2$ , a térét  $\mathbb{R}^3$  elemeivel koordinátáztuk. Később  $\mathbb{R}^n$  elemein vektorműveleteket fogunk bevezetni, és  $\mathbb{R}^n$ -ről, mint vektortérrel fogunk beszélni. Hasonlóképp,  $\mathbb{R}^n$ -t geometriai vagy ponttérnek fogjuk tekinteni, ha elemeire, mint pontokra gondolunk, és köztük geometriai műveleteket végezzünk. E sokféleség sosem fog gondot okozni:  $\mathbb{R}^n$  szerepét mindig az fogja meghatározni, hogy mit teszünk elemeivel, vagyis a szám- $n$ -esekkel.

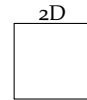
Az  $\mathbb{R}^n$  megismerésében az *analógia* fonalán haladunk, a 2- és 3-dimenziós tér fogalmait fogjuk átvinni, általánosítani  $n$  dimenzióra. Ez



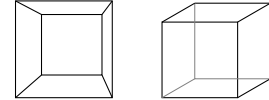
1.44. ábra: A paralelepipedon térfogata megegyezik az azt kifeszítő három vektor vegyes szorzatának abszolút értékével. A vektori szorzáshoz használt séma itt is működik: a három vektort egymás alá írjuk, az így kapott táblázat első két oszlopát utána másoljuk, végül a  $\searrow$  nyíl menti szorzatokból kivonjuk a  $\swarrow$  menti szorzatokat. A később tanulandó determinánsokra használt jelöléssel tehát a paralelepipedon előjeles térfogata:

$$\mathbf{abc} = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

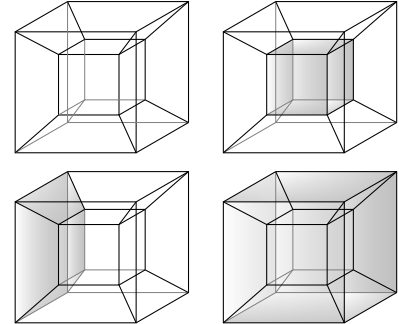
1D



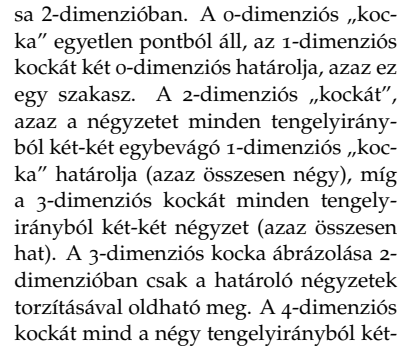
2D



3D



4D



1.45. ábra: 4-dimenziós kocka ábrázolása 2-dimenzióban. A 0-dimenziós „kocka” egyetlen pontból áll, az 1-dimenziós kockát két 0-dimenziós határolja, azaz ez egy szakasz. A 2-dimenziós „kockát”, azaz a négyzetet minden tengelyirányból két-két egybevágó 1-dimenziós „kocka” határolja (azaz összesen négy), míg a 3-dimenziós kockát minden tengelyirányból két-két négyzet (azaz összesen hat). A 3-dimenziós kocka ábrázolása 2-dimenzióban csak a határoló négyzetek torzításával oldható meg. A 4-dimenziós kockát mind a négy tengelyirányból két-két 3-dimenziós kocka határolja, összesen nyolc. Az ábrán három ilyen 3-dimenziós kockát kiszíneztünk.

az analógia fog segíteni abban, hogy valamit „lássunk”  $n$  dimenzióban is (ha nem is olyan jól, mint 3 dimenzióban). Példaként az analógiára egy 4-dimenziós kocka 2-dimenziós vetületét mutatjuk az 1.45. ábrán.

**Vektorok összeadása és skalárral szorzása  $\mathbb{R}^n$ -ben** A 2- és 3-dimenziós vektorok műveleteinek koordinátás alakja az összeadás, kivonás és skalárral szorzás esetén analóg módon átvihető az  $n$ -dimenziós vektorokra.

**1.45. DEFINÍCIÓ (VEKTORMŰVELETEK  $\mathbb{R}^n$ -BEN).** Legyen  $c \in \mathbb{R}$  egy tetszőleges valós,  $\mathbf{u} = (u_1, u_2, \dots, u_n)$  és  $\mathbf{v} = (v_1, v_2, \dots, v_n)$  az  $\mathbb{R}^n$  tér két tetszőleges vektora. E két vektor összegét és egyikük  $c$ -szeresét a következő képletekkel definiáljuk:

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n) \\ c\mathbf{u} &= (cu_1, cu_2, \dots, cu_n).\end{aligned}$$

Összefoglaljuk e műveletek legfontosabb tulajdonságait. Ezekre többször is hivatkozni fogunk a későbbiekben.

**1.46. TÉTEL (AZ ÖSSZEADÁS ÉS SKALÁRRAL SZORZÁS TULAJDONSÁGAI).** Legyen  $\mathbf{u}, \mathbf{v}$  és  $\mathbf{w}$  az  $\mathbb{R}^n$  három tetszőleges vektora, és legyen  $c, d$  két tetszőleges valós, jelölje  $\mathbf{0}$  a  $(0, 0, \dots, 0)$  vektort és  $-\mathbf{u}$  a  $(-u_1, -u_2, \dots, -u_n)$  vektort. Ekkor

|    |                                                                                   |                                                |
|----|-----------------------------------------------------------------------------------|------------------------------------------------|
| a) | $\mathbf{u} + \mathbf{v} \in \mathbb{R}^n$                                        | $a +$ művelet nem vezet ki $\mathbb{R}^n$ -ből |
| b) | $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$                               | $a$ művelet fölcserélhető (kommutatív)         |
| c) | $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$ | csoportosítható (asszociatív)                  |
| d) | $\mathbf{u} + \mathbf{0} = \mathbf{u}$                                            | zérusvektor                                    |
| e) | $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$                                         | ellentett vektor                               |
| f) | $c\mathbf{u} \in \mathbb{R}^n$                                                    | $e$ szorzás nem vezet ki $\mathbb{R}^n$ -ből   |
| g) | $c(d\mathbf{u}) = (cd)\mathbf{u}$                                                 | $a$ két szorzás kompatibilis                   |
| h) | $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$                          | disztributív                                   |
| i) | $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$                                   | disztributív                                   |
| j) | $1\mathbf{u} = \mathbf{u}$                                                        | szorzás 1-gyel                                 |

E tíz tulajdonság később kitüntetett szerepet fog játszani, ezért elkülönítjük a további felsorolandóktól:

- 1)  $0\mathbf{u} = \mathbf{0}$
- 2)  $(-1)\mathbf{u} = -\mathbf{u}$
- 3)  $\mathbf{u} - \mathbf{v} = \mathbf{u} + (-\mathbf{v})$
- 4)  $a$  skalár hátra is írható, azaz  $c\mathbf{u} = \mathbf{uc}$  és így  $\mathbf{u}/c = \frac{1}{c}\mathbf{u}$

► E tulajdonságok mindegyike könnyen visszavezethető a valós számok algebrai tulajdonságaira, ezért ezek ellenőrzését (bizonyítását) az



Olvasóra hagyjuk. Mintaként megmutatjuk a *b)* bizonyítását:

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n) \\ &= (v_1 + u_1, v_2 + u_2, \dots, v_n + u_n) \\ &= \mathbf{v} + \mathbf{u}.\end{aligned}$$

► Az *a)–e)* tulajdonságok az összeadás, az *f)–j)* tulajdonságok a skalárral szorzás tulajdonságait írják le. Mindkét csoport első tulajdonsága csak annyit mond, hogy a művelet eredménye is ugyanabba a vektorhalmazba, azaz  $\mathbb{R}^n$ -be esik, ahová a műveletben szereplő vektor való.

**1.47. PÉLDA.** Mutassuk meg, hogy az  $\mathbb{R}^n$ -beli  $\mathbf{e}_1 = (1, 0, \dots, 0)$ ,  $\mathbf{e}_2 = (0, 1, \dots, 0)$ , ...,  $\mathbf{e}_n = (0, 0, \dots, 1)$  vektorok lineárisan függetlenek, és hogy  $\mathbb{R}^n$  minden vektora egyértelműen előáll ezek lineáris kombinációjaként!

**MEGOLDÁS.** Az  $\mathbf{e}_1$  biztosan nem áll elő a többi vektor lineáris kombinációjaként, hisz a többi vektor első koordinátája 0, így azok bármely lineáris kombinációjában is 0 az első koordináta,  $\mathbf{e}_1$ -ben pedig 1. Hasonlóan igazolható, hogy egyik  $\mathbf{e}_i$  sem áll elő a többi vektor lineáris kombinációjaként ( $i = 2, 3, \dots, n$ ). A megadott vektorok tehát lineárisan függetlenek.

Mivel az  $i$ -edik koordináta egyedül csak az  $\mathbf{e}_i$  vektorban 1, a többiben 0, ezért ha egy tetszőleges  $\mathbf{v} = (v_1, v_2, \dots, v_n)$  vektor előáll az  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  vektorok lineáris kombinációjaként, akkor abban  $\mathbf{e}_i$  együtthatója csak  $v_i$  lehet. Másrészt az is világos, hogy

$$(v_1, v_2, \dots, v_n) = v_1 \mathbf{e}_1 + v_2 \mathbf{e}_2 + \dots + v_n \mathbf{e}_n.$$

Ezzel igazoltuk, hogy  $\mathbb{R}^n$  minden vektora egyértelműen áll elő az  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  vektorok lineáris kombinációjaként.  $\square$

**Lineáris kombináció, lineáris függetlenség, lineáris összefüggőség** Hiába definiáltuk vektorok lineáris függetlenségének fogalmát tetszőleges számú vektorból álló vektorhalmazra, láttuk, hogy a 3-dimenziós térben legföljebb csak 3 vektor lehet lineárisan független. Viszont az  $\mathbb{R}^n$  térben találtunk  $n$  lineárisan független vektort is.

Az 1.10. definíció szerint egy legalább kételemű vektorrendszer lineárisan független, ha mindegyik vektor független a többitől, azaz egyik sem fejezhető ki a többi lineáris kombinációjaként, egyetlen vektor pedig lineárisan független, ha nem a zérusvektor. Ez nehézkes feltétel, hisz mindegyik vektorra külön ellenőrizni kell, ezért egy könnyebben ellenőrizhető, de ekvivalens feltételt keresünk. A háromdimenziós térben láttuk, hogy ha három vektor független, akkor a tér bármely vektora egyértelműen előáll lineáris kombinációjuként. Ez igaz a nullvektorra is. A nullvektor egyféleképp biztosan előáll: a három vektor nullákkal vett lineáris kombinációjaként. Ezt nevezzük a nullvektor



triviális előállításának. A fentiek szerint más előállítása nincs is, ha a három vektor lineárisan független. Ez az alapja a következő tételnek:

- 1.48. TÉTEL (LINEÁRIS FÜGGETLENSÉG). *Tetszőleges  $\mathbb{R}^n$ -beli  $\mathcal{V} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$  vektorrendszerre az alábbi két állítás ekvivalens:*
1.  $\mathcal{V}$  lineárisan független, azaz  $k > 1$  esetén egyik vektora sem fejezhető ki a többi lineáris kombinációjaként,  $k = 1$  esetén pedig a vektor nem a zérusvektor.
  2. A zérusvektor csak egyféleképp – a triviális módon – áll elő  $\mathcal{V}$  lineáris kombinációjaként. Másként fogalmazva, a  $c_1, c_2, \dots, c_k$  skalárokkal vett lineáris kombináció csak akkor lehet a nullvektor, azaz

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_k\mathbf{v}_k = \mathbf{0}$$

csak akkor állhat fenn, ha

$$c_1 = c_2 = \dots = c_k = 0.$$

**BIZONYÍTÁS.** Először tegyük fel, hogy a vektorrendszer csak egyetlen  $\mathbf{v}$  vektorból áll. Ekkor a tétel azt állítja, hogy e vektor pontosan akkor lineárisan független, azaz pontosan akkor nem a nullvektor, ha a  $c\mathbf{v} = \mathbf{0}$  csak  $c = 0$  esetén állhat fenn. Ez nyilvánvaló, hisz ha  $\mathbf{v} \neq \mathbf{0}$  és  $c \neq 0$ , akkor  $c\mathbf{v} = \mathbf{0}$  sem állhat fenn. A továbbiakban tegyük fel, hogy a vektorrendszer legalább két vektorból áll. A következőkben kontrapozícióval bizonyítunk, azaz az  $A \Rightarrow B$  állítást a vele ekvivalens  $\neg B \Rightarrow \neg A$  állítással igazoljuk.

( $\Leftarrow$ ) Megmutatjuk, hogy ha  $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_k\mathbf{v}_k = \mathbf{0}$  csak  $c_1 = c_2 = \dots = c_k = 0$  esetén állhat fenn, akkor semelyik  $\mathbf{v}_i$  vektor sem fejezhető ki a többi lineáris kombinációjaként ( $i = 1, 2, \dots, k$ ). Tegyük fel, hogy valamelyik vektor – például a  $\mathbf{v}_1$  – kifejezhető a többi lineáris kombinációjaként, azaz

$$\mathbf{v}_1 = d_2\mathbf{v}_2 + \dots + d_k\mathbf{v}_k,$$

vagyis átrendezés után

$$(-1)\mathbf{v}_1 + d_2\mathbf{v}_2 + \dots + d_k\mathbf{v}_k = \mathbf{0}.$$

Mivel  $\mathbf{v}_1$  együtthatója nem 0, így elő tudtuk állítani a nullvektort olyan lineáris kombinációjaként, melyben nem minden együttható 0.

( $\Rightarrow$ ) Megmutatjuk, hogy ha a vektorrendszer egyik vektora sem áll elő a többi lineáris kombinációjaként, akkor egyedül csak a csupa zérus együtthatójú lineáris kombinációja lehet zérusvektor. Ismét kontrapozícióval bizonyítunk: ha van olyan – nem csupa 0 együtthatójú – lineáris kombináció, mely a nullvektorral egyenlő, azaz

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_k\mathbf{v}_k = \mathbf{0},$$

de valamelyik együttható – például a  $c_1$  – nem 0, akkor  $\mathbf{v}_1$  kifejezhető a többi vektor lineáris kombinációjaként:

$$\mathbf{v}_1 = -\frac{c_2}{c_1}\mathbf{v}_2 - \dots - \frac{c_k}{c_1}\mathbf{v}_k,$$

ami bizonyítja az állítást.  $\square$

Egy vektorrendszert *lineárisan összefüggőnek* nevezünk, ha nem független, azaz egyelemű vektorrendszer esetén ha az a vektor a zérusvektor, többelemű vektorrendszer esetén pedig ha van olyan vektora, mely kifejezhető a többi lineáris kombinációjaként. Az előző tétel szerint ez azzal ekvivalens, hogy a vektorrendszernek van olyan zérusvektort adó lineáris kombinációja, melyben nem mindegyik együttható zérus. A lineáris összefüggőség definíciója kicsit élesíthető:

**1.49. TÉTEL (LINEÁRIS ÖSSZEFÜGGŐSÉG).** Egy nullvektortól különböző elemekből álló, legalább kételemű  $\mathbb{R}^n$ -beli  $V = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$  vektorrendszer pontosan akkor lineárisan összefüggő, ha van olyan  $t \geq 2$  index, hogy  $\mathbf{v}_t$  a  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{t-1}$  vektorok lineáris kombinációja.

Másként fogalmazva, ha egy nullvektort nem tartalmazó vektorrendszerben találunk olyan vektort, mely a többi lineáris kombinációja, akkor olyat is találunk a vektorok bármely sorba rendezése mellett, mely sorrendben csak az őt megelőző vektor(ok) lineáris kombinációja.

**BIZONYÍTÁS.** Először tegyük fel, hogy a vektorrendszer összefüggő, és legyen  $t$  az a legkisebb egész, melyre a  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_t$  vektorok már összefüggők. Mivel  $\mathbf{v}_1 \neq \mathbf{0}$ , ezért az első vektor nem lehet összefüggő, ezért  $t \geq 2$ . E vektorok összefüggősége miatt vannak olyan  $c_i$  konstansok, melyekkel

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_t\mathbf{v}_t = \mathbf{0}.$$

Biztos, hogy  $c_t \neq 0$ , különben már a  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{t-1}$  vektorok is lineárisan összefüggők lennének, és ez ellentmond  $t$  definíciójának. Így

$$\mathbf{v}_t = \frac{-c_1}{c_t}\mathbf{v}_1 + \frac{-c_2}{c_t}\mathbf{v}_2 + \dots + \frac{-c_{t-1}}{c_t}\mathbf{v}_{t-1},$$

ami bizonyítja, hogy összefüggő vektorrendszerben létezik ilyen vektor.

A másik irányú implikáció definíció szerint igaz, hisz ha létezik ilyen  $\mathbf{v}_t$  vektor, akkor ez valóban lineáris kombinációja az összes többi vektornak.  $\square$

**Skaláris szorzás  $\mathbb{R}^n$ -ben** A skaláris szorzást először abból az alakból általánosítjuk, amelyet a 2- és 3-dimenziós térben ortonormált bázis esetén láttunk. A tetszőleges bázis esetére való általánosításra később térünk vissza.

**1.50. DEFINÍCIÓ (SKALÁRIS SZORZÁS  $\mathbb{R}^n$ -BEN).** Legyen  $\mathbf{u} = (u_1, u_2, \dots, u_n)$  és  $\mathbf{v} = (v_1, v_2, \dots, v_n)$  az  $\mathbb{R}^n$  tér két tetszőleges vektora. Skaláris szorzatukon a következő kifejezést értjük:

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \dots + u_n v_n.$$

A skaláris szorzásnak most csak azokat a tulajdonságait említjük, melyek a későbbi általánosításhoz szükségesek lesznek:

**1.51. TÉTEL (A SKALÁRIS SZORZÁS TULAJDONSÁGAI).** Legyen  $\mathbf{u}, \mathbf{v}$  és  $\mathbf{w}$  az  $\mathbb{R}^n$  három tetszőleges vektora, és legyen  $c$  egy tetszőleges valós. Ekkor

- 
- a)  $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$  *a művelet fölcserélhető (kommutatív)*  
b)  $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$  *disztributív*  
c)  $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v})$  *a két szorzás kompatibilis*  
d)  $\mathbf{u} \cdot \mathbf{u} \geq 0$  és  $\mathbf{u} \cdot \mathbf{u} = 0$  pontosan akkor teljesül, ha  $\mathbf{u} = \mathbf{0}$ .
- 

**BIZONYÍTÁS.** A bizonyítás itt is igen egyszerű, ezért csak az a) pontét mutatjuk meg, a többit az Olvasóra hagyjuk:

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= u_1 v_1 + u_2 v_2 + \dots + u_n v_n \\ &= v_1 u_1 + v_2 u_2 + \dots + v_n u_n \\ &= \mathbf{v} \cdot \mathbf{u}. \end{aligned}$$

□

**Távolság és szög  $\mathbb{R}^n$ -ben** Két 2- vagy 3-dimenziós vektor (végpontja) távolságának és szögének a skaláris szorzatukkal való kapcsolatát használjuk e fogalmaknak a magasabb dimenziós terekben való definíciójához.

**1.52. DEFINÍCIÓ (ABSZOLÚT ÉRTÉK, SZÖG, MERŐLEGESSÉG, TÁVOLSÁG).** Legyen  $\mathbf{u}$  és  $\mathbf{v}$  az  $\mathbb{R}^n$  tér két tetszőleges vektora.

- a) Az  $\mathbf{u}$  vektor hosszán önmagával vett skaláris szorzatának gyökét értjük, azaz

$$|\mathbf{u}| = \sqrt{\mathbf{u} \cdot \mathbf{u}}. \quad (1.3)$$

- b) Az  $\mathbf{u}$  és  $\mathbf{v}$  vektorok (hajlás)szögének koszinuszát az alábbi törttel definiáljuk:

$$\cos(\mathbf{u}, \mathbf{v})_{\angle} := \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}| |\mathbf{v}|} \quad (1.4)$$

- c) Azt mondjuk, hogy az  $\mathbf{u}$  és  $\mathbf{v}$  vektorok merőlegesek egymásra, ha

$$\mathbf{u} \cdot \mathbf{v} = 0. \quad (1.5)$$

- d) A két vektor végpontjának távolságán, amit egyszerűen a két vektor távolságának nevezünk, a

$$d(\mathbf{u}, \mathbf{v}) = |\mathbf{u} - \mathbf{v}| \quad (1.6)$$

értéket értjük.

► A fenti definíciókat érdemes megtekinteni koordinátás alakjukba átírva. Eszerint például

$$|\mathbf{u}| = \sqrt{u_1^2 + u_2^2 + \dots + u_n^2},$$

$$\cos(\mathbf{u}, \mathbf{v})_{\angle} = \frac{u_1 v_1 + u_2 v_2 + \dots + u_n v_n}{\sqrt{u_1^2 + u_2^2 + \dots + u_n^2} \sqrt{v_1^2 + v_2^2 + \dots + v_n^2}}.$$

► A fenti definíciók közül a vektorok hajlásszögének definíciója még hiányos. Egy szög koszinusza csak a  $[-1, 1]$  intervallumba eső szám lehet, ezért e definíció csak akkor értelmes, ha az (1.4) képlet szám-lálójára és nevezőjére igaz, hogy  $|\mathbf{u} \cdot \mathbf{v}| \leq |\mathbf{u}| |\mathbf{v}|$ , azaz ha  $n$ -dimenziós vektorokra is fennáll a CBS-egyenlőtlenség. Ezt hamarosan igazolni fogjuk!

**1.53. PÉLDA (VEKTOROK SZÖGE ÉS TÁVOLSÁGA).** Az  $\mathbf{u} = (2, 3, 4, 14)$  vektornak mennyi az abszolút értéke, mennyi a  $\mathbf{v} = (4, 6, -10, 10)$  vektortól való távolsága, és mennyi a  $\mathbf{w} = (0, 3, 6, -2)$  vektorral bezárt szögének koszinusza?

**MEGOLDÁS.** A válaszhoz az (1.3), az (1.6) és az (1.4) képleteket használjuk:

$$|\mathbf{u}| = \sqrt{2^2 + 3^2 + 4^2 + 14^2} = \sqrt{225} = 15,$$

$$d(\mathbf{u}, \mathbf{v}) = \sqrt{(2-4)^2 + (3-6)^2 + (4-(-10))^2 + (14-10)^2}$$

$$= \sqrt{2^2 + 3^2 + 14^2 + 4^2} = 15$$

$$\cos(\mathbf{u}, \mathbf{w})_{\angle} = \frac{2 \cdot 0 + 3 \cdot 3 + 4 \cdot 6 + 14 \cdot (-2)}{\sqrt{2^2 + 3^2 + 4^2 + 14^2} \sqrt{0^2 + 3^2 + 6^2 + (-2)^2}} = \frac{1}{21}. \quad \square$$

**1.54. TÉTEL (CAUCHY–BUNYAKOVSKIJ–SCHWARZ-EGYENLŐTLENSÉG).** Tetszőleges  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$  vektorokra

$$|\mathbf{u} \cdot \mathbf{v}| \leq |\mathbf{u}| |\mathbf{v}|. \quad (1.7)$$

Egyenlőség pontosan akkor áll fenn, ha  $\mathbf{u}$  és  $\mathbf{v}$  lineárisan összefüggők, azaz ha egyik vektor a másik skalárszorosa.

**BIZONYÍTÁS.** Tegyük fel először, hogy  $\mathbf{v} = \mathbf{0}$ . Ekkor a tétel állításának mindkét része nyilván igaz, hisz egyenlőség áll fenn, és a két vektor lineárisan összefüggő. Ha  $\mathbf{v} \neq \mathbf{0}$ , akkor legyen  $\mathbf{e} = \mathbf{v}/|\mathbf{v}|$  a  $\mathbf{v}$  irányú egységvektor. Az  $\mathbf{u}$  vektor  $\mathbf{e}$  egyenesére merőleges összetevőjének hossza,

illetve annak négyzete nyilván nem negatív, azaz

$$\begin{aligned}
 0 &\leq |\mathbf{u} - (\mathbf{u} \cdot \mathbf{e})\mathbf{e}|^2 \\
 &= |\mathbf{u}|^2 + |\mathbf{u} \cdot \mathbf{e}|^2 - 2|\mathbf{u} \cdot \mathbf{e}|^2 \\
 &= |\mathbf{u}|^2 - |\mathbf{u} \cdot \mathbf{e}|^2 \\
 &= |\mathbf{u}|^2 - \frac{|\mathbf{u} \cdot \mathbf{v}|^2}{|\mathbf{v}|^2},
 \end{aligned}$$

Innen átrendezéssel azonnal megkapjuk a bizonyítandó állítást. Másrészt az is világos, hogy  $0 = |\mathbf{u} - (\mathbf{u} \cdot \mathbf{e})\mathbf{e}|$  csak akkor állhat fenn, ha  $\mathbf{u} = (\mathbf{u} \cdot \mathbf{e})\mathbf{e}$ , azaz ha  $\mathbf{u}$  és  $\mathbf{e}$  párhuzamosak, azaz ha  $\mathbf{u}$  a  $\mathbf{v}$  skalárszorosa, vagyis ha a két vektor lineárisan összefüggő.  $\square$

**1.55. TÉTEL (HÁROMSZÖG-EGYENLŐTLENSÉG  $\mathbb{R}^n$ -BEN).** *Tetszőleges  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$  vektorokra*

$$|\mathbf{u} + \mathbf{v}| \leq |\mathbf{u}| + |\mathbf{v}|.$$

A bizonyítás megegyezik a 3-dimenziós változatra, azaz az 1.22. tételre adott bizonyítással.

A vektor abszolút értékét a skaláris szorzat segítségével definiáltuk, de fordítva, a skaláris szorzat is kifejezhető a vektor abszolút értékével.

**1.56. TÉTEL (SKALÁRIS SZORZAT ÉS ABSZOLÚT ÉRTÉK  $\mathbb{R}^n$ -BEN).** *Tetszőleges  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$  vektorokra*

$$\mathbf{u} \cdot \mathbf{v} = \frac{1}{4} (|\mathbf{u} + \mathbf{v}|^2 - |\mathbf{u} - \mathbf{v}|^2) \quad (1.8)$$

$$\mathbf{u} \cdot \mathbf{v} = \frac{1}{2} (|\mathbf{u} + \mathbf{v}|^2 - |\mathbf{u}|^2 - |\mathbf{v}|^2) \quad (1.9)$$

**BIZONYÍTÁS.** A bizonyításban az abszolút érték (1.3)-beli definícióját használjuk:

$$\begin{aligned}
 \frac{1}{4} (|\mathbf{u} + \mathbf{v}|^2 - |\mathbf{u} - \mathbf{v}|^2) &= \frac{1}{4} ((\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) - (\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v})) \\
 &= \frac{1}{4} (\mathbf{u} \cdot \mathbf{u} + \mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{v} - \mathbf{u} \cdot \mathbf{u} + \mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{u} - \mathbf{v} \cdot \mathbf{v}) \\
 &= \frac{1}{4} (4\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot \mathbf{v}.
 \end{aligned}$$

A másik formula hasonlóan bizonyítható.  $\square$

Végül egy fontos összefüggés az ortogonális vektorrendszerekről:

**1.57. ÁLLÍTÁS (ORTOGONÁLIS VEKTORRENDSZER LINEÁRIS FÜGGETLEN-SÉGE).** *Tegyük fel, hogy a zérusvektortól különböző  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  vektorok páronként ortogonálisak, azaz bármely  $i \neq j$  esetén  $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ . Ekkor a vektorok lineárisan függetlenek.*

**BIZONYÍTÁS.** Tegyük fel, hogy valamely  $c_1, c_2, \dots, c_k$  konstansokra

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_k \mathbf{v}_k = \mathbf{0}.$$

Szorozzuk be az egyenlőség mindkét oldalát skalárisan a  $\mathbf{v}_i$  vektorral. Mivel  $i \neq j$  esetén  $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ , ezért azt kapjuk, hogy

$$c_i \mathbf{v}_i \cdot \mathbf{v}_i = 0,$$

amiből  $\mathbf{v}_i \cdot \mathbf{v}_i \neq 0$  miatt következik, hogy  $c_i = 0$ . Mivel ez minden  $i = 1, 2, \dots, k$  indexre igaz, ezért a vektorok valóban lineárisan függetlenek.  $\square$

► Ha  $e$  vektorok az  $\mathbb{R}^n$  tér vektorai, akkor felvetődik a kérdés, hogy mi  $k$  és  $n$  viszonya. Később látni fogjuk, hogy  $k \leq n$ .

► Később azt is meg fogjuk mutatni, hogy bármely független vektorrendszerből kiindulva megkonstruálható egy vele azonos elemszámú ortogonális vektorrendszer. A műszaki alkalmazásokban is gyakori, hogy egy meglévő alapvektorrendszerből egy ortogonális, majd abból egy ortonormált vektorrendszert konstruálunk.

**Korrelációs együttható\*** Az  $n$ -dimenziós térben szerzett friss szemléletünk segítségünkre lesz egy fontos fogalom megértésében.

Adva van két adatsor:  $x_1, x_2, \dots, x_n$  és  $y_1, y_2, \dots, y_n$ . Átlagukat jelölje  $\bar{x}$ , illetve  $\bar{y}$ , azaz

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}, \quad \bar{y} = \frac{y_1 + y_2 + \dots + y_n}{n}.$$

Az  $r$ -rel jelölt ún. *korrelációs együttható* azt méri, hogy a két adatsor közti lineáris függvénykapcsolat milyen erős. Az erre használt képlet a következő:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}.$$

Vajon hogyan méri  $r$  a lineáris függvénykapcsolat erősségét?

Tegyük fel, hogy a két adatsor közt fönnáll az  $y_i = cx_i + d$  lineáris függvénykapcsolat minden  $i = 1, 2, \dots, n$  indexre valamely  $c, d$  konstansokkal. Ha mindkét adatsorból levonjuk az átlagukat (más szóval a két adatsort *normáljuk*), akkor az így kapott

$$a_i = x_i - \bar{x}, \quad b_i = y_i - \bar{y} \quad (i = 1, 2, \dots, n)$$

adatsorokra igaz a  $b_i = ca_i$  összefüggés. Ugyanis

$$\bar{y} = \frac{1}{n} \left( \sum_{i=1}^n cx_i + d \right) = c\bar{x} + d,$$

amiből  $b_i = y_i - \bar{y} = cx_i + d - c\bar{x} - d = c(x_i - \bar{x}) = ca_i$ .

Tehát az  $y_i = cx_i + d$  lineáris függvénykapcsolat pontosan akkor áll fenn, ha a normált adatsorokra  $b_i = ca_i$ . Az adatsorokat  $n$ -dimenziós vektorokba foglalva ez azzal ekvivalens, hogy  $\mathbf{b} = c\mathbf{a}$ , azaz ha a vektorok kollineárisak. A korrelációs együttható nem más, mint a két utóbbi vektor szögének koszinusza, ugyanis

$$\begin{aligned}\cos(\mathbf{a}, \mathbf{b})_{\angle} &= \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}| |\mathbf{b}|} = \frac{\sum_{i=1}^n a_i b_i}{\sqrt{\sum_{i=1}^n a_i^2} \sqrt{\sum_{i=1}^n b_i^2}} \\ &= \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} = r.\end{aligned}$$

Valóban, a két vektor szögének koszinusza pontosan akkor 1, ha a vektorok szöge 0, és akkor  $-1$ , ha a szög  $\pi$ . A korreláció tehát  $-1$  és  $1$  közt változik, és abszolút értéke annál kisebb, minél nagyobb az  $\mathbf{a}$  és  $\mathbf{b}$  vektorok hajlásszöge, azaz minél kevésbé kollineárisak, azaz minél kevésbé erős a két számsorozat közti lineáris függvénykapcsolat.

Ha  $r = 0$ , akkor  $\mathbf{a}$  és  $\mathbf{b}$  merőlegesek, ekkor lineáris függvénykapcsolat nincs az eredeti két mennyiség közt (más kapcsolat még lehet, tehát nem feltétlenül független a két adatsor egymástól valószínűség-számítási értelemben). A fogalom mélyebb megértése a valószínűség-számítás ismeretét is megkívánja, ezzel mi itt nem foglalkozunk.

**Bitvektorok, kódvektorok\*** A modern számítógépek memóriájában vagy háttértárolóin az adatok tárolásának legkisebb egysége a bit. Egy bittel két állapot tárolható, melyeket a 0 és 1 számokkal jelölünk, de amelyek több mindent is reprezentálhatnak: hamis/igaz, nem/igen, ki/be,... A biteket a hardver lehetőségei és a feladat igényei szerint csoportokba, sorozatokba, vektorokba gyűjtik, melyekkel különféle műveletek végezhetők. Ezek attól is függenek, hogy a bitvektorok milyen adatokat kódolnak. E műveletek közül minket azok fognak érdekelni, melyek algebrailag a korábban megismert vektorműveletekre hasonlítanak.

Az egyszerűség kedvéért a bitvektorokat gyakran a biteket jelölő számjegyek egyszerű egymás mellé írásával adjuk meg, pl. 01110101 a  $(0, 1, 1, 1, 0, 1, 0, 1)$  vektort jelöli.

A modern számítástechnika számtalan kódot használ, mely bitvektorokkal (is) leírható. Például karakterek kódolására használatos a 7-dimenziós bitvektorokból álló ASCII-kód, a decimális számok kódolására a 4-dimenziós bitvektorokból álló BCD-kód.

Az emberek által is elolvasható kódok gyakran decimális számokból állnak. Például az emberek azonosítására használt *személyi szám* egy olyan vektornak tekinthető, amelynek koordinátái a 10-elemű  $\{0, 1, \dots, 9\}$  halmazból valók.

A kódoláshoz mi a továbbiakban mindig egy rögzített, véges kód-

*Bit:* az angol *binary digit* kifejezésből képzett szó, ami magyarul bináris, azaz kettes számrendszerbeli számot jelent. A szoftver (software) szót is megalkotó John W. Tukey ötlete.

Az ASCII-kód eredendően az angol nyelvű szövegek kódolására tervezett 8-hosszú bináris kód (azaz 1 bájt). Az angol nyelv betűi, írásjelei, és néhány számítógépet vezérlő karakter mindegyikének egy olyan vektor felel meg, melynek első koordinátája 0. Tehát a lehetséges 256 darab 8-hosszú vektorból 128 tartozik a kódba. Pl. a „z” betű ASCII-kódja 01111010, decimális alakban 122.

A BCD-kód decimális számok egyik szokásos kódolása, mely a szám kettes számrendszerbe való átírása helyett a számjegyenként való kódolást választja. Több változata is van, a legegyszerűbbikben minden számjegynek 4-4 bit felel meg, így a 16 lehetséges 4-hosszú kódszó helyett csak tízet használ: a 0, 1, ..., 9 jegyek kódja rendre 0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001. Így az 561 BCD-kódja három kódvektorból áll: 0101 0110 0001. A kettes számrendszerbeli alak 1000110001.

ábécét használunk, amelynek betűi általában a 0-tól  $n - 1$ -ig terjedő egészek lesznek. Az ábécé „betűiből”, azaz elemeiből képzett vektorokat *kódvektoroknak* vagy *kódszavaknak* nevezzük. A bitvektorok is kódvektorok, ahol a kódábécé a kételemű  $\{0, 1\}$  halmaz.

A kódvektorok koordinátáinak számát, vagyis a kódvektor dimenzióját a kód *hosszának* nevezzük. Ez természetesen nem analóg fogalom a vektor abszolút értékével.

A személyi szám tehát egy 10-elemű ábécéből képzett 11-hosszú kódszó. Nem minden 11-hosszú decimális vektor lehet személyi szám, mert az utolsó koordináta egy ellenőrző jegy, vagyis a személyi szám, mint kód, matematikailag a 11-hosszú kódvektorok halmazának egy részhalmazaként írható le. Ezért általában a kódábécé betűiből képzett vektorok részhalmazait fogjuk kódnak nevezni.

**1.58. DEFINÍCIÓ (KÓD).** A kód egy közös ábécéből képzett azonos hosszúságú kódszavak egy halmaza. Kódolás során a kódolandó objektumokhoz kódszavakat rendelünk, dekódolás az ellenkező irányú folyamat.

Főként az információelméletben változó hosszú kódszavak is tartozhatnak egy kódhoz. Mi ilyenekkel nem fogunk foglalkozni, de megemlítjük, hogy a karakterek manapság elterjedt UTF-8 kódolása is ilyen, amelyben egy karakterkódja 8-, 16-, 24- vagy 32-bites lehet.

**Vektorműveletek  $\mathbb{Z}_m^n$ -ben\*** A  $\mathbb{Z}_m$ -re vonatkozó ismereteket a függelékben részletezzük. Az 1.44. definíció szerint  $\mathbb{Z}_m^n$  a  $\mathbb{Z}_m$ -beli  $n$ -hosszú vektorokból áll. E vektorok összeadása, skalárral való szorzása és skaláris szorzása a  $\mathbb{Z}_m$ -beli műveletekkel az  $\mathbb{R}^n$ -beli vektorműveletekhez hasonlóan végezhető el. Ennek következtében a lineáris kombináció, lineáris függetlenség itt is ugyanúgy definiálható és használható.

**1.59. PÉLDA (LINEÁRIS KOMBINÁCIÓ  $\mathbb{Z}_2^5$ -BEN).** Számítsuk ki a  $\mathbb{Z}_2^5$ -beli

$$\mathbf{a} = (1, 0, 0, 1, 1, 0), \mathbf{b} = (0, 1, 0, 1, 0, 1) \text{ és } \mathbf{c} = (0, 0, 1, 0, 1, 1)$$

vektorok összes lineáris kombinációját  $\mathbb{Z}_2$ -beli együtthatókkal, valamint a  $\mathbb{Z}_3^3$ -beli

$$\mathbf{u} = (1, 1, 0) \text{ és } \mathbf{v} = (0, 1, 1)$$

vektorok összes lineáris kombinációját  $\mathbb{Z}_3$ -beli együtthatókkal.

**MEGOLDÁS.** A lehetséges  $x\mathbf{a} + y\mathbf{b} + z\mathbf{c}$  alakú lineáris kombinációk száma 8, hisz  $x, y, z \in \mathbb{Z}_2$ , vagyis mindegyik együtthatónak 0 vagy 1 az értéke, és ez  $2 \cdot 2 \cdot 2 = 8$  eshetőség. Az  $x = y = z = 0$  eset a zérusvektort adja. Ha  $x, y$  és  $z$  közül csak egyikük értéke 1, a többi 0, akkor a három adott vektort kapjuk vissza. Azok az esetek maradnak, amikor legalább két vektort kell összeadni. Például  $1\mathbf{a} + 1\mathbf{b} + 0\mathbf{c} = (1, 0, 0, 1, 1, 0) + (0, 1, 0, 1, 0, 1) = (1, 1, 0, 0, 1, 1)$ . Az



összes lineáris kombináció az 1.1. (a) táblázatban látható.

Az  $x\mathbf{u} + y\mathbf{v}$  alakú lineáris kombinációk száma 9, mivel  $x, y \in \mathbb{Z}_3$ , azaz lehetséges értékük 0, 1 vagy 2, ami  $3 \cdot 3 = 9$  lehetőséget ad. Példaként egy lineáris kombináció, a többi az 1.1. (b) táblázatban látható:  $2\mathbf{u} + 1\mathbf{v} = 2(1, 1, 0) + (0, 1, 1) = (2, 2, 0) + (0, 1, 1) = (2, 0, 1)$ .  $\square$

E paragrafus további részében a vektorműveletekre két alkalmazást mutatunk.

**1.60. PÉLDA (ONE TIME PAD – A TÖKÉLETES TITKOSÍTÁS).** Az üzenet küldése előtt a küldő és a fogadó megegyezik egy titkos kulcsban, mely egy olyan hosszú véletlen bitvektor, mint amilyen az üzenet legföljebb lehet. Legyen a kulcs  $\mathbf{k} \in \mathbb{Z}_2^m$ . Legyen a titkosítandó üzenet  $\mathbf{u} \in \mathbb{Z}_2^m$ . A titkosítás során a küldő kiszámolja az  $\mathbf{u} + \mathbf{k}$  vektort, és azt küldi a fogadónak, aki a titkosított üzenethez maga is hozzáadja a kulcsot, és mivel bármely  $\mathbf{x} \in \mathbb{Z}_2^m$  vektorra  $\mathbf{x} + \mathbf{x} = \mathbf{0}$ , ezért  $(\mathbf{u} + \mathbf{k}) + \mathbf{k} = \mathbf{u} + (\mathbf{k} + \mathbf{k}) = \mathbf{u}$ , vagyis a fogadó így valóban megfejti az üzenetet.

► Példaként egy üzenet, egy kulcs és a kettő összege – a titkosított üzenet – a tömör bitvektor-jelöléssel

az üzenet:  $\mathbf{u} = 010101010000111111111111$

a kulcs:  $\mathbf{k} = 0010111000101101001011010$

a titkosított üzenet:  $\mathbf{u} + \mathbf{k} = 011110010101010110100101$

► A bitvektorok ilyen módon való összeadása megegyezik a kizáró vagy nevű logikai művelettel, melyet a XOR szóval (exclusive or), vagy a  $\oplus$  műveleti jellel is szoktak jelölni (ld. ??). példa).

► E titkosítás hátránya, hogy a  $\mathbf{k}$  kulcs csak egyszer használható fel, mert két különböző  $\mathbf{u}_1 + \mathbf{k}$  és  $\mathbf{u}_2 + \mathbf{k}$  üzenetet elcsípve és összeadva az  $(\mathbf{u}_1 + \mathbf{k}) + (\mathbf{u}_2 + \mathbf{k}) = \mathbf{u}_1 + \mathbf{u}_2$  vektorban már nem szerepel  $\mathbf{k}$ , és ebből statisztikai módszereket is használva már mindkét üzenet kinyerhető.

► Bizonyítható, hogy e kód megfejthetetlen, ha  $\mathbf{k}$  valóban véletlen bitsorozat, és csak egyetlen üzenet titkosítására használjuk.

A kódelmélet egyik célja, hogy redundáns információ hozzáadásával elérje az elküldött üzenet megérkezését zajos, veszteséges csatornán keresztül is. Ennek két gyakran alkalmazott típusa a hibajelző és a hibajavító kód: az előbbi az átvitel során bekövetkezett bizonyos hibákat jelez a fogadó számára, míg az utóbbi bizonyos hibák kijavítását is lehetővé teszi. Az 1.59. példában előállított lineáris kombinációk hibajelző kódok, egyikük hibajavító is. Az 1.39. feladat arra kérdez, hogy milyen hibát jeleznek, illetve javítanak.

Az elektronikus számítógépek adatkezelésének egyik első ötlete az adattárolás vagy továbbítás biztonságosabbá tételére a paritásbit. Ha egy  $n$ -hosszú  $\mathbf{b}$  bitvektorhoz még egy bitet csatolunk, melynek értéke

| $x \ y \ z$ | $x\mathbf{a} + y\mathbf{b} + z\mathbf{c}$ | $x \ y$ | $x\mathbf{u} + y\mathbf{v}$ |
|-------------|-------------------------------------------|---------|-----------------------------|
| 0 0 0       | (0, 0, 0, 0, 0, 0)                        | 0 0     | (0, 0, 0)                   |
| 1 0 0       | (1, 0, 0, 1, 1, 0)                        | 1 0     | (1, 1, 0)                   |
| 0 1 0       | (0, 1, 0, 1, 0, 1)                        | 2 0     | (2, 2, 0)                   |
| 0 0 1       | (0, 0, 1, 0, 1, 1)                        | 0 1     | (0, 1, 1)                   |
| 1 1 0       | (1, 1, 0, 0, 1, 1)                        | 1 1     | (1, 2, 1)                   |
| 1 0 1       | (1, 0, 1, 1, 0, 1)                        | 2 1     | (2, 0, 1)                   |
| 0 1 1       | (0, 1, 1, 1, 1, 0)                        | 0 2     | (0, 2, 2)                   |
| 1 1 1       | (1, 1, 1, 0, 0, 0)                        | 1 2     | (1, 0, 2)                   |
|             |                                           | 2 2     | (2, 1, 2)                   |

(a)

(b)

1.1. táblázat: Vektorok lineáris kombinációi (a)  $\mathbb{Z}_2$  és (b)  $\mathbb{Z}_3$  fölött.

1, ha  $\mathbf{b}$ -ben páratlan sok bit egyenlő 1-gyel, egyébként 0, akkor olyan  $(n+1)$ -hosszú vektort kapunk, melyben páros sok 1-esnek kell lenni. Ha egy bit elromlik, páratlan sok 1-es lesz, tehát ez az  $(n+1)$ -edik bit hibajelző. Ezt nevezzük *paritásbitnek*.

**1.61. PÉLDA (PARITÁSBIT).** Írjuk fel a paritásbitet skaláris szorzatként!

**MEGOLDÁS.** A paritásbit  $\mathbb{Z}_2$  fölött  $\mathbf{1} \cdot \mathbf{b}$  alakba írható, ahol  $\mathbf{1}$  a  $\mathbf{b}$ -vel azonos hosszúságú és csupa 1-esből álló vektor.  $\square$

A paritásbit általánosítása az ún. *ellenőrző összeg*, melyre számtalan példát találunk a mindennapi életben.

A magyar személyi szám a személyre jellemző 10 jegyből, és az azt követő  $e$  ellenőrző összegből áll. Az  $e$  kiszámítási képlete

$$\mathbb{Z}_{11}\text{-ben számolva: } e = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10) \cdot \mathbf{u},$$

ahol  $\mathbf{u}$  a személyi szám első 10 jegye. A személyi szám 8-10-edik jegyét úgy választják ki, hogy  $e \neq 10$ , így az ellenőrző összeg mindig egyjegyű.

A termékek EAN-kódja (European Article Number) egy 13-jegyű, a termék azonosítására szolgáló kód, melyhez egy vonalkód is tartozik. A 13-dik jegy az ellenőrző összeg. Ha az EAN kódvektort  $\mathbf{v}$  jelöli, akkor fönn kell állni

$$\mathbb{Z}_{10}\text{-ben számolva az } (1, 3, 1, 3, 1, 3, 1, 3, 1, 3, 1) \cdot \mathbf{v} = 0$$

összefüggésnek (1.46. ábra).

ISBN 978-963-545-398-6



9 789635 453986 >

1.46. ábra: Egy könyv ISBN-13 kódja, ami egyúttal az EAN kódja is. Az EAN-kódhoz tartozik egy vonalkód is. 2007 óta a könyvek ISBN-száma (ISBN-13) megegyezik EAN-kódjával (korábban az ISBN-szám 10-jegyű volt).

## Feladatok

**1.39▲** Az 1.1. (a) táblázatban egy 8 bináris kódszóból, a (b) táblázatban egy 9 ternér kódszóból álló kód kódszavai vannak felsorolva. Határozzuk meg, hogy e kódok hány hiba jelzésére és hány hiba javítására képesek! (Az, hogy egy kód képes  $k$  hibát jelezni, azt jelenti, hogy ha legfőbb  $k$  jelet megváltoztatunk bármelyik kódszóban, akkor egy kódba nem tartozó vektort kapunk. Az, hogy a kód képes  $d$  hibát javítani, azt jelenti, hogy bármely kódszóban legfőbb  $d$  jelet megváltoztatva olyan vektort kapunk, amelyből más kódszó nem kapható meg legfőbb  $d$  jel megváltoz-

tatásával.)

**1.40.** Mely bitvektorokra igaz, hogy minden bitjük a vektor maradék részének paritásbitje.

**1.41▲ ELLENŐRZŐ ÖSSZEG** Csak az ellenőrző összeget nézve érvényes személyi szám-e a 26012310018, és érvényes EAN-kód-e a 9998887776665?

**1.42▲** 2007 előtt a személyi számhoz hasonló módon számolták ki a könyvek ún. ISBN-10 kódjának ellenőrző jegyét, ami ha 10 volt, X-et írtak (római 10-es). A képlet:

$$Z_{11}\text{-ben számolva: } (10, 9, 8, 7, 6, 5, 4, 3, 2, 1) \cdot \mathbf{u} = 0$$

ahol  $\mathbf{u}$  a könyv ISBN-kódja, aminek utolsó jegye az ellenőrző összeg. Egy könyv kódjának első 9 jegye 963076198. Mi e könyv teljes ISBN-száma?

## Megoldások

1.1. „Ha az irányított szakasz a hal, akkor a vektor a halraj.”

1.2. a) Igaz. b) Hamis, például ha  $O = A$ , akkor  $\vec{OA} + \vec{OB} = \vec{AB}$ , míg ha  $O = B$ , akkor  $\vec{OA} + \vec{OB} = \vec{BA}$ . c) Igaz, az eredmény  $O$  választásától függetlenül  $\vec{BA}$ . d) Igaz. e) Hamis, lehetnek ellenkező irányúak is. f) Igaz.

$$1.9. \vec{P_1P_2} + \vec{P_2P_3} + \vec{P_3P_4} + \dots + \vec{P_{n-1}P_n} = \vec{P_1P_n}, \vec{P_1P_2} + \vec{P_2P_3} + \vec{P_3P_4} + \dots + \vec{P_{n-1}P_n} + \vec{P_nP_1} = \vec{0}.$$

1.12. a) Hamis, lehet, hogy a három közül két vektor egy egyenesbe esik, és a harmadik független tőlük: ez a harmadik nem állítható elő a másik kettő lineáris kombinációjaként. b) Igaz, például  $\mathbf{i}$ ,  $\mathbf{j}$  és  $\mathbf{i} + \mathbf{j}$  ilyenek. De bármely három egy síkba eső nemzérus-vektor ilyen, ha közülük bármely kettő lineárisan független. c) Igaz, például ha  $\mathbf{b} = \mathbf{c}$  és  $\mathbf{a}$  független  $\mathbf{b}$ -től. d) Hamis, például ha  $\mathbf{b} = \mathbf{c}$  és  $\mathbf{a}$  független  $\mathbf{b}$ -től, akkor  $\mathbf{a}$  nem fejezhető ki  $\mathbf{b}$  és  $\mathbf{c}$  lineáris kombinációjaként.

1.13. Ha  $|\vec{AP}| : |\vec{PB}| = m : n$ , akkor  $|\vec{AB}| : |\vec{PB}| = (m+n) : n$ , amiből  $\vec{BP} = \frac{n}{m+n} \vec{BA}$ . De  $\vec{OP} = \vec{OB} + \vec{BP}$  és  $\vec{BA} = \vec{OA} - \vec{OB}$ , így  $\vec{OP} = \vec{OB} + \frac{n}{m+n}(\vec{OA} - \vec{OB})$ , amiből azonnal következik a bizonyítandó formula. A felezőpontot az  $m = n = 1$  esetben kapjuk, és ekkor valóban  $\vec{OP} = \frac{1}{2}\vec{OA} + \frac{1}{2}\vec{OB}$ .

$$1.14. |\mathbf{a}||\mathbf{b}| \cos \gamma = 1 \cdot 2 \cdot \frac{1}{2} = 1.$$

$$1.15. |\mathbf{a}||\mathbf{b}| \cos \gamma = \sqrt{2} \cdot 2 \cdot \left(-\frac{1}{\sqrt{2}}\right) = -2.$$

$$1.16. 1 \cdot 2 \cdot (-1) = -2.$$

$$1.17. 0, \text{ hisz merőlegesek } (|\mathbf{a}||\mathbf{b}| \cos \gamma = \sqrt{2} \cdot 2 \cdot 0 = 0).$$

1.18. Legyenek  $\mathbf{a}$  és  $\mathbf{c}$  független vektorok,  $\mathbf{b}$  pedig tetszőleges. Ekkor az  $(\mathbf{a} \cdot \mathbf{b})\mathbf{c}$  szorzat párhuzamos a  $\mathbf{c}$  vektorral, míg az  $\mathbf{a}(\mathbf{b} \cdot \mathbf{c})$  szorzat az  $\mathbf{a}$  vektorral, tehát  $(\mathbf{a} \cdot \mathbf{b})\mathbf{c} \neq \mathbf{a}(\mathbf{b} \cdot \mathbf{c})$ .

1.19. a) igaz, b) hamis, az egységvektor egyenesére eső merőleges vetületének hosszával egyenlő, c) igaz, d) hamis (asszociativitásról nem is lehet szó, mert a két szorzás művelet egyike skaláris szorzás, a másika skalárral való szorzás az  $\mathbf{a}(\mathbf{b} \cdot \mathbf{c})$  szorzatban), e) igaz, f) igaz.

$$1.20. \mathbf{a} \cdot \mathbf{a} - \mathbf{b} \cdot \mathbf{b} = |\mathbf{a}|^2 - |\mathbf{b}|^2$$

$$1.21. \mathbf{a} \cdot \mathbf{a} + 2\mathbf{b} \cdot \mathbf{a} - 2\mathbf{a} \cdot \mathbf{b} = \mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2$$

1.22. A Pihagorász-tétel következményeként  $\mathbf{a}$  és  $\mathbf{b}$  merőlegesek.

1.23. Nem, legyen pl.  $\mathbf{a} = (1, 0, 0)$ ,  $\mathbf{b} = (0, 0, 1)$ ,  $\mathbf{c} = (1, 0, -1)$ .

1.24. Geometriai megoldás: a három egységvektor egy szabályos háromszög három oldalvektora azonos körüljárás szerint irányítva, mivel összegük  $\vec{0}$ . Így hajlásszögük  $2\pi/3 = 120^\circ$ , tehát a vektorpárok skaláris szorzata  $-\frac{1}{2}$ , így az összeg  $-\frac{3}{2}$ .

Algebrai megoldás:  $(\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3) \cdot (\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3) = 0$ , tehát  $0 = \mathbf{e}_1 \cdot \mathbf{e}_1 + \mathbf{e}_2 \cdot \mathbf{e}_2 + \mathbf{e}_3 \cdot \mathbf{e}_3 + 2(\mathbf{e}_1 \cdot \mathbf{e}_2 + \mathbf{e}_1 \cdot \mathbf{e}_3 + \mathbf{e}_2 \cdot \mathbf{e}_3)$ . Kihasználva, hogy a vektorok egységvektorok, kapjuk hogy  $\mathbf{e}_1 \cdot \mathbf{e}_2 + \mathbf{e}_1 \cdot \mathbf{e}_3 + \mathbf{e}_2 \cdot \mathbf{e}_3 = -\frac{3}{2}$ .

$$1.25. |\mathbf{a} \times \mathbf{b}| = |\mathbf{a}||\mathbf{b}| \sin \gamma = 1 \cdot 2 \cdot \frac{1}{2} = 1.$$

1.26.  $\vec{0}$ , hisz párhuzamosak ( $\sin \gamma = 0$ , így abszolút értéke 0).

$$1.27. (\mathbf{a} + \mathbf{b}) \times (\mathbf{a} - \mathbf{b}) = \mathbf{a} \times \mathbf{a} - \mathbf{a} \times \mathbf{b} + \mathbf{b} \times \mathbf{a} - \mathbf{b} \times \mathbf{b} = -\mathbf{a} \times \mathbf{b} - \mathbf{a} \times \mathbf{b} = -2\mathbf{a} \times \mathbf{b}.$$

$$1.28. (\mathbf{i} + \mathbf{j} + \mathbf{k}) \times (\mathbf{i} + \mathbf{j}) = \mathbf{i} \times \mathbf{i} + \mathbf{i} \times \mathbf{j} + \mathbf{j} \times \mathbf{i} + \mathbf{j} \times \mathbf{j} + \mathbf{k} \times \mathbf{i} + \mathbf{k} \times \mathbf{j} = \vec{0} + \mathbf{i} \times \mathbf{j} - \mathbf{i} \times \mathbf{j} + \vec{0} + \mathbf{j} - \mathbf{i} = \mathbf{j} - \mathbf{i}.$$

1.29. Jelölje  $P$  szomszédait  $Q$ ,  $R$  és  $S$ .

a) Ekkor két lapátló-vektor például a  $\vec{PQ} + \vec{PR}$  és a  $\vec{PR} + \vec{PS}$  vektorok. Ezek szorzata:

$$\begin{aligned} (\vec{PQ} + \vec{PR}) \cdot (\vec{PR} + \vec{PS}) &= \\ \vec{PQ} \cdot \vec{PR} + \vec{PQ} \cdot \vec{PS} + \vec{PR} \cdot \vec{PR} + \vec{PR} \cdot \vec{PS} &= \\ \vec{PR} \cdot \vec{PR} &= 1. \end{aligned}$$

Kihasználtuk, hogy merőleges vektorok skaláris szorzata 0.

b) Hasonlóan kapható meg egy lapátló-vektor és a testátló-vektor  $(\vec{PQ} + \vec{PR} + \vec{PS})$  szorzata:

$$(\vec{PQ} + \vec{PR}) \cdot (\vec{PQ} + \vec{PR} + \vec{PS}) = \vec{PQ} \cdot \vec{PQ} + \vec{PR} \cdot \vec{PR} = 2.$$

c) A  $Q$ ,  $R$  és  $S$  csúcsok olyan sorrendben legyenek megválasztva, hogy  $\vec{PQ}$ ,  $\vec{PR}$  és  $\vec{PS}$  ebben a sorrendben jobbrendszeret alkosszon. Ki fogjuk használni, hogy ekkor  $\vec{PQ} \times \vec{PR} = \vec{PS}$ . Egy élvektor és egy szomszédos lapátló-vektor vektori szorzata:

$$\begin{aligned} \vec{PQ} \times (\vec{PQ} + \vec{PR}) &= \\ \vec{PQ} \times \vec{PQ} + \vec{PQ} \times \vec{PR} &= \\ \vec{0} + \vec{PS} &= \vec{PS}, \end{aligned}$$

vagyis a szorzat a két vektor lapjára merőleges élvektor.

d) Legyen a lapvektor a  $\vec{PR}$ , a nem szomszédos lapátló-

vektor  $\overrightarrow{PR} + \overrightarrow{PS}$ . Ezek szorzata:

$$\begin{aligned}\overrightarrow{PQ} \times (\overrightarrow{PR} + \overrightarrow{PS}) &= \\ \overrightarrow{PQ} \times \overrightarrow{PR} + \overrightarrow{PQ} \times \overrightarrow{PS} &= \overrightarrow{PS} - \overrightarrow{PR},\end{aligned}$$

ami a lapátló-vektor síkjának másik lapátló-vektora.

**1.30.** Ha  $\mathbf{u} \perp \mathbf{v}$  és  $\mathbf{u} \perp \mathbf{w}$ , akkor  $\mathbf{u} \cdot \mathbf{v} = 0$  és  $\mathbf{u} \cdot \mathbf{w} = 0$ , így bármely  $c, d \in \mathbb{R}$  számokra  $\mathbf{u} \cdot (c\mathbf{v} + d\mathbf{w}) = \mathbf{u} \cdot (c\mathbf{v}) + \mathbf{u} \cdot (d\mathbf{w}) = c\mathbf{u} \cdot \mathbf{v} + d\mathbf{u} \cdot \mathbf{w} = c \cdot 0 + d \cdot 0 = 0$ , tehát  $\mathbf{u}$  merőleges a  $c\mathbf{v} + d\mathbf{w}$  lineáris kombinációra.

**1.31.** Három különböző dolog (így három vektor is) hatféleképp rakható sorba. Ha az  $\mathbf{a}$ ,  $\mathbf{b}$  és  $\mathbf{c}$  vektorok jobbrendszert alkotnak, akkor ugyancsak jobbrendszert alkotnak a  $\mathbf{b}$ ,  $\mathbf{c}$ ,  $\mathbf{a}$  és a  $\mathbf{c}$ ,  $\mathbf{a}$ ,  $\mathbf{b}$  vektorhármasok is. A további három esetben, azaz a  $\mathbf{c}$ ,  $\mathbf{b}$ ,  $\mathbf{a}$ , valamint a  $\mathbf{b}$ ,  $\mathbf{a}$ ,  $\mathbf{c}$  és az  $\mathbf{a}$ ,  $\mathbf{c}$ ,  $\mathbf{b}$  hármasok esetén balrendszert kapunk a vegyes szorzatról tanultak szerint.

**1.33.** Egyik lehetőség a megoldásra:  $\|\mathbf{b}\|\mathbf{a} = \|\mathbf{a}\|\mathbf{b} = \|\mathbf{a}\|\mathbf{b}\|$ , ezért a paralelogramma-módszert egy rombuszra kell alkalmazni. Egy másik lehetőség: az  $\mathbf{a}/\|\mathbf{a}\|$  és  $\mathbf{b}/\|\mathbf{b}\|$  két egységvektor, így összegük szögfelező, mivel a paralelogramma-módszer rombuszt ad. E vektor  $\|\mathbf{a}\|\mathbf{b}\|$ -szerese ugyanúgy szögfelező, és épp ez a feladatbeli vektor.

**1.34.** Használjuk az 1.13. példa eredményét!

**1.35.** Milyen irányokat cserél föl a tükör, és milyeneket nem? Nem cseréli föl a síkkal párhuzamos irányokat: minden, a tükör síkjával párhuzamos vektor tükörképe ön maga. Tehát, ha a tükör előtt állunk, és a tükör is függőleges, akkor a „fölfelé” irány a tükörképen sem változik. Viszont a tükör fölcseréli a tükörre merőleges irányokat. Mielőtt megnézzük, hogy hogy cserélődik fel a jobb és a bal, definiálnunk kell mi az, hogy jobb és bal? Egy lehetőség a definiálásra: ha értelmezve van egy viszonyítási rendszerben (pl. az emberi testhez képest, vagy a mozgó járműben...) a *föl* és az *előre*, melyek egymásra merőleges irányok, akkor a *jobb* irány az *előre*  $\times$  *föl* vektori szorzattal definiálható. Ennek képe a tükörben viszont  $(-\text{előre}) \times \text{föl} = -\text{jobb}$ , ami

épp a *bal*.

**1.36.** A feladat szerint  $c \in R_a$  pontosan akkor teljesül, ha  $a R c$ . Tegyük fel, hogy  $R_a$  és  $R_b$  nem diszjunkt. Ha  $c$  egy közös elemük, akkor  $c$  az  $a$ -val és  $b$ -vel is relációban van, azaz  $a R c$  és  $b R c$ , de a szimmetria miatt  $c R b$ , a tranzitivitás miatt pedig az  $a R c$  és  $c R b$  relációkból következik az  $a R b$ . Ekkor pedig a tranzitivitást használva bármely  $x$  elemere  $b R x$ -ből következik  $a R x$ , azaz  $x \in R_b$ -ből következik  $x \in R_a$ , azaz  $R_b \subseteq R_a$ . Az  $a$  és  $b$  szerepét megfordítva kapjuk  $R_a \subseteq R_b$ , tehát  $R_a = R_b$ . Végül be kell még látnunk, hogy e halmazok uniója kiadja az egész  $X$  halmazt. Ez igaz, hisz minden  $a$  elemre  $a R a$ , azaz  $a \in R_a$ .

**1.37.** Tekintsünk egy  $\overrightarrow{AB}$  és egy  $\overrightarrow{CD}$  irányított szakaszt! Azt mondjuk, hogy ezek relációban vannak, ha van egy olyan eltolás, mely  $A$ -t  $C$ -be,  $B$ -t  $D$ -be viszi. E reláció ekvivalenciareláció (ellenőrizzük), így egy osztályozást definiál az irányított szakaszok halmazán. Egy ilyen osztályt nevezzünk (szabad) vektornak.

**1.38.** A vektor iránya a félegyenesek, az állása az egyenesek halmazán – az előző feladathoz hasonlóan az eltolással – definiált ekvivalencia reláció egy ekvivalenciaosztálya.

**1.39.** ...

**1.40.** Minden olyan vektor, amelyben páros sok 1-es van, eleget tesz a feladatbeli feltételnek, a többi nem.

**1.41.** Ez a személyi szám érvényes, mert

$$\begin{aligned}(1, 2, 3, 4, 5, 6, 7, 8, 9, 10) \cdot (2, 6, 0, 1, 2, 3, 1, 0, 0, 1) &= \\ 1 \cdot 2 + 2 \cdot 6 + 3 \cdot 0 + 4 \cdot 1 + 5 \cdot 2 + 6 \cdot 3 + 7 \cdot 1 + 8 \cdot 0 + 9 \cdot 0 + 10 \cdot 1 \bmod 11 &= 63 \bmod 11 = 8.\end{aligned}$$

Ez az EAN-kód nem érvényes, mert

$$\begin{aligned}(1, 3, 1, 3, 1, 3, 1, 3, 1, 3, 1) \cdot (9, 9, 9, 8, 8, 8, 7, 7, 6, 6, 5) &= \\ (9 + 3 \cdot 9 + 9 + 3 \cdot 8 + 8 + 3 \cdot 8 + 7 + 3 \cdot 7 + 7 + 3 \cdot 6 + 6 + 3 \cdot 6 + 5) \bmod 10 &= 183 \bmod 10 = 3 \neq 0.\end{aligned}$$

**1.42.** Mivel  $(10, 9, 8, 7, 6, 5, 4, 3, 2, 1) \cdot (9, 6, 3, 0, 7, 6, 1, 9, 8, e) = 10 \cdot 9 + 9 \cdot 6 + 8 \cdot 3 + 6 \cdot 7 + 5 \cdot 6 + 4 \cdot 1 + 3 \cdot 9 + 2 \cdot 8 + e = 287 + e \bmod 11 = 1 + e = 0$ , ezért  $e = 10$ , azaz a teljes kód 963076198X (a könyvre olvashatóbban írva 963-07-6198-X).



## 2

# Lineáris egyenletrendszerek és megoldásuk

A lineáris egyenletrendszerek geometriai megközelítése után megismerjük a megoldás technikáit, végül a megoldások halmazának szerkezetét!

### Egyenes és sík egyenletei

A sík és a tér pontjainak és vektorainak koordinátáit használva lehetővé válik geometriai alakzatok algebrai vizsgálata, vagy algebrai problémák jobb megértése geometriai szemléltetéssel.

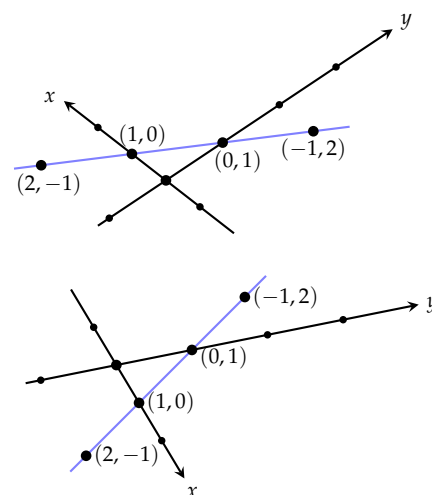
### Alakzatok és egyenletek

**2.1. PÉLDA (AZ  $x + y = 1$  EGYENLET).** Egy tetszőleges síkbeli koordinátarendszerben ábrázoljunk néhány pontot, melynek koordinátái kielégítik az  $x + y = 1$  egyenletet. Fogalmazzunk meg sejtést az egyenlet összes megoldásának megfelelő pontok halmazáról!

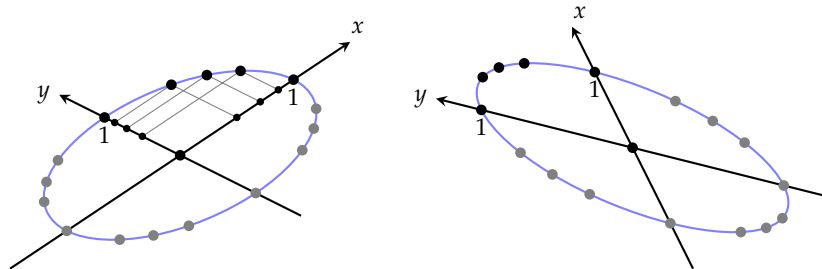
**MEGOLDÁS.** A 2.1 ábrán két különböző koordinátarendszert ábrázolunk, és azokban a fenti egyenletet kielégítő pontok közül néhányat. Ennek alapján azt sejtethetjük, hogy az  $x + y = 1$  egyenletet kielégítő pontok egy egyenesen vannak. Ezt az egyenest is berajzoltuk. A sejtést hamarosan bizonyítjuk.  $\square$

**2.2. PÉLDA (AZ  $x^2 + y^2 = 1$  EGYENLET).** Egy tetszőleges síkbeli koordinátarendszerben ábrázoljunk néhány pontot, melynek koordinátái kielégítik az  $x^2 + y^2 = 1$  egyenletet. Fogalmazzunk meg sejtést az egyenlet összes megoldásának megfelelő pontok halmazáról!

**MEGOLDÁS.** Az alábbi ábrán néhány koordinátarendszert ábrázoltunk, az  $x^2 + y^2 = 1$  egyenletet kielégítő néhány ponttal. A ??? fejezetben visszatérünk e feladatra, és meg fogjuk mutatni, hogy az egyenletet kielégítő pontok egy ellipszisen vannak.  $\square$



2.1. ábra: Az  $x + y = 1$  egyenletet kielégítő néhány pont két különböző koordinátarendszerben.



2.2. ábra: Az  $x^2 + y^2 = 1$  egyenletet kielégítő  $(x, y)$  pontok halmaza két koordináta-rendszerben.

**2.3. DEFINÍCIÓ (ALAKZAT (IMPLICIT) EGYENLETRENDSZERE).** Egy geometriai alakzat egy adott koordináta-rendszerre vonatkozó (implicit) egyenletrendszerén olyan egyenletrendszert értünk, melynek egyszerre minden egyenletét kielégítik a térnek az alakzathoz tartozó pontjai, de más pontok nem. Ha az egyenletrendszer egy egyenletből áll, az alakzat egyenletéről beszélünk. Az egyenletet vektoregyenletnek nevezzük, ha nem a pontok koordinátáira, hanem a pontokba mutató vektorokra írjuk fel. Egy alakzat  $m$  egyenletből álló egyenletrendszerének, illetve  $m$  vektoregyenletből álló egyenletrendszerének általános alakja

$$\begin{cases} F_1(x_1, x_2, \dots, x_n) = 0 \\ F_2(x_1, x_2, \dots, x_n) = 0 \\ \vdots \\ F_m(x_1, x_2, \dots, x_n) = 0 \end{cases} \quad \text{illetve} \quad \begin{cases} F_1(\mathbf{r}) = 0 \\ F_2(\mathbf{r}) = 0 \\ \vdots \\ F_m(\mathbf{r}) = 0 \end{cases}$$

ahol  $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$  a tér egy pontja, és  $\mathbf{r}$  az oda mutató vektor.

Középiskolai tanulmányainkban több példát láttunk alakzat egyenletére, például tudjuk, hogy a síkban a koordinátatengelyek szögét felező egyenes egyenlete  $y = x$ , azaz  $x - y = 0$ . Ortonormált bázist választva az origó közepű egységsugarú kör egyenlete  $x^2 + y^2 = 1$ . Az előző két egyenlet mindegyikéből kifejezhető a két koordináta egy paraméter bevezetésével. Az  $y = x$  egyenlet ekvivalens az

$$x = t$$

$$y = t$$

egyenletrendszerrel, míg az  $x^2 + y^2 = 1$  egyenlet ekvivalens az

$$x = \cos t$$

$$y = \sin t$$

egyenletrendszerrel. Mindkettő átírható vektoralakba is. Használjuk a oszlopvektoros jelölést:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} t \\ t \end{bmatrix}, \quad t \in \mathbb{R}, \quad \text{illetve} \quad \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos t \\ \sin t \end{bmatrix}, \quad t \in [0, 2\pi) \subseteq \mathbb{R}.$$

A latin eredetű *implicit* szó jelentése *nem kifejtett, rejtett*, ami az összeköt, össze-függ, összekever, körülcsavar jelentésű *implico* (implicō) szó származéka. E szó a matematikában az implicit alak, implicit függvény, stb. kifejezésekben arra utal, hogy valamely fontosnak tekintett mennyiség, változó, stb. nincs kifejezve a képletből. Ugyanennek a szónak a származéka a magába foglal, maga után von jelentésű *implikál* szó is, mely a matematikai logika „ha... akkor...” szerkezetű műveletével, az *implikációval* is kapcsolatban van.



E két példa vezet a következő általános fogalomhoz.

**2.4. DEFINÍCIÓ (ALAKZAT (EXPLICIT) EGYENLETRENDSZERE).** Egy geometriai alakzat egy adott koordináta-rendszerre vonatkozó (explicit) egyenletrendszerén olyan egyenletrendszert értünk, melyben az egyenletek bal oldalán a pontok koordinátáit megadó változók, jobb oldalán adott paraméterek függvényei szerepelnek. Általános alakja

$$\begin{aligned}x_1 &= f_1(t_1, t_2, \dots, t_k) \\x_2 &= f_2(t_1, t_2, \dots, t_k) \\&\vdots \\x_n &= f_n(t_1, t_2, \dots, t_k)\end{aligned}$$

ahol  $t_1 \in I_1, t_2 \in I_2, \dots, t_n \in I_n$ , és  $I_1, \dots, I_n \subseteq \mathbb{R}$ . Az ilyen egyenletrendszer egyetlen vektoregyenletté fogható össze:

$$\mathbf{r} = \mathbf{f}(t_1, t_2, \dots, t_k),$$

ahol  $\mathbf{f}$  egy  $\mathbb{R}^k \rightarrow \mathbb{R}^n$  függvény. Az explicit egyenletrendszereket szokás paraméteres egyenletrendszernek is nevezni.

A következő paragrafusokban egyenes és sík különböző egyenleteit, egyenletrendszereit fogjuk áttekinteni példákat adva a fenti két általános definícióra.

**Síkbeli egyenes egyenletei** Tekintsük a sík egy tetszőleges  $e$  egyenesét, és jelöljük ki a síkban az  $O$  origót. Legyen a nemzérus  $\mathbf{v}$  egy tetszőleges, az egyenessel párhuzamos vektor. Az ilyen vektorokat az egyenes *irányvektorának* nevezzük. Mutasson  $\mathbf{r}_0$  az egyenes egy tetszőleges, kijelölt pontjába. Világos, hogy az  $e$  egyenes bármely pontjába mutató  $\mathbf{r}$  vektor előáll  $\mathbf{r}_0 + t\mathbf{v}$  alakban, ahol  $t$  valós szám. Másrészt ha  $Q$  a sík egy tetszőleges, nem az  $e$  egyenesre eső pontja, akkor az  $\overrightarrow{OQ} - \mathbf{r}_0$  vektor nem párhuzamos  $\mathbf{v}$ -vel, tehát nem is konstansszoros, azaz  $\overrightarrow{OQ} - \mathbf{r}_0 \neq t\mathbf{v}$  semmilyen  $t$ -re sem, így  $\overrightarrow{OQ}$  nem áll elő  $\mathbf{r}_0 + t\mathbf{v}$  alakban. Tehát az  $e$  tetszőleges pontjába mutató  $\mathbf{r}$  vektor felírható  $\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}$  alakban, és ez csak  $e$  pontjaira igaz.

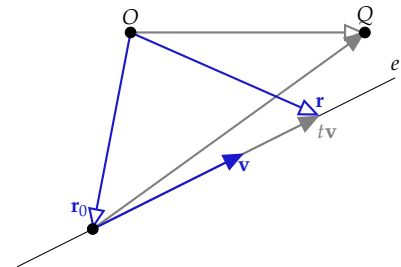
**2.5. ÁLLÍTÁS (SÍKBELI EGYENES EXPLICIT VEKTOREGYENLETE).** A sík minden egyenesének van

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}, \quad t \in \mathbb{R} \quad (2.1)$$

alakú vektoregyenlete, és minden ilyen alakú egyenlet egy egyenes egyenlete, ahol  $\mathbf{v}$  az egyenes egy irányvektora, és  $\mathbf{r}_0$  az egyenes egy tetszőleges, rögzített pontjába mutató vektor.

A síkbeli egyenesre merőleges vektorokat az egyenes *normálvektora*

A latin eredetű *explicit* szó jelentése *ki-fejtett, világosan kimondott*, ami a kibont, szétterít, kiszabadít, átvitt értelemben tisztáz, kifejti, megfejt jelentésű *explico* (explicō) szó származéka. E szó a matematikában az explicit alak, explicit függvény, stb. kifejezésekben arra utal, hogy valamely fontosnak tekintett mennyiség, változó, stb. ki van fejezve a többi segítségével.



2.3. ábra: Egyenes explicit vektoregyenlete:  $\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}$ .

rainak nevezzük. Legyen  $\mathbf{n}$  egy tetszőleges, a  $\mathbf{v}$ -re merőleges vektor, azaz legyen  $\mathbf{n}$  az  $e$  egy normálvektora. Azt, hogy  $\mathbf{r} - \mathbf{r}_0$  az  $e$  tetszőleges pontjába mutató  $\mathbf{r}$  vektorra párhuzamos  $\mathbf{v}$ -vel, úgy is kifejezhetjük, hogy  $\mathbf{r} - \mathbf{r}_0$  merőleges  $\mathbf{n}$ -re. A merőlegesség pedig kifejezhető a skaláris szorzattal. Így az egyenes egy implicit vektoregyenletéhez jutunk:  $\mathbf{r}$  pontosan akkor mutat az  $e$  egy pontjába, ha  $\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0$ . Ez az egyenlet átrendezés után  $\mathbf{n} \cdot \mathbf{r} = \mathbf{n} \cdot \mathbf{r}_0$  alakra, majd a  $C = \mathbf{n} \cdot \mathbf{r}_0$  jelölés bevezetésével  $\mathbf{n} \cdot \mathbf{r} = C$  alakra hozható.

2.6. ÁLLÍTÁS (SÍKBELI EGYENES IMPLICIT VEKTOREGYENLETE). A sík minden egyenesének van

$$\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0, \quad (2.2)$$

és vele ekvivalens

$$\mathbf{n} \cdot \mathbf{r} = C \quad (2.3)$$

alakú vektoregyenlete, és minden ilyen alakú egyenlet egy egyenes egyenlete, ahol  $\mathbf{n}$  az egyenes egy normálvektora,  $\mathbf{r}_0$  az egyenes egy tetszőleges, de rögzített pontjába mutató vektor és  $C$  konstans.

A (2.2) alakú egyenlet könnyen átírható (2.3) alakúvá a  $C = \mathbf{n} \cdot \mathbf{r}_0$  jelöléssel. Az átalakítás fordított irányban is egyszerű, hisz ha  $\mathbf{n} \cdot \mathbf{r} = C$ , akkor találunk olyan  $\mathbf{r}_0$  vektort, melyre  $\mathbf{n} \cdot \mathbf{r}_0 = C$ . Ez azért igaz, mert ha tetszőleges  $\mathbf{n}$ -re nem merőleges  $\mathbf{v}$  vektorra  $\mathbf{n} \cdot \mathbf{v} = D$ , akkor  $\mathbf{n} \cdot (\frac{C}{D}\mathbf{v}) = C$ , így az  $\mathbf{r}_0 = \frac{C}{D}\mathbf{v}$  megfelel.

Az  $\mathbf{r} = (x, y)$ ,  $\mathbf{r}_0 = (x_0, y_0)$  és  $\mathbf{v} = (a, b)$  jelöléseket használva az explicit vektoregyenlet azonnal egyenletrendszerre alakítható.

2.7. ÁLLÍTÁS (SÍKBELI EGYENES EXPLICIT EGYENLETRENDSZERE). A sík minden egyenesének van

$$\begin{aligned} x &= x_0 + at \\ y &= y_0 + bt \end{aligned} \quad (2.4)$$

alakú egyenletrendszere, ahol  $(a, b)$  az egyenes egy irányvektora, és  $(x_0, y_0)$  az egyenes egy tetszőleges rögzített pontja.

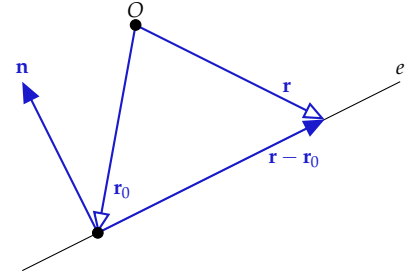
A következőkben megmutatjuk, hogy az explicit egyenletrendszerből a  $t$  paraméter kiküszöbölhető, és így egy implicit egyenletet kapunk.

2.8. ÁLLÍTÁS (SÍKBELI EGYENES (IMPLICIT) EGYENLETE). A sík minden egyenesének van

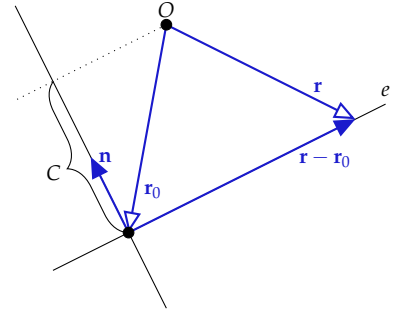
$$Ax + By = C \quad (2.5)$$

alakú egyenlete, és minden ilyen alakú egyenlet egy egyenes egyenlete, ahol  $A$  és  $B$  közül nem mindkettő nulla, és  $(-B, A)$  az egyenes egy irányvektora.

► A bizonyítás előtt érdemes megjegyezni, hogy az egyenes fenti imp-



2.4. ábra: Síkbeli egyenes implicit vektoregyenlete:  $\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0$ .



2.5. ábra: Síkbeli egyenes (implicit) vektoregyenlete:  $\mathbf{n} \cdot \mathbf{r} = C$ . Ha  $\mathbf{n}$  egységvektor, akkor az  $\mathbf{n} \cdot \mathbf{r} = C$  geometriai jelentése az, hogy az egyenes bármely pontjába mutató vektornak az  $\mathbf{n}$  egyenesére eső merőleges vetülete  $C$ . Ez az ábra is ezt az esetet szemlélteti.

licit egyenlete az egyenes  $\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0$  alakú vektoregyenletéből azonnal megkapható, de ezt egyelőre csak *ortonormált* koordinátarendszerben tudjuk könnyen igazolni. Legyen  $(A, B) = (b, -a)$ . Ez az egyenes egy normálvektora, hisz merőleges az  $(a, b)$  irányvektorra. Továbbá  $\mathbf{r} - \mathbf{r}_0 = (x - x_0, y - y_0)$ , ezért a vektoregyenlet

$$(A, B) \cdot (x - x_0, y - y_0) = 0$$

alakú lesz, ami a skaláris szorzást elvégezve az  $A(x - x_0) + B(y - y_0) = 0$  formulát adja. Ha a koordinátarendszer nem ortonormált, az  $(A, B)$  vektor nem szükségképpen normálvektor, és a skaláris szorzás képlete is más, de azért a (2.5) egyenletről mondott állítás igaz. Erre olyan bizonyítást adunk, mely az explicit egyenletrendszerre épül.

**BIZONYÍTÁS.** Ha  $a$  vagy  $b$  valamelyike 0, akkor a két egyenlet egyike felesleges, például ha  $a = 0$ , akkor az egyenletrendszer alakja

$$\begin{aligned} x &= x_0 \\ y &= y_0 + bt \end{aligned}$$

ami ekvivalens az  $x = x_0$  egyenlettel, hisz az  $y = y_0 + bt$  semmi mást nem mond, mint hogy  $y$  egy valós szám. Mivel  $(a, b) \neq (0, 0)$ , ezért csak az az eset marad, amikor  $a$  és  $b$  egyike sem 0. Ekkor mindkét egyenletből kifejezhető  $t$ , és a két értéket egyenlővé téve kapjuk, hogy

$$\frac{x - x_0}{a} = \frac{y - y_0}{b},$$

azaz

$$bx - ay = bx_0 - ay_0, \text{ vagy } b(x - x_0) - a(y - y_0) = 0.$$

Legyen a továbbiakban  $A = b$  és  $B = -a$ . Ekkor a fenti egyenlet  $Ax + By = Ax_0 + By_0$  lesz. Az egyenlet jobb oldalán lévő konstanst  $C$ -vel jelölve az egyenes egyenlete  $Ax + By = C$  alakot ölt. Másrészt könnyen látható, hogy minden ilyen alakú egyenlet egy egyenes egyenlete, mert ekvivalens egy egyenes paraméteres egyenletrendszerével. Nevezetesen az  $Ax + By = C$  egyenlet visszaírható  $Ax + By = Ax_0 + By_0$  alakba, hisz az  $Ax_0 + By_0 = C$  egyenletben  $A \neq 0$  esetén egy tetszőleges  $y_0$ -t választva, egyértelműen kifejezhető  $x_0$ . (A  $B \neq 0$  eset analóg.) Ennek alapján felírható a (2.4) egyenletrendszer.  $\square$

**2.9. PÉLDA (SÍKBELI EGYENES EGYENLETEI).** Írjuk fel annak az egyenesnek összes egyenletét vagy egyenletrendszerét, mely átmegy a  $(2, 3)$  és az  $(1, 1)$  koordinátájú pontokon.

**MEGOLDÁS.** Ha egy egyenes átmegy e két ponton, akkor irányvektora a két pontba mutató vektorok különbsége, azaz  $\mathbf{v} = (2, 3) - (1, 1) = (1, 2)$ . Legyen  $\mathbf{r}_0 = (1, 1)$ , de az  $\mathbf{r}_0 = (2, 3)$  választás is megfelelő.

Az irányvektor segítségével kapjuk, hogy

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Explicit (paraméteres) egyenletrendszer alakban:

$$\begin{aligned} x &= 1 + t \\ y &= 1 + 2t. \end{aligned}$$

Az irányvektorból  $(A, B) = (2, -1)$ , innen az egyenes egyenlete  $2x - y = 2 \cdot 1 - 1 \cdot 1$ , azaz

$$2x - y = 1.$$

Ortonormált koordináta-rendszerben a

$$(2, -1) \cdot (x - 1, y - 1) = 0$$

egyenletet kapjuk vektoregyenletként, mely kiszámolva az előző egyenletet adja.  $\square$

*Síkbeli pont egyenletei* Tekintsük a síkbeli  $(x_0, y_0)$  pontot. Ennek explicit egyenletrendszere, illetve vektoregyenlete:

$$\begin{aligned} x &= x_0 \\ y &= y_0 \end{aligned} \quad \text{illetve} \quad \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}.$$

Ez annyira nyilvánvaló, semmitmondó, hogy a gyakorlatban nem is szoktunk pont egyenleteiről beszélni, e könyvbe is csak didaktikai okokból került, ugyanis a matematikai fogalmak megértésében gyakran nagy segítségünkre van a szélső, extrémális esetek megértése, vizsgálata.

Mivel itt az alakzat csak egyetlen pontból áll, nincs szükség paraméterre, így ez az alak egyúttal implicitnek is tekinthető. Ekkor úgy tekintünk ugyanerre az egyenletrendszerre, mint két egyenes egyenletére, melyek normálvektorai  $(1, 0)$  illetve  $(0, 1)$ , és amelyek metszéspontja a keresett pont:

$$\begin{aligned} x &= x_0 \\ y &= y_0 \end{aligned} \quad \text{vagy minden együtthatót kiírva} \quad \begin{aligned} x + 0y &= x_0 \\ 0x + y &= y_0 \end{aligned}$$

Ez adja az ötletet, egy pont implicit egyenletrendszerének tekinthetnénk két egyenletet, melyek egymást az adott pontban metsző egy-egy egyenes egyenletei. Tehát mondhatjuk, hogy a pont implicit egyenletrendszerének általános alakja:

$$\begin{aligned} A_1x + B_1y &= C_1 \\ A_2x + B_2y &= C_2 \end{aligned}$$

Az azonban nem igaz, hogy minden ilyen alakú egyenletrendszer egy pont egyenletrendszere, mert két egyenes metszheti egymást egyetlen pontban, de lehet, hogy nincs közös pontjuk, és lehet végtelen sok közös pontjuk is. Épp ennek a kérdésnek a részletes vizsgálata lesz a 2. fejezet témája.

**A 3-dimenziós tér síkjainak egyenletei** Tudjuk, hogy két lineárisan független  $\mathbf{u}$  és  $\mathbf{v}$  vektor bármely lineáris kombinációja a két vektor által meghatározott síkban van, továbbá hogy e sík bármely vektora előáll a megadott két vektor lineáris kombinációjaként (ld. 1.8. és 1.11. tételek). Ebből azonnal adódik, hogy a sík egy rögzített pontjába mutató  $\mathbf{r}_0$  vektor segítségével a sík bármelyik pontjába mutató  $\mathbf{r}$  vektor felírható  $\mathbf{r} = \mathbf{r}_0 + s\mathbf{u} + t\mathbf{v}$  alakban.

2.10. ÁLLÍTÁS (SÍK EXPLICIT VEKTOREGYENLETE). *Bármely síknak van*

$$\mathbf{r} = \mathbf{r}_0 + s\mathbf{u} + t\mathbf{v} \quad (2.6)$$

*alakú vektoregyenlete, és minden ilyen alakú egyenlet egy sík egyenlete, ahol  $\mathbf{u}$  és  $\mathbf{v}$  a sík két lineárisan független vektora és  $\mathbf{r}_0$  a sík egy tetszőleges, de rögzített pontjába mutató vektor.*

Hasonlóan a síkbeli egyeneshez, a térbeli sík egyenletéből is kiküszöbölhető a paraméter a merőlegesség felhasználásával. Az 1.30. feladat állítása szerint, ha egy vektor merőleges két tetszőleges vektor mindegyikére, akkor merőleges azok lineáris kombinációjára is. Mivel az  $\mathbf{n} = \mathbf{u} \times \mathbf{v}$  merőleges  $\mathbf{u}$ -ra és  $\mathbf{v}$ -re is, ezért merőleges azok minden lineáris kombinációjára is, azaz az  $\mathbf{r} - \mathbf{r}_0 = s\mathbf{u} + t\mathbf{v}$  vektorra is. Ez az észrevétel az alapja az alábbi tételnek.

2.11. ÁLLÍTÁS (SÍK IMPLICIT VEKTOREGYENLETE). *A háromdimenziós térben minden síknak van*

$$\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0, \quad (2.7)$$

*és a vele ekvivalens*

$$\mathbf{n} \cdot \mathbf{r} = C \quad (2.8)$$

*alakú vektoregyenlete, és minden ilyen alakú egyenlet egy sík egyenlete, ahol  $\mathbf{n}$  a sík egy normálvektora,  $\mathbf{r}_0$  a sík egy tetszőleges, de rögzített pontjába mutató vektor és  $C$  konstans.*

Az állítás igazolása analóg a síkbeli egyenesnél leírtakkal (ld. 2.2. feladat).

Az  $\mathbf{r} = (x, y, z)$ ,  $\mathbf{r}_0 = (x_0, y_0, z_0)$  és  $\mathbf{u} = (a_1, b_1, c_1)$   $\mathbf{v} = (a_2, b_2, c_2)$  jelöléseket használva az explicit vektoregyenlet azonnal egyenletrendszerré alakítható.

**2.12. ÁLLÍTÁS (SÍK EXPLICIT EGYENLETRENDSZERE).** *A háromdimenziós tér minden síkjának van*

$$\begin{aligned}x &= x_0 + a_1s + a_2t \\y &= y_0 + b_1s + b_2t \\z &= z_0 + c_1s + c_2t\end{aligned}\tag{2.9}$$

*alakú egyenletrendszer, ahol  $(a_1, b_1, c_1)$  és  $(a_2, b_2, c_2)$  a sík két lineárisan független vektora, és  $(x_0, y_0, z_0)$  a sík egy tetszőleges rögzített pontja.*

Az explicit egyenletrendszerből kiküszöbölhető a két paraméter, ha például két egyenletből kifejezzük a paramétereket, és behelyettesítjük a harmadik egyenletbe. Így egy implicit egyenletet kapunk. A számításokat nem részletezzük, az eredmény

$$(b_1c_2 - b_2c_1)(x - x_0) + (c_1a_2 - c_2a_1)(y - y_0) + (a_1b_2 - a_2b_1)(z - z_0) = 0.$$

Az  $(A, B, C) = (b_1c_2 - b_2c_1, c_1a_2 - c_2a_1, a_1b_2 - a_2b_1)$  jelöléssel a sík egyenlete  $A(x - x_0) + B(y - y_0) + C(z - z_0) = 0$  alakra hozható, vagy ami vele ekvivalens,  $Ax + By + Cz = D$  alakra.

**2.13. ÁLLÍTÁS (SÍK IMPLICIT EGYENLETE).** *A háromdimenziós térben minden síknak van*

$$Ax + By + Cz = D\tag{2.10}$$

*alakú egyenlete, és minden ilyen alakú egyenlet egy sík egyenlete, ha  $A, B$  és  $C$  legalább egyike nem nulla, és  $D = Ax_0 + By_0 + Cz_0$ , ahol  $(x_0, y_0, z_0)$  a sík valamely pontja.*

A sík fenti egyenlete a sík  $\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0$  alakú vektoregyenletéből is megkapható, de ezt egyelőre csak *ortonormált* koordinátarendszerben tudjuk könnyen igazolni. Mivel

$$(A, B, C) = (b_1c_2 - b_2c_1, c_1a_2 - c_2a_1, a_1b_2 - a_2b_1),\tag{2.11}$$

ami épp az  $\mathbf{u} \times \mathbf{v}$  vektorral egyenlő, ezért  $(A, B, C)$  merőleges a sík minden vektorára, vagyis a sík egy normálvektora. Az  $\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0$  egyenletet koordináták alakba átírva kapjuk, hogy

$$(A, B, C) \cdot (x - x_0, y - y_0, z - z_0) = 0.$$

**2.14. PÉLDA (SÍK EGYENLETEI).** *Írjuk fel annak a síknak az egyenleteit, mely átmegy a  $(0, -1, 2)$ , a  $(-1, 0, 7)$  és a  $(2, 1, 4)$  pontokon.*

**MEGOLDÁS.** A három pontba mutató vektorok különbségei a síkkal párhuzamos vektorok, így azokkal felírható a sík mindegyik egyenlete. Két vektor a lehetséges háromból:

$$\mathbf{u} = (2, 1, 4) - (0, -1, 2) = (2, 2, 2), \text{ és}$$

$$\mathbf{v} = (-1, 0, 7) - (0, -1, 2) = (-1, 1, 5).$$

Ezek alapján például az  $\mathbf{r}_0 = (0, -1, 2)$  választás mellett a sík explicit vektoregyenlete

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix} + s \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} + t \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix},$$

explicit egyenletrendszere

$$\begin{aligned} x &= 2s - t \\ y &= -1 + 2s + t \\ z &= 2 + 2s + 5t. \end{aligned}$$

Mivel a (2.11) képlet szerint  $(A, B, C) = (8, -12, 4)$ , ezért a sík implicit egyenlete  $8(x - 0) - 12(y - (-1)) + 4(z - 2) = 0$ , azaz 4-gyel való osztás és átrendezés után

$$2x - 3y + z = 5.$$

Így ortonormált koordinátarendszerben a

$$(2, -3, 1) \cdot (x, y, z) = 5, \text{ vagy } (2, -3, 1) \cdot (x, y + 1, z - 2) = 0$$

a sík implicit vektoregyenlete. □

**Térbeli egyenes egyenletei** Mindaz, amit a síkbeli egyenes explicit vektoregyenletéről mondtunk a 65. oldalon, lényegében változtatás nélkül megismételhető. Jelöljük ki a térben az origót, és tekintsük azt az  $e$  egyenest, melynek irányvektora  $\mathbf{v}$ , és amely átmegy azon a ponton, melybe az  $\mathbf{r}_0$  vektor mutat. Világos, hogy az  $e$  egyenes bármely pontjába mutató  $\mathbf{r}$  vektor előáll  $\mathbf{r}_0 + t\mathbf{v}$  alakban, ahol  $t$  valós szám, és az  $e$ -re nem illeszkedő pontokra ez nem áll. Így igaz a következő állítás:

**2.15. ÁLLÍTÁS (TÉRBELI EGYENES EXPLICIT VEKTOREGYENLETE).** *A háromdimenziós tér minden egyenesének van*

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{v} \tag{2.12}$$

*alakú vektoregyenlete, és minden ilyen alakú egyenlet egy egyenes egyenlete, ahol  $\mathbf{v}$  az egyenes egy irányvektora, és  $\mathbf{r}_0$  egy tetszőleges, de rögzített pontjába mutató vektor.*

Itt nem tudjuk a paramétert egyetlen vektoregyenletben kiküszöbölni, de az explicit egyenletrendszerré való átírás megy, ha felvesszünk egy koordinátarendszert, melyben  $\mathbf{r} = (x, y, z)$ ,  $\mathbf{r}_0 = (x_0, y_0, z_0)$  és  $\mathbf{v} = (a, b, c)$ :

**2.16. ÁLLÍTÁS (TÉRBELI EGYENES EXPLICIT EGYENLETRENDSZERE).** *A tér minden egyenesének van*

$$\begin{aligned}x &= x_0 + at \\y &= y_0 + bt \\z &= z_0 + ct\end{aligned}\tag{2.13}$$

*alakú egyenletrendszere, ahol  $(a, b, c)$  az egyenes egy irányvektora, és  $(x_0, y_0, z_0)$  az egyenes egy tetszőleges rögzített pontja.*

A fenti explicit (paraméteres) egyenletrendszerből a paraméter ki-küszöbölhető. Ha az  $a$ ,  $b$  és  $c$  számok valamelyike 0, akkor a neki megfelelő fenti egyenletben már nem szerepel  $t$ , akkor nincs is mit tennünk. Ha legalább két egyenletben szerepel  $t$ , akkor mindegyikből kifejezve  $t$ -t, majd egyenlővé téve őket paraméter nélküli egyenleteket kapunk. Például ha  $a$ ,  $b$  és  $c$  egyike sem 0, akkor

$$t = \frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}.$$

A  $t$ -t elhagyva valójában három egyenletet kaptunk:

$$\frac{x - x_0}{a} = \frac{y - y_0}{b}, \quad \frac{x - x_0}{a} = \frac{z - z_0}{c}, \quad \frac{y - y_0}{b} = \frac{z - z_0}{c}.$$

Annak az egyenletnek nincs értelme, amelyikben a nevező 0, de a nevezőkkel való bővítés után kapott

$$b(x - x_0) = a(y - y_0), \quad c(x - x_0) = a(z - z_0), \quad c(y - y_0) = b(z - z_0)$$

egyenletek mindegyike korrekt akkor is, ha 0 valamelyik együttható. E három egyenlet három sík egyenlete, melyek metszészvonala az adott egyenes. Kivétel az az eset, amikor az egyik egyenlet  $0 = 0$  alakú, ilyenkor a másik két egyenlet egy-egy sík egyenlete. Egy egyenes azonban megadható két sík metszészvonalaként, így adódik a következő tétel, melynek bizonyítását feladatként tűzzük ki (ld. 2.3. feladat):

**2.17. ÁLLÍTÁS (TÉRBELI EGYENES IMPLICIT EGYENLETRENDSZERE).** *A tér minden egyenesének van két egyenletből álló egyenletrendszere. Ha az egyenes egy irányvektora  $(a, b, c)$ , akkor a két egyenlet az alábbi három közül bármelyik kettő, amelyik nem  $0 = 0$  alakú:*

$$\begin{aligned}b(x - x_0) &= a(y - y_0) \\c(x - x_0) &= a(z - z_0) \\c(y - y_0) &= b(z - z_0)\end{aligned}\tag{2.14}$$



► A (2.14) egyenletrendszer a következő alakba is átírható:

$$\begin{aligned} bx - ay &= bx_0 - ay_0 \\ cx - az &= cx_0 - az_0 \\ cy - bz &= cy_0 - bz_0 \end{aligned} \quad (2.15)$$

Erről könnyen leolvasható, hogy ha pl.  $a \neq 0$ , akkor a második egyenlet  $b$ -szereséből kivonva az első egyenlet  $c$ -szeresét, a harmadik egyenlet  $a$ -szorosát kapjuk. Hamarosan látni fogjuk, hogy eszerint a harmadik egyenlet elhagyható, anélkül, hogy az egyenletrendszert kielégítő pontok halmaza megváltozna.

**2.18. PÉLDA (TÉRBELI EGYENES EGYENLETRENDSZEREI).** Írjuk fel annak az egyenesnek az explicit és implicit egyenletrendszerét, mely átmegy az  $A(1, 3, 4)$  és  $a) a B(3, 3, 1)$ , illetve  $b) a C(5, 5, -2)$  ponton.

**MEGOLDÁS.** *a)* Az  $A$  és  $B$  pontot összekötő vektor  $= (2, 0, -3)$ . Innen az egyenes explicit egyenletrendszere

$$\begin{aligned} x &= 1 + 2t \\ y &= 3 \\ z &= 4 - 3t, \end{aligned}$$

melynek második egyenlete,  $y = 3$ , egy  $xz$ -síkkal párhuzamos sík egyenlete. A másik két egyenletből kiküszöbölve  $t$ -t, egy másik sík egyenletét kapjuk. Az egyenes ennek a két síknak a metszészvonala. Az első egyenletből  $t = \frac{1}{2}(x - 1)$ , a harmadikból  $t = -\frac{1}{3}(z - 4)$  ezért  $3x + 2z = 11$ . Így az előző egyeneshez a következő implicit (paraméter nélküli) egyenletrendszer tartozik, mely két sík egyenletéből áll:

$$\begin{aligned} 3x + 2z &= 11 \\ y &= 3. \end{aligned}$$

*b)* Az  $A$  és  $C$  pontot összekötő vektor itt  $= (4, 2, -6)$ . Innen az egyenes explicit egyenletrendszere

$$\begin{aligned} x &= 1 + 4t \\ y &= 3 + 2t \\ z &= 4 - 6t. \end{aligned}$$

Mindegyik egyenletből kifejezve  $t$ -t kapjuk, hogy

$$t = \frac{x-1}{4} = \frac{y-3}{2} = \frac{z-4}{-6}.$$

Ez a következő három sík egyenletét adja:

$$\frac{x-1}{4} = \frac{y-3}{2}, \quad \frac{z-4}{-6} = \frac{x-1}{4}, \quad \frac{y-3}{2} = \frac{z-4}{-6}.$$

Átrendezve

$$\begin{aligned}x - 2y &= -5 \\3x &+ 2z = 11 \\3y + z &= 13.\end{aligned}$$

E három sík közül bármely kettő meghatározza az adott egyenest, így e három egyenlet közül bármely kettő az egyenes (implicit) egyenletrendszere.  $\square$

**Térbeli pont egyenletei** Csak a teljesség és az analógiák megértése céljából vizsgáljuk meg a tér egy pontjának lehetséges egyenleteit. A térbeli  $(x_0, y_0, z_0)$  pont explicit egyenletrendszere, illetve vektoregyenlete:

$$\begin{aligned}x &= x_0 \\y &= y_0, \text{ illetve } \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}. \\z &= z_0\end{aligned}$$

Az explicit egyenletrendszert implicit alaknak is tekinthetjük, ekkor három – a koordinátasíkokkal párhuzamos – sík egyenletét látjuk, melyek egyetlen közös pontban metszik egymást.

$$\begin{aligned}x &= x_0 & x + 0y + 0z &= x_0 \\y &= y_0, \text{ vagy minden együtthatót kiírva} & 0x + y + 0z &= y_0. \\z &= z_0 & 0x + 0y + z &= z_0\end{aligned}$$

A síkbeli esethez hasonlóan egy pont implicit egyenletrendszerének tekinthetnénk három egyenletet, melyek egymást az adott pontban metsző egy-egy sík egyenletei. Tehát mondhatjuk, hogy a pont implicit egyenletrendszerének általános alakja:

$$\begin{aligned}A_1x + B_1y + C_1z &= D_1 \\A_2x + B_2y + C_2z &= D_2 \\A_3x + B_3y + C_3z &= D_3.\end{aligned}$$

Itt is óvatosnak kell lennünk, mert nem minden ilyen alakú egyenletrendszer egy pont egyenletrendszere. Például három sík metszheti egymást egy egyenesben, de párhuzamos síkok esetén az is előfordulhat, hogy nincs közös pontjuk. E kérdés vizsgálatára visszatérünk a 2. fejezetben.

**Egyenletek  $\mathbb{R}^n$ -ben** Az egyenes és a sík explicit vektoregyenlete  $\mathbb{R}^n$ -ben is ugyanolyan alakú, mint  $\mathbb{R}^3$ -ben, azaz az egyenes explicit vektoregyenlete  $\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}$ , a síké  $\mathbf{r} = \mathbf{r}_0 + s\mathbf{u} + t\mathbf{v}$  alakú.

**2.19. PÉLDA** (EGYENES ÉS SÍK EXPLICIT VEKTOREGYENLETE). Írjuk fel az  $A(1,1,1,1)$ ,  $B(2,3,2,4)$  pontokon átmenő egyenes, valamint az  $A$ ,  $B$  és  $C(3,2,1,0)$  pontokon átmenő sík explicit vektoregyenletét!

**MEGOLDÁS.** Az  $\overrightarrow{AB} = (1,2,1,3)$  és az  $\overrightarrow{AC} = (2,1,0,-1)$  vektorok segítségével azonnal fölírható az egyenes és a sík egyenlete is:

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 2 \\ 1 \\ 3 \end{bmatrix}, \text{ illetve } \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + s \begin{bmatrix} 1 \\ 2 \\ 1 \\ 3 \end{bmatrix} + t \begin{bmatrix} 2 \\ 1 \\ 0 \\ -1 \end{bmatrix}. \quad \square$$

A síkbeli egyenes és a térbeli sík vektoregyenlete  $\mathbf{n} \cdot \mathbf{r} = c$  alakú. E két esetben ez az egyenlet az  $n$ -dimenziós tér egy  $n-1$ -dimenziós alakzatának egyenlete ( $n=2,3$ ). A későbbiekben látni fogjuk, hogy ez általában is igaz, de e pillanatban még a dimenzió fogalmát sem definiáltuk, ezért egyelőre csak nevet adunk ennek az alakzatnak. Az  $\mathbb{R}^n$  térben  $\mathbf{n} \neq \mathbf{0}$  esetén az  $\mathbf{n} \cdot \mathbf{r} = c$  egyenletet kielégítő  $\mathbf{r}$  vektorok végpontjainak halmazát *hipersíknak* nevezzük. Koordinátás alakban

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = c,$$

ahol  $\mathbf{n} = (a_1, a_2, \dots, a_n)$  a hipersík *normálvektora* (ld. 2.1. feladat),  $\mathbf{r} = (x_1, x_2, \dots, x_n)$  a hipersík egy tetszőleges pontjába mutató vektor.

A következő táblázat összefoglalja geometriai alakzatoknak a továbbiak szempontjából legfontosabb egyenleteit.

|                     |          | Explicit<br>vektoregyenlet                              | Implicit<br>egyenlet(rendszer)                                                         |
|---------------------|----------|---------------------------------------------------------|----------------------------------------------------------------------------------------|
| Síkban              | egyenes  | $\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}$               | $Ax + By = C$                                                                          |
|                     | pont     | $\mathbf{r} = \mathbf{r}_0$                             | $A_1x + B_1y = C_1$<br>$A_2x + B_2y = C_2$                                             |
| Térben              | sík      | $\mathbf{r} = \mathbf{r}_0 + s\mathbf{u} + t\mathbf{v}$ | $Ax + By + Cz = D$                                                                     |
|                     | egyenes  | $\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}$               | $A_1x + B_1y + C_1z = D_1$<br>$A_2x + B_2y + C_2z = D_2$                               |
|                     | pont     | $\mathbf{r} = \mathbf{r}_0$                             | $A_1x + B_1y + C_1z = D_1$<br>$A_2x + B_2y + C_2z = D_2$<br>$A_3x + B_3y + C_3z = D_3$ |
| $\mathbb{R}^n$ -ben | hipersík | ???                                                     | $a_1x_1 + a_2x_2 + \dots + a_nx_n = b$                                                 |
|                     | sík      | $\mathbf{r} = \mathbf{r}_0 + s\mathbf{u} + t\mathbf{v}$ | ???                                                                                    |
|                     | egyenes  | $\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}$               | ???                                                                                    |
|                     | pont     | $\mathbf{r} = \mathbf{r}_0$                             | ???                                                                                    |

2.1. táblázat: Geometriai alakzatok egyenletei: az  $\mathbb{R}^n$ -beli egyenletek közül többet még nem ismerünk, ezeket három kérdőjel jelzi, de arra biztatjuk az Olvasót, hogy az analógia fonalán haladva fogalmazza meg sejtéseit.

*Feladatok*

**2.1.** Mutassuk meg, hogy  $\mathbb{R}^n$  egy tetszőleges  $\mathbf{n} \cdot \mathbf{r} = c$  egyenletű hipersíkjának bármely két pontját összekötő vektor merőleges  $\mathbf{n}$ -re.

**2.2.** Igazoljuk a sík implicit vektoregyenletére vonatkozó [2.11.](#) állítást.

**2.3.** Igazoljuk a térbeli egyenes implicit egyenletrendszerére vonatkozó [2.17.](#) állítást!

## A lineáris egyenletrendszer és két modellje

E szakasz témája a lineáris egyenletrendszerek fogalma és a lineáris egyenletrendszer megoldásának két geometriai interpretációja: hipersíkok metszetének meghatározása és egy vektor lineáris kombinációként való előállítása. A számítások kényelmes könyvelésére bevezetjük a mátrix fogalmát.

**Lineáris egyenlet és egyenletrendszer** Az előző rész végén láttuk, hogy a síkbeli egyenes egyenletének általános alakja  $Ax + By = C$ , ahol  $A$ ,  $B$  és  $C$  konstansok. Ennek általánosításaként jutunk a lineáris egyenlet fogalmához.<sup>1</sup>

2.20. DEFINÍCIÓ (LINEÁRIS EGYENLET). Az

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b \quad (2.16)$$

alakra hozható egyenletet az  $x_1, x_2, \dots, x_n$  ismeretlenekben lineáris egyenletnek nevezzük, ahol  $a_1, a_2, \dots$  és  $a_n$ , valamint  $b$  konstansok. Az  $a_1, a_2, \dots$  és  $a_n$  konstansokat az egyenlet együtthatóinak,  $b$ -t az egyenlet konstans tagjának nevezzük.

<sup>1</sup> *Lineáris*: a vonalas jelentésű latin lineāris szóból ered, mely a lenfonal, horgászszinór, átvitt értelemben vonal, határvonal jelentésű linea (línea) szó származéka. A matematikában egyenessel kapcsolatba hozható, illetve elsőfokú értelemben szokás használni.

► Például az alábbi egyenletek lineárisak:

$$x - 2y = 1, \quad \frac{1}{2}x_1 - \sqrt{2}x_2 + (5 - \pi)x_3 = 0, \quad a \cos 0.87 - 0.15c = 0.23.$$

► A következő egyenletek nem lineárisak az  $x$ ,  $y$  és  $z$  ismeretlenekben:

$$xz - y = 0, \quad x + 2y = 3^z, \quad x \sin z + y \cos z + y = z^2,$$

viszont mindegyikük lineáris az  $x$  és  $y$  ismeretlenekben, hisz ekkor  $z$  paraméter, melynek bármely értéke mellett lineárisak az egyenletek.

► Az

$$x = y, \quad x = 3 - y + 2z$$

egyenletek az  $x$ ,  $y$  és  $z$  ismeretlenekben lineárisak, mert azonos átalakítással a definícióbeli alakra hozhatók:

$$x - y + 0z = 0, \quad x + y - 2z = 3.$$

► Másrészt az

$$\frac{x}{z} + \frac{y}{z} + 2 = 0$$

egyenlet nem lineáris, mert a  $z$ -vel való beszorzás nem azonos átalakítás, tehát a lineáris  $x + y + 2z = 0$  egyenlettel nem ekvivalens.

Lineáris egyenletek egy véges halmazát *lineáris egyenletrendszernek* nevezzük. Az egyenletrendszer ismeretlenek mindegyike az ismeretlenek, amelyek legalább egy egyenletben szerepelnek. Ha egy ismeretlen egy egyenletben nem szerepel, akkor úgy tekintjük, hogy 0 az együtthatója. A jobb áttekinthetőség kedvéért az egyenletrendszereket úgy írjuk fel, hogy az ismeretlenek mindegyik egyenletben ugyanabban a sorrendben szerepeljenek. Egy egyenletrendszer egy egyenletből is állhat.

► Lineáris egyenletrendszerek például a következők:

$$\begin{array}{rclcl} 3x - y = 2 & x_1 & = 3 & & \\ -x + 2y = 6 & x_2 & = 1 & 2x - 3y + z - w = 6. & (2.17) \\ x + y = 6 & x_3 & = 4 & & \end{array}$$

► Elképzelhető, hogy egy egyenletrendszer azonos átalakítása közben olyan egyenletet kapunk, melyben minden együttható 0, azaz amely  $0 = b$  alakú. Az is lehet, hogy egy egyenletrendszerben egyes együtthatók paraméterek. Ilyenkor tudnunk kell, mely változók az ismeretlenek, melyek a paraméterek. Így a következő egyenletrendszerek is lineárisak az  $x$  és  $y$  ismeretlenekben:

$$\begin{array}{rclcl} ax + y = 2a & 3x - y = 0 & x + y = 1 & & \\ x - \frac{1}{a}y = 0 & -x + 2y = 0 & 0 = 2. & & (2.18) \\ & 0 = 0 & & & \end{array}$$

Mindezek után fölírjuk az egyenletrendszer általános alakját:

**2.21. DEFINÍCIÓ (LINEÁRIS EGYENLETRENDSZER).** Lineáris egyenletrendszeren ugyanazokban a változóiban lineáris egyenletek egy véges halmazát értjük. Általános alakja  $m$  egyenlet és  $n$  ismeretlen esetén

$$\begin{array}{rclcl} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n & = & b_1 & & \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n & = & b_2 & & \\ \vdots & & \vdots & & \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n & = & b_m, & & (2.19) \end{array}$$

ahol  $x_1, x_2, \dots, x_n$  az ismeretlenek,  $a_{ij}$  az  $i$ -edik egyenletben az  $x_j$  ismeretlen együtthatóját jelöli, és  $b_i$  az  $i$ -edik egyenlet konstans tagja. Ha mindegyik egyenlet konstans tagja 0, a lineáris egyenletrendszer *homogén*, ha csak egy is különbözik 0-tól, *inhomogén*.

► A (2.17) egyenletrendszerei mind inhomogének, míg a (2.18) középső egyenletrendszere homogén.

A *konzisztens* szó jelentése: belső ellentmondástól mentes. Egyéb jelentései: szilárd, sűrű, tömött, tömör, tartalmas, egyetemes, következetes. A latin *consistens* szóból ered, melynek jelentése helytálló.

**2.22. DEFINÍCIÓ (LINEÁRIS EGYENLETRENDSZER MEGOLDÁSA).** Azt mondjuk, hogy a rendezett  $(u_1, u_2, \dots, u_n)$  szám- $n$ -es megoldása a (2.19) egyenletrendszernek, ha megoldása minden egyenletnek, azaz ha minden egyenletet kielégít az  $x_1 = u_1, x_2 = u_2, \dots, x_n = u_n$  helyettesítéssel. Ha a szám- $n$ -est vektornak tekintjük, megoldásvektorról beszélünk. Az összes megoldás halmazát az egyenletrendszer megoldáshalmazának nevezzük. Egy egyenletrendszert konzisztensnek (vagy megoldhatónak) nevezzük, ha megoldáshalmaza nem üres. Ellenkező esetben az egyenletrendszer inkonzisztens (nem megoldható).

► A (2.17) egyenletrendszerének egy-egy megoldása:  $(x, y) = (2, 4)$ ,  $(x_1, x_2, x_3) = (3, 1, 4)$ ,  $(x, y, z, w) = (2, 0, 2, 0)$ . A harmadik egyenletrendszernek több megoldása is van, például egy másik megoldás az  $(x, y, z, w) = (3, 0, 0, 0)$ .

► A (2.18) első egyenletrendszerének megoldása  $(x, y) = (1, a)$ , a második  $(x, y) = (0, 0)$ . A harmadik egyenletrendszernek nincs megoldása, hisz nincs olyan  $x$  és  $y$  érték, melyre fennállna a  $0x + 0y = 2$  egyenlőség.

► Általában, a

$$0x_1 + 0x_2 + \dots + 0x_n = 0$$

egyenletnek minden szám- $n$ -es megoldása, míg a

$$0x_1 + 0x_2 + \dots + 0x_n = b, \quad (b \neq 0)$$

egyenletnek egyetlen megoldása sincs.

**Ekvivalens lineáris egyenletrendszerek** Tekintsük az alábbi három egyenletrendszert:

$$\begin{array}{rcl} x + y = 3 & x + y = 3 & x = 2 \\ x + 2y = 4 & y = 1 & y = 1 \end{array} \quad (2.20)$$

Mindháromnak  $(x, y) = (2, 1)$  az egyetlen megoldása.

**2.23. DEFINÍCIÓ (EKVIVALENS EGYENLETRENDSZEREK).** Azonos ismeretlenekkel felírt két egyenletrendszert ekvivalensnek nevezzük, ha megoldásaik halmaza azonos.

**2.24. TÉTEL (EKVIVALENS ÁTALAKÍTÁSOK).** Az alábbi transzformációk minden egyenletrendszert ekvivalens egyenletrendszerbe visznek át:

1. két egyenlet felcserélése;
2. egy egyenlet nem nulla számmal való szorzása;
3. egy egyenlet konstansszorosának egy másikhoz adása.

Ezen kívül

4. egy  $0 = 0$  alakú egyenlet elhagyása  
is ekvivalens átalakítás, de ez egyet csökkenti az egyenletek számát.

Ha egy egyenletrendszer több egyenletből áll, mint ahány ismeretlene van, *túlhatalozottnak* nevezzük, míg ha kevesebb egyenletből áll, *alulhatalozottnak*. E fogalmak időnként félrevezető megfogalmazásokhoz és téves következtetésekre vezetnek, ha az az elképzelés alakul ki, hogy a túlhatalozottság azt jelenti: az egyenletek (a feltételek) már „túl sokan” vannak ahhoz, hogy akár csak egy szám- $n$ -es is kielégítse. Később látni fogjuk, hogy ezzel ellentétben nem a „túl sok” egyenlet, hanem az egymásnak ellentmondó egyenletek okozzák az inkonzisztenciát. Hasonlóképp az alulhatalozottság nem jelenti azt, hogy szükségképpen több megoldás is van. Alulhatalozott egyenletrendszer is lehet inkonzisztens. Egyedül annyi mondható: alulhatalozott egyenletrendszernek nem lehet csak egyetlen megoldása.

**BIZONYÍTÁS.** Az első kettő és a negyedik átalakítás nyilvánvalóan nem változtatja meg a megoldások halmazát (a negyedikkel kapcsolatban lásd a 2.11. feladatot). Nézzük a harmadik átalakítást. Tekintsük az *eredeti* egyenletrendszer egy megoldását, és azt az *új* egyenletrendszert, melyet az  $i$ -edik egyenlet  $c$ -szeresének a  $j$ -edikhez adásával kapunk. Világos, az átalakítás előtt is elvégezhetjük a behelyettesítést, akkor viszont egy kielégített egyenlőség konstansszorosát adjuk egy másikhoz, ami így ugyancsak ki lesz elégítve. Tehát az eredeti egyenletrendszer minden megoldása az újnak is megoldása. Másrészt viszont az új egyenletrendszer minden megoldása az eredetinek is megoldása, hisz az visszakapható az újból az  $i$ -edik egyenlet  $-c$ -szeresének a  $j$ -edikhez adásával. Vagyis a két megoldáshalmaz megegyezik. Tehát ez az átalakítás is ekvivalens.  $\square$

**Mátrixok** Az egyenletrendszer megoldásában az ekvivalens átalakítások során a műveleteket csak az egyenletrendszer együtthatóival és konstans tagjaival végezzük, ezért az egyenletrendszer megoldásainak lépéseit elég csak egy olyan táblázaton elvégezni, mely az együtthatókat és a konstansokat tartalmazza. Az ilyen számtáblázatokat *mátrixoknak* nevezzük, ezekkel később külön fejezetben foglalkozunk. A mátrixba írt számokat a *mátrix elemeinek* nevezzük.

A mátrix méretének jellemzéséhez mindig előbb a sorok, majd az oszlopok számát adjuk meg, tehát egy  $m \times n$ -es mátrixnak  $m$  sora és  $n$  oszlopa van. Egy ilyen mátrix általános alakja:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

A mátrixokat<sup>2</sup> általában nagy betűvel jelöljük, e könyvben – követve a műszaki nyelv szokásait – félkövér nagy betűvel. A matematikában elterjedt az a szokás, hogy a mátrixot jelölő nagy betűvel azonos kis betűk jelölik a mátrix elemeit, tehát  $\mathbf{A}$  elemei  $a_{11}, a_{12}, \dots$ . A fenti mátrixra szokás még a tömörebb

$$\mathbf{A} = [a_{ij}]_{m \times n} \text{ vagy egyszerűen az } \mathbf{A} = [a_{ij}]$$

jelölést használni. Sokan használnak szögletes helyett kerek zárójelet a mátrixok jelölésére.

Mindig az első index jelöli a sor, a második az oszlop számát, tehát  $a_{23}$  a 2-dik sor 3-adik oszlop kereszteződésében álló elemet jelöli. Időnként, a félreérthetőség elkerülésére  $a_{ij}$  helyett  $a_{i,j}$ -t írunk (pl.  $a_{n,n-1}$ ).

A mátrix *főátlójába* azok az elemek tartoznak, amelyek ugyanannyi-adik sorban vannak, mint ahányadik oszlopban, azaz a például a fenti mátrixban a főátló elemei  $a_{11}, a_{22}, \dots$

*Mátrix:* a latin mater (māter) (*anya, szülőanya, forrás*) szó származéka a matrix (mātrix), melynek jelentése az európai nyelvekben a következő változásokon ment át: *anyaállat, vemhes állat, anyaméh, bezárt hely, ahonnan valami kifejlődik, bezárt, körülzárt dolgok sokasága, tömbje*. Jelentése az élettanban méh, a geológiában finomszemcsés kő, melybe fossziliák, kristályok, drágakövek vannak zárva, az anatómiában a körmöt, fogat kialakító szövet.

<sup>2</sup> A programnyelvekben – ellentétben a matematikával – a kisbetűvel/nagybetűvel való jelölésnek nincs a mátrixot az elemétől való megkülönböztető szerepe. A legtöbb magasszintű nyelvben az  $\mathbf{A}$ -val jelölt mátrix (informatikai szóhasználatnál *tömb*)  $i$ -edik sorának  $j$ -edik elemét  $\mathbf{A}[i,j]$  vagy  $\mathbf{A}(i,j)$  jelöli. Az alacsonyabb szintű C-típusú nyelvekben nincs 2-dimenziós tömb, a mátrixot egy olyan 1-dimenziós tömb reprezentálja, melynek minden eleme 1-dimenziós tömb, így  $\mathbf{A}[i]$  az  $i$ -edik sort,  $\mathbf{A}[i][j]$  az  $i$ -edik sor  $j$ -edik elemét jelöli. A mátrix alapú nyelvekben egy mátrix egy sorvektorra vagy oszlopvektorra könnyen kiemelhető, pl. az  $\mathbf{A}$  mátrix 2. sorát az  $\mathbf{A}(2, :)$ , 3. oszlopát a  $\mathbf{A}(:, 3)$  kóddal érhetjük el. Sok programnyelvben a tömbök elemeit nem 1-től, hanem 0-tól indexelik, ilyen például a C és a Python is.



A vektorokat is szokás *mátrix jelöléssel, mátrix alakban*, azaz egy 1-soros vagy 1-oszlopos mátrixszal leírni – ahogy azt az első fejezetben mi is tettük. Az  $n \times 1$ -es mátrixot *oszlopvektornak* vagy *oszlopmátrixnak*, az  $1 \times n$ -es mátrixot *sorvektornak* vagy *sormátrixnak* is szokás nevezni. Annak a kérdésnek az eldöntése, hogy egy  $n$ -dimenziós vektort sor- vagy oszlopvektorral reprezentáljunk, döntés (szokás, ízlés) kérdése. Manapság jobban el van terjedve a vektorok oszlopvektoros jelölése, ezért e könyvben alapértelmezésként mi is ezt a jelölést fogjuk használni, de egyes témáknál a másik használatát is bemutatjuk. Így tehát az  $(1, 2)$  vektornak megfelelő sorvektor és oszlopvektor alakja

$$\begin{bmatrix} 1 & 2 \end{bmatrix}, \text{ illetve } \begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

amelyek közül, ha mást nem mondunk, az utóbbit fogjuk a vektor mátrixos jelöléseként használni.

Általában  $\mathbf{a}_j$  jelöli az  $\mathbf{A}$  mátrixból kiválasztható  $j$ -edik oszlopvektort, ha csak oszlopvektorokkal dolgozunk. Ha sor- és oszlopvektorok is együtt szerepelnek, az  $i$ -edik sorvektort  $\mathbf{a}_{i*}$ , a  $j$ -edik oszlopvektort  $\mathbf{a}_{*j}$  jelöli összhangban az elemek indexelésével. Ehhez hasonló jelölést használnak a mátrix alapú nyelvek is (ld. a széljegyzetet). Az  $\mathbf{A}$  mátrix  $i$ -edik sorára az  $(\mathbf{A})_{i*}$ ,  $j$ -edik oszlopára az  $(\mathbf{A})_{*j}$ , elemére az  $(\mathbf{A})_{ij}$  jelölés is használatos.

2.25. PÉLDA (MÁTRIXOK ÉS ELEMEIK). Ha

$$\mathbf{C} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 7 \end{bmatrix}, \text{ akkor } c_{23} = 7, \mathbf{c}_2 = \mathbf{c}_{*2} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}, \mathbf{c}_{2*} = \begin{bmatrix} 4 & 5 & 7 \end{bmatrix}.$$

**Egyenletrendszer mátrixa és bővített mátrixa** Az egyenletrendszer *együtthatómátrixa* az egyenletek együtthatóit, míg *bővített mátrixa*, vagy egyszerűen csak *mátrixa* az egyenletek együtthatóit és konstans tagjait tartalmazza. Az áttekinthetőség érdekében a bővített mátrixban egy függőleges vonallal választhatjuk el az együtthatókat a konstans tagoktól.

A 2.21. definícióbeli általános alak együttható- és bővített mátrixa:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad \left[ \begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right].$$

A gyakorlatban nagy méretű egyenletrendszereket, s így nagy méretű mátrixokat is kezelni kell. Ha elemeik nagy része 0, *ritka mátrixoknak* nevezzük. A nem ritka mátrixokat *sűrűnek* nevezzük. Előbb a kis méretű sűrű mátrixokra hatékony módszerekkel ismerkedünk meg.

VEKTOROK MAGYAR IRODAI és általános iskolában használt jelölése – a tizedes vessző használata miatt – pontosvesszőt tesz a vektor koordinátái közé *elválasztó-jelként*. Magyar nyelvű felsőbb matematika szövegekben ez nem szokás, mi is elkerüljük, és tizedespontot, vektor koordinátái közt vesszőt használunk. Vegyük észre, hogy vektorok sorvektorral (sormátrixszal) való megadásánál írásjelet nem használunk, csak szóközzel választjuk el a koordinátákat!

**2.26. PÉLDA (MÁTRIX HASZNÁLATA A MEGOLDÁSHOZ).** *Oldjuk meg a következő egyenletrendszert!*

$$2x + 3y + 2z = 7$$

$$x + y + z = 3$$

$$2x + 2y + 3z = 6$$

**MEGOLDÁS.** Két lehetséges megoldást mutatunk. A (2.20) egyenletrendszereinek látott utolsó két egyszerűbb alak elérése a cél. Először írjuk fel az egyenletrendszer bővített mátrixát!

$$\begin{array}{l} 2x + 3y + 2z = 7 \\ x + y + z = 3 \\ 2x + 2y + 3z = 6 \end{array} \quad \left[ \begin{array}{ccc|c} 2 & 3 & 2 & 7 \\ 1 & 1 & 1 & 3 \\ 2 & 2 & 3 & 6 \end{array} \right]$$

Kicseréljük az első két egyenletet:

$$x + y + z = 3$$

$$2x + 3y + 2z = 7$$

$$2x + 2y + 3z = 6$$

Kicseréljük az első két sort:

$$\left[ \begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 2 & 3 & 2 & 7 \\ 2 & 2 & 3 & 6 \end{array} \right]$$

Az első egyenlet 2-szeresét kivonjuk a második, majd a harmadik egyenletből (azaz  $-2$ -szeresét hozzáadjuk a második majd a harmadik egyenlethez).

$$x + y + z = 3$$

$$y = 1$$

$$z = 0$$

Az első sor kétszeresét kivonjuk a második és harmadik sorból (azaz az első sor  $-2$ -szeresét hozzáadjuk a második majd a harmadik sorhoz).

$$\left[ \begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array} \right]$$

Az egyenletrendszerről azonnal leolvasható  $y$  és  $z$  értéke. Ezeket az első egyenletbe helyettesítve megkapjuk  $x$  értékét is, nevezetesen  $x + y + z = 3$ , azaz  $y = 1$  és  $z = 0$  behelyettesítése után:  $x + 1 + 0 = 3$ , vagyis  $x = 2$ . Másik megoldási módszerhez jutunk, ha a visszahelyettesítés helyett folytathatjuk az ekvivalens átalakítások sorozatát:

Kivonjuk a második, majd a harmadik egyenletet az elsőből:

$$x = 2$$

$$y = 1$$

$$z = 0$$

Kivonjuk a második, majd a harmadik sort az elsőből:

$$\left[ \begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array} \right]$$

Így olyan alakra hoztuk az egyenletrendszert, illetve a bővített mátrixot, amiből azonnal leolvasható a megoldás:  $(x, y, z) = (2, 1, 0)$ .  $\square$

**Sormodell: hipersíkok metszete** A lineáris egyenletrendszerek szemléltetésére két geometriai modellt mutatunk, melyek segíteni fognak az általánosabb fogalmak megértésében, szemléltetésében.

Tudjuk, hogy a kétváltozós lineáris  $ax + by = c$  egyenletet kielégítő pontok halmaza egyenest alkot, ha  $a$  és  $b$  legalább egyike nem 0. (Ha  $a = b = c = 0$ , akkor az egyenlet alakja  $0x + 0y = 0$ , azaz  $0 = 0$ , ami minden  $(x, y)$  számpárra fennáll, tehát a megoldások halmaza a sík összes pontjának halmazával azonos. Ha  $a = b = 0$ , de  $c \neq 0$ , akkor az egyenletnek nincs megoldása, a megoldáshalmaz üres.)

**2.27. PÉLDA (SORMODELL KÉT KÉTISMERETLENES EGYENLETTEL).** Ábrázoljuk az alábbi egyenletrendszereket és megoldásukat a sormodellben!

$$\begin{array}{rcl} x + y = 3 & \text{az} & x + 2y = 3 \\ x + 2y = 4 & \text{és az} & 2x + 4y = 6 \end{array}$$

**MEGOLDÁS.** Az első egyenletrendszer szerinti ábra egy metsző egyenespárt tartalmaz. Metszéspontjuk a megoldás. Ezt a 2.6 ábra felső rajza mutatja. Oldjuk meg az egyenletrendszert! A megoldás közben két újabb egyenletrendszert kapunk:

$$\begin{array}{rcl} x + y = 3 & \Rightarrow & x + y = 3 \\ x + 2y = 4 & \Rightarrow & y = 1 \end{array} \Rightarrow \begin{array}{l} x = 2 \\ y = 1 \end{array}$$

A 2.6 ábrán az így kapott egyenletrendszerek sormodell szerinti ábráit is megrajzoltuk.

A második egyenletrendszerben két párhuzamos egyenes egyenlete szerepel. Ezeknek nincs közös pontjuk, így az egyenletrendszer nem oldható meg. Ha az első egyenlet kétszeresét kivonjuk a másodikból, az ellentmondó  $0 = 1$  egyenletet kapjuk, vagyis így is arra jutottunk, hogy az egyenletrendszer nem oldható meg. Valóban, a  $0x + 0y = 1$  egyenletet kielégítő pontok üres halmazt adnak.

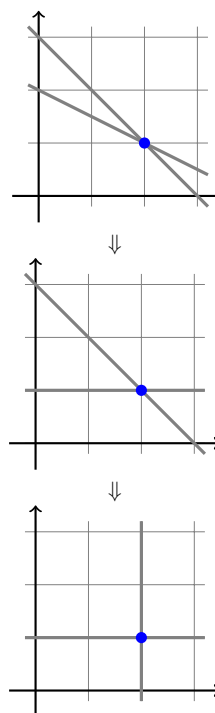
A harmadik egyenletrendszer egyenleteihez két egybeeső egyenes tartozik. Az egyenletrendszer megoldáshalmaza tehát ennek az egyenesnek a pontjaiból áll.

Ha az első egyenlet kétszeresét kivonjuk a másodikból, itt egy  $0 = 0$  egyenletet kapjuk, amely így elhagyható. A megmaradó  $x + 2y = 3$  egyenlet összes megoldása paraméteres alakba írva például  $(x, y) = (3 - 2t, t)$ . □

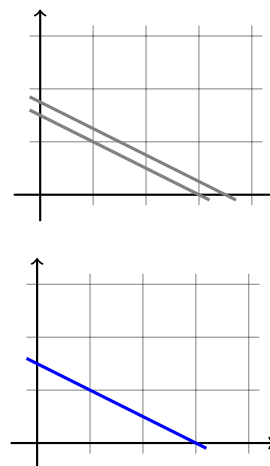
► Hasonlóan szemléltethető a 3-dimenziós térben a háromismeretlenes egyenletrendszerek megoldása. Föl soroljuk a három háromismeretlenes egyenletből álló egyenletrendszerre vonatkozó lehetőségeket:

- Ha a három egyenlettel meghatározott három sík általános helyzetű, akkor az egyenletrendszernek egyetlen megoldása van (ld. 2.8.

Egyenletrendszer megoldásának szemléltetése a sormodellben jól nyomon követhető a SagePlayer sormodell című demonstrációján. Ott saját bővített mátrixokkal is lehet kísérletezni.



2.6. ábra: Egyenletrendszer megoldásának szemléltetése



2.7. ábra: A megoldás szemléltetése, ha a két egyenlet egyikének bal oldala nullává tehető

(a) ábra). Például a 2.26. példában szereplő egyenletrendszernek egyetlen megoldása van:  $(x, y, z) = (2, 1, 0)$ .

- Lehet, hogy a három sík metszete egy egyenes. Ilyen például a

$$\begin{aligned} 2x + y + 2z &= 5 \\ x + y + z &= 3 \\ 3x + 2y + 3z &= 8 \end{aligned} \quad (2.21)$$

egyenletrendszer. Ekkor a három normálvektor egy síkba (de nem egy egyenesbe) esik. Itt például a normálvektorok közt fennáll a

$$(2, 1, 2) + (1, 1, 1) - (3, 2, 3) = \mathbf{0}$$

összefüggés. Ugyanez a lineáris kapcsolat áll fenn az egyenletek közt is, vagyis a harmadik egyenlet az első kettő összege, ami azt jelenti, hogy a harmadik egyenlet el is hagyható.

- Változtassuk meg a fenti egyenletrendszer harmadik egyenletének konstans tagját:

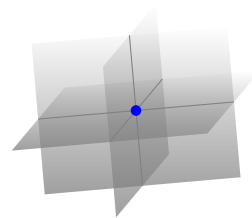
$$\begin{aligned} 2x + y + 2z &= 5 \\ x + y + z &= 3 \\ 3x + 2y + 3z &= 9 \end{aligned} \quad (2.22)$$

Ennek sormodell szerinti ábrája három olyan síkot tartalmaz, melyek párhuzamosak egy egyenessel, de nincs közös pontjuk, mint az a 2.9. (b) ábrán látható. Ha kivonjuk az első és második egyenletet a harmadikból, akkor az ellentmondó  $0 = 1$  egyenletre jutunk, míg az imént a 2.21 egyenletrendszer esetén az elhagyható  $0 = 0$  egyenletet kaptuk. Ennek oka, hogy bár a normálvektorok közt ugyanaz a lineáris kapcsolat van mint az imént (egy síkba esnek), az egyenletek lineárisan függetlenek! Általában is igaz, nincs megoldása az olyan egyenletrendszereknek, ahol bizonyos normálvektorok lineárisan összefüggők, de a hozzájuk tartozó egyenletek már nem. Gondoljuk meg, ilyen eseteket mutat a 2.9. (a) ábra is.

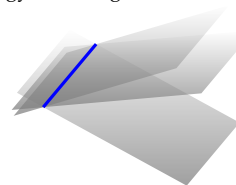
- Végül tekintsük az

$$\begin{aligned} x + y + z &= 3 \\ 2x + 2y + 2z &= 6 \\ 3x + 3y + 3z &= 9 \end{aligned} \quad (2.23)$$

egyenletrendszert! Látható, hogy a második és a harmadik egyenlet az első konstansszorososa, azaz ugyanannak a síknak az egyenletei, az egyenletrendszer tehát ekvivalens az egyetlen  $x + y + z = 3$  egyenletből álló egyenletrendszerrel. Az  $y$ -nak és  $z$ -nek tetszőleges értéket választunk, például legyen  $y = s$ ,  $z = t$ , akkor  $x = 3 - y - z$ , azaz  $x = 3 - s - t$ . Így az összes megoldás:  $(x, y, z) = (3 - s - t, s, t)$ . Ezt



(a) Három általános helyzetű sík: egyetlen megoldás

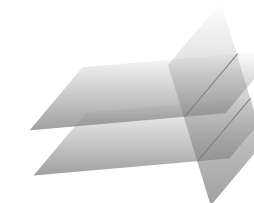


(b) Egy egyenesen átmenő, de nem csupa azonos sík: végtelen sok megoldás, a megoldások egy egyenest alkotnak



(c) Azonos síkok: végtelen sok megoldás, a megoldások egy síkot alkotnak

2.8. ábra: Konzisztens (megoldható) egyenletrendszerek ábrázolása (a megoldáshalmazt kék szín jelzi)



(a) A síkok közül legalább kettő párhuzamos, de nem azonos.



(b) Egy egyenessel párhuzamos, de egymással nem párhuzamos és közös egyenest sem tartalmazó három sík.

2.9. ábra: Nem megoldható egyenletrendszerek szemléltetése

oszlopvektorokkal fölírva kapjuk, hogy a megoldás

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Ez a 2.8. (c) ábra szerint eset.

► A 2- és 3-dimenziós esetek analógiájára kialakíthatunk egy elképzelést az  $n$ -dimenziós esetre is. Itt minden egyenlet megoldáshalmaza a tér egy hipersíkja, kivéve a  $0 = 0$  egyenletet, mert annak megoldáshalmaza az egész tér, és a  $0 = 1$  egyenletet, mert annak megoldáshalmaza az üreshalmaz.

**2.28. ÁLLÍTÁS (SORMODELL).** *Ha egy  $n$ -ismeretlenes egyenlet bal oldalán nem minden együttható 0, akkor az egyenletet kielégítő pontok (azaz az egyenlet megoldásai) egy hipersíkot alkotnak  $\mathbb{R}^n$ -ben. Ha egy  $n$ -ismeretlenes egyenletrendszer  $m$  ilyen egyenletből áll, akkor az egyenletrendszer megoldása a nekik megfelelő  $m$  hipersík közös része  $\mathbb{R}^n$ -ben.*

Az  $m$  egyenlet a skaláris szorzás segítségével tömörebb alakban is fölírható. Az  $m \times n$ -es  $\mathbf{A}$  együtthatómátrixú lineáris egyenletrendszer  $i$ -edik egyenletének alakja

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i.$$

Ha  $\mathbf{a}_{i*}$  jelöli az  $\mathbf{A}$  mátrix  $i$ -edik sorvektorát, és  $\mathbf{x}$  az ismeretlenek vektorát, akkor az előző egyenlet a következő alakot ölti:

$$\mathbf{a}_{i*} \cdot \mathbf{x} = b_i. \quad (2.24)$$

Ez különösen akkor lesz érdekes, ha homogén lineáris egyenletrendszereket fogunk vizsgálni, mert ott mindegyik egyenlet  $\mathbf{a}_{i*} \cdot \mathbf{x} = 0$  alakot ölt, ami azt jelenti, hogy olyan  $\mathbf{x}$  vektort keresünk, mely merőleges az  $\mathbf{a}_{i*}$  vektorok mindegyikére.

**Oszlopmodell: vektor előállítása lineáris kombinációként** E modellben az egyenletrendszerre úgy tekintünk, mint egy olyan vektoregyenletre, amelyben egy vektort kell előállítani adott vektorok lineáris kombinációjaként. Például a 2.27. példabeli

$$\begin{aligned} x + y &= 3 \\ x + 2y &= 4 \end{aligned}$$

egyenletrendszer ekvivalens az

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} x + \begin{bmatrix} 1 \\ 2 \end{bmatrix} y = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

Az oszlopmodell lépései jól nyomon követhetők a SagePlayer oszlopmodell című demonstrációján. Ott saját bővített mátrixokkal is lehet kísérletezni.

vektoregyenlettel. Itt az a feladat, hogy megkeressük az  $(1, 1)$  és  $(1, 2)$  vektoroknak azt a lineáris kombinációját, amely egyenlő a  $(3, 4)$  vektorral.

2.29. PÉLDA (OSZLOPMODELL). Ábrázoljuk a 2.27. példában megadott

$$\begin{array}{rcl} x + y = 3 & x + 2y = 3 & x + 2y = 3 \\ x + 2y = 4 & 2x + 4y = 7 & \text{és} & 2x + 4y = 6 \end{array}$$

egyenletrendszereket az oszlopmodellben!

**MEGOLDÁS.** Az első egyenletrendszer esetén két lineárisan független vektor lineáris kombinációjaként kell előállítani egy harmadik vektort. Ezt szemlélteti a 2.10 ábra. Érdekességként itt is megmutatjuk, hogy az egyenletrendszer megoldásának lépései hogy mutatnak e modellben. Az ekvivalens átalakítások lépései:

$$\begin{array}{rcl} x + y = 3 & \Rightarrow & x + y = 3 \Rightarrow x = 2 \\ x + 2y = 4 & \Rightarrow & y = 1 \Rightarrow y = 1 \end{array}$$

Vektoros alakban:

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} x + \begin{bmatrix} 1 \\ 2 \end{bmatrix} y = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \end{bmatrix} y = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} y = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

A második és harmadik egyenletrendszer vektoros alakja

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} x + \begin{bmatrix} 2 \\ 4 \end{bmatrix} y = \begin{bmatrix} 3 \\ 7 \end{bmatrix}, \quad \text{illetve} \quad \begin{bmatrix} 1 \\ 2 \end{bmatrix} x + \begin{bmatrix} 2 \\ 4 \end{bmatrix} y = \begin{bmatrix} 3 \\ 6 \end{bmatrix}.$$

A 2.11 ábráról szemléletesen is látható, hogy az egyik vektoregyenletnek nincs megoldása, míg a másiknak végtelen sok is van. □

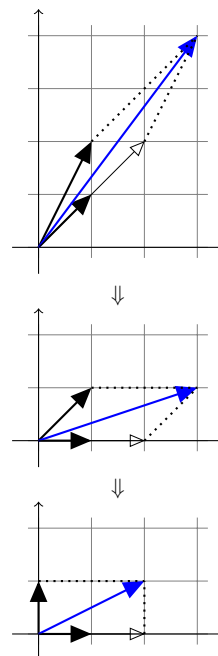
Általánosan kimondható a következő:

2.30. ÁLLÍTÁS (OSZLOPMODELL). A 2.21. definícióban megadott (2.19) egyenletrendszer a következő vektoregyenlettel ekvivalens:

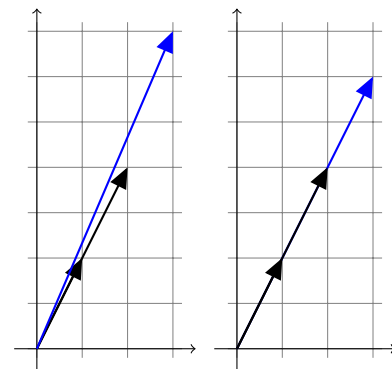
$$\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} x_1 + \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} x_2 + \dots + \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} x_n = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Az egyenletrendszer megoldása ekvivalens egy vektoregyenlet megoldásával, ahol az egyenletrendszer konstans tagjaiból álló vektort kell az együtthatómátrix oszlopvektorainak lineáris kombinációjaként előállítani.

E modell szerint egy egyenletrendszer pontosan akkor oldható meg, ha az együtthatómátrix oszlopvektorainak összes lineáris kombinációjából álló halmazban a konstans tagokból álló vektor is szerepel (ld. 2.13. feladat).



2.10. ábra: A megoldás lépései a oszlopmodellben.



2.11. ábra: Oszlopmodell lineárisan összefüggő vektorok esetén.

## Feladatok

### Lineáris egyenletek és egyenletrendszerek

2.4.\* Melyek lineáris egyenletek az  $x$ ,  $y$  és  $z$  változóiban az alábbiak közül?

- a)  $3x - (\ln 2)y + e^3z = 0.4$     b)  $a^2x - b^2y = 0$   
 c)  $xy - yz - zx = 0$     d)  $(\sin 1)x + y - \pi z = 0$   
 e)  $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$     f)  $\frac{1}{x} + \frac{1}{y} + \frac{1}{z} = 1$

Igazoljuk, hogy az alábbi egyenletrendszerek ekvivalensek!

$$2.5. \quad \begin{cases} x + 3y = 5 \\ y = 1 \end{cases} \quad \begin{cases} x + y = 3 \\ x = 2 \end{cases}$$

$$2.6. \quad \begin{cases} 2x + 3y = 2 \\ 0x + 0y = 3 \end{cases} \quad \begin{cases} x + y = 2 \\ x + y = 7 \end{cases}$$

Oldjuk meg (fejben számolva) az alábbi lineáris egyenletrendszereket az  $a = 1$ ,  $b = 2$ ,  $c = 3$  paraméterválasztás esetén!

$$2.7. \quad \begin{cases} (2a - b)x + (3a - c)y = 0 \\ (3b - 2c)x + (b - 2a)y = 0 \end{cases}$$

$$2.8. \quad \begin{cases} (b - a)x + (3a - c)y = 1 \\ (3b - 2c)x + (b - 2a)y = 0 \end{cases}$$

$$2.9. \quad \begin{cases} (b - a)x + (3a - c)y = 1 \\ (3b - 2c)x + (b - 2a)y = 1 \end{cases}$$

$$2.10. \quad \begin{cases} (b - a)x + (3a - c)y = 1 \\ (3b - 2c)x + (c - b)y = 2 \end{cases}$$

2.11. **EGYENLETRENDSZEREK KÖZÖS MEGOLDÁSA** Tekintsük az azonos ismeretleneket tartalmazó  $\mathcal{E}_1$  és  $\mathcal{E}_2$  egyenletrendszereket. Legyen ezek megoldáshalmaza  $\mathcal{M}_1$ , illetve  $\mathcal{M}_2$ . Mutassuk meg, hogy ha  $\mathcal{E}$  az  $\mathcal{E}_1$  és  $\mathcal{E}_2$  egyenletrendszerek egyesítése, azaz  $\mathcal{E} = \mathcal{E}_1 \cup \mathcal{E}_2$ , és  $\mathcal{M}$  az  $\mathcal{E}$  megoldáshalmaza, akkor  $\mathcal{M}$  az  $\mathcal{M}_1$  és  $\mathcal{M}_2$  közös része, azaz  $\mathcal{M} = \mathcal{M}_1 \cap \mathcal{M}_2$ .

Vizsgáljuk meg ezt az állítást az alábbi esetekben:

- a)  $\mathcal{E}_1 = \{x + y = 2\}$ ,  $\mathcal{E}_2 = \{x - y = 0\}$ ;  
 b)  $\mathcal{E}_1 = \{x + y = 2, x - y = 0\}$ ,  $\mathcal{E}_2 = \{x - y = 0\}$ ;  
 c)  $\mathcal{E}_1 = \{x + y = 2, x - y = 0\}$ ,  $\mathcal{E}_2 = \{x - y = 1\}$ ;  
 d)  $\mathcal{E}_1 = \{x + y = 2, x - y = 0\}$ ,  $\mathcal{E}_2 = \{0x + 0y = 0\}$ ;  
 e)  $\mathcal{E}_1$  tetszőleges egyenletrendszer,  $\mathcal{E}_2 = \{0 = 0\}$ .

### Sormodell, oszlopmodell

2.12.\* **SOR ÉS OSZLOPMODELL** Rajzoljuk fel a következő két egyenletrendszerhez tartozó sormodell és oszlopmodell szerinti ábrát!

$$a) \quad \begin{array}{l} 2x + 3y = 7 \\ 3x - 2y = 4 \end{array} \quad b) \quad \begin{array}{l} 2x + 4y = 3 \\ 3x + 6y = 4 \end{array}$$

2.13. **SOR- ÉS AZ OSZLOPMODELL 3D-BEN** Vizsgáljuk meg az alábbi két – azonos együtthatómátrixú – egyenletrendszer megoldhatóságát a sor- és az oszlopmodellben:

$$\begin{array}{ll} x + y + 2z = 3 & x + y + 2z = 3 \\ x + 2y + 4z = 3 & x + 2y + 4z = 3 \\ 3x + 4y + 8z = 9 & 3x + 4y + 8z = 1 \end{array}$$

2.14. **SOR ÉS OSZLOPMODELL  $m \neq n$  ESETÉN** Vizsgáljuk meg az alábbi három egyenletrendszer megoldhatóságát a sor- és az oszlopmodellben:

$$\begin{array}{lll} x + y = 3 & x + y = 3 & x + y = 3 \\ a) \quad x + y = 4 & b) \quad x + 2y = 4 & c) \quad x + 2y = 3 \\ x + 3y = 5 & x + 3y = 5 & x + 3y = 5 \end{array}$$

2.15.\* **IGAZ – HAMIS** Mely állítások igazak, melyek hamisak az alábbiak közül?

- a) Ha egy  $n$ -ismeretlenes egyenletrendszer olyan hipersíkok egyenleteiből áll, melyek közt van két párhuzamos, akkor az egyenletrendszer nem oldható meg.  
 b) Ha egy  $n$ -ismeretlenes egyenletrendszer nem oldható meg, akkor az egyenletek olyan hipersíkok egyenletei, melyek közt van két párhuzamos, de nem azonos hipersík.  
 c) Ha egy  $n$ -ismeretlenes egyenletrendszer csak két egyenletből áll, akkor az oszlopmodell szerint pontosan akkor oldható meg tetszőleges jobb oldal esetén, ha a vektor-egyenlet bal oldalán szereplő vektorok közt van kettő lineárisan független.

2.16.\* Egészítsük ki az alábbi állításokat úgy, hogy igazak legyenek!

- a) Egy két egyenletből álló háromismeretlenes egyenletrendszer sormodellje szerinti ábra  $a(z)$   $\dots$ -dimenziós térben  $\dots$  darab  $\dots\dots\dots$ ból/ből áll, melyek ha  $\dots\dots\dots$ , akkor az egyenletrendszernek nincs megoldása, egyébként megoldásainak száma  $\dots$ . Oszlopmodellje  $a(z)$   $\dots$ -dimenziós térben  $\dots$  darab  $\dots\dots\dots$ ból/ből áll.  
 b) Egy három egyenletből álló kétismeretlenes egyenletrendszer sormodellje szerinti ábra  $a(z)$   $\dots$ -dimenziós térben  $\dots$  darab  $\dots\dots\dots$ ból/ből áll, míg az oszlopmodellje  $a$   $\dots$ -dimenziós térben  $\dots$  darab  $\dots\dots\dots$ ból/ből áll.  
 c) Egy négy egyenletből álló ötismeretlenes egyenletrendszer sormodellje szerinti ábra  $a(z)$   $\dots$ -dimenziós térben  $\dots$  darab  $\dots\dots\dots$ ból/ből áll. Oszlopmodellje  $a(z)$   $\dots$ -dimenziós térben  $\dots$  darab  $\dots\dots\dots$ -ból/ből áll.



## Megoldás kiküszöböléssel

**Elemi sorműveletek** A lineáris egyenletrendszerek egyik megoldási módszerének lényege, hogy ekvivalens átalakításokkal olyan alakra hozzuk az egyenletrendszert, melyből visszahelyettesítések után, vagy azok nélkül azonnal leolvasható az eredmény. Az átalakításokat a bővített mátrixon hajtjuk végre úgy, hogy a nekik megfelelő átalakítások az egyenletrendszeren ekvivalens átalakítások legyenek. A 2.24. tételben felsorolt első három ekvivalens átalakítás nem változtatja meg az egyenletrendszer egyenleteinek számát sem. Az egyenletrendszer bővített mátrixán az ezeknek megfelelő átalakításokat elemi sorműveleteknek nevezzük.<sup>3</sup>

**2.31. DEFINÍCIÓ (ELEMI SORMŰVELETEK).** Egy mátrix sorain végzett alábbi műveleteket elemi sorműveleteknek nevezzük:

1. Sorcsere: két sor cseréje.
2. Beszorzás: egy sor beszorzása egy nemnulla számmal.
3. Hozzáadás: egy sorhoz egy másik sor konstansszorosának hozzáadása.

Természetesen egy sort el is oszthatunk egy nemnulla  $c$  számmal, hisz az az  $1/c$ -vel való beszorzással egyenértékű. Hasonlóképp levonhatjuk egy sorból egy másik sor  $c$ -szeresét, hisz az a  $-c$ -szeresének hozzáadásával ekvivalens.

Az elemi átalakításokra a következő jelöléseket fogjuk használni:

1.  $S_i \leftrightarrow S_j$ : az  $i$ -edik és a  $j$ -edik sorok cseréje.
2.  $cS_i$ : az  $i$ -edik sor beszorzása  $c$ -vel.
3.  $S_i + cS_j$ : a  $j$ -edik sor  $c$ -szeresének az  $i$ -edik sorhoz adása.

Az elemi sorműveletek mintájára elemi oszlopműveletek is definiálhatók, de azokat ritkán használjuk. Jelölésükre értelemszerűen az  $O_i \leftrightarrow O_j$ ,  $cO_i$ ,  $O_i + cO_j$  formulákat használjuk.

**Lépcsős alak** Az eddig megoldott egyenletrendszerekben igyekeztünk átlós, vagy legalább átló alatt kinullázott alakra hozni az egyenletrendszert, mint azt például a 2.26. példában tettük. Ez nem mindig sikerül, mert néha megtartani kívánt elemek is kinullázódnak, de a következőkben definiált lépcsős alakhoz mindig el tudunk jutni.

**2.32. DEFINÍCIÓ (LÉPCSŐS ALAK).** Egy mátrix lépcsős, vagy sorlépcsős alakú, ha kielégíti a következő két feltételt:

1. a csupa 0-ból álló sorok (ha egyáltalán vannak) a mátrix utolsó sorai;
2. bármely két egymás után következő nem-0 sorban az alsó sor elején (legalább egyfel) több 0 van, mint a fölötte lévő sor elején.

A nemnulla sorok első zérustól különböző elemét főelemnek, vezérelemnek vagy pivotelemnek hívjuk. Egy főelem oszlopának főoszlop vagy bázisoszlop a neve.

<sup>3</sup> Lineáris egyenletrendszerek felírása és megoldása már időszámításunk előtt 300 körül babiloni iratokban szerepelt. Az első századra teszik a kínai Jiùzhāng Suànshù (tradicionális jelekkel: 九章算術, egyszerűsített jelekkel: 九章算术) című mű megjelenését, mely az előző ezer évben összegyűlt matematikai tudást foglalja össze (címének magyar fordítása „A matematikai művészet kilenc fejezete” vagy „Kilenc fejezet a matematikai eljárásokról” lehet). E műben már a kiküszöbölés (azaz a Gauss-elimináció) néven ismert technikát alkalmazzák lineáris egyenletrendszer megoldására. A két fenti műben szereplő egyenletrendszerek, és további történeti részletek olvashatók a [The MacTutor History of Mathematics archive](#) című weboldalon.



2.33. PÉLDA (LÉPCSŐS ALAK). A következő mátrixok lépcsős alakúak:

$$\begin{bmatrix} 3 & 2 \\ 0 & 4 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & -2 & 3 & -4 \\ 0 & 0 & -5 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

**Gauss-módszer** A Gauss-módszer, más néven Gauss-kiküszöbölés vagy Gauss-elimináció a lineáris egyenletrendszerek megoldásának egy módszere. Lényege, hogy a lineáris egyenletrendszer bővített mátrixát elemi sorműveletekkel lépcsős alakra hozzuk, és abból visszahelyettesítéssel meghatározzuk a megoldás általános alakját. A Gauss-módszer könnyen algoritmizálható, ha sorban haladunk az oszlopokon. A módszerre láttunk már példát, ilyen volt a 2.26. példa első megoldása. Most lássunk két további példát.

2.34. PÉLDA (GAUSS-MÓDSZER, EGY MEGOLDÁS). Oldjuk meg az alábbi egyenletrendszert Gauss-módszerrel:

$$\begin{aligned} x + y + 2z &= 0 \\ 2x + 2y + 3z &= 2 \\ x + 3y + 3z &= 4 \\ x + 2y + z &= 5 \end{aligned}$$

**MEGOLDÁS.** Írjuk fel az egyenletrendszer bővített mátrixát, és oszloponként haladva küszöböljük ki – nullázzuk ki – a főelemek alatti elemeket!

$$\left[ \begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 2 & 2 & 3 & 2 \\ 1 & 3 & 3 & 4 \\ 1 & 2 & 1 & 5 \end{array} \right] \xrightarrow{\substack{S_2 - 2S_1 \\ S_3 - S_1 \\ S_4 - S_1}} \left[ \begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 0 & -1 & 2 \\ 0 & 2 & 1 & 4 \\ 0 & 1 & -1 & 5 \end{array} \right] \xrightarrow{S_2 \leftrightarrow S_3} \left[ \begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 2 & 1 & 4 \\ 0 & 0 & -1 & 2 \\ 0 & 1 & -1 & 5 \end{array} \right] \xrightarrow{S_4 - \frac{1}{2}S_2}$$

$$\left[ \begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 2 & 1 & 4 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & -\frac{3}{2} & 3 \end{array} \right] \xrightarrow{S_4 - \frac{3}{2}S_3} \left[ \begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 2 & 1 & 4 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right] \longrightarrow \begin{aligned} x + y + 2z &= 0 \\ 2y + z &= 4 \\ -z &= 2 \end{aligned}$$

A harmadik egyenletből  $z = -2$ , ezt a másodikba helyettesítve  $y = 3$ , ezeket az elsőbe helyettesítve kapjuk, hogy  $x = 1$ , azaz az egyetlen megoldás  $(x, y, z) = (1, 3, -2)$ .  $\square$

Mit csinálunk akkor, ha a lépcsős alak szerint kevesebb a főelemek, mint az oszlopok száma? Egyelőre bevezetünk két elnevezést, melyek jelentése hamarosan világos lesz: az egyenletrendszer azon változó-

it, melyek főelemek oszlopaihoz tartoznak, *kötött változóknak*, míg az összes többi változót *szabad változónak* nevezzük.

2.35. PÉLDA (GAUSS-MÓDSZER, VÉGTELEN SOK MEGOLDÁS). Oldjuk meg az alábbi egyenletrendszert Gauss-módszerrel:

$$\begin{aligned}x_1 + 2x_2 + x_3 + 2x_4 + x_5 &= 1 \\x_1 + 2x_2 + 3x_3 + 3x_4 + x_5 &= 0 \\3x_1 + 6x_2 + 7x_3 + 8x_4 + 3x_5 &= 1\end{aligned}$$

**MEGOLDÁS.** Írjuk fel az egyenletrendszer bővített mátrixát, és oszloponként haladva küszöböljük ki a főelemek alatti elemeket!

$$\begin{aligned}&\left[\begin{array}{ccccc|c}1 & 2 & 1 & 2 & 1 & 1 \\1 & 2 & 3 & 3 & 1 & 0 \\3 & 6 & 7 & 8 & 3 & 1\end{array}\right] \xrightarrow{\substack{s_2-s_1 \\ s_3-3s_1}} \left[\begin{array}{ccccc|c}1 & 2 & 1 & 2 & 1 & 1 \\0 & 0 & 2 & 1 & 0 & -1 \\0 & 0 & 4 & 2 & 0 & -2\end{array}\right] \xrightarrow{s_3-2s_2} \\&\left[\begin{array}{ccccc|c}1 & 2 & 1 & 2 & 1 & 1 \\0 & 0 & 2 & 1 & 0 & -1 \\0 & 0 & 0 & 0 & 0 & 0\end{array}\right] \longrightarrow \begin{aligned}x_1 + 2x_2 + x_3 + 2x_4 + x_5 &= 1 \\2x_3 + x_4 &= -1\end{aligned}$$

Az egyenletrendszer kötött változói a lépcsős alak főoszlopaihoz tartozó változók, azaz  $x_1$  és  $x_3$ . A szabad változók:  $x_2, x_4, x_5$ . A szabad változóknak tetszőleges értékeket adhatunk, a kötöttek értéke kifejezhető velük. Legyen például a szabad változók értéke  $x_2 = s, x_4 = t, x_5 = u$ . Ezek behelyettesítése után a fenti egyenletek közül először a másodikból kifejezzük  $x_3$ -at, majd azt behelyettesítjük az elsőbe, ahonnan kifejezzük az  $x_1$ -et, azaz a fenti egyenletekből kifejezzük a kötött változókat:

$$\begin{aligned}x_1 &= \frac{3}{2} - 2s - \frac{3}{2}t - u \\x_3 &= -\frac{1}{2} - \frac{1}{2}t\end{aligned}$$

Innen az egyenletrendszer megoldása:

$$(x_1, x_2, x_3, x_4, x_5) = \left(\frac{3}{2} - 2s - \frac{3}{2}t - u, s, -\frac{1}{2} - \frac{1}{2}t, t, u\right),$$

vagy mátrixjelöléssel

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} \frac{3}{2} - 2s - \frac{3}{2}t - u \\ s \\ -\frac{1}{2} - \frac{1}{2}t \\ t \\ u \end{bmatrix} = \begin{bmatrix} \frac{3}{2} \\ 0 \\ -\frac{1}{2} \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -\frac{3}{2} \\ 0 \\ -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + u \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Később különösen ez utóbbi felírásmód lesz hasznos, melyben vektorok lineáris kombinációja szerepel.  $\square$

Világos, hogy ha a szabad változóknak tetszőleges értéket adhatunk, melyből a kötött változók egyértelműen kifejezhetők, akkor a fenti példában mutatott módszerrel az egyenletrendszer összes megoldását leírtuk. Az ilyen módon megadott megoldást az egyenletrendszer *általános megoldásának*, a konkrét paraméterértékekhez tartozó megoldásokat *partikuláris megoldásoknak* nevezzük. Például az előző példabeli egyenletrendszer egy partikuláris megoldása az  $s = 0$ ,  $t = 1$ ,  $u = 2$  értékekhez tartozó

$$(x_1, x_2, x_3, x_4, x_5) = (-2, 0, -1, 1, 2).$$

Lényeges e megoldási módban, hogy a bővített mátrixot lépcsős alakra tudtuk hozni. Vajon ez mindig sikerül?

**2.36. TÉTEL (LÉPCSŐS ALAKRA HOZÁS).** *Bármely mátrix elemi sorműveletekkel lépcsős alakra hozható.*

**BIZONYÍTÁS.** Tekintsünk egy tetszőleges  $m \times n$ -es mátrixot. A következő eljárás egyes lépéseiben e mátrixnak le fogjuk takarni egy-egy sorát vagy oszlopát. Az egyszerűség kedvéért a letakarás után keletkezett mátrix sorainak és oszlopainak számát ismét  $m$  és  $n$  fogja jelölni,  $a_{ij}$  pedig a letakarások után maradt mátrix  $i$ -edik sorának  $j$ -edik elemét.

1. Ha az első oszlopban csak 0 elemek állnak, takarjuk le ezt az oszlopot, és tekintsük a maradék mátrixot. Ha ennek első oszlopában ismét csak 0 elemek vannak, azt is takarjuk le, és ezt addig folytassuk, míg egy olyan oszlopot nem találunk, amelyben van nem 0 elem. Ha ilyen oszlopot *nem találunk*, az eljárásnak vége, a mátrix lépcsős alakú.
2. Ha az első oszlop első sorában álló elem 0, akkor cseréljük ki e sort egy olyanal, melynek első eleme nem 0. Így olyan mátrixot kaptunk, amelyben  $a_{11} \neq 0$ .
3. Tekintsük az  $i$ -edik sort  $i = 2$ -től  $i = m$ -ig. Ha az  $i$ -edik sor első eleme  $a_{i1} \neq 0$ , akkor az első sor  $-a_{i1}/a_{11}$ -szeresét adjuk hozzá, azaz hajtsuk végre az  $S_i - \frac{a_{i1}}{a_{11}} S_1$  elemi átalakítást. Mivel  $a_{i1} - \frac{a_{i1}}{a_{11}} a_{11} = 0$ , ezért e lépés után az  $a_{11}$  alatti elemek mind 0-k lesznek.
4. A fenti átalakítás után takarjuk le az első sort és az első oszlopot. Ha ekkor *nem marad* a mátrixban több sor, vége az eljárásnak, a korábban letakart részeket feltárva megkaptuk a lépcsős alakot.

Egyébként ugorjunk vissza az 1. lépéshez, és folytassuk az eljárást. Világos, hogy ez az eljárás véges sok lépésben véget ér, melynek eredményeként eljutunk az eredeti mátrix egy lépcsős alakjához.  $\square$

Egy *inhomogén lineáris egyenletrendszerhez tartozó homogén lineáris egyenletrendszeren* azt a homogén egyenletrendszert értjük, melyet az inhomogénból a konstans tagok 0-ra változtatásával kapunk.

**2.37. PÉLDA (HOMOGÉN LINEÁRIS EGYENLETRENDSZER MEGOLDÁSA).**

Oldjuk meg a 2.35. példabeli egyenletrendszerhez tartozó

$$x_1 + 2x_2 + x_3 + 2x_4 + x_5 = 0$$

$$x_1 + 2x_2 + 3x_3 + 3x_4 + x_5 = 0$$

$$3x_1 + 6x_2 + 7x_3 + 8x_4 + 3x_5 = 0$$

homogén lineáris egyenletrendszert.

**MEGOLDÁS.** Mivel homogén lineáris egyenletrendszerről van szó, a megoldáshoz szükségtelen a bővített mátrixot használni, hisz annak utolsó oszlopa csak nullákból áll, így az elemi sorműveletek közben nem változik. Az együtthatómátrix lépcsős alakja ugyanazokkal a sorműveletekkel megkapható, mint a 2.35. példa megoldásában, azaz

$$\begin{bmatrix} 1 & 2 & 1 & 2 & 1 \\ 1 & 2 & 3 & 3 & 1 \\ 3 & 6 & 7 & 8 & 3 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 1 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \longrightarrow$$

$$\begin{aligned} x_1 + 2x_2 + x_3 + 2x_4 + x_5 &= 0 \\ 2x_3 + x_4 &= 0 \end{aligned}$$

Innen a megoldás is ugyanúgy kapható meg, sőt, ugyanaz a lineáris kombináció szerepel benne a konstans tagok nélkül:

$$(x_1, x_2, x_3, x_4, x_5) = \left(-2s - \frac{3}{2}t - u, s, -\frac{1}{2}t, t, u\right),$$

vagy mátrixjelöléssel

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2s - \frac{3}{2}t - u \\ s \\ -\frac{1}{2}t \\ t \\ u \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -\frac{3}{2} \\ 0 \\ -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + u \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

A homogén és inhomogén egyenletrendszerek e példából sejtethető kapcsolatára még visszatérünk a 3.17. tételben.  $\square$

Végül egy alkalmazás:

**2.38. PÉLDA (SÍKOK METSZÉSVONALÁNAK MEGHATÁROZÁSA).**

Határozzuk meg az alábbi két sík metszésvonalának explicit (paraméteres) alakját!

$$x + y + z = 1$$

$$3x + 4y = 2$$

**MEGOLDÁS.** A fenti egyenletekkel megadott két sík metszésvonalának meghatározásához, pontosabban a metszésvonal explicit, paraméteres egyenletrendszerének felírásához egyszerűen meg kell oldani a két

egyenletből álló egyenletrendszert:

$$\left[ \begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 3 & 4 & 0 & 2 \end{array} \right] \xrightarrow{s_2 - 3s_1} \left[ \begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -3 & -1 \end{array} \right] \longrightarrow \begin{array}{l} x + y + z = 1 \\ y - 3z = -1 \end{array}$$

Ebből  $z = t$  paraméterválasztással  $y = -1 + 3t$  és  $x = 2 - 4t$ , azaz

$$(x, y, z) = (-4t + 2, 3t - 1, t) = (2, -1, 0) + t(-4, 3, 1),$$

vagy mátrixjelöléssel

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -4 \\ 3 \\ 1 \end{bmatrix}. \quad \square$$

**Redukált lépcsős alak** A 2.26. példa második megoldási módszerében átlós alakra hoztuk az együtthatómátrixot, azaz nem elégedtünk meg azzal, hogy a főelemek alatt kinulláztunk minden együtthatót, hanem a főelemeket 1-re változtattuk a sor beszorzásával, és a főelemek fölött is kinulláztunk (elimináltunk) minden elemet.

2.39. DEFINÍCIÓ (REDUKÁLT LÉPCSŐS ALAK). Egy mátrix redukált lépcsős, vagy redukált sorlépcsős alakú, ha kielégíti a következő feltételeket:

1. lépcsős alakú;
2. minden főelem egyenlő 1-gyel;
3. a főelemek oszlopaiban a főelemeken kívül minden elem 0;

A főelemet itt vezéregyesnek vagy vezető egyesnek is szokás nevezni.

2.40. PÉLDA (REDUKÁLT LÉPCSŐS ALAK). A következő mátrixok redukált lépcsős alakúak:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & -2 & 0 & -4 \\ 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Minden valós, vagy racionális elemű mátrix redukált lépcsős alakra hozható, azonban az egészegyütthatós mátrixok általában nem, ha az egészekben belül akarunk maradni. Azonban az egészegyütthatós mátrixok is redukált lépcsős alakra hozhatók a racionálisok számkörében. A kiküszöbölés lépéseinek követésére használható a SagePlayer redukált lépcsős alak című szemléltető munkalapja.

2.41. PÉLDA (REDUKÁLT LÉPCSŐS ALAKRA HOZÁS). Hozzuk redukált lépcsős alakra az

$$\begin{bmatrix} 1 & 3 & 0 \\ 1 & 1 & 2 \\ 2 & 2 & 4 \end{bmatrix}$$

mátrixot!

MEGOLDÁS. Egy lehetséges megoldás:

$$\begin{aligned} & \begin{bmatrix} 1 & 3 & 0 \\ 1 & 1 & 2 \\ 2 & 2 & 4 \end{bmatrix} \xrightarrow{\substack{S_2 - S_1 \\ S_3 - 2S_1}} \begin{bmatrix} 1 & 3 & 0 \\ 0 & -2 & 2 \\ 0 & -4 & 4 \end{bmatrix} \xrightarrow{-\frac{1}{2}S_2} \\ & \begin{bmatrix} 1 & 3 & 0 \\ 0 & 1 & -1 \\ 0 & -4 & 4 \end{bmatrix} \xrightarrow{S_3 + 4S_2} \begin{bmatrix} 1 & 3 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{S_1 - 3S_2} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

Egy másik lehetséges megoldás:

$$\begin{aligned} & \begin{bmatrix} 1 & 3 & 0 \\ 1 & 1 & 2 \\ 2 & 2 & 4 \end{bmatrix} \xrightarrow{S_1 \leftrightarrow S_2} \begin{bmatrix} 1 & 1 & 2 \\ 1 & 3 & 0 \\ 2 & 2 & 4 \end{bmatrix} \xrightarrow{\substack{S_2 - S_1 \\ S_3 - 2S_1}} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & -2 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{\frac{1}{2}S_2} \\ & \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{S_1 - S_2} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

Bár különböző úton, de mindkétszer azonos eredményre jutottunk. Valóban, hamarosan be fogjuk látni, hogy a redukált lépcsős alak egyértelmű, míg e példából is látjuk, hogy a lépcsős alak nem: a megadott mátrixnak a megoldás során négy különböző lépcsős alakját is előállítottuk (az első megoldás utolsó két mátrixa és a második megoldás utolsó három mátrixa lépcsős alakú).  $\square$

**Gauss–Jordan-módszer** A Gauss–Jordan-módszer (Gauss–Jordan-kiküszöbölés, Gauss–Jordan-elimináció) a lineáris egyenletrendszerek egy megoldási módszere. Lényege, hogy a lineáris egyenletrendszer bővített mátrixát elemi sorműveletekkel redukált lépcsős alakra hozzuk. Ebből az alakból azonnal leolvasható a megoldás. Adjunk új megoldást a Gauss-módszernél bemutatott egyenletrendszerekre.

2.42. PÉLDA (GAUSS–JORDAN-MÓDSZER, EGY MEGOLDÁS). Oldjuk meg a 2.34. példában felírt egyenletrendszert Gauss–Jordan-módszerrel!

MEGOLDÁS. Felírjuk az egyenletrendszer bővített mátrixát, és a 2.34. példában látott módon eljutunk a lépcsős alakhoz, majd folytatjuk, először beszorozzuk a sorokat a főátlóbeli elem reciprokával, majd a har-

madik oszlopot, végül a másodikat kinullázzuk:

$$\begin{aligned} \left[ \begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 2 & 2 & 3 & 2 \\ 1 & 3 & 3 & 4 \\ 1 & 2 & 1 & 5 \end{array} \right] & \xrightarrow{\dots} \left[ \begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 2 & 1 & 4 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right] \xrightarrow{\substack{1/2 S_2 \\ -S_3}} \left[ \begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 1 & \frac{1}{2} & 2 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{array} \right] \xrightarrow{\substack{S_2 - \frac{1}{2} S_3 \\ S_1 - 2S_3}} \\ \left[ \begin{array}{ccc|c} 1 & 1 & 0 & 4 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{array} \right] \xrightarrow{S_1 - S_2} \left[ \begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{array} \right] \longrightarrow \begin{array}{l} x = 1 \\ y = 3 \\ z = -2 \end{array} \end{aligned}$$

Tehát az egyenletrendszer egyetlen megoldása  $(x, y, z) = (1, 3, -2)$ .  $\square$

#### 2.43. PÉLDA (GAUSS–JORDAN-MÓDSZER, VÉGTELEN SOK MEGOLDÁS).

Oldjuk meg a 2.35. példabeli

$$\begin{aligned} x_1 + 2x_2 + x_3 + 2x_4 + x_5 &= 1 \\ x_1 + 2x_2 + 3x_3 + 3x_4 + x_5 &= 0 \\ 3x_1 + 6x_2 + 7x_3 + 8x_4 + 3x_5 &= 1 \end{aligned}$$

egyenletrendszert Gauss–Jordan-módszerrel!

**MEGOLDÁS.** A 2.35. példában eljutottunk egy lépcsős alakig. Az eljárást folytatjuk, míg a redukált lépcsős alakra nem jutunk.

$$\begin{aligned} \left[ \begin{array}{ccccc|c} 1 & 2 & 1 & 2 & 1 & 1 \\ 1 & 2 & 3 & 3 & 1 & 0 \\ 3 & 6 & 7 & 8 & 3 & 1 \end{array} \right] & \xrightarrow{\dots} \left[ \begin{array}{ccccc|c} 1 & 2 & 1 & 2 & 1 & 1 \\ 0 & 0 & 2 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right] \xrightarrow{\substack{1/2 S_2 \\ S_1 - S_2}} \\ \left[ \begin{array}{ccccc|c} 1 & 2 & 0 & 3/2 & 1 & 3/2 \\ 0 & 0 & 1 & 1/2 & 0 & -1/2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right] & \longrightarrow \begin{array}{l} x_1 + 2x_2 + \frac{3}{2}x_4 + x_5 = \frac{3}{2} \\ x_3 + \frac{1}{2}x_4 = -\frac{1}{2} \end{array} \end{aligned}$$

Végezzük el az  $x_2 = s$ ,  $x_4 = t$ ,  $x_5 = u$  helyettesítést, az  $x_1$  és  $x_3$  változók azonnal kifejezhetők. Így a megoldás:

$$(x_1, x_2, x_3, x_4, x_5) = \left( \frac{3}{2} - 2s - \frac{3}{2}t - u, s, -\frac{1}{2} - \frac{1}{2}t, t, u \right),$$

vagy mátrixjelöléssel

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} \frac{3}{2} - 2s - \frac{3}{2}t - u \\ s \\ -\frac{1}{2} - \frac{1}{2}t \\ t \\ u \end{bmatrix} = \begin{bmatrix} \frac{3}{2} \\ 0 \\ -\frac{1}{2} \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -\frac{3}{2} \\ 0 \\ -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + u \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

$\square$

*A redukált lépcsős alak egyértelműsége* Fontos következményei vannak a következő tételnek:

2.44. TÉTEL (A REDUKÁLT LÉPCSŐS ALAK EGYÉRTELMŰ). Minden mátrix redukált lépcsős alakra hozható, amely egyértelmű.

**BIZONYÍTÁS.** A redukált lépcsős alak létezését már beláttuk, az egyértelműsége indirekt bizonyítást adunk. Tegyük fel, hogy van egy olyan mátrix, mely elemi sorműveletekkel két különböző redukált lépcsős alakra hozható. Jelölje ezeket  $\mathbf{R}$  és  $\mathbf{S}$ . Mivel mindketten ugyanazzal a mátrixszal ekvivalensek, elemi sorműveletekkel egymásba alakíthatók, vagyis egymással is ekvivalensek. Válasszuk ki oszlopaik közül azt a balról első oszlopot, melyben különböznek, valamint az összes előttük álló vezéroszlopot. Az így kapott mátrixokat jelölje  $\hat{\mathbf{R}}$  és  $\hat{\mathbf{S}}$ . Tehát  $\hat{\mathbf{R}} \neq \hat{\mathbf{S}}$ , mert különböznek az utolsó oszlopukban. Például, ha

$$\mathbf{R} = \begin{bmatrix} 1 & 2 & 0 & 4 & 5 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{és} \quad \mathbf{S} = \begin{bmatrix} 1 & 2 & 0 & 4 & 5 \\ 0 & 0 & 1 & 9 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

akkor

$$\hat{\mathbf{R}} = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{és} \quad \hat{\mathbf{S}} = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 9 \\ 0 & 0 & 0 \end{bmatrix}.$$

Ez az oszlop, melyben különböznek, nem lehet az első oszlop, mert ha az a zérusvektor az egyik mátrixban, akkor a sorkvivalencia miatt a másikban is az lenne, egyébként pedig ez az oszlop mindenképp az első helyen 1-est, alatta 0-kat tartalmaz.

Tekintsük az így kapott  $\hat{\mathbf{R}}$ ,  $\hat{\mathbf{S}}$  mátrixokat egy-egy egyenletrendszer bővített együtthatómátrixának. Ezek általános alakja tehát a következő:

$$\hat{\mathbf{R}} = \left[ \begin{array}{cccc|c} 1 & 0 & \dots & 0 & r_1 \\ 0 & 1 & \dots & 0 & r_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & r_k \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \end{array} \right] \quad \text{vagy} \quad \hat{\mathbf{R}} = \left[ \begin{array}{cccc|c} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \end{array} \right]$$



és

$$\hat{\mathbf{S}} = \left[ \begin{array}{cccc|c} 1 & 0 & \dots & 0 & s_1 \\ 0 & 1 & \dots & 0 & s_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & s_k \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \end{array} \right] \quad \text{vagy} \quad \hat{\mathbf{S}} = \left[ \begin{array}{cccc|c} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \end{array} \right]$$

Mivel oszlopok kihagyása nem változtat a sorkvivalencián – hisz elemi sorműveletekben műveletet csak egy oszlopon belül végzünk –, ezért az  $\hat{\mathbf{R}}$  és  $\hat{\mathbf{S}}$  mátrixok ekvivalensek, azaz a hozzájuk tartozó két egyenletrendszernek ugyanaz a megoldása. Ez csak úgy lehet, ha vagy minden  $i = 1, \dots, k$  indexre  $r_i = s_i$ , vagy egyik egyenletrendszer sem oldható meg, azaz mindkét esetben azt kaptuk, hogy  $\hat{\mathbf{R}} = \hat{\mathbf{S}}$ , ami ellentmondás. Ez az ellentmondás bizonyítja, hogy a kiinduló  $\mathbf{R} \neq \mathbf{S}$  feltevés helytelen volt, tehát  $\mathbf{R} = \mathbf{S}$ . (E bizonyítás Holzmann<sup>4</sup> interneten publikált cikkén alapul).  $\square$

<sup>4</sup> Wolf Holzmann. Uniqueness of reduced row echelon form. <http://www.cs.uleth.ca/~holzmann/notes/reduceduniq.pdf>, 2002

Mivel a redukált lépcsős alak egyértelmű, definiálhatunk egy függvényt, mely minden mátrixhoz annak ezt az alakját rendeli. Az  $\text{rref}(\mathbf{A})$  jelölést mi arra a függvényre fogjuk alkalmazni, mely egy  $m \times n$ -es mátrixhoz a redukált lépcsős alakjának a zérussorok elhagyásával kapott alakját rendeli. Például

$$\text{rref} \left[ \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{array} \right] = \left[ \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \end{array} \right].$$

**Szimultán egyenletrendszerek** Gyakori feladat az alkalmazásokban, hogy sok olyan egyenletrendszert kell megoldani, amelyek csak a konstans tagokban térnek el egymástól. A kiküszöböléses módszerekkel ezek egyszerre is megoldhatók alig több erőforrás felhasználásával, mint ami egyetlen egyenletrendszer megoldásához szükséges.

**2.45. DEFINÍCIÓ (SZIMULTÁN EGYENLETRENDSZEREK).** Több egyenletrendszer halmazát szimultán egyenletrendszernek nevezzük, ha együtthatómátrixaik azonosak.

2.46. PÉLDA (SZIMULTÁN EGYENLETRENDSZER MEGOLDÁSA). Oldjuk meg az alábbi egyenletrendszereket!

$$\begin{array}{lll} x + y + z = 3 & u + v + w = 3 & r + s + t = 0 \\ 2x + 3y + 2z = 7 & 2u + 3v + 2w = 7 & 2r + 3s + 2t = 0 \\ 2x + 2y + 3z = 6 & 2u + 2v + 3w = 7 & 2r + 2s + 3t = 1 \end{array}$$

**MEGOLDÁS.** Mivel e három egyenletrendszer együtthatómátrixa azonos, a bal oldal átalakítását elég egyszer elvégezni, a jobb oldalak átalakítását pedig vele együtt. Ehhez a szimultán egyenletrendszerre a következő bővített mátrixot érdemes képezni:

$$\left[ \begin{array}{ccc|ccc} 1 & 1 & 1 & 3 & 3 & 0 \\ 2 & 3 & 2 & 7 & 7 & 0 \\ 2 & 2 & 3 & 6 & 7 & 1 \end{array} \right]$$

A megoldáshoz használjuk a Gauss–Jordan-módszert:

$$\begin{array}{l} \left[ \begin{array}{ccc|ccc} 1 & 1 & 1 & 3 & 3 & 0 \\ 2 & 3 & 2 & 7 & 7 & 0 \\ 2 & 2 & 3 & 6 & 7 & 1 \end{array} \right] \xrightarrow[\begin{smallmatrix} S_3-2S_1 \\ S_2-2S_1 \end{smallmatrix}]{S_2-2S_1} \left[ \begin{array}{ccc|ccc} 1 & 1 & 1 & 3 & 3 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{array} \right] \xrightarrow[\begin{smallmatrix} S_1-S_2 \\ S_1-S_3 \end{smallmatrix}]{S_1-S_2} \\ \left[ \begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & 1 & -1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{array} \right]. \end{array}$$

Ebből leolvasható mindhárom egyenletrendszer megoldása:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} r \\ s \\ t \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}. \quad \square$$

Ha tudjuk, hogy több egyenletrendszerből álló szimultán egyenletrendszerről van szó, mindegyik egyenletrendszerben használhatjuk ugyanazokat a változókat.

2.47. PÉLDA (SZIMULTÁN EGYENLETRENDSZER BŐVÍTETT MÁTRIXA). Oldjuk meg azt a szimultán egyenletrendszert, melynek bővített mátrixa

$$\left[ \begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 5 & 3 & 0 & 1 \end{array} \right]$$

**MEGOLDÁS.** A Gauss–Jordan-módszer lépései:

$$\begin{aligned} \left[ \begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 5 & 3 & 0 & 1 \end{array} \right] &\xrightarrow{s_2 - \frac{5}{2}s_1} \left[ \begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 0 & \frac{1}{2} & -\frac{5}{2} & 1 \end{array} \right] \xrightarrow{2s_2} \left[ \begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 0 & 1 & -5 & 2 \end{array} \right] \xrightarrow{s_1 - s_2} \\ &\left[ \begin{array}{cc|cc} 2 & 0 & 6 & -2 \\ 0 & 1 & -5 & 2 \end{array} \right] \xrightarrow{\frac{1}{2}s_1} \left[ \begin{array}{cc|cc} 1 & 0 & 3 & -1 \\ 0 & 1 & -5 & 2 \end{array} \right]. \quad \square \end{aligned}$$

**Kiküszöbölés  $\mathbb{Z}_p$ -ben\*** Ha  $p$  prím, akkor a modulo  $p$  maradékosztályok közti műveletek rendelkeznek minden olyan tulajdonsággal, melyet a kiküszöbölés során a valós számok körében használtunk. Ennek következtében a Gauss- és Gauss–Jordan-módszerek minden további nélkül használhatók  $\mathbb{Z}_p$  fölötti egyenletrendszerekre is. (Lásd még a ?? oldalon az algebrai tesztől írtakat.)

**2.48. PÉLDA (EGYENLETRENDSZER  $\mathbb{Z}_2$  FÖLÖTT).** 4-bites kódszavakat küldünk, bitjeit jelölje  $a, b, c$  és  $d$ . Hibajavító kódot készítünk úgy, hogy minden kódszó végére három paritásbitet teszünk, nevezetesen  $a + b + c + d$ ,  $a + c + d$  és  $a + b + d$  bitet. Az összeadás itt természetesen  $\mathbb{Z}_2$  fölött értendő. Például a 0110 kódszó helyett a 0110011 kódszót küldjük. Egy üzenetben az egyik ilyen 7-bites kódszó első 4 bitjét a vevő szerkezet bizonytalanul érzékeli, amit kapunk, az  $(?, ?, ?, ?, 1, 0, 1)$  kódvektor. Mi lehetett az eredeti üzenet, ha az utolsó 3 bit biztosan jó?

**MEGOLDÁS.** Az  $a, b, c$  és  $d$  bitek ismeretlenek, csak annyit tudunk, hogy

$$\begin{aligned} b + c + d &= 1 \\ a + \quad c + d &= 0 \\ a + b + \quad d &= 1 \end{aligned}$$

Oldjuk meg ezt az egyenletrendszert Gauss–Jordan kiküszöböléssel  $\mathbb{Z}_2$  fölött. Ne felejtsük, hogy  $\mathbb{Z}_2$ -ben  $1 + 1 = 0$ , így  $1 = -1$ , azaz a kivonás nem különbözik az összeadástól.

$$\begin{aligned} \left[ \begin{array}{cccc|c} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \end{array} \right] &\xrightarrow{s_1 \leftrightarrow s_2} \left[ \begin{array}{cccc|c} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \end{array} \right] \xrightarrow{s_3 + s_1} \left[ \begin{array}{cccc|c} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{array} \right] \xrightarrow{s_3 + s_2} \\ \left[ \begin{array}{cccc|c} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{array} \right] &\xrightarrow{\substack{s_2 + s_3 \\ s_1 + s_3}} \left[ \begin{array}{cccc|c} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{array} \right] \longrightarrow \begin{aligned} a + c &= 0 \\ b + c &= 1 \\ d &= 0 \end{aligned} \end{aligned}$$

Az utolsó egyenletből  $d = 0$ . A szabad változó  $c$ , legyen  $c = s$ . Így a második egyenletből  $b = 1 + c$ , azaz  $b = 1 + s$  és az elsőből  $a = c$ , azaz  $a = s$ . A megoldás általános alakban  $(a, b, c, d) = (s, 1 + s, s, 0)$ , azaz

$(a, b, c, d) = (0, 1, 0, 0) + s(1, 1, 1, 0)$ . Az  $s = 0$  és az  $s = 1$  értékekhez tartozó megoldások tehát:  $(0, 1, 0, 0)$  és  $(1, 0, 1, 0)$ .

Ha az egyenletrendszert vektoregyenletnek tekintjük, akkor az első megoldás azt mutatja, hogy az együtthatómátrix második oszlopa megegyezik a jobb oldallal (és valóban), a második megoldás pedig azt, hogy az első és a harmadik oszlop összege a jobb oldalt adja.

Megjegyezzük e kódról, melyet  $[7, 4, 3]_2$  bináris *Hamming-kódnak* neveznek, hogy a kód 16 szóból áll, bármely szavának bármely 4 bite egyértelműen meghatározza a maradék hármat. Így ha legföljebb 3 bit megváltozik egy szóban, akkor az kimutatható, és ha csak egy bit változik meg, az kijavítható.  $\square$

**2.49. PÉLDA** (EGYENLETRENDSZER  $\mathbb{Z}_5$  FÖLÖTT). Oldjuk meg az alábbi két egyenletrendszert  $\mathbb{Z}_5$  fölött.

$$\begin{array}{rcl} 2x + 3y & = & 1 \\ 3x + 2y & = & 4 \end{array} \quad \begin{array}{rcl} 2x + 3y & = & 1 \\ 3x + 4y & = & 3 \end{array}$$

**MEGOLDÁS.** A számolás megkönnyítésére vagy készítsünk osztási táblát, vagy használjuk a ?? oldalon található ?? ábra szorzástábláját.

$$\left[ \begin{array}{cc|c} 2 & 3 & 1 \\ 3 & 2 & 4 \end{array} \right] \xrightarrow{3S_1} \left[ \begin{array}{cc|c} 1 & 4 & 3 \\ 3 & 2 & 4 \end{array} \right] \xrightarrow{S_2 - 3S_1} \left[ \begin{array}{cc|c} 1 & 4 & 3 \\ 0 & 0 & 0 \end{array} \right],$$

azaz az egyenletrendszernek több megoldása van. Itt ez nem azt jelenti, hogy végtelen sok, hanem azt, hogy legalább egy paraméter végigfut  $\mathbb{Z}_5$  összes elemén. Szabad változó az  $y$ , legyen  $y = s$ , így  $x = 3 - 4s = 3 + s$ , tehát  $(x, y) = (3 + s, s)$ , azaz a vektorok mátrixjelölésével:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix} + s \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad s \in \mathbb{Z}_5$$

Mivel  $\mathbb{Z}_5$ -nek öt eleme van, ezért  $s$ -nek is ennyi értéke lehet, azaz az első egyenletrendszer összes megoldása  $(3, 0)$ ,  $(4, 1)$ ,  $(0, 2)$ ,  $(1, 3)$ ,  $(2, 4)$ . A másik egyenletrendszer megoldása:

$$\left[ \begin{array}{cc|c} 2 & 3 & 1 \\ 3 & 4 & 3 \end{array} \right] \xrightarrow{3S_1} \left[ \begin{array}{cc|c} 1 & 4 & 3 \\ 3 & 4 & 3 \end{array} \right] \xrightarrow{S_2 - 3S_1} \left[ \begin{array}{cc|c} 1 & 4 & 3 \\ 0 & 2 & 4 \end{array} \right] \xrightarrow{3S_2} \left[ \begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 2 \end{array} \right].$$

Így a megoldás  $(x, y) = (0, 2)$ .  $\square$

## Feladatok

### Mátrix lépcsős és redukált lépcsős alakja

**2.17.\* LÉPCSŐS ALAK: IGAZ – HAMIS** Melyek igazak, melyek hamisak az alábbi állítások közül?

- Egy mátrix minden lépcsős alakjában ugyanannyi nem-zérus sor van.
- Egy mátrix minden lépcsős alakjában ugyanannyi fő-oszlop (bázisoszlop) van.
- Minden valós mátrixnak van lépcsős alakja, ami egyértelmű.
- Különböző mátrixoknak különböző a redukált lépcsős alakjuk.
- Ha egy mátrix elemi sorműveletekkel egy másikba vihető, akkor redukált lépcsős alakjuk megegyezik.

Határozzuk meg valamely lépcsős alakját, majd a redukált lépcsős alakját az alábbi mátrixoknak!

$$2.18. \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 3 & 2 & 3 & 4 \\ 1 & 2 & 1 & 2 & 3 \end{bmatrix}$$

$$2.19. \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 3 & 2 & 3 & 4 \\ 3 & 2 & 1 & 2 & 3 \end{bmatrix}$$

**2.20.\*** A három elemi sorművelet egyike elvégezhető a másik kettő segítségével is. Melyik és hogyan?

### Egyenletrendszer megoldása Gauss-módszerrel

Oldjuk meg az alábbi lineáris egyenletrendszereket Gauss-módszerrel!

$$2.21. \begin{cases} x_1 + x_2 + x_3 = 1 \\ x_2 + x_3 + x_4 = 2 \\ x_3 + x_4 + x_5 = 2 \\ x_4 + x_5 + x_1 = 1 \end{cases}$$

$$2.22. \begin{cases} x_1 + x_2 + x_3 = 4 \\ -x_1 + x_2 - x_3 = 2 \\ 2x_1 + x_2 + 2x_3 = 1 \\ 4x_1 + 4x_2 + 4x_3 = 1 \end{cases}$$

$$2.23.* \begin{cases} 7x + 14y - 21z = 7 \\ x + 2y - 3z = 1 \\ 5x + 10y + 15z = 5 \\ 3x + 6y - 9z = 3 \end{cases}$$

$$2.24. \begin{cases} x + 3y + z = 1 \\ 2x + 7y + 2z = 0 \\ x + 4y + 4z = 1 \\ x + 4y + z = -1 \\ 4x + 15y + 10z = 2 \end{cases}$$

$$2.25. \begin{cases} x + y = 4 \\ 3x - y = 2 \\ -3x + 5y = 2 \\ x + 2y = 1 \end{cases}$$

$$2.26. \begin{cases} x_1 + x_2 + 4x_4 = 3 \\ x_2 - x_3 + 3x_4 = 1 \\ x_1 - 2x_2 + 3x_3 - 5x_4 = 0 \\ 3x_1 - x_2 + 4x_3 = 5 \end{cases}$$

### Egyenletrendszer megoldása Gauss–Jordan-módszerrel

#### Egyenletrendszerek

**2.27.\* EGYENLETRENDSZEREK: IGAZ – HAMIS** Melyek igazak, melyek hamisak az alábbi állítások közül?

- A bővített mátrixon végrehajtott elemi sorműveletek közben az egyenletrendszer megoldáshalmaza nem változik.
- Egy lineáris egyenletrendszer nem konzisztens, ha több egyenletből áll, mint ahány ismeretlenes.
- Ha egy valóegyűtthetős lineáris egyenletrendszernek van két különböző megoldása, akkor végtelen sok is van.
- Egy homogén lineáris egyenletrendszer mindig konzisztens.

Ekvivalensek-e az alábbi egyenletrendszerek?

$$2.28. \begin{cases} 3x + 2y - 2z = 1 \\ 2x + 3y - 3z = -1 \\ 4x + 2y = 8 \end{cases} \quad \begin{cases} 2x + 2y - 2z = 0 \\ 3x + 3y - 2z = 3 \\ 5x - 3y + 2z = 5 \end{cases}$$

$$2.29. \begin{cases} 2x + 3y + 5z = 0 \\ 3x + 2y + 2z = 3 \\ 5x = -4z = 9 \end{cases} \quad \begin{cases} x - y - 3z = 3 \\ 5x + 5y + 7z = 3 \end{cases}$$

**2.30.** Csak egész számokkal számolva megoldható-e az az egyenletrendszer, melynek bővített mátrixa a következő:

$$a) \begin{bmatrix} 3 & 4 & 1 & 1 \\ 7 & 8 & 3 & 7 \\ 11 & 7 & -2 & 2 \end{bmatrix} \quad b) \begin{bmatrix} 3 & 4 & 1 & 1 \\ 7 & 8 & 3 & 7 \\ 11 & 7 & 2 & 2 \end{bmatrix}$$

**2.31.** Egy legalább 2-ismeretlenes lineáris egyenletrendszerről annyit tudunk, hogy egyértelműen megoldható, és hogy bővített mátrixának elemei sorfolytonosan olvasva számtani sorozatot adnak. Mi a megoldása?

**2.32. LINEÁRISRA VISSZAVEZETHETŐ EGYENLETRENDSZEREK**

Oldjuk meg az alábbi, nem lineáris egyenletrendszereket!

a)  $2\sqrt{x} + 2\sqrt{y} = 8$

$3\sqrt{x} + \sqrt{y} = 4$

b)  $2x^3 + 2y^2 = 8$

$3x^3x + y^2 = 4$

c)  $2e^x + 2e^y = 8$

$3e^x + e^y = 4$

d)  $2\cos x + 2\cos y = 8$

$3\cos x + \cos y = 4$

**2.33. EGYENLETRENDSZER POZITÍV EGÉSZ MEGOLDÁSOKKAL**

Egy érmegyűjteményben régi 1, 5 és 10Ft-osok vannak, összesen 11 darab, 53Ft összértékben. Melyik érméből hány darab van?

## Megoldás a gyakorlatban

Bár e szakasz tartalma elsősorban nem a lineáris algebra, hanem a numerikus analízis témakörébe tartozik, ismerete elengedhetetlen annak, aki a gyakorlatban lineáris algebrai eszközöket alkalmaz. Először a Gauss- és Gauss–Jordan-kiküszöbölés műveletigényét, majd numerikus megbízhatóságának kérdését vizsgáljuk. Ezután az iterációs módszerek lényegét vázoljuk, melyek alkalmazásakor az együtthatómátrix nem változik, így a számítási hibák sem halmozódnak. Ráadásul e módszerek a ritka mátrixokat sem „rontják el”, mint a Gauss-módszer, mely sok zérust ír fölül.

**A kiküszöbölés műveletigénye** Ahhoz, hogy a lineáris egyenletrendszerek különböző megoldási módszereit össze tudjuk hasonlítani, azt is tudnunk kell, mennyi a műveletigényük. A flop mértékegységről részletesen a függelékben írunk a ?? oldalán.

**2.50. TÉTEL (A KIKÜSZÖBÖLÉS MŰVELETIGÉNYE).** A Gauss- és a Gauss–Jordan-módszer műveletigénye egy  $n$ -ismeretlenes,  $n$  egyenletből álló egyenletrendszer esetén egyaránt

$$\frac{n^3}{3} + \frac{n^2}{2} - \frac{5n}{6} \text{ összeadás/kivonás, } \frac{n^3}{3} + n^2 - \frac{n}{3} \text{ szorzás/osztás.}$$

azaz összesen

$$\frac{2}{3}n^3 + \frac{3}{2}n^2 - \frac{7}{6}n \text{ flop,}$$

azaz jó közelítéssel  $2n^3/3$  flop.

**BIZONYÍTÁS.** Először felelevenítünk két elemi összefüggést, amire a bizonyításban szükség van:

$$1 + 2 + \dots + n = \frac{n(n+1)}{2}, \quad 1^2 + 2^2 + \dots + n^2 = \frac{n(n+1)(2n+1)}{6}.$$

A továbbiakban feltételezzük, hogy a kiküszöbölés során a főátlóba kerülő elemek egyike sem 0. A Gauss-módszernél a főátló alatti elemek eliminálásához  $\frac{1}{3}n^3 - \frac{1}{3}n$  összeadás és  $\frac{1}{3}n^3 + \frac{1}{2}n^2 - \frac{5}{6}n$  szorzás szükséges. A visszahelyettesítés  $\frac{1}{2}n^2 - \frac{1}{2}n$  összeadásból és  $\frac{1}{2}n^2 + \frac{1}{2}n$  szorzásból áll. Ha a Gauss–Jordan-módszernél a főátló alatti elemek kiküszöbölése mellett a főátló elemeit is 1-re változtatjuk, az  $\frac{1}{3}n^3 - \frac{1}{3}n$  összeadás mellett  $\frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$  szorzás szükséges. A főátló feletti elemek eliminálásához  $\frac{1}{2}n^2 - \frac{1}{2}n$  összeadás és ugyanennyi szorzás kell. A számítások részletezését az olvasóra hagyjuk.  $\square$

**Numerikusan instabil egyenletrendszerek** A gyakorlati feladatokban gyakran mérési eredményekkel, így nem pontos adatokkal dolgozunk.

2.51. PÉLDA (INSTABIL EGYENLETRENDSZER). Oldjuk meg a következő egyenletrendszert!

$$6.73x - 8.97y = 5.61$$

$$4.79x - 6.39y = 3.99$$

Mutassuk meg, hogy az együtthatók 0.01-dal való megváltoztatása a megoldások nagy megváltozását okozhatja, sőt az is elérhető, hogy az egyenletrendszernek ne legyen, vagy épp végtelen sok megoldása legyen!

**MEGOLDÁS.** Az egyenletrendszer megoldása:  $x = 1.5$ ,  $y = 0.5$ . Az első egyenletben az  $x$  együtthatóját újra mérjük, és másodszorra egy századdal kevesebbnek, 6.72-nek adódik. Oldjuk meg ezt az egyenletrendszert is! Az eredmény meglepő módon nagyon messze van az előzőtől:  $x \approx -2.26$ ,  $y \approx -2.32$ . Újabb mérés az  $y$  együtthatóját  $-8.96$ -nak mutatja. E két együttható egy századdal való megváltozása után a megoldás messze van mindkét előző eredménytől:  $x \approx 4.35$ ,  $y \approx 2.64$ . Ha végül az első egyenlet konstans tagját is megváltoztatjuk egy századdal 5.62-re, akkor  $x \approx 7.21$ ,  $y \approx 4.78$  lesz az eredmény, ha pedig 5.60-ra, akkor – csemegeként – ismét a kerek  $x = 1.5$ ,  $y = 0.5$  értékeket kapjuk.

A fenti egyenletrendszeren tovább változtatva az együtthatókat az is elérhető, hogy végtelen sok megoldása legyen:

$$6.72x - 8.96y = 5.60$$

$$4.80x - 6.40y = 4.00$$

ugyanis itt a két egyenlet egymás konstansszorososa. Ha pedig a második egyenlet konstans tagját visszaírjuk 3.99-re, egy ellentmondó egyenletrendszert kapunk.  $\square$

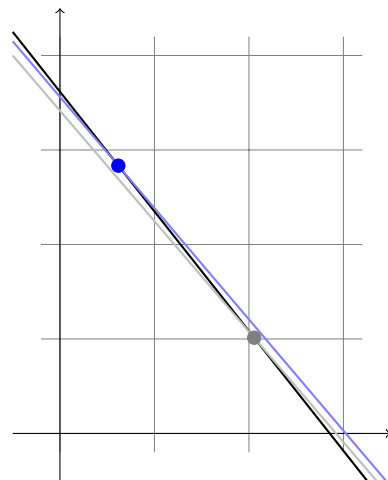
Ilyen megbízhatatlan eredményekkel a gyakorlatban semmit nem lehet kezdeni!

Az olyan egyenletrendszert, melyben az együtthatók vagy a konstans tagok kis változása a megoldásban nagy változást okoz, *numerikusan instabillnak* vagy *rosszul kondicionáltnak* nevezzük. Egyébként *numerikusan stabil*, illetve *jól kondicionált* egyenletrendszerről beszélünk.

Világos, hogy a fentiek nem precíz matematikai fogalmak. Később precízen definiálva számmal fogjuk mérni a kondicionáltság fokát, de azt, hogy egy adott egyenletrendszer megoldásai elfogadhatóak-e vagy nem, csak a feladat döntheti el.

A numerikus instabilitás okát szemlélteti a 2.12. ábra. Kétváltozós egyenletrendszerek esetén, ha a két egyenes grafikonja „közel” van egymáshoz, azaz majdnem egybe esnek, akkor kis változások az egyeneseken messze vihetik a metszéspontot, de párhuzamossá is tehetik a két egyenest.

Ha a gyakorlatban numerikusan instabil egyenletrendszerrel találkozunk, vizsgáljuk meg, hogy az egyenleteink közti „majdnem” lineá-



2.12. ábra: Instabil egyenletrendszer, melyben az egyenletek együtthatóinak kis megváltoztatása a megoldás nagy megváltozását okozza.



ris összefüggőség mögött nem valódi lineáris összefüggőség van-e kis mérési hibával.

**Részleges főelemkiválasztás** A következő példákban lebegőpontos aritmetikát használunk. A számításokat úgy kell elvégezni, hogy az adott pontosságnak megfelelően minden részeredményt  $p$  értékes jegyre kerekítünk.

**2.52. PÉLDA (GAUSS-MÓDSZER LEBEGŐPONTOS SZÁMOKKAL).** Oldjuk meg az alábbi – numerikusan stabil – egyenletrendszert pontosan, majd 3 értékes jegy pontossággal számolva.

$$\begin{aligned} 10^{-4}x + y &= 2 \\ x - y &= 0 \end{aligned}$$

**MEGOLDÁS.** Pontosán számolva

$$\left[ \begin{array}{cc|c} 10^{-4} & 1 & 2 \\ 1 & -1 & 0 \end{array} \right] \xrightarrow{S_2 - 10^4 S_1} \left[ \begin{array}{cc|c} 10^{-4} & 1 & 2 \\ 0 & -1 - 10^4 & -2 \cdot 10^4 \end{array} \right]$$

amiből az eredmény  $x = y = \frac{2 \cdot 10^4}{1 + 10^4}$ . Igazolható, hogy az egyenletrendszer numerikusan stabil, ami azt jelenti, hogy például  $10^{-4}$  helyébe 0-t helyettesítve, vagyis kicsit változtatva egy együtthatót, a kapott

$$\left[ \begin{array}{cc|c} 0 & 1 & 2 \\ 1 & -1 & 0 \end{array} \right]$$

egyenletrendszer megoldása csak kicsit különbözik az előzőtől:  $x = y = 2$ . Végezzük most el a Gauss-kiküszöbölést 3 értékes jeggyel számolva:

$$\left[ \begin{array}{cc|c} 10^{-4} & 1 & 2 \\ 1 & -1 & 0 \end{array} \right] \xrightarrow{S_2 - 10^4 S_1} \left[ \begin{array}{cc|c} 10^{-4} & 1 & 2 \\ 0 & -1 - 10^4 & -2 \cdot 10^4 \end{array} \right] \approx \left[ \begin{array}{cc|c} 10^{-4} & 1 & 2 \\ 0 & -10^4 & -2 \cdot 10^4 \end{array} \right],$$

ahol a közelítésnél a  $\text{fl}(-1 - 10^4) = -10^4$  összefüggést használtuk. Az így kapott egyenletrendszernek viszont  $x = 0, y = 2$  a megoldása, ami nagyon messze van az eredeti egyenletrendszer megoldásától! Most végezzünk egy apró változtatást: először cseréljük fel a két egyenletet!

$$\left[ \begin{array}{cc|c} 1 & -1 & 0 \\ 10^{-4} & 1 & 2 \end{array} \right] \xrightarrow{S_2 - 10^{-4} S_1} \left[ \begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 1 + 10^{-4} & 2 \end{array} \right] \approx \left[ \begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 1 & 2 \end{array} \right],$$

amelynek megoldása  $x = y = 2$ , ami nagyon közel van a pontos megoldáshoz!  $\square$

Mi az oka a két megoldás közti különbségnek, és fel tudnánk-e használni minél jobb megoldás megtalálásában?

Mindkét megoldásban az első egyenlet konstansszorosát hozzáadtuk a második egyenlethez, de az első esetben a kisebb, a másodikban az első oszlop nagyobb elemét választottuk főelemnek. Amikor a kisebbet választottuk, akkor az első sort egy kis számmal osztottuk, vagyis reciprokával – egy nagy számmal – szoroztuk, és ezt adtuk a második sorhoz. A nagy számmal való beszorzás következtében a második egyenlet együtthatóit „elnyomták” e nagy számok, nagyon megváltoztatva az egyenletet, aminek következtében a megoldások is nagyon megváltoztak! A  $\text{fl}(-1 - 10^4) = -10^4$  kerekítés hatása, vagyis a  $-1$  „eltüntetése”, ekvivalens azzal, mintha az eredeti egyenletrendszer helyett a következőt kéne megoldani:

$$\left[ \begin{array}{cc|c} 10^{-4} & 1 & 2 \\ 1 & 0 & 0 \end{array} \right].$$

És ennek valóban  $x = 0, y = 2$  a megoldása! Amikor viszont az első oszlop nagyobbik elemét választottuk főelemnek, a sort egy kis számmal kellett beszorozni, és ezt hozzáadni a másik sorhoz, vagyis kerekítéskor az eredeti egyenlet együtthatói megmaradtak, az egyenletrendszer kevésbé torzult. Ennek alapján megfogalmazható egy széles körben elterjedt szabály: a Gauss-féle kiküszöbölési eljárás során, lebegőpontos adatokkal dolgozva minden oszlopban a szóbajöhető elemek közül – sorcserék segítségével – mindig a legnagyobb abszolút értékűt válasszuk főelemnek! E módszert *részleges főelemkiválasztásnak*, illetve *részleges pivotálásnak* nevezik.

Bizonyos – a gyakorlatban ritkán előforduló – esetekben jobb eredmény kapható a *teljes főelemkiválasztás* módszerével. Ekkor főelemnek az összes még hátralévő elem abszolút értékben legnagyobbikát választjuk. Itt oszlopcserékre is szükség van, és műveletigényesebb is ez az eljárás, ezért ritkán alkalmazzák.

**2.53. PÉLDA (RÉSZLEGES FŐELEMKIVÁLASZTÁS).** *Részleges főelemkiválasztással hozzuk lépcsős alakra az alábbi mátrixot!*

$$\begin{bmatrix} 1.8 & 3.0 & 3.0 & 3.7 & 7.5 \\ 3.6 & 3.2 & 3.6 & 6.2 & 7.8 \\ 7.2 & 3.6 & 4.8 & 2.4 & 1.2 \\ 2.4 & 5.4 & 5.2 & 2.6 & 5.2 \end{bmatrix}$$

**MEGOLDÁS.** Az első oszlop legnagyobb eleme a harmadik sorban van,

így az első és a harmadik sor cseréjével kezdünk:

$$\begin{array}{l}
 \xrightarrow{S_1 \leftrightarrow S_3} \begin{bmatrix} 7.2 & 3.6 & 4.8 & 2.4 & 1.2 \\ 3.6 & 3.2 & 3.6 & 6.2 & 7.8 \\ 1.8 & 3.0 & 3.0 & 3.7 & 7.5 \\ 2.4 & 5.4 & 5.2 & 2.6 & 5.2 \end{bmatrix} \xrightarrow{\begin{array}{l} S_2 - S_1/2 \\ S_3 - S_1/4 \\ S_4 - S_1/3 \end{array}} \begin{bmatrix} 7.2 & 3.6 & 4.8 & 2.4 & 1.2 \\ 0.0 & 1.4 & 1.2 & 5.0 & 7.2 \\ 0.0 & 2.1 & 1.8 & 3.1 & 7.2 \\ 0.0 & 4.2 & 3.6 & 1.8 & 4.8 \end{bmatrix} \\
 \xrightarrow{S_2 \leftrightarrow S_4} \begin{bmatrix} 7.2 & 3.6 & 4.8 & 2.4 & 1.2 \\ 0.0 & 4.2 & 3.6 & 1.8 & 4.8 \\ 0.0 & 2.1 & 1.8 & 3.1 & 7.2 \\ 0.0 & 1.4 & 1.2 & 5.0 & 7.2 \end{bmatrix} \xrightarrow{\begin{array}{l} S_3 - S_2/2 \\ S_4 - S_2/3 \end{array}} \begin{bmatrix} 7.2 & 3.6 & 4.8 & 2.4 & 1.2 \\ 0.0 & 4.2 & 3.6 & 1.8 & 4.8 \\ 0.0 & 0.0 & 0.0 & 2.2 & 4.8 \\ 0.0 & 0.0 & 0.0 & 4.4 & 5.6 \end{bmatrix} \\
 \xrightarrow{S_3 \leftrightarrow S_4} \begin{bmatrix} 7.2 & 3.6 & 4.8 & 2.4 & 1.2 \\ 0.0 & 4.2 & 3.6 & 1.8 & 4.8 \\ 0.0 & 0.0 & 0.0 & 4.4 & 5.6 \\ 0.0 & 0.0 & 0.0 & 2.2 & 4.8 \end{bmatrix} \xrightarrow{S_4 - S_3/2} \begin{bmatrix} 7.2 & 3.6 & 4.8 & 2.4 & 1.2 \\ 0.0 & 4.2 & 3.6 & 1.8 & 4.8 \\ 0.0 & 0.0 & 0.0 & 4.4 & 5.6 \\ 0.0 & 0.0 & 0.0 & 0.0 & 2.0 \end{bmatrix} \quad \square
 \end{array}$$

**Skálázás** A részleges főelemkiválasztásban mindig az oszlop legnagyobb elemét választottuk. Nem lehet egy elemet nagyobbá tenni, és ezzel az egész módszert elrontani úgy, hogy egy sorát egyszerűen beszorozzuk?

**2.54. PÉLDA (SOR SZORZÁSA).** A 2.52. példában szorozzuk meg az első egyenletet  $10^5$ -nel, azaz a kisebb elemből csináljunk nagyot, és ezt az egyenletrendszert is oldjuk meg részleges főelemkiválasztással.

$$\begin{array}{rcl}
 10x + 10^5y & = & 2 \cdot 10^5 \\
 x - y & = & 0
 \end{array}$$

**MEGOLDÁS.** Egy egyenlet beszorozása egy nemzérus számmal ekvivalens átalakítás, így ennek az egyenletrendszernek is  $x = y = \frac{2 \cdot 10^4}{1+10^4}$  a pontos megoldása. Ha 3 értékes jegyre számolunk, és alkalmazzuk a részleges főelemkiválasztás módszerét, akkor ismét rossz eredményt kapunk:

$$\left[ \begin{array}{cc|c} 10 & 10^5 & 2 \cdot 10^5 \\ 1 & -1 & 0 \end{array} \right] \xrightarrow{S_2 - \frac{1}{10} S_1} \left[ \begin{array}{cc|c} 10 & 10^5 & 2 \cdot 10^5 \\ 0 & -1 - 10^4 & -2 \cdot 10^4 \end{array} \right] \approx \left[ \begin{array}{cc|c} 10 & 10^5 & 2 \cdot 10^5 \\ 0 & -10^4 & -2 \cdot 10^4 \end{array} \right],$$

amiből  $x = 0$  és  $y = 2$ . □

Hasonlóképp zavart okozhat az együtthatómátrix egy oszlopának beszorozása is, ami az egyenletrendszeren például úgy valósítható meg, ha egyik változó mértékegységét megváltoztatjuk. (Ha például a korábban kilométerben meghatározott ismeretlent milliméterben keressük, együtthatóját minden egyenletben  $10^6$ -al kell osztani.)

Az együtthatók ilyen „egyenetlenségeiből” származó számítási hibák csökkentésére a *skálázás* nevű gyakorlati módszer ajánlható. Ez a

következő két skálázási szabály követéséből áll, mely a tapasztalatok szerint a gyakorlati feladatok nagy részében nagyon jó eredményt ad a részleges főelemkiválasztással együtt alkalmazva:

1. *Oszlopok skálázása:* Válasszunk a feladatban szereplő mennyiségeknek természetes mértékegységet, ezzel általában elkerülhetők az együtthatók közti tetemes nagyságrendi különbségek. Ezen kívül nincs szükség az oszlopok elemeinek beszorzására.
2. *Sorok skálázása:* Az egyenletrendszer  $[A|b]$  bővített mátrixának minden sorát osszuk el az  $A$  együtthatómátrix adott sorbeli legnagyobb abszolút értékű elemével. Így  $A$  minden sorának 1 a legnagyobb eleme.

Nem ismeretes olyan módszer, mely a lebegőpontos ábrázolás korlátai mellett hatékonyan megtalálná a lehető legpontosabb eredményt. Az elmélet és a tapasztalatok alapján sűrű, nem túlzottan nagy méretű egyenletrendszerekre a skálázott főelem kiválasztásos Gauss-módszer ajánlható. A ritka egyenletrendszerekre a következőkben ismertetendő iteratív módszerek általában jobb eredményt adnak.

*Iteratív módszerek* A továbbiakban is csak olyan egyenletrendszerekkel foglalkozunk, melyek  $n$ -ismeretlenesek és  $n$  egyenletből állnak, tehát melyek együtthatómátrixa négyzetes.

Az iteratív módszerek lényege, hogy olyan

$$x_0, x_1, \dots, x_k, \dots$$

vektorsorozatot generálunk, mely az adott egyenletrendszer megoldásvektorához konvergál. Első pillanatra meglepőnek tűnhet egy végtelen sorozat generálásával keresni a megoldást, de mivel számításaink eleve csak véges pontosságúak, gyakran igen kevés lépésben elérhetjük a megkívánt pontosságot. Ráadásul a kerekítési hibák még növelhetik is a konvergencia sebességét.

A kiindulási pont – a matematika több más területén is gyümölcsöző módszer – a fixpontkeresés. Ennek lényegét egy egyváltozós függvény példáján mutatjuk be. Legyen  $f$  egy minden valós helyen értelmezett függvény, mely bármely két  $a$  és  $b$  pontot két olyan pontba visz, melyek távolsága  $a$  és  $b$  távolságának legfeljebb a fele. Képletben:

$$|f(b) - f(a)| \leq \frac{1}{2}|b - a|, \text{ azaz } \frac{|f(b) - f(a)|}{|b - a|} \leq \frac{1}{2}.$$

Ez azt jelenti, hogy  $f$  összes különbségi hányadosa legfeljebb  $1/2$ . A sokkal általánosabban megfogalmazható Banach-féle fixpont tétel szerint ekkor egyetlen olyan  $\bar{x}$  pont létezik, hogy  $\bar{x} = f(\bar{x})$ , és ez megkapható úgy, hogy egy tetszőleges  $x_0$  pontból kiindulva képezzük az

$$x_0, x_1 = f(x_0), x_2 = f(x_1), \dots, x_{k+1} = f(x_k), \dots$$

sorozatot, és vesszük a határértékét. Ekkor

$$\bar{x} = \lim_{k \rightarrow \infty} x_k.$$

A 2.13. ábra szemlélteti a fenti állítást. Az  $1/2$ -es szorzó kicserélhető tetszőleges  $0$  és  $1$  közé eső konstansra.

A **Banach fixponttétel** könnyen szemléltethető hétköznapi módon: képzeljük el, hogy egy nagyobb gumilapot néhányan körbeállva egy kerek asztal tetején széthúznak az asztal széléig, majd (most jön a leképezés!) visszaengedik eredeti állapotába. Ekkor igaz az, hogy az asztalon pontosan egy olyan pont van, mely fölött a gumilap helyben marad. E pont megkapható, ha kiválasztunk az asztalon egy tetszőleges  $P_0$  pontot, és megnézzük, hogy a kinyújtott gumilap e fölötti pontja összehúzódáskor hová ugrik, legyen ez a  $P_1$  pont az asztalon. A kinyújtott gumilap  $P_1$  fölötti pontja összehúzódáskor az  $P_2$  pont fölé ugrik, stb. Az így kapott pontsorozat a fixponthoz konvergál.

**Jacobi-iteráció** Az előző paragrafusban leírtakat követve megpróbáljuk az egyenletrendszert átrendezni úgy, hogy az  $\mathbf{x} = f(\mathbf{x})$  alakú legyen, ahol  $\mathbf{x}$  jelöli az ismeretlenek vektorát.

2.55. PÉLDA (JACOBI-ITERÁCIÓ). Oldjuk meg a

$$4x - y = 2$$

$$2x - 5y = -8$$

egyenletrendszert, majd hozzuk  $\mathbf{x} = f(\mathbf{x})$  alakra, és egy tetszőleges  $\mathbf{x}_0$  vektorból indulva végezzünk az  $\mathbf{x}_{k+1} = f(\mathbf{x}_k)$  formulával iterációt. Számoljunk 3 tizedes pontossággal. Hová tart az így kapott vektorsorozat?

**MEGOLDÁS.** Az egyenletrendszert kiküszöböléssel megoldva kapjuk, hogy  $\mathbf{x} = (1, 2)$  az egyetlen megoldás.

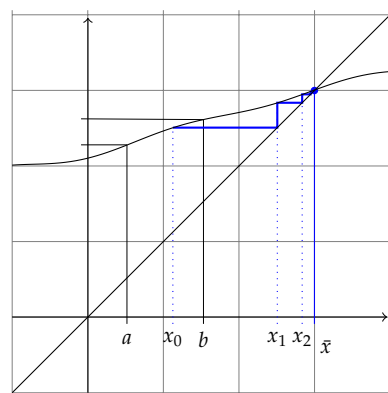
Hozzuk az egyenletrendszert  $\mathbf{x} = f(\mathbf{x})$ , azaz  $\begin{bmatrix} x \\ y \end{bmatrix} = f\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)$  alakra. Erre több lehetőség is adódik. Ezek közül talán az a legkézenfekvőbb, hogy az első egyenletből az  $x$ -et, a másodikból  $y$ -t kifejezzük:

$$x = \frac{y+2}{4}$$

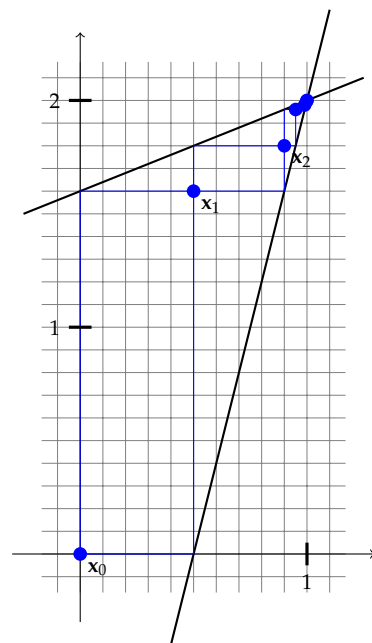
$$y = \frac{2x+8}{5}$$

Válasszunk egy  $\mathbf{x}_0$  vektort tetszőlegesen, legyen pl.  $\mathbf{x}_0 = (0, 0)$ , azaz  $x = y = 0$ . A fenti képletekbe helyettesítve kapjuk, hogy  $\mathbf{x}_1 = \left(\frac{0+2}{4}, \frac{0+8}{5}\right) = (0.5, 1.6)$ . A további értékeket egy táblázatban adjuk meg:

|     | $\mathbf{x}_0$ | $\mathbf{x}_1$ | $\mathbf{x}_2$ | $\mathbf{x}_3$ | $\mathbf{x}_4$ | $\mathbf{x}_5$ | $\mathbf{x}_6$ | $\mathbf{x}_7$ | $\mathbf{x}_8$ |
|-----|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| $x$ | 0              | 0.5            | 0.9            | 0.95           | 0.99           | 0.995          | 0.999          | 1.000          | 1.000          |
| $y$ | 0              | 1.6            | 1.8            | 1.96           | 1.98           | 1.996          | 1.998          | 2.000          | 2.000          |



2.13. ábra: Egy függvény, mely bármely  $a$  és  $b$  pontot két olyan pontba visz, melyek távolsága  $a$  és  $b$  távolságának legfölbbe a fele, így a függvény minden különbségi hányadosa abszolút értékben legfölbbe  $1/2$ . E függvénynek pontosan egy fixpontja van, mely megkapható egy tetszőleges  $x_0$  pontból induló  $x_k = f(x_{k-1})$  sorozat határértékéént.



2.14. ábra: A Jacobi-iteráció szemléltetése

E példa esetén tehát a végtelen sorozat konvergensenek mutatkozott, de a kerekítési hiba folytán véges sok lépés után megtalálta a konvergenciapontot.  $\square$

Az általános eset hasonlóan írható le. Tegyük fel, hogy az

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ \vdots & \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= b_n \end{aligned}$$

egyenletrendszer egyértelműen megoldható, és főátlójának minden eleme különbözik 0-tól. A *Jacobi-iteráció* menete tehát a következő. A  $k$ -adik egyenletből fejezzük ki az  $x_k$  változót:

$$\begin{aligned} x_1 &= \frac{1}{a_{11}}(b_1 - a_{12}x_2 - \dots - a_{1,n-1}x_{n-1} - a_{1n}x_n) \\ x_2 &= \frac{1}{a_{22}}(b_2 - a_{21}x_1 - \dots - a_{2,n-1}x_{n-1} - a_{2n}x_n) \\ &\vdots \\ x_n &= \frac{1}{a_{nn}}(b_n - a_{n1}x_1 - a_{n2}x_2 - \dots - a_{n,n-1}x_{n-1}). \end{aligned} \quad (2.27)$$

Válasszunk az ismeretlenek  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  vektorának egy  $\mathbf{x}_0$  kezdőértéket, pl. legyen  $\mathbf{x}_0 = (0, 0, \dots, 0)$ . A (2.27) egyenletrendszer jobb oldalába helyettesítsük be  $\mathbf{x}_0$  koordinátáinak értékét, a bal oldal adja  $\mathbf{x}_1$  koordinátáit. Ezt a lépést ismételjük meg, generálva az  $\mathbf{x}_2, \mathbf{x}_3, \dots$  vektorokat addig, míg el nem érjük a megfelelő pontosságot.

**Gauss–Seidel-iteráció** A Jacobi-iteráció gyorsaságán könnyen javíthatunk, ha a (2.27) egy egyenletének jobb oldalába való behelyettesítés után a bal oldalon kapott változó értékét azonnal fölhasználjuk, nem várunk vele a ciklus végéig. Ezt a módosított algoritmust nevezzük *Gauss–Seidel-iterációnak*.

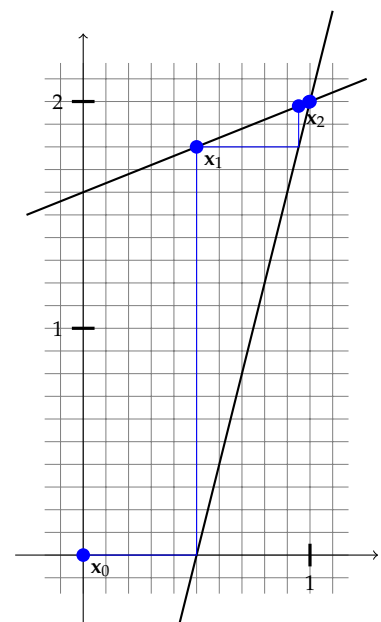
2.56. PÉLDA (GAUSS–SEIDEL-ITERÁCIÓ). Oldjuk meg a

$$\begin{aligned} 4x - y &= 2 \\ 2x - 5y &= -8 \end{aligned}$$

egyenletrendszert Gauss–Seidel-iterációval.

**MEGOLDÁS.** Itt is, mint a Jacobi-iterációnál az

$$\begin{aligned} x &= \frac{y+2}{4} \\ y &= \frac{2x+8}{5} \end{aligned}$$



2.15. ábra: A Gauss–Seidel-iteráció szemléltetése

egyenleteket használjuk, de míg a Jacobi-iterációnál  $\mathbf{x}_0 = (0, 0)$  kezdőérték után az  $x = \frac{0+2}{4} = \frac{1}{2}$ ,  $y = \frac{0+8}{5} = \frac{8}{5}$  értékek következtek, a Gauss–Seidel-iterációnál a második egyenletben 0 helyett már az első egyenletben kiszámolt  $x = \frac{1}{2}$  értéket helyettesítjük, azaz  $y = \frac{2\frac{1}{2}+8}{5} = \frac{9}{5}$ . A további értékeket egy táblázatban adjuk meg de úgy, hogy jelezzük a kiszámítás sorrendjét:

|     | $\mathbf{x}_0$ | $\mathbf{x}_1$ | $\mathbf{x}_2$ | $\mathbf{x}_3$ | $\mathbf{x}_4$ |
|-----|----------------|----------------|----------------|----------------|----------------|
| $x$ | 0              | 0.5            | 0.95           | 0.995          | 1.000          |
| $y$ | 0              | 1.8            | 1.98           | 1.998          | 2.000          |

Hasonlítsuk össze az eredményt a Jacobi iterációnál készített táblázattal. □

Mindkét iteráció jól szemléltethető a 2-ismeretlenes esetben. A 2.14. és a 2.15. ábrák ezt mutatják. Ha pontosan számolunk, végtelen sok pontot kapunk, az ábrák csak néhány pontot mutatnak.

**Az iterációk konvergenciája** A fenti példákból nem látszik, hogy vajon a Jacobi- és a Gauss–Seidel-iterációk mindig konvergálnak-e.

2.57. PÉLDA (DIVERGENS ITERÁCIÓ). Oldjuk meg Jacobi- és Gauss–Seidel-iterációval a következő egyenletrendszert:

$$\begin{aligned}x - y &= 2 \\ 2x - y &= 5\end{aligned}$$

**MEGOLDÁS.** Alakítsuk át az egyenletrendszert:

$$\begin{aligned}x &= y + 2 \\ y &= 2x - 5\end{aligned}$$

Először próbálkozzunk Jacobi-iterációval:

|     | $\mathbf{x}_0$ | $\mathbf{x}_1$ | $\mathbf{x}_2$ | $\mathbf{x}_3$ | $\mathbf{x}_4$ | $\mathbf{x}_5$ | $\mathbf{x}_6$ | $\mathbf{x}_7$ | $\mathbf{x}_8$ |
|-----|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| $x$ | 0              | 2              | -3             | 1              | -9             | -1             | -21            | -5             | -45            |
| $y$ | 0              | -5             | -1             | -11            | -3             | -23            | -7             | -47            | -15            |

Úgy tűnik, nem konvergens a vektorsorozat, mint ahogy nem tűnik annak a Gauss–Seidel-iterációnál sem:

|     | $\mathbf{x}_0$ | $\mathbf{x}_1$ | $\mathbf{x}_2$ | $\mathbf{x}_3$ | $\mathbf{x}_4$ |
|-----|----------------|----------------|----------------|----------------|----------------|
| $x$ | 0              | 2              | 1              | -1             | -5             |
| $y$ | 0              | -1             | -3             | -7             | -15            |

A divergencia leolvasható az iterációkat szemléltető ábrákról is! □

**2.58. DEFINÍCIÓ (SORONKÉNT DOMINÁNS FŐÁTLÓJÚ MÁTRIX).** Azt mondjuk, hogy az  $n \times n$ -es  $\mathbf{A}$  mátrix soronként (szigorúan) domináns főátlóval rendelkezik, vagy soronként (szigorúan) domináns főátlójú, ha a főátló minden eleme abszolút értékben nagyobb a sorában lévő többi elem abszolút értékeinek összegénél, azaz képletben

$$\begin{array}{rcl} |a_{11}| & > & |a_{12}| + \dots + |a_{1,n-1}| + |a_{1n}| \\ |a_{22}| & > & |a_{21}| + \dots + |a_{2,n-1}| + |a_{2n}| \\ \vdots & & \ddots \\ |a_{n-1,n-1}| & > & |a_{n-1,1}| + |a_{n-1,2}| + \dots + |a_{n-1,n}| \\ |a_{nn}| & > & |a_{n1}| + |a_{n2}| + \dots + |a_{n,n-1}| \end{array}$$

Hasonlóan definiálható az oszloponként domináns főátlójú mátrix is.

Világos, hogy az alábbi mátrixok soronként domináns főátlójúak:

$$\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} -10 & 1 & 2 \\ 1 & 10 & -3 \\ -1 & -2 & 10 \end{bmatrix}, \quad \begin{bmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -5 \end{bmatrix}, \quad \begin{bmatrix} 1 & .25 & .25 & .25 \\ .25 & 1 & .25 & .25 \\ .25 & .25 & 1 & .25 \\ .25 & .25 & .25 & 1 \end{bmatrix}.$$

Az alábbi mátrixok nem soronként domináns főátlójúak, de sorcserékkel azzá tehetők:

$$\begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}, \quad \begin{bmatrix} -1 & -2 & 10 \\ -10 & 1 & 2 \\ 1 & 10 & -3 \end{bmatrix}, \quad \begin{bmatrix} 0 & -3 & 0 \\ 0 & 0 & -5 \\ 2 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} .25 & .25 & .25 & 1 \\ 1 & .25 & .25 & .25 \\ .25 & .25 & 1 & .25 \\ .25 & 1 & .25 & .25 \end{bmatrix}.$$

Az alábbi egyenletrendszer együtthatómátrixa soronként domináns főátlójú:

$$\begin{array}{rcrcrcr} 4x & - & y & = & 11 \\ 2x & - & 5y & = & -17 \end{array}$$

**2.59. TÉTEL (ELÉGSÉGES FELTÉTEL AZ ITERÁCIÓK KONVERGENCIÁJÁRA).** Ha az  $n$  egyenletből álló  $n$ -ismeretlenes egyenletrendszer együtthatómátrixa soronként domináns főátlójú, akkor bármely indulóvektor esetén a Jacobi- és a Gauss–Seidel-iteráció is konvergens.

E tételt nem bizonyítjuk. Megjegyezzük, hogy a tételbeli feltétel nem szükséges, csak elégséges, azaz olyan egyenletrendszeren is konvergens lehet valamelyik iteráció, melynek nem domináns főátlójú az együtthatómátrixa. Hasonló tétel igaz oszloponként domináns főátlójú együtthatómátrixok esetén is.

A domináns főátlójú mátrixokon a Gauss–Seidel-iteráció sosem lassabb, mint a Jacobi-iteráció, sőt, gyakran érezhetően gyorsabb. Az



viszont előfordulhat, hogy a Gauss–Seidel-iteráció divergens, míg a Jacobi-iteráció konvergens (ld. 2.35. feladat).

A gyakorlatban ezeknek az iterációknak különböző, hatékonyabb javításait használják. E témában az olvasó figyelmébe ajánljuk a „numerikus módszerek” témában írt könyveket, web-oldalakat.

## Feladatok

### Jacobi-iteráció

2.34. Oldjuk meg a

$$4x - y = 8$$

$$2x - 5y = -5$$

egyenletrendszer Jacobi-iterációval! Számoljunk 3, majd 4 értékes jegyre!

### Vegyes feladatok

2.35. **JACOBI-ITERÁCIÓ KONVERGÁL, GAUSS-SEIDEL-ITERÁCIÓ NEM** Írjunk programot annak az állításnak az ellenőrzésére, hogy a

$$x + z = 0$$

$$-x + 5/6y = 0$$

$$x + 2y - 3z = 1$$

egyenletrendszeren a Jacobi-iteráció konvergál, a Gauss-Seidel-iteráció nem.

2.36. **GAUSS-ELIMINÁCIÓ DOMINÁNS FŐÁTLÓJÚ MÁTRIXON** Bizonyítsuk be, hogy ha az  $A$  mátrix főátlója soronként domináns, akkor végrehajtható rajta a főelemkiválasztásos Gauss-elimináció sorcsere nélkül!

2.37. **AZ ITERÁCIÓK SZEMLÉLTETÉSE** Az  $A$  városból elindul egy  $A$  jelű vonat a  $B$  város felé, vele egy időben a  $B$  városból egy  $B$  jelű  $A$  felé. A  $B$  vonat indulásával egy időben a  $B$

vonat orráról elindul egy légy is  $A$  felé, de amint találkozik az  $A$  vonattal megfordul, és addig repül, míg a  $B$  vonattal nem találkozik, amikor ismét megfordul, stb. Mindhármuk sebessége konstans, de a légy sebessége nagyobb mindkét vonaténál.

1. Egy táblázatban megadjuk mindkét vonat távolságát az indulási helyüktől km-ben mérve azokban a pillanatokban, amikor a légy épp a  $B$  vonattal találkozik.

|                         | $(x_0, y_0)$ | $(x_1, y_1)$ | $(x_2, y_2)$ | $(x_3, y_3)$ |
|-------------------------|--------------|--------------|--------------|--------------|
| $x$ : távolság $A$ -tól | 0            | 40           | 48           | 49.6         |
| $y$ : távolság $B$ -től | 0            | 80           | 96           | 99.2         |

Számítsuk ki a táblázat egy-két további oszlopát! Milyen messze van  $A$  város  $B$ -től?

2. Most egy másik táblázatban megadjuk annak a vonatnak a távolságát az indulási helyétől, amelyik épp találkozik a légygel:

|                         | $y_0$ | $x_1$ | $y_1$ | $x_2$ | $y_2$ | $x_3$ | $y_3$ |
|-------------------------|-------|-------|-------|-------|-------|-------|-------|
| $x$ : távolság $A$ -tól |       | 30    |       | 46    |       | 49.2  |       |
| $y$ : távolság $B$ -től | 0     |       | 80    |       | 96    |       | 99.2  |

Számítsuk ki a táblázat egy-két további oszlopát! Milyen messze van  $A$  város  $B$ -től?

3. Mi köze van e feladatnak a Jacobi- és a Gauss-Seidel-iterációhoz?

## Megoldások

2.4. a) igen, b) igen, c) nem, d) igen, e) igen, f) nem.

2.5. Könnyen látható, hogy mindkét egyenletrendszer egyetlen megoldása:  $x = 2$ ,  $y = 1$ , tehát a két egyenletrendszer ekvivalens.

2.6. Az első egyenletrendszer nem oldható meg a  $0 = 3$  alakú egyenlet miatt, de a második sem, mivel nincs olyan  $x$  és  $y$ , melyre  $x + y = 2$  és  $x + y = 7$  lenne, hisz  $2 \neq 7$ .

2.7. Behelyettesítés után mindkét egyenlet  $0 = 0$  alakú, amit tetszőleges  $x$  és  $y$  kielégít, így az összes  $(x, y)$  számpár megoldása az egyenletrendszernek.

2.8.  $x = 1$ ,  $y$  tetszőleges, azaz az összes  $(1, y)$  alakú számpár megoldás.

2.9. A második egyenlet behelyettesítés után  $0 = 1$  alakú, így az egyenletrendszernek nincs megoldása.

2.10.  $x = 1$ ,  $y = 2$ , azaz  $(x, y) = (1, 2)$  az egyetlen megoldás.

2.12. a) A sormodellben két metsző egyenest kell megrajzolni ( $y = 7/3x - 2/3x$ ,  $y = -2 + 3/2x$ ), melyek a  $(2, 1)$  pontban metszik egymást, míg az oszlopmodellben a  $(2, 3)$ , a  $(3, -2)$  vektorokat és azok lineáris kombinációjaként előállított  $(7, 4) = 2(2, 3) + (3, -2)$  vektort!

b) A sormodellben két párhuzamos egyenest kell megrajzolni, míg az oszlopmodellben a  $(2, 3)$  és a  $(4, 6)$  vektorokat, melyek egy egyenesbe esnek, és semmilyen lineáris kombinációjuk sem adja ki a  $(3, 4)$  vektort!

2.13. A három sík közül semelyik kettő nem párhuzamos, másrészt a normálvektorai egy síkba esnek, ugyanis  $2(1, 1, 2) + (1, 2, 4) = (3, 4, 8)$ . Ez azt jelenti, hogy van olyan vektor, mely mindhárom síkkal párhuzamos. Az első esetben a három sík egy egyenesen megy át, mivel van a síkoknak közös pontjuk, pl. a  $(3, 0, 0)$  pont, így végtelen sok megoldása is van, míg a második esetben a síkoknak nincs közös pontjuk.

Az egyenletrendszerek ekvivalensek a következő vektoregyenletekkel:

$$\begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} x + \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} y + \begin{bmatrix} 2 \\ 4 \\ 8 \end{bmatrix} z = \begin{bmatrix} 3 \\ 3 \\ 9 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} x + \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} y + \begin{bmatrix} 2 \\ 4 \\ 8 \end{bmatrix} z = \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix}.$$

Itt a közös együtthatómátrix minden oszlopvektora benne van a

$$2x + y - z = 0$$

egyenletű síkban, (ez könnyen ellenőrizhető a vektorok koordinátáinak a sík egyenletébe való helyettesítésével), és ki

is feszítik a síkot, mert a három vektor nem kollineáris. Másrészt a  $(3, 3, 9)$  vektor is benne van e síkban, a  $(3, 3, 1)$  vektor viszont nem. Tehát az első egyenletrendszer megoldható, a második nem.

2.14. A sormodell szerinti ábra az a) esetben 3 síkbeli egyenest tartalmaz, melyek közt van két párhuzamos, így az egyenletrendszer nem oldható meg. A b) esetben a három egyenes egy ponton megy át, ez a megoldás:  $x = 2$ ,  $y = 1$ . A c) esetben ugyan nincsenek párhuzamos egyenesek, de nincs közös pontjuk sem, így az egyenletrendszer nem oldható meg. Az oszlopmodell szerint az a)

$$x + y = 3$$

$$x + y = 4$$

$$x + 2y = 4$$

egyenletrendszer ekvivalens a következővel:

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} y = \begin{bmatrix} 3 \\ 4 \\ 4 \end{bmatrix}.$$

Az  $(1, 1, 1)$  és az  $(1, 1, 2)$  vektorok benne fekszenek az  $x = y$  egyenletű síkban, mivel első két koordinátájuk megegyezik, ezért minden lineáris kombinációjuk is ebbe a síkba esik. A  $(3, 4, 4)$  vektor viszont nem esik e síkba, így *független* az előbbi kettőtől, tehát nem áll elő azok lineáris kombinációjaként. Vagyis ez az egyenletrendszer nem oldható meg. A b) és c) egyenletrendszerekben a bal oldali két vektor, az  $(1, 1, 1)$  és az  $(1, 2, 3)$  az  $x - 2y + z = 0$  egyenletű síkban van, melyben a  $(3, 4, 5)$  vektor benne van, míg a  $(3, 3, 5)$  vektor nincs benne, tehát b) megoldható, c) nem.

2.15. a) hamis, az állítás csak úgy igaz, ha a párhuzamos hipersíkok különbözőek is (két azonos hipersíkot párhuzamosnak tekintünk), b) hamis, például a 2.9. (a) ábrán látható esetben nincsenek párhuzamos síkok, és mégis megoldás, c) igaz, mert akkor a jobb oldalon álló bármely vektor kifejezhető e két kétdimenziós vektor lineáris kombinációjaként, tehát az egyenletrendszer megoldható.

2.16.

a) Egy két egyenletből álló háromismeretlenes egyenletrendszer sormodellje szerinti ábra a **három**dimenziós térben **két** darab **síkból** áll, melyek ha **párhuzamosak**, **de nem azonosak**, akkor az egyenletrendszernek nincs megoldása, egyébként megoldásainak száma **végtelen**. Oszlopmodellje a **kétdimenziós** térben **négy** darab **vektorból** áll (három lineáris kombinációja a negyedik).

b) Egy három egyenletből álló kétismeretlenes egyenletrendszer sormodellje szerinti ábra a **kétdimenziós** térben **három egyenesből** áll, míg az oszlopmodellje a **három**dimenziós térben **három** darab **vektorból**.

c) Egy négy egyenletről álló ötismeretlenes egyenletrendszer sormodellje szerinti ábra az ötdimenziós térben **négy** darab **hipersíkból** áll. Oszlopmodellje a **négydimenziós** térben **hat** darab **vektorból** áll.

2.17. a) Igaz, ez a szám megegyezik a főelemek számával. b) Igaz, ez a szám megegyezik a főelemek számával. c) Hamis, van lépcsős alakja minden mátrixnak, de csak a redukált lépcsős alak egyértelmű. d) Hamis. e) Igaz.

$$2.18. \begin{bmatrix} 1 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$2.19. \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & -1 \end{bmatrix}$$

2.20. A sorcsere előállítható a beszorzás és a hozzáadás segítségével, nevezetesen az  $S_i \leftrightarrow S_j$  sorcsere ekvivalens az

$$S_i + S_j, S_j - S_i, S_i + S_j, -S_j,$$

elemi sorműveletekkel. Ellenőrzésül a mátrixokon való hatásukat is megadjuk:

$$\begin{bmatrix} \vdots \\ \mathbf{s}_i \\ \vdots \\ \mathbf{s}_j \\ \vdots \end{bmatrix} \xrightarrow{S_i + S_j} \begin{bmatrix} \vdots \\ \mathbf{s}_i + \mathbf{s}_j \\ \vdots \\ \mathbf{s}_j \\ \vdots \end{bmatrix} \xrightarrow{S_j - S_i} \begin{bmatrix} \vdots \\ \mathbf{s}_i + \mathbf{s}_j \\ \vdots \\ -\mathbf{s}_i \\ \vdots \end{bmatrix} \xrightarrow{S_i + S_j} \begin{bmatrix} \vdots \\ \mathbf{s}_j \\ \vdots \\ -\mathbf{s}_i \\ \vdots \end{bmatrix} \xrightarrow{-S_j} \begin{bmatrix} \vdots \\ \mathbf{s}_j \\ \vdots \\ \mathbf{s}_i \\ \vdots \end{bmatrix}$$

2.21. Az egyenletrendszernek és lépcsős alakjának bővített mátrixa:

$$\left[ \begin{array}{cccc|c} 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 & 1 & 2 \\ 1 & 0 & 0 & 1 & 1 & 1 \end{array} \right] \rightarrow \left[ \begin{array}{cccc|c} 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 2 & 1 & 2 \end{array} \right]$$

Legyen  $x_5 = t$ , az utolsó egyenletből  $x_4 = 1 - \frac{1}{2}t$ , a harmadik egyenletbe való helyettesítés után  $x_3 = 1 - \frac{1}{2}t$ , a második egyenletből  $x_2 = t$ , végül az első egyenletből  $x_1 = -\frac{1}{2}t$ . Tehát a megoldás

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}t \\ t \\ 1 - \frac{1}{2}t \\ 1 - \frac{1}{2}t \\ t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -\frac{1}{2} \\ 1 \\ -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix}.$$

Az  $x_4 = t$  paraméterválasztás esetén az eredmény törtmen-

tes alakot ölt:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} t-1 \\ -2t+2 \\ t \\ t \\ -2t+2 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 0 \\ 0 \\ 2 \end{bmatrix} + t \begin{bmatrix} 1 \\ -2 \\ 1 \\ 1 \\ -2 \end{bmatrix}.$$

2.22. Egyetlen elemi sorművelet világossá teszi, hogy az egyenletrendszer inkonzisztens:

$$\left[ \begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ -1 & 1 & -1 & 2 \\ 2 & 1 & 2 & 1 \\ 4 & 4 & 4 & 1 \end{array} \right] \xrightarrow{S_4 - 4S_1} \left[ \begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ -1 & 1 & -1 & 2 \\ 2 & 1 & 2 & 1 \\ 0 & 0 & 0 & -15 \end{array} \right]$$

2.23. Az elemi sorműveletek:

$$\begin{bmatrix} 7 & 14 & -21 & 7 \\ 1 & 2 & -3 & 1 \\ 5 & 10 & 15 & 5 \\ 3 & 6 & -9 & 3 \end{bmatrix} \xrightarrow{\frac{1}{7}S_1, \frac{1}{5}S_3, \frac{1}{3}S_4} \begin{bmatrix} 1 & 2 & -3 & 1 \\ 1 & 2 & -3 & 1 \\ 1 & 2 & 3 & 1 \\ 1 & 2 & -3 & 1 \end{bmatrix} \xrightarrow{\substack{S_2 - S_1 \\ S_3 - S_1 \\ S_4 - S_1}} \begin{bmatrix} 1 & 2 & -3 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & -3 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Innen a megoldás:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1-2t \\ t \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}.$$

E feladat arra példa, hogy abból, hogy az egyenletek száma több az ismeretleneknél, nem következik, hogy az egyenletrendszer inkonzisztens. Sőt, mint látjuk, akár végtelen sok megoldása is lehet!

2.24.  $(x, y, z) = (19/3, -2, 2/3)$  az egyetlen megoldás. E feladat arra példa, hogy abból, hogy az egyenletek száma több az ismeretleneknél, nem következik, hogy az egyenletrendszer inkonzisztens.

2.25. Az egyenletrendszer inkonzisztens.

$$2.26. \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2-s-t \\ 1+s-3t \\ s \\ t \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ -3 \\ 0 \\ 1 \end{bmatrix}.$$

2.27. a) Igaz. b) Hamis, az egyenletrendszer megoldhatósága nem függ az egyenletek számától. Az ismeretlenek számánál akár kevesebb, akár több egyenletről álló rendszer akár konzisztens, akár inkonzisztens is lehet. c) Igaz. d) Igaz, a nullvektor mindig megoldás.

2.28. Igen, mindkettőnek  $(x, y, z) = (1, 2, 3)$  az egyetlen megoldása, azaz megoldáshalmazaik megegyeznek.

2.29. Igen, mindkettő bővített együtthatómátrixának redukált lépcsős alakja a zérussor nélkül

$$\begin{bmatrix} 1 & 0 & -0.8 & 1.8 \\ 0 & 1 & 2.2 & -1.2 \end{bmatrix}$$

2.30. a) Igen. b) Nem.

2.31. Ha az egyenletrendszernek van legalább három sora, akkor az első és a harmadik számtani közepe a középső, így nem tudunk  $n > 2$  esetén fölírni  $n$  független egyenletről álló  $n$ -ismeretlenes egyenletrendszert e feltételekkel. Ha  $n = 2$ , akkor a bővített mátrix és annak lépcsős alakja

$$\left[ \begin{array}{cc|c} a & a+d & a+2d \\ a+3d & a+4d & a+5d \end{array} \right] \rightarrow \left[ \begin{array}{ccc} 1 & 1 & 1 \\ 0 & 1 & 2 \end{array} \right],$$

ahol  $a$  tetszőleges,  $d \neq 0$ . Innen a megoldásvektor  $(-1, 2)$ . (Két ismeretlen, de kettőnél több egyenlet esetén,  $d \neq 0$  mellett két független egyenlet lesz, melyek ugyanezt a megoldást adják).

2.32. Mind a négy egyenlet megfelelő helyettesítéssel a következő lineáris egyenletrendszerre vezet:

$$2X + 2Y = 8$$

$$3X + Y = 4$$

Ennek megoldása  $X = 0$ ,  $Y = 4$ . Innen a)  $(x, y) = (0, 16)$ , b) két megoldás van:  $(0, 2)$  és  $(0, -2)$ , c) ez az egyenletrendszer nem oldható meg, mivel a  $e^x = 0$  egyenlet nem oldható meg, d) ez az egyenletrendszer sem oldható meg, mivel a  $\cos x = 4$  egyenlet nem oldható meg (a valós számok körében).

2.33. Jelölje az 1, 5 és 10 Ft-osok számát rendre  $x$ ,  $y$  és  $z$ . A feladat a következő egyenletrendszerre vezet:

$$x + y + z = 11 \quad (2.25)$$

$$x + 2y + 5z = 53 \quad (2.26)$$

Ennek megoldásai  $(x, y, z) = (\frac{1}{2} + \frac{5}{4}t, \frac{21}{2} - \frac{9}{4}t, t)$ . A 10 Ft-os érmék száma legfőljebb 5, így elég a  $t = 1, 2, \dots, 5$  értékeket kipróbálni. Az egyetlen megoldás, amely pozitív egészekből áll, azaz ahol minden érme darabszáma pozitív egész:  $x = 3$ ,  $y = 6$ ,  $z = 2$ . (Ha észrevesszük, hogy csak úgy kaphatunk egész megoldásokat, ha  $t$  néggyel osztva kettő

maradékot ad, akkor a  $t = 2$  eseten kívül más megoldás nem is jöhet szóba.)

2.34. Legyen  $x_0 = (0, 0)$ . Az iteráció képletei:

$$x = \frac{y+8}{4}, \quad y = \frac{2x+5}{5}.$$

Az iteráció lépéseinek táblázata 3 értékes jeggyel számolva:

|     | $x_0$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ |
|-----|-------|-------|-------|-------|-------|-------|-------|
| $x$ | 0     | 2     | 2.25  | 2.45  | 2.48  | 2.50  | 2.50  |
| $y$ | 0     | 1     | 1.80  | 1.90  | 1.98  | 1.99  | 2.00  |

Az iteráció lépéseinek táblázata 4 értékes jeggyel számolva:

|     | $x_0$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ | $x_8$ |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $x$ | 0     | 2     | 2.25  | 2.45  | 2.475 | 2.495 | 2.498 | 2.500 | 2.5   |
| $y$ | 0     | 1     | 1.80  | 1.90  | 1.980 | 1.990 | 1.998 | 1.999 | 2.0   |

2.36.

2.37. A Jacobi-iteráció szerinti módon, a vonatok valamilyének és a légynek a  $k$ -adik találkozásából kiszámítva a  $k+1$ -edik találkozáshoz jellemző távolságokat, az  $x_{k+1} = ay_k + b$ ,  $y_{k+1} = cx_k + d$  egyenletekre jutunk. Az első táblázat adatait behelyettesítve, és  $a$ ,  $b$ ,  $c$  és  $d$  értékre megoldva az  $a = 1/10$ ,  $b = 40$ ,  $c = 2/5$ ,  $d = 80$  értékeket kapjuk, amiből a táblázat további értékei számolhatók, és az eredeti egyenletrendszer is fölírható:

$$\begin{aligned} x - \frac{1}{10}y &= 40 \\ -\frac{2}{5}x + y &= 80. \end{aligned}$$

Ennek megoldása  $(x, y) = (50, 100)$ , vagyis a vonatok találkozásáig az  $A$  vonat 50 km-t, a  $B$  vonat 100 km-t tesz meg. Eszerint a két város 150 km-re van egymástól.

A Gauss-Seidel-iteráció szerinti  $x_{k+1} = ay_k + b$ ,  $y_{k+1} = cx_{k+1} + d$  egyenletek épp illeszkednek a feladat második táblázatához az  $a = 1/5$ ,  $b = 30$ ,  $c = 1$ ,  $d = 50$  értékekkel.

Ez az

$$\begin{aligned} x - \frac{1}{5}y &= 30 \\ -x + y &= 50 \end{aligned}$$

egyenletrendszerhez tartozik, melynek megoldása ismét  $(x, y) = (50, 100)$ .



### 3

## Megoldhatóság és a megoldások tere

E fejezetet az egyenletrendszer megoldásainak jellemzésére szánjuk. Ennek során megismerkedünk az altér fogalmával és a velük végezhető legegyszerűbb műveletekkel. Végül megmutatjuk, hogy minden konzisztens lineáris egyenletrendszer origóhoz legközelebbi megoldása az egyetlen, mely a sortérbe esik, és bármely két megoldásának különbsége merőleges rá.

### Homogén és inhomogén egyenletrendszerek megoldásai

---

*Az előzőekben a megoldások megtalálásának módszereit tanulmányoztuk. E szakaszban a megoldhatóság kérdését és a megoldások halmazának legfontosabb tulajdonságait vizsgáljuk. A vizsgálatokban a lineáris egyenletrendszerek mindkét geometriai interpretációja fontos szerepet kap.*

---

**Kötött változók száma, mátrix rangja** A redukált lépcsős alak egyértelműségének egy nyilvánvaló, de fontos folyománya az alábbi eredmény:

**3.1. KÖVETKEZMÉNY (FŐELEMENK OSZLOPAI).** *Egy valós mátrix bármely lépcsős alakjában a főelemek ugyanazokban az oszlopokban vannak, tehát ezek száma is független a lépcsős alaktól.*

A bizonyítás azonnal adódik abból, hogy bármely lépcsős alak főelemeiből kapjuk a redukált lépcsős alak vezéregyeseit, így bármely lépcsős alak főelemei ugyanott vannak, ahol a vezéregyese, a redukált lépcsős alak pedig egyértelmű.

Ebből az is következik, hogy bármely valós mátrix esetén

|                                              |   |                                                    |   |                                                        |
|----------------------------------------------|---|----------------------------------------------------|---|--------------------------------------------------------|
| bármely lépcsős<br>alak főelemeinek<br>száma | = | bármely lépcsős<br>alak nemzérus<br>sorainak száma | = | a redukált lépcsős<br>alak vezéregye-<br>seinek száma. |
|----------------------------------------------|---|----------------------------------------------------|---|--------------------------------------------------------|

Ez a következő definícióhoz vezet.

**3.2. DEFINÍCIÓ (MÁTRIX RANGJA).** Egy mátrix valamely lépcsős alakjában a nemnulla sorok számát a mátrix rangjának nevezzük. Az  $\mathbf{A}$  mátrix rangját  $r(\mathbf{A})$  jelöli.

**3.3. PÉLDA (MÁTRIX RANGJÁNAK KISZÁMÍTÁSA).** Számítsuk ki az alábbi mátrixok rangját!

$$\begin{bmatrix} 2 & 3 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 3 & 2 & 1 \\ 0 & 0 & 4 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}.$$

**MEGOLDÁS.** Az első és második mátrix lépcsős alakú, rangjuk azonnal leolvasható: 1 és 2. A harmadik és negyedik mátrix elemi sorműveletekkel

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & -2 & 0 & -2 \\ 0 & 0 & -2 & 2 \\ 0 & 0 & 0 & -4 \end{bmatrix}, \text{ illetve } \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

alakra hozható, tehát a rang 4, illetve 2. □

**3.4. ÁLLÍTÁS (KÖTÖTT ÉS SZABAD VÁLTOZÓK SZÁMA).** Ha egy  $n$ -ismeretlenes egyenletrendszer megoldható, és együtthatómátrixának rangja  $r$ , akkor a Gauss- vagy a Gauss–Jordan-módszerrel kapott megoldásában a kötött változók száma  $r$ , a szabad változók száma  $n - r$ .

► Megjegyezzük, hogy egyelőre csak annyit látunk, hogy ha az együtthatómátrix és a bővített mátrix rangja  $r$ , akkor az egyenletrendszernek van olyan megoldása, amelyben a kötött változók száma  $r$ , a szabad változóké  $n - r$ , és egy ilyen megoldás megkapható a Gauss- vagy a Gauss–Jordan-módszerrel. Arról még nem tudunk semmit, hogy a változók sorrendjének felcserélésével, vagy más megoldási módszerrel nem kaphatjuk-e meg ugyanezt a megoldás több vagy épp kevesebb kötött változóval. A következő fejezetben be fogjuk látni, hogy a kötött és szabad változók száma független a változók sorrendjétől, és a meghatározásuk módszerétől is.

► Például a

$$\left[ \begin{array}{cccccc|c} 1 & 3 & 2 & 6 & 0 & 4 & 1 & 2 \\ 0 & 0 & 3 & 1 & 2 & 3 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

bővített mátrixhoz tartozó egyenletrendszerben 3 a kötött és 4 a szabad változók száma.



**Egyenletrendszer megoldhatóságának feltétele** Tudjuk, hogy egy lineáris egyenletrendszer pontosan akkor *nem* oldható meg, ha a bővített mátrix lépcsős alakjának van olyan sora, melyben csak a legutolsó elem nem nulla. Ez ugyanis egyenletté visszaírva  $0 = c$  alakú, ahol  $c \neq 0$ , és ennek az egyenletnek nincs megoldása. Ez viszont azt jelenti, hogy ilyenkor a bővített mátrix rangja nagyobb az együtthatómátrix rangjánál. E megállapítás azonnali következménye a következő tétel.

**3.5. TÉTEL (A MEGOLDHATÓSÁG MÁTRIXRANGOS FELTÉTELE).** *Legyen egy  $n$ -ismeretlenes lineáris egyenletrendszer együtthatómátrixa  $\mathbf{A}$ , a konstans tagokból álló vektora  $\mathbf{b}$ .*

1. *Ez az egyenletrendszer pontosan akkor oldható meg, ha együtthatómátrixának és bővített mátrixának rangja megegyezik, azaz*

$$r(\mathbf{A}) = r(\mathbf{A}|\mathbf{b}).$$

2. *Ez az egyenletrendszer pontosan akkor oldható meg egyértelműen, ha együtthatómátrixának és bővített mátrixának rangja megegyezik az ismeretlenek számával, azaz*

$$r(\mathbf{A}) = r(\mathbf{A}|\mathbf{b}) = n.$$

**BIZONYÍTÁS.** Az egyenletrendszer pontosan akkor oldható meg, ha bővített mátrixának lépcsős alakjában nincs olyan sor, melynek csak az utolsó eleme nem 0. Ez épp azt jelenti, hogy  $r(\mathbf{A}) = r(\mathbf{A}|\mathbf{b})$ .

A második állítás abból következik, hogy az egyenletrendszer akkor oldható meg egyértelműen, ha megoldható, és nincs szabad változója, vagyis az együtthatómátrix rangja megegyezik az ismeretlenek számával.  $\square$

Az előzőekből az is adódik, hogy egy lineáris egyenletrendszernek pontosan akkor van egynél több megoldása, ha

$$r(\mathbf{A}) = r(\mathbf{A}|\mathbf{b}) < n.$$

(Miért nem lehet  $r(\mathbf{A}) > n$ ?)

Egy valós együtthatós egyenletrendszernek csak úgy lehet egynél több megoldása, ha van szabad változója. Viszont, annak minden értékéhez egy-egy másik megoldás tartozik, vagyis ekkor az egyenletrendszernek végtelen sok megoldása van. Így, ha  $\mathbf{A}$  valós mátrix, akkor a megoldások száma, a két rang és az ismeretlenek száma közt a következő a kapcsolat:

| Feltétel                                       | Megoldások száma |
|------------------------------------------------|------------------|
| $r(\mathbf{A}) < r(\mathbf{A} \mathbf{b})$     | 0                |
| $r(\mathbf{A}) = r(\mathbf{A} \mathbf{b}) = n$ | 1                |
| $r(\mathbf{A}) = r(\mathbf{A} \mathbf{b}) < n$ | $\infty$         |

Ha az egyenletrendszer homogén lineáris, azaz mindegyik konstans tag 0, akkor az elemi sorműveletek közben a bővített mátrix utolsó oszlopában minden elem 0 marad, így ebben az oszlopban biztosan nem lesz főelem. Eszerint a homogén lineáris egyenletrendszerek mindig megoldhatók, hisz ekkor a  $r(\mathbf{A}) = r(\mathbf{A}|\mathbf{b})$  összefüggés mindig fennáll. A megoldhatóság persze e feltétel ellenőrzése nélkül is látszik, hisz az  $x_1 = x_2 = \dots = x_n = 0$  mindig megoldás! Mivel  $r(\mathbf{A})$  megegyezik a redukált lépcsős alak főelemeinek számával, ezért  $r(\mathbf{A}) \leq m$  és  $r(\mathbf{A}) \leq n$  is fennáll, ahol  $m$  az egyenletek,  $n$  az ismeretlenek száma. Így viszont  $m < n$  esetén  $r(\mathbf{A}) = n$  nem állhat fenn, tehát a homogén lineáris egyenletrendszernek van az  $\mathbf{x} = \mathbf{0}$  vektoron kívül is megoldása. Ezzel bizonyítottuk az alábbi tételt:

**3.6. TÉTEL (HOMOGÉN LINEÁRIS EGYENLETRENDSZER MEGOLDHATÓSÁGA).** Az  $\mathbf{A}$  együtthatómátrixú homogén lineáris egyenletrendszer mindig megoldható, mert a nullvektor – az ún. triviális megoldás – mindig megoldás. Pontosán akkor van nemtriviális, vagyis a  $\mathbf{0}$ -vektortól különböző megoldása is, ha

$$r(\mathbf{A}) < n,$$

ahol  $n$  az ismeretlenek – azaz  $\mathbf{A}$  oszlopainak – számát jelöli. Speciálisan, az  $m$  egyenletből álló homogén lineáris egyenletrendszernek  $m < n$  esetén mindig van nemtriviális megoldása.

Valós együtthatós homogén lineáris egyenletrendszerekre az előző táblázat a következő alakot ölti:

| Feltétel            | Megoldások száma |
|---------------------|------------------|
| $r(\mathbf{A}) = n$ | 1                |
| $r(\mathbf{A}) < n$ | $\infty$         |

**3.7. PÉLDA (EGYENLETRENDSZER MEGOLDÁSAINAK SZÁMA).** Az  $a$  paraméter mely értékei mellett van az alábbi egyenletrendszernek 0, 1, illetve  $\infty$  sok megoldása?

$$\begin{aligned} x_1 + x_2 + ax_3 &= 1 \\ x_1 + ax_2 + x_3 &= a \\ ax_1 + x_2 + x_3 &= a^2 \end{aligned}$$

**MEGOLDÁS.** Hozzuk a bővített mátrixot lépcsős alakra:

$$\begin{aligned} \left[ \begin{array}{ccc|c} 1 & 1 & a & 1 \\ 1 & a & 1 & a \\ a & 1 & 1 & a^2 \end{array} \right] &\xrightarrow[S_3 - aS_1]{S_2 - S_1} \left[ \begin{array}{ccc|c} 1 & 1 & a & 1 \\ 0 & a-1 & 1-a & a-1 \\ 0 & 1-a & 1-a^2 & a^2-a \end{array} \right] \xrightarrow{S_3 + S_2} \\ \left[ \begin{array}{ccc|c} 1 & 1 & a & 1 \\ 0 & a-1 & 1-a & a-1 \\ 0 & 0 & -(a-1)(a+2) & (a+1)(a-1) \end{array} \right] \end{aligned}$$

Látható, hogy  $a = 1$  esetén az utolsó két sorban minden elem 0, tehát az együtthatómátrix és a bővített mátrix rangja is 1, így az egyenletrendszer az  $x_1 + x_2 + x_3 = 1$  egyenlettel ekvivalens. Ennek megoldása:  $(x_1, x_2, x_3) = (1 - s - t, s, t)$ , azaz oszlopvektor alakba írva:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Ha  $a = -2$ , akkor az együtthatómátrix rangja 2, a bővített mátrix rangja 3, tehát az egyenletrendszer nem oldható meg (az utolsó sor egyenletté visszaírva  $0 = 3$  alakú). Minden egyéb esetben, azaz ha  $a \neq 1$  és  $a \neq -2$ , akkor a két rang 3, ami megegyezik az ismeretlenek számával, tehát egyetlen megoldás van. Ez ki is fejezhető:

$$x_1 = \frac{(a+1)^2}{a+2}, \quad x_2 = \frac{1}{a+2}, \quad x_3 = -\frac{a+1}{a+2}. \quad \square$$

**Homogén lineáris egyenletrendszer megoldásai** Tekintsünk egy tetszőleges homogén lineáris egyenletrendszert. Mint a 3.6. tételben láttuk, ez biztosan megoldható, és a megoldások halmazában a nullvektor benne van. Mit mondhatunk a megoldások halmazáról, ha több megoldása is van a homogén egyenletrendszernek?

**3.8. ÁLLÍTÁS (MEGOLDÁSOK LINEÁRIS KOMBINÁCIÓJA).** Egy homogén lineáris egyenletrendszer megoldásainak bármely lineáris kombinációja is megoldás.

**BIZONYÍTÁS.** Elég az állítást két megoldásra bizonyítani. Jelölje  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$  az egyenletrendszer együtthatómátrixának oszlopvektorait. Legyen  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  és  $\mathbf{y} = (y_1, y_2, \dots, y_n)$  két tetszőleges megoldás, azaz

$$\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \dots + \mathbf{a}_n x_n = \mathbf{0}$$

$$\mathbf{a}_1 y_1 + \mathbf{a}_2 y_2 + \dots + \mathbf{a}_n y_n = \mathbf{0},$$

és  $c, d$  legyen két tetszőleges skalár. Megmutatjuk, hogy ekkor  $c\mathbf{x} + d\mathbf{y}$  is megoldás, ugyanis

$$\begin{aligned} & \mathbf{a}_1(cx_1 + dy_1) + \mathbf{a}_2(cx_2 + dy_2) + \dots + \mathbf{a}_n(cx_n + dy_n) = \\ & (c\mathbf{a}_1 x_1 + d\mathbf{a}_1 y_1) + (c\mathbf{a}_2 x_2 + d\mathbf{a}_2 y_2) + \dots + (c\mathbf{a}_n x_n + d\mathbf{a}_n y_n) = \\ & c(\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \dots + \mathbf{a}_n x_n) + d(\mathbf{a}_1 y_1 + \mathbf{a}_2 y_2 + \dots + \mathbf{a}_n y_n) = \\ & \mathbf{0} + \mathbf{0} = \mathbf{0}. \end{aligned}$$

azaz  $c\mathbf{x} + d\mathbf{y}$  is megoldás. Ez bizonyítja állításunkat.

E bizonyítás az oszlopmodellre épült, de hasonlóan egyszerű bizonyítás adható a sormmodellben is (ld. 3.10. feladat).  $\square$

**Vektortér** Eddig vektortéren az összes rendezett szám- $n$ -esek halmazát értettük, ahol  $n$  egy rögzített pozitív egész. A következőkben kiterjesztjük a vektortér fogalmát.

Egyelőre csak a valós vektorokkal foglalkozunk, azaz vektoron rendezett valós szám- $n$ -est értünk, ahol  $n$  tetszőleges pozitív egész.

**3.9. DEFINÍCIÓ (VEKTORTÉR).** Vektortéren vektorok olyan nem üres  $\mathcal{V}$  halmazát értjük, melyre igaz, hogy  $\mathcal{V}$  bármely két vektora összeadható, és összegük is  $\mathcal{V}$ -beli, valamint  $\mathcal{V}$  bármely valós  $c$  számmal vett szorzata is  $\mathcal{V}$ -beli. Másként fogalmazva  $\mathcal{V}$  vektortér, ha zárt a vektorösszeadás és a skalárral szorzás műveletére.

**3.10. DEFINÍCIÓ (ALTÉR).** Ha  $\mathcal{U}$  és  $\mathcal{V}$  két vektortér és  $\mathcal{U} \subseteq \mathcal{V}$ , akkor azt mondjuk, hogy az  $\mathcal{U}$  vektortér a  $\mathcal{V}$  vektortér altere. Jelölése:  $\mathcal{U} \leq \mathcal{V}$ .

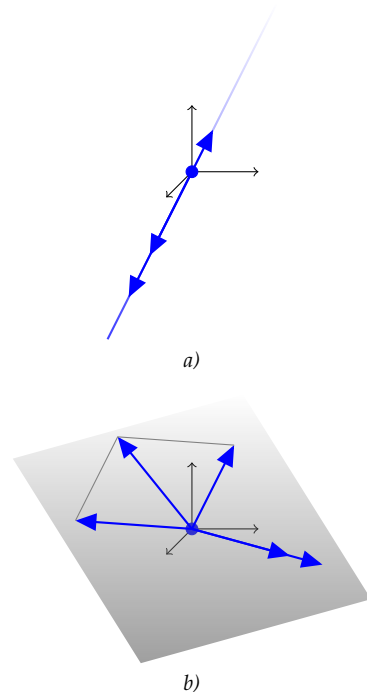
- Az  $\mathcal{A}$  vektorhalmaz pontosan akkor vektortér, ha az  $\mathcal{A}$ -beli vektorokból képzett lineáris kombinációk is mind  $\mathcal{A}$ -ban vannak (ld. a 7.12. feladatot).
- Minden pozitív  $n$  egész esetén  $\mathbb{R}^n$  vektortér.
- A síkban ( $\mathbb{R}^2$ -ben) egy origón átmenő egyenes vektorai (az egyenes pontjaiba mutató helyvektorok) alteret alkotnak.
- A térben ( $\mathbb{R}^3$ -ben) bármely origón átmenő sík vagy egyenes vektorai alteret alkotnak (ld. a 3.1. ábrát).
- Az  $\mathbb{R}^3$  imént felsorolt alterei – az origón átmenő egyenes és sík – „olyanok”, mint az  $\mathbb{R}$  és az  $\mathbb{R}^2$ . E ködös megfogalmazást a vektortér absztrakt definíciója és a vektorterek izomorfizmusának fogalma fogja tisztázni (ld. a 13 fejezetet). Akkor fogjuk igazolni, hogy  $\mathbb{R}^n$  alterei valóban mind „olyanok”, mint  $\mathbb{R}^k$ , ahol  $k \leq n$ .

**Alterek tulajdonságai és szemléltetésük** A halmazok szemléltetésére használt Venn-diagramok példáját követve a vektortereket olyan formában fogjuk ábrázolni, mely annak néhány tulajdonságát egyszerűen szemlélteti.

A *levéldiagram* egy egyszerű levélformával szemlélteti a vektorteret, melynek a levél száránál lévő csúcsa (tőve) jelzi a nullvektort. Az altereket olyan kisebb levelek szemléltetik, melyek tőve – azaz a nullvektor – közös (ld. 3.2. ábra). A levelek felső csúcsába a tér nevét, esetleg dimenzióját írhatjuk.

Felsoroljuk az alterek néhány egyszerűen belátható tulajdonságát, néhányukat e diagrammal szemléltetve:

- Minden altérnek eleme a nullvektor, hisz bármely altérbeli vektorral együtt annak 0-szorosa is, vagyis a  $\mathbf{0}$ -vektor is eleme az altérnek.
- Minden altérbeli  $\mathbf{x}$  vektorral együtt annak ellentettje ( $-1$ -szerese), a  $-\mathbf{x}$  vektor is eleme az altérnek.



3.1. ábra: a) Egy origón átmenő egyenes bármely vektorának konstansszorzosa és bármely két vektorának összege az egyenesbe esik, b) egy origón átmenő sík bármely vektorának konstansszorzosa és bármely két vektorának összege a síkba esik.



3.2. ábra: Egy  $\mathcal{W}$  vektortér és annak  $\mathcal{U}$  és  $\mathcal{V}$  alterei, valamint a mindannyiukban közös nullvektor

- Minden vektortér maga is altér (saját maga altere), hisz bármely két vektorának összes lineáris kombinációját is tartalmazza.
- A nullvektor önmagában alteret alkot, ez a *zérustér*, amit  $\mathcal{Z}$  jelöl. A nulltér kifejezést másra használjuk, ne keverjük össze. A zérustert és a teljes teret szokás triviális altereknek nevezni (ld. 3.3. ábra).
- Altér altere altér, azaz ha  $\mathcal{U} \leq \mathcal{V}$ , és  $\mathcal{W} \leq \mathcal{U}$ , akkor  $\mathcal{W} \leq \mathcal{V}$ . (ld. 3.4. ábra).
- Két altér metszete is altér. Ha  $\mathcal{U}$  és  $\mathcal{V}$  egy vektortér két altere, és  $\mathcal{W}$  a közös részük, akkor  $\mathcal{W}$  nem üres, hisz a nullvektor benne van. Másrészt bármely két  $\mathbf{x}, \mathbf{y} \in \mathcal{W}$  vektor összes lineáris kombinációja benne van  $\mathcal{U}$ -ban és  $\mathcal{V}$ -ben is, így metszetükben is. Alterek metszetére is a  $\cap$  jelet használjuk, tehát az előbbi alterekre  $\mathcal{U} \cap \mathcal{V} = \mathcal{W}$  (ld. 3.5. ábra).
- Az előző gondolat tetszőleges számú alterre is megismételhető, tehát egy vektortér tetszőleges számú alterének közös része is altér. Az alterek száma végtelen is lehet.
- Két altér egyesítése csak akkor altér, ha egyik altere a másiknak. Például a térben egy origón átmenő egyenes vektorait és egy origón átmenő sík vektorait egyesítve csak akkor kapunk alteret, ha az egyenes a síkba esik.

3.11. PÉLDA (ALTÉR). Alteret alkotnak-e az alábbi vektorhalmazok  $\mathbb{R}^3$ -ben?

- $\{(x, y, z) \mid x = y, z = xy\}$ ,
- $\{(s + 2t, s - 1, 2s + t) \mid s, t \in \mathbb{R}\}$ ,
- $\{(x, y, z) \mid 2x - y + z = 0\}$ ,
- $\{(x, y, z) \mid x = 2t, y = -t, z = t, t \in \mathbb{R}\}$ .

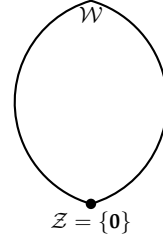
**MEGOLDÁS.** Az *a)* pontbeli halmaz nem altér, mert nem elégíti ki a definícióbeli feltételeket. Például az  $(1, 1, 1)$  vektor benne van e halmazban, azonban kétszerese nem, mert nem elégíti ki a  $z = xy$  egyenlőséget!

A *b)* pontban szereplő halmazban nincs benne a nullvektor, ugyanis az  $s + 2t = 0, s - 1 = 0, 2s + t = 0$  egyenletrendszernek nincs megoldása, így ez a halmaz sem alkot alteret!

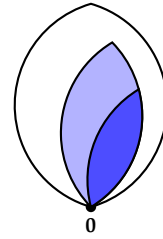
A *c)*-beli halmaz az  $\mathbf{n} = (2, -1, 1)$  normálvektorú sík pontjaiba mutató helyvektorokból áll. A sík vektoregyenlete  $\mathbf{n} \cdot \mathbf{r} = 0$ . Ha  $\mathbf{x}$  és  $\mathbf{y}$  a sík két vektora, azaz  $\mathbf{n} \cdot \mathbf{x} = 0$  és  $\mathbf{n} \cdot \mathbf{y} = 0$ , akkor  $\mathbf{n} \cdot (\mathbf{x} + \mathbf{y}) = 0$  és  $\mathbf{n} \cdot (c\mathbf{x}) = 0$  is fennáll bármely  $c \in \mathbb{R}$  valós számra, tehát valóban alteret kaptunk.

A *d)*-beli halmazba a  $\mathbf{v} = (2, -1, 1)$  vektor skalárszorosai tartoznak, melyek közül bármely kettő összege és bármelyik skalárszorosa is e halmazba tartozik, tehát ezek a vektorok is alteret alkotnak. E vektorok végpontjai egy origón áthaladó egyenes pontjait adják.  $\square$

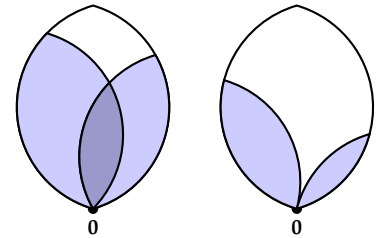
Könnyen belátható, hogy  $\mathbb{R}^2$  alterei az alábbiak:



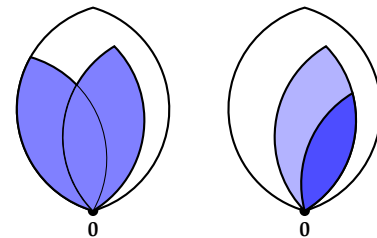
3.3. ábra: Egy  $\mathcal{W}$  vektortér két triviális alterével: az egyik maga  $\mathcal{W}$ , a másik a nullvektort tartalmazó  $\mathcal{Z}$  tér



3.4. ábra: Altér altere is altér



3.5. ábra: Alterek metszete is altér, de az megeshet, hogy ez a metszet csak az egyetlen nullvektorból álló zérustér.



3.6. ábra: Két altér egyesítése csak akkor altér, ha egyik a másik altere

1. a zérusvektorból álló egyelemű halmaz, azaz a zérustér,
2. egy origón átmenő egyenes összes vektora,
3. a sík összes vektora.

Hasonlóképp  $\mathbb{R}^3$  alterei:

1. a zérusvektorból álló egyelemű halmaz,
2. egy origón átmenő egyenes összes vektora,
3. egy origón átmenő sík összes vektora,
4. a tér összes vektora.

Mivel egy egyetlen  $n$ -ismeretlenes egyenletből álló homogén egyenletrendszer megoldáshalmaza egy  $\mathbb{R}^n$ -beli hipersík, ezért az origón átmenő hipersíkok is alterek. Ennek általánosításaként a 3.8. állítás az altér fogalmát használva a következő alakot ölti:

3.12. ÁLLÍTÁS (MEGOLDÁSOK ALTERE). Egy  $n$ -ismeretlenes homogén lineáris egyenletrendszer megoldáshalmaza alteret alkot  $\mathbb{R}^n$ -ben.

3.13. DEFINÍCIÓ (NULLTÉR). Az  $\mathbf{A}$  együtthatómátrixú homogén lineáris egyenletrendszer megoldásainak alterét az  $\mathbf{A}$  mátrix nullterének nevezzük és  $\mathcal{N}(\mathbf{A})$ -val jelöljük.

Kérdés: általában hogyan adhatjuk meg  $\mathbb{R}^n$  egy tetszőleges alterét?

**Kifeszített altér** A homogén lineáris egyenletrendszer összes megoldását néhány vektor lineáris kombinációjaként állítottuk elő. A megoldások alterét tehát „generálja” vagy geometrikusabb szóhasználatlaltal „kifeszíti” néhány megoldásvektor.

3.14. DEFINÍCIÓ (KIFESZÍTETT ALTÉR). Adva van egy  $\mathcal{V}$  vektortér. A  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in \mathcal{V}$  vektorok

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_k\mathbf{v}_k$$

alakú lineáris kombinációinak halmazát a  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  vektorok által kifeszített altérnek nevezzük, és  $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$ -val jelöljük.

Megmutatjuk, hogy a „kifeszített altér” elnevezésben az altér szó használata jogos:

3.15. ÁLLÍTÁS (A KIFESZÍTETT ALTÉR ALTÉR). A  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in \mathcal{V}$  vektorok által kifeszített  $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$  vektorhalmaz  $\mathcal{V}$  egy altere.

**BIZONYÍTÁS.** Be kell látni, hogy  $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$  bármely vektorának skalárszorosa és bármely két vektorának összege is ide tartozik. Legyen

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_k\mathbf{v}_k, \text{ és } \mathbf{v} = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \dots + d_k\mathbf{v}_k$$

a  $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$  két tetszőleges vektora, és legyen  $x \in \mathbb{R}$  tetszőleges valós. Ekkor

$$x\mathbf{u} = (xc_1)\mathbf{v}_1 + (xc_2)\mathbf{v}_2 + \dots + (xc_k)\mathbf{v}_k \in \text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k),$$

és

$$\mathbf{u} + \mathbf{v} = (c_1 + d_1)\mathbf{v}_1 + \dots + (c_k + d_k)\mathbf{v}_k \in \text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k). \quad \square$$

3.16. PÉLDA (NULLTÉR). *Határozzuk meg az*

$$\begin{bmatrix} 1 & 2 & 1 & 2 & 1 \\ 1 & 2 & 3 & 3 & 1 \\ 3 & 6 & 7 & 8 & 3 \end{bmatrix}$$

*mátrix nullterét. Keressünk véges sok olyan vektort, melyek kifeszítik e teret!*

**MEGOLDÁS.** Az adott mátrix a 2.37. példabeli homogén lineáris egyenletrendszer együtthatómátrixa. A homogén lineáris egyenletrendszer megoldáshalmaza épp a nulltér vektoraiból álló halmaz. A nulltér vektorai tehát

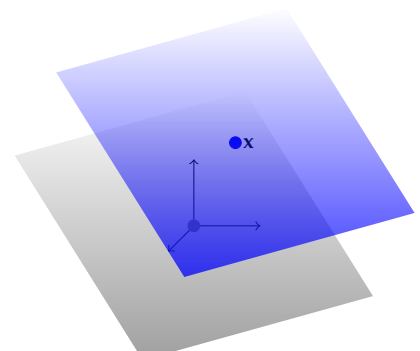
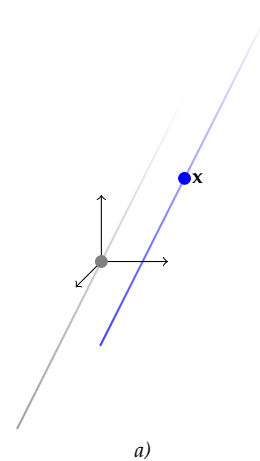
$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2s - \frac{3}{2}t - u \\ s \\ -\frac{1}{2}t \\ t \\ u \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -\frac{3}{2} \\ 0 \\ -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + u \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

alakba írhatók, amiből leolvasható, hogy e teret a megoldásban megadott három vektor feszíti ki.  $\square$

*Az inhomogén lineáris egyenletrendszer megoldásai* Az inhomogén lineáris egyenletrendszer megoldásai nem alkotnak alteret. Legegyszerűbben ez onnan látszik, hogy a zérusvektor minden alternek eleme, viszont egyetlen inhomogén egyenletrendszernek sem megoldása! Ugyanakkor az inhomogén lineáris egyenletrendszer és a hozzá tartozó homogén egyenletrendszer megoldásai közt egy igen fontos kapcsolat van.

3.17. TÉTEL (HOMOGÉN ÉS INHOMOGÉN EGYENLETRENDSZER MEGOLDÁSAI). Az inhomogén lineáris  $[\mathbf{A}|\mathbf{b}]$  mátrixú egyenletrendszer általános megoldása megkapható úgy, hogy egy partikuláris megoldásához hozzáadjuk a hozzá tartozó homogén  $[\mathbf{A}|\mathbf{0}]$  mátrixú egyenletrendszer általános megoldását.

|                                     |   |                                            |   |                                   |
|-------------------------------------|---|--------------------------------------------|---|-----------------------------------|
| inhomogén<br>általános<br>megoldása | = | inhomogén egy<br>partikuláris<br>megoldása | + | homogén<br>általános<br>megoldása |
|-------------------------------------|---|--------------------------------------------|---|-----------------------------------|



**BIZONYÍTÁS.** Jelölje az egyenletrendszer együtthatómátrixát  $\mathbf{A}$ , annak sorvektorait  $\mathbf{a}_{1*}, \mathbf{a}_{2*}, \dots, \mathbf{a}_{m*}$ , és jelölje  $\mathbf{b} = (b_1, b_2, \dots, b_m)$ , a konstans tagok vektorát. Legyen  $\mathbf{x}$  az inhomogén egyenletrendszer egy partikuláris megoldása, és jelölje  $\mathcal{H}$  a homogén,  $\mathcal{I}$  az inhomogén egyenletrendszer megoldáshalmazát. Megmutatjuk, hogy  $\mathbf{x} + \mathcal{H} = \mathcal{I}$ , ahol a bal oldali összeadást elemeenként értjük.

$\mathbf{x} + \mathcal{H} \subseteq \mathcal{I}$ : Meg kell mutatnunk, hogy ha  $\mathbf{x}$ -hez adjuk a  $\mathcal{H}$  egy tetszőleges  $\mathbf{y}$  elemét, az inhomogén egyenletrendszer egy megoldását kapjuk. Valóban,  $\mathbf{x}$ , illetve  $\mathbf{y}$  eleget tesz az

$$\begin{aligned} \mathbf{a}_{i*} \cdot \mathbf{x} &= b_i, \\ \mathbf{a}_{i*} \cdot \mathbf{y} &= 0, \quad (i = 1, 2, \dots, m) \end{aligned}$$

egyenleteknek. Ebből

$$\mathbf{a}_{i*} \cdot (\mathbf{x} + \mathbf{y}) = \mathbf{a}_{i*} \cdot \mathbf{x} + \mathbf{a}_{i*} \cdot \mathbf{y} = b_i + 0 = b_i.$$

tehát  $\mathbf{x} + \mathbf{y}$  megoldása az inhomogén egyenletrendszernek, azaz  $\mathbf{x} + \mathbf{y} \in \mathcal{I}$ .

$\mathbf{x} + \mathcal{H} \supseteq \mathcal{I}$ : Meg kell mutatnunk, hogy ha  $\mathbf{z}$  az inhomogén egy tetszőleges megoldása, azaz  $\mathbf{z} \in \mathcal{I}$ , akkor található olyan  $\mathbf{y} \in \mathcal{H}$ , hogy  $\mathbf{z} = \mathbf{x} + \mathbf{y}$ . Valóban, az  $\mathbf{y} = \mathbf{z} - \mathbf{x}$  megteszi, mert

$$\mathbf{a}_{i*} \cdot (\mathbf{z} - \mathbf{x}) = \mathbf{a}_{i*} \cdot \mathbf{z} - \mathbf{a}_{i*} \cdot \mathbf{x} = b_i - b_i = 0.$$

fennáll minden  $i = 1, 2, \dots, m$  indexre, azaz  $\mathbf{z} - \mathbf{x} \in \mathcal{H}$ . □

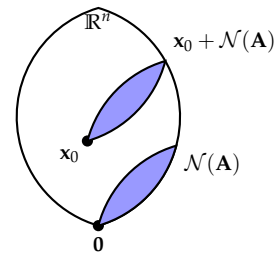
E tétel azt jelenti, hogy ugyan az inhomogén lineáris egyenletrendszer megoldásainak halmaza *nem* *altér*, de egy *altér eltoltja*. Az ilyen halmazokat geometriai nyelven *affin altereknek* is szokás nevezni. Eltolt altereket mutat a 3.7. ábra.

E tételt szemléltetik a 2.35. és a 2.37. példák is.

Ha általában szeretnénk ábrázolni az inhomogén egyenletrendszer megoldását a levéldiagramon, azt könnyen megtehetjük egy eltolt altér szemléltetésével, ahogy azt a 3.8. ábra mutatja.

Az **előző** tétel szerint az inhomogén egyenletrendszer összes megoldása a homogén összes megoldásának – azaz  $\mathcal{N}(\mathbf{A})$ -nak – az inhomogén valamelyik megoldásával való eltoltja. Fontos látnunk, hogy mindegy melyik megoldást választjuk az inhomogén megoldásai közül, bár az eltolás mértéke változik, az eredmény ugyanaz lesz. Ezt jól szemlélteti a 3.7. ábra: ha az origón átmenő egyenes origónál lévő pontját nem  $\mathbf{x}$ -be, hanem az eltolt egyenes egy másik pontjába toljuk, akkor két eltolt egyenes fedt egymást, vagyis a két affin altér azonos.

Az inhomogén  $[\mathbf{A}|\mathbf{b}]$  egyenletrendszer megoldásai nem alkotnak altér, de azok a  $\mathbf{b}$  vektorok, melyre az egyenletrendszer megoldható, igen. Ezek ugyanis az oszlopmodell szerint épp azok a vektorok, melyek az együtthatómátrix oszlopvektorainak lineáris kombinációjaként



3.8. ábra: Az  $\mathbf{A}$  együtthatómátrixú homogén egyenletrendszer megoldása a nulltér, azaz  $\mathcal{N}(\mathbf{A})$ , az inhomogéné e tér egy  $\mathbf{x}_0 + \mathcal{N}(\mathbf{A})$  eltoltja, ahol  $\mathbf{x}_0$  az inhomogén egyenletrendszer egy megoldása.



előállnak. Az ilyen vektorok pedig alteret alkotnak, mégpedig az  $\mathbf{A}$  oszlopvektorai által kifeszített alteret. Ezt nevezzük oszloptérnek.

**3.18. DEFINÍCIÓ (SORTÉR, OSZLOPTÉR).** Egy mátrix oszlopvektorai által kifeszített alteret oszloptérnek, a sorvektorai által kifeszített alteret sortérnek nevezzük.

Az  $m \times n$ -es valós  $\mathbf{A}$  mátrix sortere  $\mathbb{R}^n$  altere, oszloptere  $\mathbb{R}^m$  altere. Az  $\mathbf{A}$  sorterét  $\mathcal{S}(\mathbf{A})$ -val, oszlopterét  $\mathcal{O}(\mathbf{A})$ -val jelöljük (ld. 3.10 ábra).

Az oszlopmodellt használva az is látható, hogy egy adott vektort előállító lineáris kombináció együtthatóinak meghatározása egy lineáris egyenletrendszer megoldásával ekvivalens.

**3.19. KÖVETKEZMÉNY (INHOMOGÉN EGYENLETRENDSZER MEGOLDHATÓSÁGA).** Az  $[\mathbf{A}|\mathbf{b}]$  mátrixú egyenletrendszer pontosan akkor oldható meg, ha  $\mathbf{b}$  előáll az  $\mathbf{A}$  oszlopainak lineáris kombinációjaként, azaz  $\mathbf{b}$  benne van az  $\mathbf{A}$  oszloptérében. A lineáris kombináció együtthatói megegyeznek a megoldásvektor koordinátaival.

**3.20. PÉLDA (KIFESZÍTETT ALTÉR VEKTORAI).** Határozzuk meg, hogy a  $\mathbf{v}_1 = (1, 0, 1, 2)$ ,  $\mathbf{v}_2 = (-1, 2, -2, 1)$  és  $\mathbf{v}_3 = (1, 1, 1, 1)$  vektorok által kifeszített altérnek eleme-e az  $\mathbf{u} = (-1, 2, -3, 6)$  vektor! Ha igen, adjunk meg egy ezt bizonyító lineáris kombinációt is! Mutassuk meg, hogy a  $\mathbf{w} = (-1, 2, -3, 4)$  vektor nem eleme az altérnek!

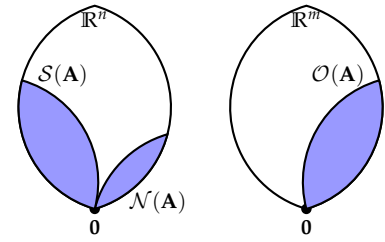
**MEGOLDÁS.** Olyan  $x_1, x_2, x_3$  valósokat keresünk, melyekkel  $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{u}$  fennáll. Ez egy négy egyenletből álló egyenletrendszerrel ekvivalens, melynek bővített mátrixa a  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$  és  $\mathbf{u}$  oszlopvektorokból áll. Az  $\mathbf{u}$  helyett  $\mathbf{w}$  vektorral egy másik hasonló egyenletrendszert is meg kell oldani. A két egyenletrendszert egyetlen bővített mátrixszal oldjuk meg, melynek oszlopvektorai  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{u}$  és  $\mathbf{w}$ . Ennek lépcsős alakja:

$$\left[ \begin{array}{ccc|cc} 1 & -1 & 1 & -1 & -1 \\ 0 & 2 & 1 & 2 & 2 \\ 1 & -2 & 1 & -3 & -3 \\ 2 & 1 & 1 & 6 & 4 \end{array} \right] \Rightarrow \left[ \begin{array}{ccc|cc} 1 & -1 & 1 & -1 & -1 \\ 0 & 1 & 0 & 2 & 2 \\ 0 & 0 & 1 & -2 & -2 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right]$$

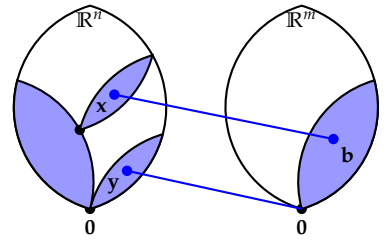
amiből  $(x_1, x_2, x_3) = (3, 2, -2)$ , és  $\mathbf{w}$  valóban nem áll elő lineáris kombinációként, mert a  $\mathbf{w}$ -t tartalmazó egyenletrendszer ellentmondásos.  $\square$

**Lineáris függetlenség és összefüggőség** A lineáris egyenletrendszerek megoldása és vektorok lineáris függetlenségével vagy összefüggőségével kapcsolatos kérdések szoros kapcsolatban vannak egymással.

Az előző 3.20. példa tanulsága úgy is összefoglalható, hogy egy  $\mathbf{w}$  vektor pontosan akkor független az  $\mathbf{A}$  mátrix oszlopvektoraitól, vagyis az  $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$  vektorrendszertől, ha az  $[\mathbf{A}|\mathbf{w}]$  egyenletrendszer nem oldható meg.



3.9. ábra: Az  $\mathbf{A}$  mátrix sortere ( $\mathcal{S}(\mathbf{A})$ ), oszloptere ( $\mathcal{O}(\mathbf{A})$ ) és nulltere ( $\mathcal{N}(\mathbf{A})$ ).



3.10. ábra: A nulltér, a sortér, és az oszloptér, valamint a homogén  $\mathbf{A}\mathbf{y} = \mathbf{0}$  és az inhomogén  $\mathbf{A}\mathbf{x} = \mathbf{b}$  egyenletrendszer egy-egy megoldása a levéldiagramban.

Egy  $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$  vektorrendszer lineáris függetlenségének eldöntéséhez meg kell oldani az

$$x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \dots + x_n \mathbf{a}_n = \mathbf{0}$$

homogén lineáris egyenletrendszert. Ha van nemtriviális megoldása, akkor a vektorrendszer lineárisan összefüggő, egyébként lineárisan független. Ez igazolja az alábbi ekvivalenciákat:

3.21. KÖVETKEZMÉNY (LINEÁRIS FÜGGETLENSÉG ELDÖNTÉSE). Tekintsük

az  $\mathbf{A} = [\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_k]$  mátrixot! Az alábbi állítások ekvivalensek:

- a) az  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k$  vektorok lineárisan függetlenek;
- b) az  $\mathbf{A}$  együtthatómátrixú homogén lineáris egyenletrendszernek a triviálison kívül nincs más megoldása;
- c) az  $\mathbf{A}$  lépcsős alakjának minden oszlopában van főelem, azaz  $r(\mathbf{A}) = k$ .

3.22. PÉLDA (VEKTOROK LINEÁRIS FÜGGETLENSÉGÉNEK ELDÖNTÉSE).

Mutassuk meg, hogy a 4-dimenziós  $(1, 2, 3, 4)$ ,  $(0, 1, 0, 1)$  és  $(1, 1, 1, 0)$  vektorok lineárisan függetlenek.

**MEGOLDÁS.** A vektorokból képzett mátrix és lépcsős alakja

$$\begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 3 & 0 & 1 \\ 4 & 1 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix},$$

ami azt mutatja, hogy a homogén lineáris egyenletrendszernek csak egyetlen megoldása van, azaz az oszlopvektorok lineárisan függetlenek.  $\square$

## Feladatok

**3.1. IGAZ – HAMIS** Melyek igazak, melyek hamisak az alábbi állítások közül?

- Ha egy 10-ismeretlenes egyenletrendszer csak 6 egyenletből áll, azaz alulhatározott, akkor végtelen sok megoldása van.
- Ha egy 10-ismeretlenes egyenletrendszer csak 6 egyenletből áll, azaz alulhatározott, akkor lehet, hogy végtelen sok megoldása van, de az is lehet, hogy csak egy.
- Ha egy 10-ismeretlenes egyenletrendszer 20 egyenletből áll, azaz túlhatározott, akkor biztosan nem oldható meg!
- Ha egy 10-ismeretlenes egyenletrendszer 20 egyenletből áll, akkor nem lehet végtelen sok megoldása.

**3.2. ALTEREK TULAJDONSÁGAI: IGAZ – HAMIS**

- $\mathbb{R}^n$  bármely három alterének metszete altér.
- Ha az  $\mathcal{U}$  altér altere a  $\mathcal{V}$  és a  $\mathcal{W}$  altérnek is, akkor altere metszetüknek is.
- Alterek egyesítése altér.
- Minden altérnek eleme a zérusvektor.
- Minden altérnek van legalább egy nemzérus vektora.

**3.3. ALTEREK ÉS EGYENLETRENDSZEREK: IGAZ – HAMIS**

- Egy lineáris egyenletrendszer megoldásai alteret alkotnak.
- Egy homogén lineáris egyenletrendszer megoldásai alteret alkotnak.
- Rögzített  $\mathbf{A}$  mátrix mellett alteret alkotnak azok a  $\mathbf{b}$  vektorok, melyekre az  $[\mathbf{A}|\mathbf{b}]$  egyenletrendszer konzisztens.
- Egy egyenletrendszer megoldásvektorainak különbségeként kapott vektorok halmaza alteret alkot.

**3.4. MEGOLDHATÓSÁG: IGAZ – HAMIS**

- Az  $[\mathbf{A}|\mathbf{b}]$  mátrixú egyenletrendszer pontosan akkor oldható meg, ha  $\mathbf{b}$  előáll  $\mathbf{A}$  oszlopainak lineáris kombinációjaként.
- Az  $[\mathbf{A}|\mathbf{b}]$  mátrixú egyenletrendszer bármely két megoldásának különbsége megoldása a homogén  $[\mathbf{A}|\mathbf{0}]$  egyenletrendszernek.
- Az  $[\mathbf{A}|\mathbf{b}]$  mátrixú egyenletrendszer bármely megoldása előáll a homogén  $[\mathbf{A}|\mathbf{0}]$  mátrixú egyenletrendszer két megoldásának különbségeként.
- Az  $[\mathbf{A}|\mathbf{b}]$  mátrixú egyenletrendszer pontosan akkor oldható meg, ha  $r(\mathbf{A}|\mathbf{b}) \leq r(\mathbf{A})$ .
- Az  $n$ -ismeretlenes  $[\mathbf{A}|\mathbf{b}]$  mátrixú egyenletrendszer pontosan akkor oldható meg egyértelműen, ha  $r(\mathbf{A}) = n$ .

**3.5.** Alteret alkotnak-e az alábbi vektorhalmazok  $\mathbb{R}^3$ -ben?

- $\{\mathbf{x} \in \mathbb{R}^3 : |\mathbf{x}| = 1\}$
- $\{(x, y, z) : x + 2y - 3z = 0\}$
- $\{(x, y, z) : x + 2y - 3z = 1\}$
- $\{(x, y, z) : x = 2t, y = t, z = 0, t \in \mathbb{R}\}$
- $\{(x, y, z) : x^2 + y^2 + z^2 = 0\}$
- $\{(x, y, z) : x^3 + y^3 + z^3 = 0\}$

**3.6.** Igazoljuk, hogy a  $\mathcal{V}$  vektortér vektorainak egy  $\mathcal{W}$  hal-

maza pontosan akkor altér  $\mathcal{V}$ -ben, ha  $\mathcal{W}$ -beli vektorok bármely lineáris kombinációja  $\mathcal{W}$ -beli.

**3.7.** Mennyi lehet az  $r(\mathbf{A}|\mathbf{b})$  rang, ha az  $[\mathbf{A}|\mathbf{b}]$  kibővített mátrixú egyenletrendszerről tudjuk, hogy

- 2-ismeretlenes és megoldásainak száma végtelen;
- inkonzisztens, és  $r(\mathbf{A}) = 4$ ;
- egyetlen megoldása van és  $\mathbf{A}$   $5 \times 3$ -as;
- inkonzisztens,  $n$ -ismeretlenes és 2 egyenletből áll.

**3.8.** Egy lineáris egyenletrendszerről tudjuk, hogy  $(1, 2, 3)$  és  $(0, 1, 3)$  is megoldásvektora. Adjuk meg további két megoldásvektorát! Mekkora lehet az együtthatómátrix rangja? És mekkora lehet e rang, ha az egyenletrendszer homogén?

**3.9. INHOMOGÉN LINEÁRIS EGYENLETRENDSZER MEGOLDÁSAI** Egy négyismeretlenes lineáris egyenletrendszer megoldását számítógéppel próbálom ellenőrizni, de más jön ki. A saját eredményem ez:

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + s \begin{bmatrix} 1 \\ 0 \\ -2 \\ -1 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \\ -3 \\ -2 \end{bmatrix},$$

a számítógépé ez:

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 2 \\ 2 \end{bmatrix} + s \begin{bmatrix} -2 \\ 1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 3 \\ -2 \\ 0 \\ 1 \end{bmatrix}.$$

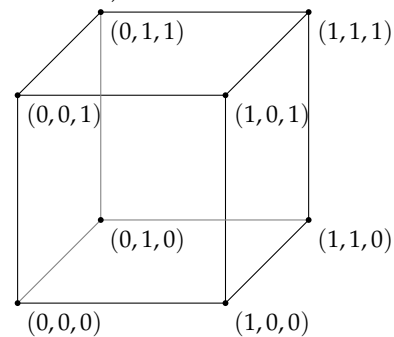
Lehet-e mindkét eredmény jó?

**3.10. MEGOLDÁSOK LINEÁRIS KOMBINÁCIÓJA** Adjunk új bizonyítást a 3.8. tételre a sormodellt használva.

**3.11. HOMOGÉN ÉS INHOMOGÉN EGYENLETRENDSZER MEGOLDÁSAI** Adjunk az oszlopmodellben megfogalmazott új bizonyítást a 3.17. tételre.

*Véges testek feletti terek*

**3.12.  $\mathbb{F}_2^3$  ALTEREI** Soroljuk fel  $\mathbb{F}_2^3$  összes alterét (ehhez segítségül hívhatjuk az alábbi ábrát, mely az  $\mathbb{F}_2^3$  vektortér vektorait szemlélteti).



## Alterek tulajdonságai és az egyenletrendszerek

E szakaszban az alterek tulajdonságait, és az egyenletrendszerek kapcsán felmerülő alterek viszonyát vizsgáljuk. Különösen fontos az együtthatómátrixhoz tartozó négy kitiüntetett altér kapcsolata.

**Sor- és oszloptér** Az előző feladatokban a homogén lineáris egyenletrendszer együtthatómátrixát lépcsős alakra hoztuk. Ebből az alakból azonban több minden leolvasható.

**3.23. TÉTEL (ELEMI SORMŰVELETEK HATÁSA A SOR- ÉS OSZLOPVEKTOROKRA).** *Elemi sorműveletek közben a sortér nem változik, az oszlopvektorok pedig olyan vektorokba transzformálódnak, melyek megőrzik az eredeti lineáris kapcsolatokat.*

**BIZONYÍTÁS.** Megmutatjuk, hogy egy mátrix sortere nem változik az elemi sorműveletek közben. A sorcserére ez nyilvánvaló. Legyenek  $\mathbf{A}$  sorvektorai  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$  és legyen  $\mathbf{u}$  a sortér egy tetszőleges vektora, ami azt jelenti, hogy vannak olyan  $c_1, c_2, \dots, c_m$  skalárok, hogy

$$\mathbf{u} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_m \mathbf{v}_m.$$

Ha egy sort (mondjuk az elsőt) beszorozzuk egy  $d \neq 0$  skalárral, akkor  $\mathbf{u}$  az új mátrix sorterében is benne van, melyet a  $d\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$  vektorok feszítenek ki, hisz

$$\mathbf{u} = \frac{c_1}{d} (d\mathbf{v}_1) + c_2 \mathbf{v}_2 + \dots + c_m \mathbf{v}_m.$$

Ez azt jelenti, hogy a sortér nem csökken, azaz minden vektor, ami eddig benne volt a sortérben, benne lesz az új mátrix sorterében is. A hozzáadás műveleténél ugyanezt tapasztaljuk: az általánosság megszorítása nélkül feltehetjük, hogy az első sor  $d$ -szeresét adjuk a második sorhoz. Ekkor az új sorteret kifeszítő vektorok:  $\mathbf{v}_1, \mathbf{v}_2 + d\mathbf{v}_1, \dots, \mathbf{v}_m$ . Az  $\mathbf{u}$  vektor e térben is benne van, ugyanis egy egyszerű átalakítás után

$$\mathbf{u} = (c_1 - c_2 d) \mathbf{v}_1 + c_2 (\mathbf{v}_2 + d\mathbf{v}_1) + \dots + c_m \mathbf{v}_m.$$

Mivel minden sorművelet inverze is egy sorművelet, az inverz sorműveletben sem csökken a sortér, ami csak úgy lehet, ha a sortér az elemi sorműveletek során változatlan marad.

Megmutatjuk, hogy az elemi sorműveletek közben az oszlopvektorok közt nem változnak a lineáris kapcsolatok. Az egyszerűség kedvéért tegyük fel, hogy például  $c\mathbf{v}_1 + d\mathbf{v}_2 = \mathbf{0}$ . Mivel a skalárral való szorzás és a vektorösszeadás is koordinátánként végezhető, könnyen látható, hogy sorcseré, egy koordináta nemnulla skalárral való szorzása, vagy az  $i$ -edik koordináta konstansszorosának a  $j$ -edikhez adása

után is fenn fog állni a két új vektor közt a fenti összefüggés. Tetszőleges számú vektor lineáris kombinációjára az állítás hasonlóan adódik.  $\square$

3.24. KÖVETKEZMÉNY (MÁTRIX LÉPCSŐS ALAKJÁNAK VEKTORAI). *Legyen  $\mathbf{B}$  az  $\mathbf{A}$  mátrix egy lépcsős alakja. Ekkor*

1.  $\mathbf{A}$  és  $\mathbf{B}$  sortere megegyezik,
2. az  $\mathbf{A}$  oszlopvektorai között lévő lineáris kapcsolatok azonosak a  $\mathbf{B}$  ugyanolyan sorszámú oszlopai közti lineáris kapcsolatokkal,
3.  $\mathbf{B}$  nemzérus sorvektorai lineárisan függetlenek,
4. a főelemek oszlopvektorai  $\mathbf{A}$ -ban és  $\mathbf{B}$ -ben is lineárisan függetlenek.

**BIZONYÍTÁS.** Az első két állítás közvetlen következménye az előző tételnek.

A harmadik állítás bizonyításához megmutatjuk, hogy egy lépcsős alak egy nemzérus sorvektora nem fejezhető ki a többi sorvektor lineáris kombinációjaként. Tekintsük a lépcsős alak  $k$ -edik sorvektorát. Főeleme legyen a  $j$ -edik oszlopban. E főelem nem állítható elő a  $k$ -nál nagyobb indexű sorok lineáris kombinációjával, mert azokban a  $j$ -edik koordináta 0. A  $k$ -nál kisebb indexű sorvektorok pedig nem szerepelhetnek a lineáris kombinációban, mivel a legkisebb indexű vektor főelemét a többi vektor nem eliminálhatja, pedig a  $k$ -edik sorban azon a helyen 0 áll.

Annak bizonyítása, hogy a főelemek oszlopai  $\mathbf{B}$ -ben lineárisan függetlenek, ugyanúgy megy, mint a sorvektorok esetén. Innen pedig az előző tétellel adódik, hogy az ilyen indexű oszlopok  $\mathbf{A}$ -ban is lineárisan függetlenek.  $\square$

► Fontos megjegyezni, hogy míg a lépcsős alak sortere megegyezik az eredeti mátrix sorterével, addig az oszloptér az elemi sorműveletek alatt megváltozik, tehát a mátrix és lépcsős alakjának oszloptere különbözik!

**Bázis** Az elemi sorműveleteket alkalmazva, egy mátrix sorterében és oszloptérében is találtunk olyan lineárisan független vektorokat, melyek kifeszítik az adott teret. Azt már az 1.9. tételben megmutattuk, hogy a háromdimenziós tér tetszőleges három lineárisan független vektorának lineáris kombinációjaként a tér minden vektora előáll. Más szavakkal ez azt jelenti, hogy a tér három lineárisan független vektora kifeszíti a teret. Az ilyen vektorhármasokat, melyeket egy koordináta-rendszer alapvektorainak vettünk, bázisnak neveztük. Ezek vezetnek a következő definícióhoz.

3.25. DEFINÍCIÓ (BÁZIS). *A  $\mathcal{V}$  vektortér bázisán olyan vektorrendszert értünk, mely*

1. lineárisan független és

2. *kifeszíti a  $\mathcal{V}$  teret (azaz generátorrendszer).*

Az  $\mathbf{e}_1 = (1, 0, \dots, 0)$ ,  $\mathbf{e}_2 = (0, 1, \dots, 0)$ , ...,  $\mathbf{e}_n = (0, 0, \dots, 0, 1)$  vektorból álló halmazt az  $\mathbb{R}^n$  vektortér standard bázisának nevezzük.

- Világos, hogy a zérustérnek nincs bázisa, hisz abban nincs egyetlen lineárisan független vektor sem.
- A standard bázis  $\mathbb{R}^n$  egy  $n$ -elemű bázisa.
- Hamarosan meg fogjuk mutatni, hogy  $\mathbb{R}^n$  minden bázisa  $n$ -elemű, és hogy bármely alterének bázisa legfeljebb  $n$ -elemű.

3.26. PÉLDA (ALTÉR BÁZISÁNAK MEGHATÁROZÁSA). *Határozzuk meg az  $(1, 1, 0, -2)$ ,  $(2, 3, 3, -2)$ ,  $(1, 2, 3, 0)$  és  $(1, 3, 6, 2)$  vektorok által kifeszített altér egy bázisát!*

**MEGOLDÁS.** *Első megoldás:* A megadott vektorokból, mint sorvektorokból képzett mátrix valamely sorlépcsős alakjának nemnulla sorai az altér egy bázisát adják:

$$\begin{bmatrix} 1 & 1 & 0 & -2 \\ 2 & 3 & 3 & -2 \\ 1 & 2 & 3 & 0 \\ 1 & 3 & 6 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 0 & -2 \\ 0 & 1 & 3 & 2 \\ 0 & 1 & 3 & 2 \\ 0 & 2 & 6 & 4 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 0 & -2 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

A bázis vektorai  $(1, 1, 0, -2)$ ,  $(0, 1, 3, 2)$ .

*Második megoldás:* Ha a bázist az adott vektorokból akarjuk kiválasztani, akkor képezzünk egy mátrixot e vektorokból, mint oszlopvektorokból. Lépcsős alakjában a főelemek oszlopai lineárisan független vektorok. A nekik megfelelő oszlopvektorok az eredeti mátrixban az oszloptér bázisát alkotják (ld. a 3.23. tételt és a 3.24. következmény 4. pontjának állítását).

$$\begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 3 & 2 & 3 \\ 0 & 3 & 3 & 6 \\ -2 & -2 & 0 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 3 & 3 & 6 \\ 0 & 2 & 2 & 4 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Tehát az adott négy vektor közül az első kettő, azaz az  $(1, 1, 0, -2)$  és  $(2, 3, 3, -2)$  vektorok bázist alkotnak.

Ha a megadott vektorokat más sorrendben írjuk a mátrixba, másik bázist kaphatunk. □

3.27. PÉLDA (VEKTOR FELÍRÁSA A BÁZISVEKTOROK LINEÁRIS KOMBINÁCIÓJAKÉNT). *Az előző feladatban megadott négy vektor mindegyikét fejezzük ki az általuk kifeszített altér bázisvektorainak lineáris kombinációjaként!*

**MEGOLDÁS.** Az előző feladat második megoldásában találtunk egy bázist a megadott vektorok közül. Mivel az oszlopvektorokkal dolgoztunk, a vektorok közti lineáris kapcsolat leolvasható bármelyik lépcsős

alakból: legkényelmesebben a *redukált* lépcsős alakból. Folytatjuk tehát az előző példabeli eliminációs lépéseket:

$$\begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 3 & 2 & 3 \\ 0 & 3 & 3 & 6 \\ -2 & -2 & 0 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & -1 & -3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \quad (3.1)$$

A redukált lépcsős alakból látjuk, hogy például a harmadik oszlop a második és az első különbsége. Ezek alapján az eredeti vektoroknak a bázisvektorok lineáris kombinációiként való felírása:

$$\begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix} = - \begin{bmatrix} 1 \\ 1 \\ 0 \\ -2 \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \\ 3 \\ -2 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 3 \\ 6 \\ 2 \end{bmatrix} = -3 \begin{bmatrix} 1 \\ 1 \\ 0 \\ -2 \end{bmatrix} + 2 \begin{bmatrix} 2 \\ 3 \\ 3 \\ -2 \end{bmatrix}.$$

Ha a megadott vektorokat más sorrendben írjuk a mátrixba, más bázisra juthatunk (ld. a 3.15. feladatot).  $\square$

**3.28. ÁLLÍTÁS (BÁZIS EKVIVALENS DEFINÍCIÓI).** Legyen  $\mathcal{V}$  egy tetszőleges vektortér, és legyen  $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\} \subseteq \mathcal{V}$  vektorok egy halmaza. A következő állítások ekvivalensek:

1.  $\mathcal{B}$  lineárisan független vektorokból áll és kifeszíti az  $\mathcal{V}$  vektorteret, azaz  $\mathcal{B}$  a  $\mathcal{V}$  egy bázisa;
2.  $\mathcal{B}$  minimális méretű halmaz, mely kifeszíti  $\mathcal{V}$ -t;
3.  $\mathcal{B}$  maximális méretű, független vektorokból álló halmaz  $\mathcal{V}$ -ben.

**BIZONYÍTÁS.** Elég belátnunk, hogy egy minimális méretű generátorrendszer független vektorokból áll, és hogy egy maximális méretű független rendszer generátor.

Legyen  $\mathcal{B}$  minimális méretű generátor. Ha nem volna független, akkor belőle kiválasztva független vektorokat, mely ugyanazt a teret generálja egy még kisebb méretű generátort kapnánk.

Legyen most  $\mathcal{B}$  egy maximális független rendszer. Ha nem volna generátor, akkor hozzávehetnénk tőle független vektort, vagyis volna nála nagyobb méretű független halmaz.  $\square$

**Vektor egy bázisra vonatkozó koordinátás alakja** A koordinátarendszer bevezetésénél ugyanazt tettük, mint itt az előző példában: minden vektor előállítható egy bázis elemeinek lineáris kombinációjaként, és e vektor koordinátás alakja erre a bázisra vonatkozóan a lineáris kombináció konstansából áll.

Egy vektortérben több bázist is vizsgálhatunk, és a vektorok koordinátás alakjai különbözhetnek a különböző bázisokban. Félreértések elkerülésére a bázis jelét a koordinátás alak indexében jelöljük. Például ha egy  $\mathbf{v}$  vektor standard bázisbeli és  $\mathcal{B}$  bázisbeli koordinátás alakjai

$(4, 3)$ , illetve  $(0, 5)$ , akkor azt írjuk, hogy

$$\mathbf{v} = (4, 3) = (0, 5)_B, \text{ vagy mátrixjelöléssel } \mathbf{v} = \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}_B.$$

Ha általában akarunk utalni – a konkrét koordináták nélkül – egy  $\mathbf{v}$  vektor  $B$  bázisbeli koordinátás alakjára, akkor a  $[\mathbf{v}]_B$  vagy a  $(\mathbf{v})_B$  alakot használjuk. Így írhatjuk azt is, hogy

$$[\mathbf{v}]_B = \begin{bmatrix} 0 \\ 5 \end{bmatrix}_B, \text{ vagy egyszerűbben, hogy } [\mathbf{v}]_B = \begin{bmatrix} 0 \\ 5 \end{bmatrix}.$$

**3.29. PÉLDA (VEKTOR KOORDINÁTÁS ALAKJA A  $B$  BÁZISBAN).** Tekintsük a 3.26. és a 3.27. példákban is szereplő  $\mathbf{v}_1 = (1, 1, 0, -2)$ ,  $\mathbf{v}_2 = (2, 3, 3, -2)$ ,  $\mathbf{v}_3 = (1, 2, 3, 0)$  és  $\mathbf{v}_4 = (1, 3, 6, 2)$  vektorok által kifeszített alteret. Ennek  $B = \{\mathbf{v}_1, \mathbf{v}_2\}$  egy bázisa. Írjuk fel a négy vektornak a bázisra vonatkozó koordinátás alakját!

**MEGOLDÁS.** Az előző példában a (3.1) képletbeli redukált lépcsős alak nemzérus soraiból álló

$$\begin{bmatrix} 1 & 0 & -1 & -3 \\ 0 & 1 & 1 & 2 \end{bmatrix}$$

mátrix azt mutatja, hogy az  $B$  bázisban e négy vektor koordinátái rendre

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}_B, \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}_B, \mathbf{v}_3 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}_B, \mathbf{v}_4 = \begin{bmatrix} -3 \\ 2 \end{bmatrix}_B.$$

Ez a 3.24. állítás 2. pontjából következik, mely szerint a redukált lépcsős alak oszlopai közti lineáris kapcsolatok megegyeznek az eredeti mátrix oszlopai közti lineáris kapcsolatokkal.  $\square$

**Dimenzió és rang** Az előzőekben bázist kerestünk egy vektortérhez. Azt tapasztaltuk, hogy a bázis mindig ugyanannyi vektorból állt. Ez nem véletlen.

**3.30. TÉTEL (BÁZIS-TÉTEL).** Ha a  $\mathcal{V}$  vektortérnek van véges sok vektorból álló bázisa, akkor bármely két bázisa azonos számú vektorból áll.

**BIZONYÍTÁS.** Tegyük fel, hogy a  $\mathcal{V}$  vektortérnek

$$B = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}, \text{ és } C = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_r\},$$

két bázisa, melyek nem ugyanannyi vektorból állnak, azaz például  $k < r$ . Mivel  $B$  bázis  $\mathcal{V}$ -ben, ezért a  $C$  bázis vektorai is kifejezhetők lineáris kombinációikként, azaz léteznek olyan  $a_{ij}$  skalárok, hogy

$$\mathbf{w}_i = a_{i1}\mathbf{v}_1 + a_{i2}\mathbf{v}_2 + \dots + a_{ik}\mathbf{v}_k, \quad (i = 1, \dots, r). \quad (3.2)$$



Mivel a  $\mathcal{C}$  bázis vektorai lineárisan függetlenek, ezért a

$$c_1 \mathbf{w}_1 + c_2 \mathbf{w}_2 + \dots + c_r \mathbf{w}_r = \mathbf{0} \quad (3.3)$$

egyenlőség csak a  $c_1 = c_2 = \dots = c_k = 0$  konstansokra áll fenn. A (3.2) egyenlőségeit a (3.3) egyenletbe helyettesítve

$$c_1(a_{11}\mathbf{v}_1 + a_{12}\mathbf{v}_2 + \dots + a_{1k}\mathbf{v}_k) + c_2(a_{21}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + \dots + a_{2k}\mathbf{v}_k) + \dots + c_r(a_{r1}\mathbf{v}_1 + a_{r2}\mathbf{v}_2 + \dots + a_{rk}\mathbf{v}_k) = \mathbf{0},$$

aminek  $\mathcal{B}$  vektorai szerinti rendezése után kapjuk, hogy

$$(a_{11}c_1 + a_{21}c_2 + \dots + a_{r1}c_r)\mathbf{v}_1 + (a_{12}c_1 + a_{22}c_2 + \dots + a_{r2}c_r)\mathbf{v}_2 + \dots + (a_{1k}c_1 + a_{2k}c_2 + \dots + a_{rk}c_r)\mathbf{v}_k = \mathbf{0}.$$

Ez azt jelenti, hogy a homogén lineáris

$$\begin{aligned} a_{11}c_1 + a_{21}c_2 + \dots + a_{r1}c_r &= 0 \\ a_{12}c_1 + a_{22}c_2 + \dots + a_{r2}c_r &= 0 \\ \vdots & \\ a_{1k}c_1 + a_{2k}c_2 + \dots + a_{rk}c_r &= 0 \end{aligned}$$

egyenletrendszernek a  $c_1 = c_2 = \dots = c_k = 0$  az egyetlen megoldása. Ez viszont a 3.6. tétel szerint nem teljesülhet, mivel a fenti homogén egyenletrendszer egyenleteinek száma kisebb ismeretlenjei számánál ( $k < r$ ). Ez az ellentmondás bizonyítja, hogy indirekt feltevésünk helytelen volt, tehát a két bázis azonos számú vektorból áll.  $\square$

Ha a háromdimenziós térben tekintünk egy origón átmenő síkot, látjuk, hogy bármely két független vektora kifeszíti. Azaz minden bázisa kételemű. Ha egy origón átmenő egyenest tekintünk, azt minden nemnulla vektora, mint egyelemű bázisa, kifeszíti. A hétköznapi életben is használt fogalom, a dimenzió, megegyezik a bázis elemszámával. Miután egy altér minden bázisa azonos elemszámú, értelmes a következő definíció:

**3.31. DEFINÍCIÓ (DIMENZIÓ).** Ha a  $\mathcal{V}$  vektortérnek van véges bázisa, akkor dimenzióján egy bázisának elemszámát értjük, és e számot  $\dim \mathcal{V}$ -vel jelöljük.

Az  $\mathbb{R}^n$  altere saját magának, és standard bázisa épp  $n$  vektorból áll, így  $\dim \mathbb{R}^n = n$ . A nullvektorból álló altérben nincsenek független vektorok, így dimenziója 0.

Az  $m \times n$ -es  $\mathbf{A}$  mátrix *transzponáltján* azt az  $\mathbf{A}^T$ -vel jelölt  $n \times m$ -es mátrixot értjük, amelyet az  $\mathbf{A}$  sorainak és oszlopainak felcserélésével kapunk. Azaz

$$\mathbf{A}^T = [a_{ij}]^T := [a_{ji}].$$

3.32. PÉLDA (MÁTRIX TRANSZPONÁLTJA). Az

$$\mathbf{A} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 5 & 6 \end{bmatrix}$$

mátrixok transzponáltja

$$\mathbf{A}^T = \begin{bmatrix} 1 & 2 \end{bmatrix}, \quad \mathbf{B}^T = \begin{bmatrix} 0 & 2 \\ 1 & 3 \end{bmatrix}, \quad \mathbf{C}^T = \begin{bmatrix} 0 & 3 \\ 1 & 4 \\ 2 & 5 \end{bmatrix}, \quad \mathbf{D}^T = \begin{bmatrix} 5 \\ 6 \end{bmatrix}.$$

Adott véges sok  $\mathbb{R}^n$ -beli vektor által kifeszített altér dimenzióját úgy határozhatjuk meg, hogy meghatározzuk a vektorokból képzett mátrix rangját. Igaz ugyanis a következő állítás:

3.33. ÁLLÍTÁS (DIMENZIÓ = RANG). Egy mátrix rangja, sorterének dimenziója és oszlopterének dimenziója megegyezik. Ebből következőleg  $r(\mathbf{A}) = r(\mathbf{A}^T)$ .

**BIZONYÍTÁS.** A mátrix rangja megegyezik a lépcsős alakjában lévő nem-zérus sorainak számával. A 3.24. tétel szerint viszont e sorok lineárisan függetlenek és kifeszítik a sorteret, tehát bázist alkotnak, így számuk a sortér dimenzióját adja. Az oszloptérről láttuk, hogy a főelemeknek megfelelő oszlopok az eredeti mátrixban lineárisan függetlenek és kifeszítik az oszlopteret, tehát e tér dimenziója is a mátrix rangjával egyezik meg. Az utolsó állítás abból következik, hogy  $\mathbf{A}$  sortere megegyezik  $\mathbf{A}^T$  oszlopterével.  $\square$

3.34. DEFINÍCIÓ (VEKTORRENDSZER RANGJA). Egy  $\mathbb{R}^n$ -beli vektorokból álló vektorrendszer rangján a vektorokból képzett mátrix rangját, vagy ami ezzel egyenlő, az általuk kifeszített altér dimenzióját értjük.

3.35. PÉLDA (DIMENZIÓ KISZÁMÍTÁSA). Határozzuk meg az  $\mathbf{A}$  mátrix sorterének és nullterének dimenzióját!

$$\mathbf{A} = \begin{bmatrix} 3 & 3 & 3 & 3 & 3 \\ 3 & 4 & 5 & 4 & 3 \\ 3 & 2 & 1 & 2 & 3 \\ 3 & 3 & 3 & 3 & 3 \end{bmatrix}$$

**MEGOLDÁS.** Az  $\mathbf{A}$  redukált lépcsős alakja

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & -1 & 0 & 1 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Innen leolvasható, hogy a mátrix rangja 2, így sortérének dimenziója is 2. A nulltér dimenziója megegyezik az egyenletrendszer megoldásterének dimenziójával, ami megegyezik a szabad változók számával, esetünkben ez 3. Vegyük észre, hogy a sortér és a nulltér dimenziójának összege megegyezik a változók számával, azaz a mátrix oszlopainak számával, jelen példában 5-tel.  $\square$

**3.36. TÉTEL (DIMENZIÓTÉTEL).** Bármely  $\mathbf{A} \in \mathbb{R}^{m \times n}$  mátrix esetén a sortér dimenziójának és a nulltér dimenziójának összege  $n$ . Képlettel:

$$\dim(\mathcal{S}(\mathbf{A})) + \dim(\mathcal{N}(\mathbf{A})) = n.$$

**BIZONYÍTÁS.** A mátrix sortérének dimenziója megegyezik a mátrix rangjával, azaz az  $[\mathbf{A}|\mathbf{0}]$  mátrixú egyenletrendszerben a kötött változók számával. Megmutatjuk, hogy a nulltér dimenziója megegyezik a szabad változók számával, így e két szám összege valóban  $n$ , ami bizonyítja az állítást (ld. még a 3.4. állítást).

Elég tehát megmutatnunk, hogy egy homogén lineáris egyenletrendszer redukált lépcsős alakkal előállított megoldásában a szabad változók száma megegyezik a nulltérből kiválasztható bázis elemszámával. Először lássunk egy ilyen megoldást konkrétan. Például a 2.37. példabeli homogén lineáris egyenletrendszer megoldása

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2s - \frac{3}{2}t - u \\ s \\ -\frac{1}{2}t \\ t \\ u \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -\frac{3}{2} \\ 0 \\ -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + u \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

ahol  $x_2 = s$ ,  $x_4 = t$  és  $x_5 = u$  a három szabad változó. A nullteret kifesztítő három vektor közül az elsőben  $x_2 = 1$ , de az összes többiben  $x_2 = 0$ , így az első vektor független a többitől. Hasonlóképp általában is igaz, hogy a redukált lépcsős alakból való származtatás következtében a nullteret kifesztítő minden megoldásvektorban az összes szabad változóhoz tartozó koordináta 0, azt az egy koordinátát kivéve, amelyikhez a vektor tartozik. Így viszont mindegyik vektor független a többitől, vagyis e vektorok függetlenek, és mivel kifesztítik a nullteret, számuk megadja a nulltér dimenzióját.  $\square$

**A sortér és a nulltér merőlegessége** A valós  $\mathbf{A}$  mátrix sorvektorai érdekes tulajdonsággal rendelkeznek. Egy homogén lineáris egyenletrendszer minden egyenletének bal oldala az együtthatómátrix egy sorvektorának és a megoldásvektornak a skaláris szorzata. Mivel e szorzat értéke 0, ezért e két vektor merőleges egymásra.

3.37. PÉLDA (VEKTOROKRA MERŐLEGES ALTÉR). *Határozzuk meg az összes olyan vektort  $\mathbb{R}^4$ -ben, mely merőleges a  $\mathbf{v}_1 = (1, 0, 1, 2)$  és  $\mathbf{v}_2 = (-1, 2, -2, 1)$  vektorok mindegyikére!*

**MEGOLDÁS.** Olyan  $\mathbf{x}$  vektort keresünk, melyre  $\mathbf{v}_1 \cdot \mathbf{x} = 0$  és  $\mathbf{v}_2 \cdot \mathbf{x} = 0$ . Ezt koordinátákkal felírva két egyenletet kapunk, melynek együtthatómátrixa és annak egy lépcsős alakja a következő:

$$\begin{bmatrix} 1 & 0 & 1 & 2 \\ -1 & 2 & -2 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 2 & -1 & 3 \end{bmatrix},$$

amiből  $\mathbf{x} = (-s - 2t, (s - 3t)/2, s, t)$ , azaz

$$\mathbf{x} = s \begin{bmatrix} -1 \\ 1/2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -2 \\ -3/2 \\ 0 \\ 1 \end{bmatrix}.$$

A megoldás tehát két vektor által kifeszített altér összes vektora.

Másként fogalmazva, e feladatban meghatároztuk az összes olyan vektort, mely egy mátrix sorvektorainak mindegyikére merőleges.  $\square$

Az  $m \times n$ -es  $\mathbf{A}$  együtthatómátrixú homogén lineáris egyenletrendszer  $i$ -edik egyenletének alakja

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = 0, \text{ azaz } \mathbf{a}_{i*} \cdot \mathbf{x} = 0.$$

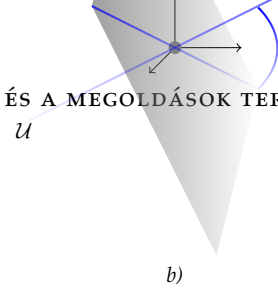
Eszerint a homogén lineáris egyenletrendszer minden megoldása merőleges az  $\mathbf{A}$  mátrix minden sorvektorára. A merőlegesség azonban a sortér összes vektorára is fennáll a következő állítás következtében.

3.38. ÁLLÍTÁS (A SORTÉR ÉS A NULLTÉR MERŐLEGESSÉGE). *A valós  $\mathbf{A}$  mátrix sortérének bármely  $\mathbf{s}$  vektora és nullterének tetszőleges  $\mathbf{x}$  vektora merőleges egymásra, azaz  $\mathbf{s} \cdot \mathbf{x} = 0$ .*

**BIZONYÍTÁS.** A sortér minden vektora az  $\mathbf{A}$  sorvektorainak valamely  $c_1, \dots, c_m$  skalárokkal vett lineáris kombinációja. Ezt felhasználva

$$\begin{aligned} \mathbf{s} \cdot \mathbf{x} &= (c_1 \mathbf{a}_{1*} + c_2 \mathbf{a}_{2*} + \cdots + c_m \mathbf{a}_{m*}) \cdot \mathbf{x} \\ &= c_1 \mathbf{a}_{1*} \cdot \mathbf{x} + c_2 \mathbf{a}_{2*} \cdot \mathbf{x} + \cdots + c_m \mathbf{a}_{m*} \cdot \mathbf{x} \\ &= c_1 0 + c_2 0 + \cdots + c_m 0 = 0. \end{aligned} \quad \square$$

Ez az állítás a következő definíciókra vezet: egy vektortér két altere *merőleges*, ha bárhogy választva egy vektort az egyik altérből, és egy másikat a másik altérből, azok merőlegesek egymásra. Eszerint az előző állítás úgy fogalmazható meg, hogy bármely valós mátrix sortere és nulltere merőleges egymásra. Ennél azonban több is igaz, a nulltér nem csak merőleges a sortérre, de az összes olyan vektort tartalmazza,



b)

3.11. ábra: a) Két merőleges alter  $U$  és  $V$ ; b) Egy alter és merőleges kiegészítő altere:  $U$  és  $U^\perp$ .

mely merőleges a sortérre. Az  $\mathbb{R}^n$  egy  $W$  alterére merőleges vektorok alterét a  $W$  merőleges kiegészítő alterének nevezzük és  $W^\perp$ -vel jelöljük (amit olvashatunk „ $W$  perp”-nek).

A két fogalom közti különbséget mutatja a 3.11. ábra, melynek a) része a 3-dimenziós tér két, egymásra merőleges 1-dimenziós  $U$  és  $V$  alterét szemlélteti, míg a b) rész az  $U$  alteret, és annak 2-dimenziós merőleges kiegészítő  $U^\perp$  alterét szemlélteti.

Vegyük az  $A$  mátrix transzponáltját! Az  $A^T$  együtthatómátrixú homogén lineáris egyenletrendszer megoldásai merőlegesek  $A^T$  sorvektoraira, azaz az  $A$  oszlopvektoraira. E két-két alter merőlegességét szemlélteti a 3.12. ábra. E négy alter igen fontos lesz a továbbiakban is, ezért ezeket a mátrixhoz tartozó négy kitüntetett alternek vagy négy fundamentális alternek nevezzük. Ezek tehát  $S(A)$ ,  $N(A)$ ,  $O(A) = S(A^T)$ ,  $N(A^T)$ .

**A lineáris algebra alaptétele** A nulltér az összes olyan vektort tartalmazza, mely merőleges a sorvektorokra, azaz a sortérre. Vajon a nulltér minden vektorára merőleges vektorok egybeesnek a sortér vektorai-val? A válasz igen, melyet a lineáris algebra alaptétele ad meg.

3.39. TÉTEL (A LINEÁRIS ALGEBRA ALAPTÉTELE). Minden valós mátrix sortere és nulltere merőleges kiegészítő alterei egymásnak.

**BIZONYÍTÁS.** Láttuk, hogy a sortér merőleges kiegészítő altere a nulltér. Később látni fogjuk, hogy általánosan is igaz az, hogy bármely  $V$  alterre  $(V^\perp)^\perp = V$  (ld. 7.38. tétel). Ezt használva kész is a bizonyítás. Csak eddigi tudásunkra építő megoldást találunk a feladatok között (ld. 3.17. feladat)!  $\square$

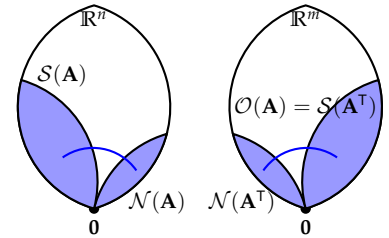
► A tétel állítása képletben kifejezve azt mondja, hogy  $S(A)^\perp = N(A)$ , ami egyúttal azt is jelenti, hogy  $N(A)^\perp = S(A)$ .

► Ha a lineáris algebra alaptételét az  $A$  mátrix transzponáltjára alkalmazzuk, és fölhasználjuk, hogy  $A$  oszloptere megegyezik  $A^T$  sortérével, akkor azt kapjuk, hogy  $O(A)^\perp = N(A^T)$ .

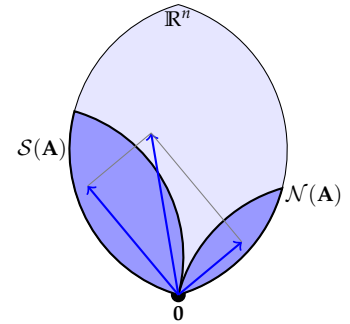
► A tételben bizonyítottuk, hogy az  $\{s_1, s_2, \dots, s_k\}$  és az  $\{e_1, e_2, \dots, e_{n-k}\}$  együtt a tér bázisát adják. Mivel minden  $x$  vektor egyértelműen áll elő  $e$  vektorok lineáris kombinációjaként, egyúttal egyértelmű az  $x$  vektornak egy sortérbe és egy nulltérbe eső vektor összegére való bontása is:

$$x = \underbrace{c_1 s_1 + \dots + c_r s_r}_c + \underbrace{d_1 e_1 + \dots + d_{n-r} e_{n-r}}_d.$$

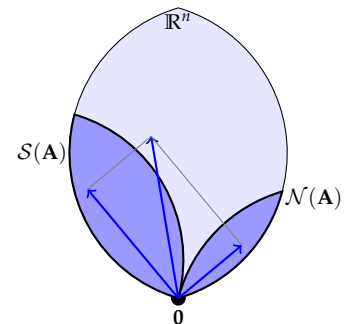
Az előző megjegyzések és a lineáris algebra alaptételének következménye az alábbi tétel:



3.12. ábra: Az  $A$  mátrix sortere merőleges nullterére, oszloptere az  $A^T$  nullterére. A berajzolt két ív az alterek merőlegességét jelöli.



3.13. ábra: A lineáris algebra alaptétele: az  $A$  mátrix sortere és nulltere merőleges kiegészítő alterek. Eszerint bármely vektor, ami merőleges a nulltérre, a sortérben van. Ennek az is következménye, hogy  $\mathbb{R}^n$  bármely vektora egyértelműen felbomlik egy sortérbe és egy nulltérbe eső vektor összegére.



3.14. ábra: A lineáris algebra alaptétele: az  $A$  mátrix sortere és nulltere merőleges kiegészítő alterek. Eszerint a sortér bármely vektora merőleges a nulltér bármely vektorára, és  $\mathbb{R}^n$  bármely vektora egyértelműen felbomlik egy sortérbe és egy nulltérbe eső vektor összegére.

3.40. TÉTEL (A NÉGY KITÜNTETETT ALTÉR). Tekintsük az  $\mathbf{A} \in \mathbb{R}^{m \times n}$  mátrixot. Akkor a következő állítások teljesülnek:

- a)  $\mathcal{S}(\mathbf{A})^\perp = \mathcal{N}(\mathbf{A})$ ,  $\mathcal{O}(\mathbf{A})^\perp = \mathcal{N}(\mathbf{A}^\top)$ .
- b)  $\mathbb{R}^n$  minden vektora egyértelműen felbomlik egy  $\mathcal{S}(\mathbf{A})$ - és egy  $\mathcal{N}(\mathbf{A})$ -beli vektor összegére,
- c)  $\mathbb{R}^m$  minden vektora egyértelműen felbomlik egy  $\mathcal{O}(\mathbf{A})$ - és egy  $\mathcal{N}(\mathbf{A}^\top)$ -beli vektor összegére.

**A lineáris egyenletrendszer megoldásainak jellemzése** E fejezet végén eljuttunk oda, hogy nagyon szép leírását tudjuk adni a lineáris egyenletrendszerek megoldásainak.

3.41. TÉTEL (LINEÁRIS EGYENLETRENDSZER MEGOLDÁSAI). Minden valószínűleg együtthatós megoldható (konzisztens) lineáris egyenletrendszerre igazak a következő állítások:

- a) egyetlen megoldása esik az együtthatómátrix sortérbe;
- b) a sortérbe eső megoldás az összes megoldás közül a legkisebb abszolút értékű;
- c) az összes megoldás előáll úgy, hogy a sortérbe eső megoldáshoz hozzáadjuk a homogén rész összes megoldását.

**BIZONYÍTÁS.** A tétel a homogén lineáris egyenletrendszerekre semmitmondó, hisz ekkor a megoldások a nullteret adják, és mivel annak metszete a sortérrel csak a nullvektorból áll, ezért csak a nullvektor esik a sortérbe, mely természetesen a legkisebb abszolút értékű megoldás. Ráadásul a nullvektort hozzáadva a nulltérhez, valóban a nullteret kapjuk, vagyis az összes megoldások terét. Így ezután csak az inhomogén esettel foglalkozunk.

a) Tegyük fel, hogy  $\mathbf{x}_1$  és  $\mathbf{x}_2$  két megoldása az  $[\mathbf{A}|\mathbf{b}]$  mátrixú egyenletrendszernek, és mindkettő a sortérbe esik. Az  $i$ -edik egyenlet alakja  $\mathbf{a}_{i*} \cdot \mathbf{x} = b_i$ , így  $\mathbf{a}_{i*} \cdot \mathbf{x}_1 = b_i$  és  $\mathbf{a}_{i*} \cdot \mathbf{x}_2 = b_i$  is fennáll minden  $i = 1, 2, \dots, m$  értékre. A két megoldás különbsége is a sortérbe esik, hisz sortérbeli vektorok lineáris kombinációja a sortérbe esik. Ekkor viszont minden  $i$  esetén

$$\mathbf{a}_{i*} \cdot (\mathbf{x}_1 - \mathbf{x}_2) = b_i - b_i = 0,$$

vagyis  $\mathbf{x}_1 - \mathbf{x}_2$  megoldása a homogén egyenletrendszernek, tehát a nulltérbe esik. Annak metszete a sortérrel csak a nullvektort tartalmazza, így  $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0}$ , vagyis  $\mathbf{x}_1 = \mathbf{x}_2$ .

Meg kell még mutatnunk, hogy mindig van a sortérbe eső megoldás. Legyen  $\mathbf{x}$  egy tetszőleges megoldás, és tekintsük az egyértelműen létező felbontását egy sortérbeli és egy nulltérbeli vektor összegére, azaz legyen

$$\mathbf{x} = \mathbf{x}_S + \mathbf{x}_N.$$

E megoldásvektort beírva az  $i$ -edik egyenletbe kapjuk, hogy

$$b_i = \mathbf{a}_{i*} \cdot \mathbf{x} = \mathbf{a}_{i*} \cdot (\mathbf{x}_S + \mathbf{x}_N) = \mathbf{a}_{i*} \cdot \mathbf{x}_S + \mathbf{a}_{i*} \cdot \mathbf{x}_N = \mathbf{a}_{i*} \cdot \mathbf{x}_S.$$

Ez azt jelenti, hogy bármely megoldás sortérbeli összetevője is megoldása az egyenletrendszernek! Találtunk tehát egy megoldást a sortérben. Egyúttal azt is beláttuk, hogy az összes megoldás e sortérbeli megoldás és a homogén egy megoldásának összege. Az előző egyenlőségekből az is kiolvasható, hogy az  $\mathbf{x}_S$  megoldáshoz bármely nulltérbeli vektort adva az egyenletrendszer egy megoldását kapjuk, igazoltuk tehát a  $c$ ) állítást is.

A sortér és a nulltér merőlegessége miatt az  $\mathbf{x} = \mathbf{x}_S + \mathbf{x}_N$  felbontás vektorai merőlegesek, azaz  $\mathbf{x}_S \perp \mathbf{x}_N$ . Használhatjuk tehát Pithagorász-tételét:

$$\mathbf{x}^2 = \mathbf{x}_S^2 + \mathbf{x}_N^2 \geq \mathbf{x}_S^2, \text{ azaz } |\mathbf{x}| \geq |\mathbf{x}_S|.$$

Így tehát minden megoldás abszolút értéke nagyobb vagy egyenlő a sortérbeli megoldás abszolút értékénél, ami bizonyítja a  $b$ ) állítást is.  $\square$

A sortérbe eső egyetlen megoldás létezése azt sugallja, hogy minden megoldható egyenletrendszer további egyenletek hozzávételével kiegészíthető olyan egyenletrendszerré, melynek már csak egyetlen megoldása van. Ez valóban igaz.

**3.42. PÉLDA (LINEÁRIS EGYENLETRENDSZER SORTÉRBE ESŐ MEGOLDÁSA).** Határozzuk meg az

$$\begin{aligned} x + y + z + 3u + 2w &= 4 \\ x + 2y + z + 5u + 2w &= 5 \\ 2x + 3y + z + 8u + 3w &= 7 \\ 2x + 3y + 2z + 8u + 4w &= 9 \end{aligned}$$

egyenletrendszer minimális abszolút értékű megoldását! Adjunk az egyenletrendszerhez olyan további egyenlet(ek)et, hogy az így kapott egyenletrendszernek csak ez legyen az egyetlen megoldása!

**MEGOLDÁS.** Először oldjuk meg az egyenletrendszert! A bővített mátrixból annak redukált lépcsős alakja könnyen adódik:

$$\begin{bmatrix} 1 & 1 & 1 & 3 & 2 & 4 \\ 1 & 2 & 1 & 5 & 2 & 5 \\ 2 & 3 & 1 & 8 & 3 & 7 \\ 2 & 3 & 2 & 8 & 4 & 9 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 2 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 2 \end{bmatrix}$$

Így a megoldás:

$$(x, y, z, u, w) = (1, 1, 2, 0, 0) + (-1, -2, 0, 1, 0)u + (-1, 0, -1, 0, 1)w.$$

Mivel a sortér merőleges a nulltérre, és mi egy sortérbe eső megoldást keresünk, ezért e megoldásnak merőlegesnek kell lennie a nullteret kifeszítő vektorokra, vagyis a  $(-1, -2, 0, 1, 0)$  és a  $(-1, 0, -1, 0, 1)$  vektorra. Így a következő két egyenletet kell az eredeti egyenletrendszerhez, vagy az egyszerűség kedvéért inkább a redukált lépcsős alak szerinti egyenletrendszerhez adni:

$$\begin{aligned} -x - 2y + u &= 0 \\ -x - z + w &= 0 \end{aligned}$$

Így a kiegészített egyenletrendszer bővített mátrixa és annak redukált lépcsős alakja

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 2 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 2 \\ -1 & -2 & 0 & 1 & 0 & 0 \\ -1 & 0 & -1 & 0 & 1 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & -4/17 \\ 0 & 1 & 0 & 0 & 0 & 5/17 \\ 0 & 0 & 1 & 0 & 0 & 19/17 \\ 0 & 0 & 0 & 1 & 0 & 6/17 \\ 0 & 0 & 0 & 0 & 1 & 15/17 \end{bmatrix},$$

tehát a keresett megoldás  $(-4/17, 5/17, 19/17, 6/17, 15/17)$ .  $\square$

**Elemi bázistranszformáció\*** Az előző paragrafusokban azt láttuk, hogy az elemi sorműveletek eredményeként az eredeti mátrix oszlopainak egy másik bázisban felírt koordinátás alakját kapjuk meg. Ez adja az ötletet ahhoz, hogy az altérből választott bázisok segítségével szemléltessük, és értsük meg mi történik akkor, amikor a mátrix egy oszlopában főelemet (pivotalemet) választunk, és oszlopának többi elemét elimináljuk. Így egy másik megközelítéshez jutunk, melyet a továbbiakban nem használunk, ezért e paragrafus átugorható.

A folyamat lényege egy kétoszlopos mátrixon is jól szemléltethető. Az egyszerűség kedvéért a két oszlop legyen az **a** és **b** vektor, a bázis, melyben e vektorok meg vannak adva, legyen a standard bázis. Tegyük fel, hogy  $a_i \neq 0$ . Ekkor az  $a_i$  pozícióját választva, a kiküszöbölés eredményeként a következőket kapjuk.

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_i & b_i \\ \vdots & \vdots \\ a_m & b_m \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & b_1 - \frac{b_i}{a_i} a_1 \\ 0 & b_2 - \frac{b_i}{a_i} a_2 \\ \vdots & \vdots \\ 1 & \frac{b_i}{a_i} \\ \vdots & \vdots \\ 0 & b_m - \frac{b_i}{a_i} a_m \end{bmatrix}$$

Tudjuk, az oszlopok az elemi sorműveletek után az eredeti vektorokat adják egy másik bázisban. Az **a** vektor nyilvánvalóan egy olyan bázisban lett felírva, amelyben szerepel az **a** vektor is, mégpedig ez az



$i$ -edik bázisvektor. Megmutatjuk, hogy mindkét oszlop az

$$\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{i-1}, \mathbf{a}, \mathbf{e}_{i+1}, \dots, \mathbf{e}_m$$

bázisban lett felírva. Az  $\mathbf{a}$  vektorra ez nyilván igaz. Nézzük a  $\mathbf{b}$  vektort! Fejezzük ki az  $\mathbf{e}_i$  vektort az  $\mathbf{a} = a_1\mathbf{e}_1 + \dots + a_i\mathbf{e}_i + \dots + a_m\mathbf{e}_m$  felírásból:

$$\mathbf{e}_i = -\frac{1}{a_i}a_1\mathbf{e}_1 - \frac{1}{a_i}a_2\mathbf{e}_2 - \dots + \frac{1}{a_i}\mathbf{a} - \dots - \frac{1}{a_i}a_m\mathbf{e}_m.$$

Ezt behelyettesítjük a  $\mathbf{b} = b_1\mathbf{e}_1 + \dots + b_i\mathbf{e}_i + \dots + b_m\mathbf{e}_m$  kifejezésbe:

$$\mathbf{b} = (b_1 - \frac{b_i}{a_i}a_1)\mathbf{e}_1 + (b_2 - \frac{b_i}{a_i}a_2)\mathbf{e}_2 + \dots + \frac{b_i}{a_i}\mathbf{a} + \dots + (b_m - \frac{b_i}{a_i}a_m)\mathbf{e}_m.$$

Tehát valóban, a  $\mathbf{b}$  koordinátás alakja a módosított bázisban épp az, amit az eredeti mátrix eliminálása után kaptunk a második oszlopban. Az imént tárgyalt lépést elemi bázistranszformációnak nevezzük, mert egy másik bázisra való áttérés egy elemi lépésének tekintjük, amikor egyetlen bázisvektort cserélünk ki. A történetek szemléltetésére a mátrixot fejléccel együtt egy táblázatba írjuk, a sorok elé az  $\mathbf{e}_1, \dots, \mathbf{e}_m$  bázisvektorok, az oszlopok fölé az oszlopvektorok neve kerül.

$$\begin{array}{c|cc} & \mathbf{a} & \mathbf{b} \\ \hline \mathbf{e}_1 & a_1 & b_1 \\ \mathbf{e}_2 & a_2 & b_2 \\ \vdots & \vdots & \vdots \\ \mathbf{e}_i & a_i & b_i \\ \vdots & \vdots & \vdots \\ \mathbf{e}_m & a_m & b_m \end{array} \Rightarrow \begin{array}{c|cc} & \mathbf{a} & \mathbf{b} \\ \hline \mathbf{e}_1 & 0 & b_1 - \frac{b_i}{a_i}a_1 \\ \mathbf{e}_2 & 0 & b_2 - \frac{b_i}{a_i}a_2 \\ \vdots & \vdots & \vdots \\ \mathbf{a} & 1 & \frac{b_i}{a_i} \\ \vdots & \vdots & \vdots \\ \mathbf{e}_m & 0 & b_m - \frac{b_i}{a_i}a_m \end{array}$$

Összefoglalva és egyúttal általánosabban megfogalmazva a fentieket:

**3.43. TÉTEL (ELEMI BÁZISTRANSZFORMÁCIÓ).** Tegyük fel, hogy az  $\mathbf{a}$  vektor  $E = \{\mathbf{e}_1, \dots, \mathbf{e}_m\}$  bázisra vonatkozó  $i$ -edik koordinátája  $a_i \neq 0$ . Ekkor az  $E$  által generált  $\mathcal{E}$  altérnek az

$$\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{i-1}, \mathbf{a}, \mathbf{e}_{i+1}, \dots, \mathbf{e}_m$$

vektorok is bázisát alkotják. Az  $\mathcal{E}$  egy tetszőleges  $\mathbf{b}$  vektorának koordinátás alakja megkapható a bázisban elemi sorműveletekkel, ha  $a_i$ -t választjuk főelemnek.

Az elemi bázistranszformáció alkalmas arra, hogy a bázisok változásán keresztül egy más nézőpontból világítsa meg a redukált lépcsős alakra hozással megoldható feladatokat. Példaként vizsgáljuk meg, mi történik egy egyenletrendszer megoldásakor. Megjegyezzük, hogy itt nincs szükség sorcserére, mert egy oszlopból szabadon választhatunk olyan sort, amelynek fejlécében még az eredeti bázisvektor szerepel.

3.44. PÉLDA (EGYENLETRENDSZER MEGOLDÁSA ELEMI BÁZISTRANSZFORMÁCIÓVAL). Oldjuk meg a 2.34. és a 2.42. példában megoldott egyenletrendszert elemi bázistranszformációval.

**MEGOLDÁS.** A táblázatokat egybefűzzük, a sorok fejlécein mindig jelezzük az aktuális bázist, az oszlopok fejléceit a jobb érthetőség végett mindig kiírjuk, a kiválasztott főelemeket külön jelöljük:

|       | $a_1$ | $a_2$ | $a_3$ | $b$ |       | $a_1$ | $a_2$ | $a_3$ | $b$ |       | $a_1$ | $a_2$ | $a_3$ | $b$ |       | $a_1$ | $a_2$ | $a_3$ | $b$ |
|-------|-------|-------|-------|-----|-------|-------|-------|-------|-----|-------|-------|-------|-------|-----|-------|-------|-------|-------|-----|
| $e_1$ | 1     | 1     | 2     | 0   | $a_1$ | 1     | 1     | 2     | 0   | $a_1$ | 1     | 0     | 3     | -5  | $a_1$ | 1     | 0     | 0     | 1   |
| $e_2$ | 2     | 2     | 3     | 2   | $e_2$ | 0     | 0     | -1    | 2   | $e_2$ | 0     | 0     | -1    | 2   | $e_2$ | 0     | 0     | 0     | 0   |
| $e_3$ | 1     | 3     | 3     | 4   | $e_3$ | 0     | 2     | 1     | 4   | $e_3$ | 0     | 0     | 3     | -6  | $a_3$ | 0     | 0     | 1     | -2  |
| $e_4$ | 1     | 2     | 1     | 5   | $e_4$ | 0     | 1     | -1    | 5   | $a_2$ | 0     | 1     | -1    | 5   | $a_2$ | 0     | 1     | 0     | 3   |

A táblázaton kicsit lehet egyszerűsíteni, azt az oszlopot, amelyben már csak egy standard egységvektor van, felesleges kiírni, az oszlopok és a sorok fejléceibe pedig elég csak azt a változót írni, amelyik a bázisba vett oszlopvektorhoz tartozik. Így a következőt kapjuk:

| $x$ | $y$ | $z$ | $b$ |     | $y$ | $z$ | $b$ |     | $z$ | $b$ |     | $b$ |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1   | 1   | 2   | 0   | $x$ | 1   | 2   | 0   | $x$ | 3   | -5  | $x$ | 1   |
| 2   | 2   | 3   | 2   |     | 0   | -1  | 2   |     | -1  | 2   |     | 0   |
| 1   | 3   | 3   | 4   |     | 2   | 1   | 4   |     | 3   | -6  | $z$ | -2  |
| 1   | 2   | 1   | 5   |     | 1   | -1  | 5   | $y$ | -1  | 5   | $y$ | 3   |

Az egyenletrendszer megoldása tehát  $x = 1, y = 3, z = -2$ . □

## Feladatok

### 3.13. BÁZIS: IGAZ – HAMIS

- a) A  $\mathcal{V}$  altérben a  $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$  vektorrendszer bázis, ha tetszőleges  $\mathbf{v} \in \mathcal{V}$  vektor egyértelműen felírható e vektorok lineáris kombinációjaként.
- b) Van olyan altér, melynek bármely nemnulla vektora bázis.
- c) Van olyan altér, melynek van kételemű bázisa, és van ettől különböző három lineárisan független vektora.

3.14. Egy lineáris egyenletrendszerrel tudjuk, hogy együtthatómátrixának rangja 2, és hogy  $(1, 2, 3)$  és  $(0, 1, 3)$  is megoldásvektora. Adjuk meg az összes megoldását!

3.15. Határozzuk meg a 3.27. példabeli vektorok által kifeszített altér egy másik bázisát úgy, hogy a vektorokat más sorrendben írjuk a mátrixba. Legyen például a sorrend  $\mathbf{w}_1 = (1, 1, 2, 1)$ ,  $\mathbf{w}_2 = (3, 1, 3, 2)$ ,  $\mathbf{w}_3 = (6, 0, 3, 3)$ ,  $\mathbf{w}_4 = (2, -2, -2, 0)$ . Fejezzük ki mind a négy vektort ezek lineáris kombinációjaként! Végül írjuk fel mind a négy vektor koordinátás alakját e bázisban!

### 3.16. GRAM-MÁTRIX Igazoljuk, hogy a

$$\begin{bmatrix} \mathbf{v}_1 \cdot \mathbf{v}_1 & \mathbf{v}_1 \cdot \mathbf{v}_2 & \dots & \mathbf{v}_1 \cdot \mathbf{v}_k \\ \mathbf{v}_2 \cdot \mathbf{v}_1 & \mathbf{v}_2 \cdot \mathbf{v}_2 & \dots & \mathbf{v}_2 \cdot \mathbf{v}_k \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{v}_k \cdot \mathbf{v}_1 & \mathbf{v}_k \cdot \mathbf{v}_2 & \dots & \mathbf{v}_k \cdot \mathbf{v}_k \end{bmatrix}$$

mátrix – az ún. *Gram-mátrix* rangja pontosan akkor  $k$ , ha az  $\mathbb{R}^n$ -beli  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  vektorok lineárisan függetlenek.

### 3.17. LINEÁRIS ALGEBRA ALAPTÉTELE Igazoljuk a lineáris algebra alaptételét!

**A SORTÉRBE ESŐ MEGOLDÁS MEGHATÁROZÁSA** Keressük meg az alábbi egyenletrendszerek sortérbe eső egyetlen megoldását, és annak segítségével írjuk fel összes megoldását!

$$3.18. \quad x + y + z = 3$$

$$2x + y - z = 2$$

$$3x + 2y = 5$$

$$3.19. \quad x + 4y + 8z + 12w = 225$$

$$3.20. \quad x + y + z + w = 3$$

$$x + y - z - w = 1$$

## Megoldások

3.1. Mindegyik állítás hamis.

3.2. a) Igaz. b) Igaz. c) Hamis, csak akkor igaz, ha egyik a másik altere. d) Igaz. e) Hamis, a zérustér az egyetlen zérusvektorból áll.

3.3. a) Hamis, csak a homogén lineáris egyenletrendszer megoldásai alkotnak alteret, az inhomogéné eltolt alteret. b) Igaz. c) Igaz. Ez épp az oszloptér, ugyanis csak az oszloptérből való  $\mathbf{b}$  vektorokra oldható meg az egyenletrendszer. d) Igaz. Ez épp a nulltér vektorait adja, ami alter.

3.4. a) Igaz. b) Igaz. c) Hamis. d) Igaz, ugyanis az állításbeli  $r(\mathbf{A}|\mathbf{b}) \leq r(\mathbf{A})$  feltétel pontosan akkor teljesül, ha  $r(\mathbf{A}|\mathbf{b}) = r(\mathbf{A})$ , és ez pontosan akkor teljesül, ha az egyenletrendszer megoldható. e) Hamis, ha  $r(\mathbf{A}) = n$ , és az egyenletrendszer több, mint  $n$  egyenletről áll, akkor előfordulhat, hogy  $r(\mathbf{A}|\mathbf{b}) = n + 1$ , és ekkor az egyenletrendszer nem oldható meg!

3.5. a) Nem, egységvektor konstansszorosai nem egységvektorok. b) Igen (origón átmenő sík). c) Nem (eltolt sík). d) Igen, ez egy origón átmenő egyenes vektoraiból áll. e) Igen, ez a zérustér. f) Nem, az  $(1, -1, 0)$  és az  $(1, 0, -1)$  vektor benne van, de az összegük nincs e halmazban.

3.7. a)  $r(\mathbf{A}|\mathbf{b}) \leq 1$ . A 0 rang csak úgy fordulhat elő, ha az összes egyenlet  $0 = 0$  alakú – nem egy érdekes eset. Ha a rang 1, akkor a kötött és a szabad változók száma is 1. b)  $r(\mathbf{A}|\mathbf{b}) = 5$ . c)  $r(\mathbf{A}|\mathbf{b}) = 3$ . d)  $r(\mathbf{A}|\mathbf{b}) = 2$ .

3.8. Két megoldásvektor különbsége, azaz az  $(1, 2, 3) - (0, 1, 3) = (1, 1, 0)$  vektor biztosan megoldása az egyenletrendszer homogén részének. Akkor viszont ennek minden skalárszorosa is megoldás, így azokat bármelyik fenti megoldáshoz adjuk, újabb megoldásokat kapunk. Például megoldás az  $(1, 2, 3) + (1, 1, 0) = (2, 3, 3)$  és az  $(1, 2, 3) + 2(1, 1, 0) = (3, 4, 3)$  vektor is.

Mivel az ismeretlenek száma 3, és azok legalább egyike szabad változó, ezért a rang legfeljebb 2. Ha viszont e megoldások egy homogén lineáris egyenletrendszer megoldásai, akkor a megoldások közt van legalább két lineárisan független megoldás, így a szabad változók száma legalább kettő, vagyis a kötöttek legfeljebb 1, tehát az együtthatómátrix rangja is legfeljebb 1.

3.9. Első ránézésre csak annyi látszik, hogy mindkét megoldás egy kétdimenziós altert eltoltja. Először megvizsgáljuk, hogy a két alter – vagyis az egyenletrendszer homogén részére adott két megoldás – egybeesik-e. Elég megmutatni, hogy az egyik alterben benne van a másikat generáló két vektor. Ha igen, a két alter megegyezik. Ezesetben el

kell dönteni, hogy az inhomogén két partikuláris megoldása az alternek ugyanabban az eltoltjában van-e. Vagy egyszerűbben, hogy a két partikuláris megoldás különbsége benne van-e az alterben. E kérdéseket egyetlen mátrix lépcsős alakra hozásával is megoldhatjuk. Az első két oszlop az első, a második két oszlop a második alter generátorait tartalmazza, az ötödik oszlop a két partikuláris megoldás különbsége.

$$\left[ \begin{array}{cc|cc|c} 1 & 0 & -2 & 3 & 1 \\ 0 & 1 & 1 & -2 & -1 \\ -2 & -3 & 1 & 0 & 1 \\ -1 & -2 & 0 & 1 & 1 \end{array} \right] \Rightarrow \left[ \begin{array}{cc|cc|c} 1 & 0 & -2 & 3 & 1 \\ 0 & 1 & 1 & -2 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

Az eredményből látszik, hogy a két megoldás azonos.

3.10. Ha  $\mathbf{a}_{i*}$  jelöli az együtthatómátrix  $i$ -edik sorát és  $\mathbf{x}$ , illetve  $\mathbf{y}$  a homogén egyenletrendszer egy-egy megoldását, azaz  $\mathbf{a}_{i*} \cdot \mathbf{x} = 0$ ,  $\mathbf{a}_{i*} \cdot \mathbf{y} = 0$  ( $i = 1, 2, \dots, m$ ), akkor

$$\mathbf{a}_{i*} \cdot (c\mathbf{x} + d\mathbf{y}) = c\mathbf{a}_{i*} \cdot \mathbf{x} + d\mathbf{a}_{i*} \cdot \mathbf{y} = 0 + 0 = 0,$$

tehát a két megoldásvektor bármely lineáris kombinációja is megoldás. Másként fogalmazva a homogén lineáris egyenletrendszerek megoldásainak bármely lineáris kombinációja is megoldás, tehát a megoldások alteret alkotnak.

3.11. Legyen  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  az inhomogén egy partikuláris megoldása, és jelölje  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$  az  $\mathbf{A}$  oszlopvektorait,  $\mathcal{H}$  a homogén,  $\mathcal{I}$  az inhomogén egyenletrendszer általános megoldását. Megmutatjuk, hogy  $\mathbf{x} + \mathcal{H} = \mathcal{I}$ , ahol a bal oldali összeadást elemenként értjük.

$\mathbf{x} + \mathcal{H} \subseteq \mathcal{I}$ : Meg kell mutatnunk, hogy ha  $\mathbf{x}$ -hez adjuk a  $\mathcal{H}$  egy tetszőleges  $\mathbf{y} = (y_1, y_2, \dots, y_n) \in \mathcal{H}$  elemét, az inhomogén egyenletrendszer egy megoldását kapjuk. Valóban,  $\mathbf{x}$ , illetve  $\mathbf{y}$  eleget tesz az

$$\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \dots + \mathbf{a}_n x_n = \mathbf{b}, \text{ illetve}$$

$$\mathbf{a}_1 y_1 + \mathbf{a}_2 y_2 + \dots + \mathbf{a}_n y_n = \mathbf{0}$$

egyenletnek. Ebből

$$\begin{aligned} \mathbf{a}_1(x_1 + y_1) + \mathbf{a}_2(x_2 + y_2) + \dots + \mathbf{a}_n(x_n + y_n) &= \\ (\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \dots + \mathbf{a}_n x_n) + (\mathbf{a}_1 y_1 + \mathbf{a}_2 y_2 + \dots + \mathbf{a}_n y_n) &= \\ \mathbf{b} + \mathbf{0} &= \mathbf{b}, \end{aligned}$$

tehát  $\mathbf{x} + \mathbf{y}$  megoldása az inhomogén egyenletrendszernek, azaz  $\mathbf{x} + \mathbf{y} \in \mathcal{I}$ .

$\mathbf{x} + \mathcal{H} \supseteq \mathcal{I}$ : Meg kell mutatnunk, hogy ha  $\mathbf{z}$  az inhomogén egy tetszőleges megoldása, azaz  $\mathbf{z} \in \mathcal{I}$ , akkor található olyan  $\mathbf{y} \in \mathcal{H}$ , hogy  $\mathbf{z} = \mathbf{x} + \mathbf{y}$ . Valóban, az  $\mathbf{y} = \mathbf{z} - \mathbf{x}$  meg-

teszi, mert

$$\begin{aligned} \mathbf{a}_1(z_1 - x_1) + \mathbf{a}_2(z_2 - x_1) + \dots + \mathbf{a}_n(z_n - x_1) = \\ (\mathbf{a}_1 z_1 + \mathbf{a}_2 z_2 + \dots + \mathbf{a}_n z_n) - (\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \dots + \mathbf{a}_n x_n) = \\ \mathbf{b} - \mathbf{b} = \mathbf{0}, \end{aligned}$$

azaz  $\mathbf{z} - \mathbf{x} \in \mathcal{H}$ .

**3.12.** Összesen 16 altere van  $\mathbb{F}_2^3$ -nek. Van egy 0-dimenziós, a  $\mathcal{Z} = \{\mathbf{0}\}$  tér. Az egydimenziós alterek a nullvektorból és egyetlen tőle különböző további vektorból állnak (7 ilyen alter van). A kétdimenziós alterek mindegyike a nullvektorból, két további egymástól is különböző vektorból és azok összegéből áll. Ezeket felsoroljuk:

$$\begin{aligned} &\{(0, 0, 0), (1, 0, 0), (0, 1, 0), (1, 1, 0)\}, \\ &\{(0, 0, 0), (0, 1, 0), (0, 0, 1), (0, 1, 1)\}, \\ &\{(0, 0, 0), (0, 0, 1), (1, 0, 0), (1, 0, 1)\}, \\ &\{(0, 0, 0), (1, 0, 0), (0, 1, 1), (1, 1, 1)\}, \\ &\{(0, 0, 0), (0, 1, 0), (1, 0, 1), (1, 1, 1)\}, \\ &\{(0, 0, 0), (0, 0, 1), (1, 1, 0), (1, 1, 1)\}, \\ &\{(0, 0, 0), (0, 1, 1), (1, 0, 1), (1, 1, 0)\}. \end{aligned}$$

Végül alter maga  $\mathbb{F}_2^3$  is.

**3.13.** 1. Igaz. 2. Igaz, bármely 1-dimenziós alter ilyen. 3. Hamis, ha van kételemű bázis, akkor a lineárisan független vektorrendszerek elemszáma legfölbbebb 2.

**3.14.** Mivel az egyenletrendszer 3-ismeretlenes, és a rang 2, ezért a kötött változók száma 2, a szabad változóké 1, és így a nulltér dimenziója is 1. A két vektor független egymástól, tehát az egyenletrendszer nem lehet homogén, akkor ugyanis legalább kettő lenne a nulltér dimenziója. Az egyenletrendszer tehát inhomogén, és a megadott két megoldás különbsége a homogén rész egy megoldását adja, annak összes skalárszorosa pedig az összes megoldását. Így az inhomogén összes megoldása:  $(1, 2, 3) + t(1, 1, 0)$ .

**3.15.** A mátrix és annak redukált lépcsős alakja:

$$\begin{bmatrix} 1 & 1 & 2 & 1 \\ 3 & 1 & 3 & 2 \\ 6 & 0 & 3 & 3 \\ 2 & -2 & -2 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & \frac{3}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Eszerint bázisvektoroknak választhatjuk a  $\mathbf{w}_1 = (1, 3, 6, 2)$  és a  $\mathbf{w}_2 = (1, 1, 0, -2)$  vektorokat. A többi vektor kifejezhető ezek lineáris kombinációjaként:

$$\begin{bmatrix} 2 \\ 3 \\ 3 \\ -2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 3 \\ 6 \\ 2 \end{bmatrix} + \frac{3}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ -2 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 3 \\ 6 \\ 2 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ -2 \end{bmatrix}.$$

A redukált lépcsős alak nemzérus soraiból álló

$$\begin{bmatrix} 1 & 0 & 1/2 & 1/2 \\ 0 & 1 & 3/2 & 1/2 \end{bmatrix}$$

mátrixból kiolvasható, hogy a fenti alternek  $\mathcal{B} = \{\mathbf{w}_1, \mathbf{w}_2\}$  bázisa, és ebben a bázisban a négy vektor koordinátás alakja rendre

$$\mathbf{v}_4 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}_B, \quad \mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}_B, \quad \mathbf{v}_2 = \begin{bmatrix} 1/2 \\ 3/2 \end{bmatrix}_B, \quad \mathbf{v}_3 = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}_B.$$

**3.16.** E mátrix rangja pontosan akkor  $k$ , ha oszlopvektorai lineárisan függetlenek, azaz ha az oszlopvektorok bármely lineáris kombinációja csak úgy lehet a nullvektor, ha minden együttható 0. Tekintsünk az oszlopvektorok egy  $c_1, \dots, c_k$  skalárokkal vett, nullvektort adó lineáris kombinációját. Ennek  $i$ -edik koordinátája

$$\begin{aligned} 0 &= c_1 \mathbf{v}_i \cdot \mathbf{v}_1 + c_2 \mathbf{v}_i \cdot \mathbf{v}_2 + \dots + c_k \mathbf{v}_i \cdot \mathbf{v}_k \\ &= \mathbf{v}_i \cdot (c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_k \mathbf{v}_k). \end{aligned}$$

Tehát azt kaptuk, hogy az  $\mathbf{x} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_k \mathbf{v}_k$  vektor olyan, hogy a  $\mathbf{v}_1, \dots, \mathbf{v}_k$  vektorok mindegyikével vett skaláris szorzata 0, így ezek bármelyik lineáris kombinációjával vett skaláris szorzata is 0, tehát például az  $\mathbf{x}$  vektorral vett szorzat is 0, azaz  $\mathbf{x} \cdot \mathbf{x} = 0$ . Ez viszont csak  $\mathbf{x} = \mathbf{0}$  esetén állhat fenn, és mivel a  $\mathbf{v}_i$  vektorok lineárisan függetlenek, csak a  $c_i = 0$  konstansokkal vett lineáris kombinációjuk lehet 0, ahol  $i = 1, 2, \dots, k$ .

**3.17.** Megmutatjuk, hogy a nulltér merőleges kiegészítő altere a sortér. Legyen a valós  $\mathbf{A}$  mátrix sortere  $\mathcal{S}$ , nulltere  $\mathcal{N}$ , és ezek egy-egy bázisa  $\{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_k\}$ , illetve  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{n-k}\}$ .  $\mathcal{S}$  és  $\mathcal{N}$  merőlegessége miatt  $\mathbf{s}_i \cdot \mathbf{e}_j = 0$  minden  $i = 1, 2, \dots, k$  és  $j = 1, 2, \dots, n - k$  esetén. E két bázis együtt  $\mathbb{R}^n$  egy bázisát adja, hisz  $n$  elemű és független vektorokból áll. A függetlenség abból következik, hogy nullvektort előállító tetszőlegesen

$$\underbrace{c_1 \mathbf{s}_1 + \dots + c_r \mathbf{s}_r}_{\mathbf{c}} + \underbrace{d_1 \mathbf{e}_1 + \dots + d_{n-r} \mathbf{e}_{n-r}}_{\mathbf{d}} = \mathbf{0} \quad (3.4)$$

lineáris kombináció csak úgy állhat fenn, ha  $\mathbf{c}$  és  $\mathbf{d}$  a két alter metszetében van, így azok mindketten a nullvektorok, és így  $c_1 = \dots = c_r = d_1 = \dots = d_{n-r} = 0$ .

Ha  $\mathbf{x}$  egy olyan vektor, mely merőleges  $\mathcal{N}$  minden vektorára, akkor  $\mathbf{x} \cdot \mathbf{e}_i = 0$  ( $i = 1, 2, \dots, n - r$ ). Ha

$$\mathbf{x} = y_1 \mathbf{s}_1 + \dots + y_r \mathbf{s}_r + x_1 \mathbf{e}_1 + \dots + x_{n-r} \mathbf{e}_{n-r},$$

akkor az  $\mathbf{e}_i$  vektorokkal való beszorzás a következő homo-

gén lineáris egyenletrendszerre vezet:

$$\begin{aligned}(\mathbf{e}_1 \cdot \mathbf{e}_1)x_1 + (\mathbf{e}_1 \cdot \mathbf{e}_2)x_2 + \dots + (\mathbf{e}_1 \cdot \mathbf{e}_r)x_n &= 0 \\(\mathbf{e}_2 \cdot \mathbf{e}_1)x_1 + (\mathbf{e}_2 \cdot \mathbf{e}_2)x_2 + \dots + (\mathbf{e}_2 \cdot \mathbf{e}_r)x_n &= 0 \\ \vdots & \\(\mathbf{e}_{n-r} \cdot \mathbf{e}_1)x_1 + (\mathbf{e}_{n-r} \cdot \mathbf{e}_2)x_2 + \dots + (\mathbf{e}_{n-r} \cdot \mathbf{e}_{n-r})x_n &= 0\end{aligned}$$

Ez pedig egyértelműen megoldható, mert együtthatómátrixának rangja  $r$ . Ennek bizonyítását az Olvasóra hagyjuk. Egy bizonyítás látható a 3.16. feladatban, egy másik, egyszerűbb a ?? feladatban.

**3.18.** Az egyenletrendszer bővített mátrixának redukált lépcsős alakja

$$\begin{bmatrix} 1 & 0 & -2 & -1 \\ 0 & 1 & 3 & 4 \end{bmatrix}$$

így a megoldása  $(x, y, z) = (-1, 4, 0) + (2, -3, 1)t$ . A nullteret a  $(2, -3, 1)$  vektor feszíti ki, a sortérbe eső vektornak erre merőlegesnek kell lennie, tehát fönn kell állnia a

$$2x - 3y + z = 0$$

egyenletnek is. Ezt az egyenletet a redukált lépcsős alakból származó egyenletrendszerhez (vagy akár az eredetihez) adva egy egyetlen megoldást adó egyenletrendszert kapunk. Ennek bővített mátrixa és annak redukált lépcsős alakja:

$$\begin{bmatrix} 1 & 0 & -2 & -1 \\ 0 & 1 & 3 & 4 \\ 2 & -3 & 1 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Innen a sortérbe eső megoldás  $(1, 1, 1)$ .

**3.19.** A sortérbe eső megoldás meghatározása egyetlen egyenlet esetén egyszerű. Mivel a sorteret az  $(1, 4, 8, 12)$  vektor feszíti ki, ennek egy skalárszorosa keressük, mellyel vett skalárszorzata 225. Mivel  $1^2 + 4^2 + 8^2 + 12^2 = 15^2 = 225$ , ezért a sortérbe eső egyetlen megoldás  $(x, y, z, w) = (1, 4, 8, 12)$ . A homogén egyenletrendszer összes megoldását meghatározva majd hozzáadva kapjuk,

hogyan

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 8 \\ 12 \end{bmatrix} + \begin{bmatrix} -4 \\ 1 \\ 0 \\ 0 \end{bmatrix} t + \begin{bmatrix} -8 \\ 0 \\ 1 \\ 0 \end{bmatrix} s + \begin{bmatrix} -12 \\ 0 \\ 0 \\ 1 \end{bmatrix} u$$

az összes megoldás.

**3.20.** A bővített mátrix és redukált lépcsős alakja:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 3 \\ 1 & 1 & -1 & -1 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

Az egyenletrendszer megoldása:

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 2-s \\ s \\ 1-t \\ t \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} s + \begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \end{bmatrix} t.$$

Tehát a nullteret a  $(-1, 1, 0, 0)$  és a  $(0, 0, -1, 1)$  vektork feszítik ki. A sortérbe eső megoldásvektor ezekre merőleges, tehát az eredeti egyenleteken kívül kielégíti a következő két egyenletet is:

$$\begin{aligned}-x + y &= 0 \\ -z + w &= 0\end{aligned}$$

Ezek mátrixával kibővítve a redukált lépcsős alakot, majd azt redukált lépcsős alakra hozva kapjuk, hogy

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 1 & 1 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1/2 \\ 0 & 0 & 0 & 1 & 1/2 \end{bmatrix}$$

tehát a sortérbe eső megoldás  $(1, 1, 1/2, 1/2)$ , az összes megoldás

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1/2 \\ 1/2 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} s + \begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \end{bmatrix} t.$$

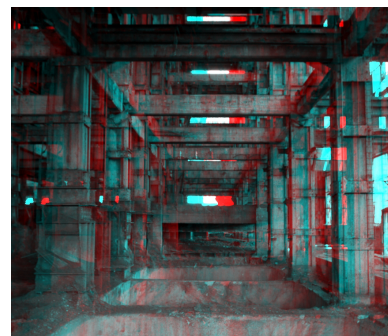
## **II. rész**

# **Mátrixok algebrája és geometriája**





Eddig a mátrixokat csak egyszerű jelölésnek tekintettük, mely az egyenletrendszer együtthatóinak tárolására, és az egyenletrendszer megoldása közbeni számítások egyszerűsítésére való. E részt a számok közti műveletek számtáblázatokra való kiterjesztésével kezdjük, majd ezeket átültetjük mátrixokra, és megvizsgáljuk algebrai tulajdonságait. E műveletek segítségével újragvizsgáljuk az egyenletrendszerek megoldhatóságának és a megoldások kiszámításának kérdését. A mátrixok „számtani” fejezetei után a „mértaniak” következnek: a determináns, mint a négyzetes mátrixhoz rendelt előjeles mérték, majd a mátrixleképezések geometriája lesz e rész tárgya.



Enter The Matrix – 3D picture (CC) on flickr by Grégory Tonon



## 4

# Mátrixműveletek definíciói

Az egyenletrendszerek megoldásához és vizsgálatához hatékony eszközökhöz jutunk a mátrixműveletek bevezetésével. E műveletek szám-talan egyéb alkalmazásban játszanak fontos szerepet, melyekkel a könyv további részében mindenütt találkozni fogunk.

## Műveletek táblázatokkal – műveletek mátrixokkal

A valós számok közti műveletek természetes módon kiterjeszthetők mátrixokkal való műveletekké. Ezek definícióihoz az összeadás és a szorzás hétköznapi alkalmazásainak táblázatokra való kiterjesztésén keresztül fogunk eljutni.

TÁBLÁZATTAL nap, mint nap találkozunk. Számszerű adatok téglalap alakban elrendezett, áttekinthető összefoglalására való. Általában a sorok előtt és az oszlopok fölött van egy *fejléc*, melyben az adott sor, illetve oszlop adatait jellemző valamely információ áll (például az oszlop számadatainak közös mértékegysége). Táblázatok használata kikerülhetetlen a gazdasági adatok kezelésében, így minden irodai szoftvercsomag tartalmaz táblázatkezelőt, de a mérnöki vagy természettudományos közleményekben is nélkülözhetetlen eszköz.

A mátrixra úgy is tekinthetünk, mint amelyet egy olyan absztrakció során kapunk a táblázatból, melyben azt megfosztjuk fejléceitől, az adatokból pedig csak a számokat őrizzük meg, azok jelentésétől, mértékegységétől eltekin-tünk.

**Táblázatok összeadása** 3 alma meg 2 alma az 5 alma. Az összeadás művelete természetes módon kiterjeszthető gyümölcsök számadatait tartalmazó táblázatokra. Ha két gyümölcskosárban piros és zöld alma és szőlő van az alábbi táblázatok szerint, akkor összeöntésük után számukat így számolhatjuk:

|       | alma<br>(db) | szőlő<br>(fürt) |   |       |   |   | alma<br>(db) | szőlő<br>(fürt) |   |   |  | alma<br>(db) | szőlő<br>(fürt) |
|-------|--------------|-----------------|---|-------|---|---|--------------|-----------------|---|---|--|--------------|-----------------|
|       |              |                 | + |       |   | = |              |                 |   |   |  |              |                 |
| piros | 3            | 2               |   | piros | 2 | 2 |              | piros           | 5 | 4 |  |              |                 |
| zöld  | 2            | 1               |   | zöld  | 0 | 1 |              | zöld            | 2 | 2 |  |              |                 |

Azonos méretű, azonos fejlécű táblázatok összeadásának egy lehetséges módja az, ha az azonos pozícióiban lévő elemek összeadásával képezzük az összeget.

**Táblázat szorzása számmal** Az asztalon 2 alma van. Ha számukat megháromszorozzuk, összeszorozunk egy mértékegység nélküli számot (3) egy mértékegységgel rendelkezővel (2 darab), és az eredmény mértékegysége is ez. Ezt megtehetjük egy kosár egész tartalmával is:

$$3 \cdot \begin{array}{cc} \hline & \text{alma} & \text{szőlő} \\ & (\text{db}) & (\text{fürt}) \\ \hline \text{piros} & 3 & 2 \\ \text{zöld} & 2 & 1 \\ \hline \end{array} = \begin{array}{cc} \hline & \text{alma} & \text{szőlő} \\ & (\text{db}) & (\text{fürt}) \\ \hline \text{piros} & 9 & 6 \\ \text{zöld} & 6 & 3 \\ \hline \end{array}$$

**Táblázatok szorzása** Egy adag (e példában a továbbiakban mindig 10 dkg) alma energiatartalma 30 kcal. Mennyi 5 adag energiatartalma? A választ ismét szorzással kapjuk meg – most mindkét mennyiség mértékegységgel is rendelkezik:

$$5 \text{ adag} \cdot 30 \frac{\text{kcal}}{\text{adag}} = 150 \text{ kcal}.$$

Több gyümölcsből (alma, banán, narancs) többféle (A, B, C) gyümölcssalátát készítünk, és a szénhidrát- és energiatartalmukat szeretnénk összehasonlítani. Két táblázatot készítünk, egyikbe a gyümölcssaláták összetételét, a másikba az összetevők szénhidrát- és energiatartalmát írjuk. Mindkét táblázatban a sorokba kerülnek azok a tételek, melyek összetételét/összetevőit részletezzük, az oszlopokba pedig az összetevők.

|          | <i>Alma</i><br>(adag) | <i>Banán</i><br>(adag) | <i>Narancs</i><br>(adag) |                | <i>Szénhidrát</i><br>(g/adag) | <i>Energia</i><br>(kcal/adag) |
|----------|-----------------------|------------------------|--------------------------|----------------|-------------------------------|-------------------------------|
| <i>A</i> | 5                     | 1                      | 4                        | <i>Alma</i>    | 7                             | 30                            |
| <i>B</i> | 4                     | 4                      | 2                        | <i>Banán</i>   | 24                            | 105                           |
| <i>C</i> | 4                     | 2                      | 4                        | <i>Narancs</i> | 8                             | 40                            |

A következőképp tudjuk az A saláta energiatartalmát kiszámítani:

$$5 \text{ adag} \cdot 30 \frac{\text{kcal}}{\text{adag}} + 1 \text{ adag} \cdot 105 \frac{\text{kcal}}{\text{adag}} + 4 \text{ adag} \cdot 40 \frac{\text{kcal}}{\text{adag}} = 415 \text{ kcal},$$

vagyis az első táblázat egy sorának és a második táblázat egy oszlopának kellett a skaláris szorzatát venni. Végezzük el e számításokat mindhárom gyümölcssaláta szénhidrát és energiatartalmára is, és az eredményt ismét egy olyan táblázatba tegyük, melynek soraiba a részletezendő tételek (A, B, C saláta), oszlopaiba a tartalmi összetevők (szénhidrát-, energiatartalom) kerüljenek.

|   | Szénhidrát<br>(g) | Energia<br>(kcal) |
|---|-------------------|-------------------|
| A | 91                | 415               |
| B | 140               | 620               |
| C | 108               | 490               |

Az áttekinthetőség kedvéért a két összeszorozandó mátrixot és az eredményt úgy helyezzük el, hogy az elvégzett műveletek jobban követhetők legyenek. Az A saláta energiatartalmát kiemeljük:

|         | Szénhidrát<br>(g/adag) | Energia<br>(kcal/adag) |
|---------|------------------------|------------------------|
| Alma    | 7                      | 30                     |
| Banán   | 24                     | 105                    |
| Narancs | 8                      | 40                     |

|   | Alma<br>(adag) | Banán<br>(adag) | Narancs<br>(adag) |   | Szénhidrát<br>(g) | Energia<br>(kcal) |
|---|----------------|-----------------|-------------------|---|-------------------|-------------------|
| A | 5              | 1               | 4                 | A | 91                | 415               |
| B | 4              | 4               | 2                 | B | 140               | 620               |
| C | 4              | 2               | 4                 | C | 108               | 490               |

Érdeemes megfigyelni, hogy ha csak az A és C gyümölcssalátákra vagyunk kíváncsiak, elég az első táblázat és a végeredmény második sorát elhagyni, hasonlóképp ha csak az energiatartalmat figyeljük, elég a második táblázat és a végeredmény második oszlopát megtartani. Az is látszik, hogy az első táblázat oszlopainak és a második táblázat sorainak száma megegyezik. Általában az igaz, hogy (a fejléceket nem számolva) egy  $m \times n$ -es táblázat csak olyan  $p \times k$ -as táblázattal szorozható össze, ahol  $p = n$ , és az eredmény  $m \times k$ -as lesz.

**Lineáris helyettesítés** A lineáris algebra alapvető fogalmai megfogalmazhatóak a lineáris helyettesítés nyelvén.

**4.1. DEFINÍCIÓ (LINEÁRIS HELYETTESÍTÉS).** *Lineáris helyettesítésről akkor beszélünk, ha változók egy halmazát más változók konstansszorosainak összegeként állítjuk elő, azaz változókat más változók lineáris kifejezéseivel helyettesítünk.*

Tekintsük például a következő két lineáris helyettesítést:

$$\begin{aligned}
 a &= 5x + y + 4z & x &= 7s + 30k \\
 b &= 4x + 4y + 2z & \text{és} & \quad y = 24s + 105k \\
 c &= 4x + 2y + 4z & z &= 8s + 40k
 \end{aligned}
 \tag{4.1}$$

Egy kis időre e két lineáris helyettesítést is táblázatokkal írjuk le:

$$\begin{array}{c|ccc} & x & y & z \\ \hline a & 5 & 1 & 4 \\ b & 4 & 4 & 2 \\ c & 4 & 2 & 4 \end{array} \qquad \begin{array}{c|cc} & s & k \\ \hline x & 7 & 30 \\ y & 24 & 105 \\ z & 8 & 40 \end{array} \qquad (4.2)$$

**4.2. PÉLDA (LINEÁRIS HELYETTESÍTÉSEK KOMPOZÍCIÓJA).** Tekintsük a (4.1)-ben megadott két lineáris helyettesítést! Az első majd a második helyettesítés egymás után való elvégzése – más szóval kompozíciója – milyen helyettesítéssel egyenértékű, és az lineáris-e?

**MEGOLDÁS.** Az  $a = 5x + y + 4z$  kifejezésben helyettesítsük  $x$ ,  $y$  és  $z$  helyébe a második lineáris helyettesítés szerinti kifejezéseket, azaz

$$a = 5x + y + 4z = 5(7s + 30k) + (24s + 105k) + 4(8s + 40k) = 91s + 415k.$$

Emeljük ki pl. a  $k$  együtthatójának kiszámítását:

$$5 \cdot 30 + 1 \cdot 105 + 4 \cdot 40 = 415.$$

Mint látjuk ez az első helyettesítés táblázata első sorának és második táblázat második oszlopának skaláris szorzata. Hasonló módon  $b$  és  $c$  is kifejezhető  $s$  és  $k$  segítségével, így kapjuk, hogy a két lineáris helyettesítés egymásutánja ekvivalens az

$$a = 91s + 415k$$

$$b = 140s + 620k$$

$$c = 108s + 490k$$

helyettesítéssel. Ez lineáris, hisz a számolás közben lineáris kifejezéseket csak konstanssal szoroztunk, és ezeket adtuk össze. Ennek a helyettesítésnek a táblázata

|     | $s$ | $k$ |
|-----|-----|-----|
| $a$ | 91  | 415 |
| $b$ | 140 | 620 |
| $c$ | 108 | 490 |

vagyis azt kaptuk, hogy a lineáris helyettesítések kompozíciójának táblázataik szorzata felel meg.  $\square$

|     | $s$ | $k$ |
|-----|-----|-----|
| $x$ | 7   | 30  |
| $y$ | 24  | 105 |
| $z$ | 8   | 40  |

|     | $x$ | $y$ | $z$ |     | $s$ | $k$ |
|-----|-----|-----|-----|-----|-----|-----|
| $a$ | 5   | 1   | 4   | $a$ | 91  | 415 |
| $b$ | 4   | 4   | 2   | $b$ | 140 | 620 |
| $c$ | 4   | 2   | 4   | $c$ | 108 | 490 |

**Mátrixok** Az eddigiek alapján megfogalmazhatjuk: az  $m$  sorba és  $n$  oszlopba rendezett  $mn$  darab szám rendszerét  $m \times n$ -típusú mátrixnak nevezzük. A mátrixban szereplő számokat a mátrix elemeinek nevezzük. Eleinte csak valamely algebrai struktúra (egész, valós, komplex) elemeit fogjuk a mátrixba írni, később pl. függvények mátrixait is vizsgálni fogjuk.

**4.3. DEFINÍCIÓ (ADOTT TÍPUSÚ MÁTRIXOK TERE).** Legyen  $S$  egy tetszőleges halmaz (pl.  $S = \mathbb{R}, \mathbb{Q}, \mathbb{N}, \mathbb{Z}, \dots$ ). Az  $S$  elemeiből képzett összes  $m \times n$ -es mátrixok halmazát jelölje

$$S^{m \times n}.$$

Azt mondjuk, hogy  $S^{m \times n}$  az  $S$  fölötti  $m \times n$  típusú mátrixok tere.

Például az  $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$  mátrix eleme az  $\mathbb{N}^{2 \times 2}$ ,  $\mathbb{Z}^{2 \times 2}$ ,  $\mathbb{Q}^{2 \times 2}$ ,  $\mathbb{R}^{2 \times 2}$  terek mindegyikének.

Két mátrixot akkor tekintünk *egyenlőnek*, ha azonos típusúak, és az azonos indexű elemek egyenlők. Például az  $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & x \end{bmatrix}$  egyenlőség pontosan akkor áll fenn, ha  $x = 4$ . Fontos felidézni, hogy minden vektornak megfeleltethetünk egy sor- vagy oszlopvektor alakba írt mátrixot, azok azonban nem egyenlők egymással. Például

$$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \neq \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},$$

mert nem azonos típusúak, így el kell dönteni, hogy e két mátrix közül melyik reprezentálja a  $(1, 2, 3)$  vektort. Mint említettük, a modern matematika legtöbb területén alapértelmezésben az oszlopvektorokat rendeljük a vektorokhoz. E könyvben mi is így teszünk (kivéve a kód-elméleti alkalmazásokat).

Egy mátrix *négyzetes*, ha sorainak és oszlopainak száma megegyezik. Az  $\mathbf{A}$  mátrix főátlójának elemei  $a_{11}, a_{22}, a_{33}, \dots$ . Ez nem csak négyzetes mátrixra értelmezhető. Az olyan négyzetes mátrixot, melynek főátlón kívüli elemei mind nullák, *diagonális mátrixnak* nevezzük. Az ilyen mátrixok egyszerű megadására a `diag` függvényt használjuk, melynek argumentumába a főátló elemei vannak felsorolva. Például

$$\text{diag}(1, 2, 3) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

**Elemenkénti mátrixműveletek** Több olyan mátrixműveletet ismerünk, melyben a számok közt már értelmezett műveletet úgy általánosítjuk mátrixokra, hogy azt a mátrix vagy mátrixok minden elemén végrehajtjuk. Ilyen pl. a mátrixok összeadása és skalárral való szorzása.

**4.4. DEFINÍCIÓ (MÁTRIXOK ÖSSZEGE, KÜLÖNBSÉGE).** Az  $m \times n$ -es típusú  $\mathbf{A} = [a_{ij}]$  és  $\mathbf{B} = [b_{ij}]$  mátrixok összegén azt az ugyancsak  $m \times n$ -es típusú, és  $\mathbf{A} + \mathbf{B}$ -vel jelölt mátrixot értjük, melynek  $i$ -edik sorában a  $j$ -edik elem  $a_{ij} + b_{ij}$ , ahol  $i = 1, \dots, m$ ,  $j = 1, \dots, n$ . Képletben:

$$\mathbf{A} + \mathbf{B} = [a_{ij}] + [b_{ij}] := [a_{ij} + b_{ij}].$$

Hasonlóan definiálható  $\mathbf{A}$  és  $\mathbf{B}$  különbsége is, azaz  $\mathbf{A} - \mathbf{B} := [a_{ij} - b_{ij}]$ .

```
OCTAVE a = [1 2 3
>           4 5 7]
a =
    1    2    3
    4    5    7
```

```
OCTAVE b = [1 2; 3 4]
b =
    1    2
    3    4
```

```
OCTAVE diag([1,2,3])
ans =
    1    0    0
    0    2    0
    0    0    3
```

```
OCTAVE a(2,3)
ans = 7
OCTAVE a(2,:)
ans =
    4    5    7
```

```
OCTAVE a(:,3)
ans =
    3
    7
```

```
OCTAVE v = [1 2 3]
v =
    1    2    3
```

```
OCTAVE w = [1;2;3]
w =
    1
    2
    3
```

```
OCTAVE size(v)
ans =
    1    3
```

```
OCTAVE size(w)
ans =
    3    1
```

4.1. ábra: Mátrix megadása, elemeinek, sorainak és oszlopainak és azok számának lekérdezése mátrix alapú nyelvekben.

A mátrixalapú nyelvekben mátrixok közötti elemenkénti művelet definiálható a műveleti jel elé tett ponttal. Így az  $\mathbf{A}$  és  $\mathbf{B}$  mátrixok elemenkénti szorzata az

$$\mathbf{A} \cdot \mathbf{B}$$

paranccsal kapható meg. Eszerint az  $\mathbf{A} \cdot \mathbf{B}$  és  $\mathbf{A} * \mathbf{B}$  kódok az eredményt tekintve ekvivalensek.





**MEGOLDÁS.** Először a skalárral való szorzásokat végezzük el:

$$2 \begin{bmatrix} 0 & 1 \\ 2 & 1 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 4 & 2 \\ 0 & -2 \end{bmatrix}, \quad -3 \begin{bmatrix} 1 & 0 \\ -1 & -2 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -3 & 0 \\ 3 & 6 \\ 3 & 0 \end{bmatrix},$$

majd az összeadást:

$$\begin{bmatrix} 0 & 2 \\ 4 & 2 \\ 0 & -2 \end{bmatrix} + \begin{bmatrix} -3 & 0 \\ 3 & 6 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} -3 & 2 \\ 7 & 8 \\ 3 & -2 \end{bmatrix}.$$

A műveletek természetesen elemenként is elvégezhetők, pl. a második sor első eleme így is megkapható:  $2 \cdot 2 - 3 \cdot (-1) = 7$ .  $\square$

Látjuk, a mátrixok az összeadásra és a skalárral való szorzásra nézve a vektorokhoz hasonlóan viselkednek. Mondhatjuk, hogy az  $\mathbb{R}^{m \times n}$ -beli  $m \times n$ -es mátrixok e két műveletre nézve úgy viselkednek, mint  $\mathbb{R}^{mn}$  vektorai (csak másként vannak elrendezve). Beszélhetünk ezért arról, hogy  $\mathbb{R}^{m \times n}$  egy  $mn$ -dimenziós vektortér. Lásd erről pl. a 4.7. és a 4.8. feladatokat.

**Mátrixszorzás** A táblázatok szorzásánál látott szabályt követjük a következő definícióban:

**4.9. DEFINÍCIÓ (MÁTRIXOK SZORZÁSA).** Egy  $m \times t$ -es  $\mathbf{A}$  és egy  $t \times n$ -es  $\mathbf{B}$  mátrix szorzatán azt az  $\mathbf{AB}$ -vel jelölt  $m \times n$ -es  $\mathbf{C}$  mátrixot értjük, amelynek  $i$ -edik sorában és  $j$ -edik oszlopában álló eleme

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{ik}b_{kj} + \dots + a_{it}b_{tj}.$$

A definícióbeli összefüggés több módon is kifejezhető. Szummával fölírva:

$$c_{ij} = \sum_{k=1}^t a_{ik}b_{kj},$$

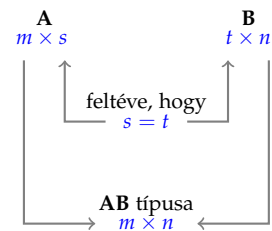
de mondhatjuk azt is, hogy  $c_{ij}$  az  $\mathbf{A}$  mátrix  $i$ -edik sorának és a  $\mathbf{B}$  mátrix  $j$ -edik oszlopának skaláris szorzata, azaz

$$c_{ij} = \mathbf{a}_{i*} \cdot \mathbf{b}_{*j}.$$

Fontos kiemelni, hogy egy  $m \times s$ -es  $\mathbf{A}$  és egy  $t \times n$ -es  $\mathbf{B}$  mátrix csak akkor szorozható össze, ha  $s = t$ , és ekkor a szorzat  $m \times n$  típusú.

Nyilvánvaló, hogy a szorzás sorrendje fontos. Lehet, hogy az  $\mathbf{AB}$  szorzás elvégezhető, de a  $\mathbf{BA}$  nem, és lehet, hogy elvégezhető, de különböző eredményt kapunk (ld. a 4.10. feladatot). Mivel a mátrixszorzás nem felcserélhető, ha szükséges, az „ $\mathbf{A}$ -t balról szorozzuk  $\mathbf{B}$ -vel”, vagy az „ $\mathbf{A}$ -t jobbról szorozzuk  $\mathbf{B}$ -vel” kifejezésekkel teszünk különbséget a  $\mathbf{BA}$  és az  $\mathbf{AB}$  szorzatok közt.

$$\begin{bmatrix} a_{i1} & a_{i2} & \dots & a_{it} \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{tj} \end{bmatrix} = c_{ij}$$



4.10. PÉLDA (MÁTRIXOK SZORZÁSA). Legyen

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 0 & 3 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 0 \end{bmatrix}$$

Számítsuk ki az  $\mathbf{AB}$  szorzat második sorának elemeit, azaz az  $(\mathbf{AB})_{2*}$  sor-mátrixot!

**MEGOLDÁS.** A szorzat második sorának elemei az  $\mathbf{A}$  második sorának  $\mathbf{B}$  oszlopaival való szorzatából kaphatók meg:

$$\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} * & * & * \\ 4 & 3 & 4 \\ * & * & * \end{bmatrix} \quad \square$$

Fontosak azok a mátrixszorzatok, amelyekben az egyik mátrixnak csak egy sora vagy oszlopa van, tehát sor- vagy oszlopvektor. Egy  $1 \times m$ -es mátrixot, azaz egy sorvektort jobbról, egy  $n \times 1$ -es mátrixot, azaz egy oszlopvektort balról lehet beszorozni egy  $m \times n$ -es mátrixszal. Például

$$\begin{bmatrix} 2 & 3 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 7 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

**Műveletek blokkmátrixokkal** Hatalmas méretű mátrixokkal végzett műveletek párhuzamosíthatók, és a memóriakezelésük is hatékonyabbá válik, ha a mátrixokat blokkokra osztjuk, és a műveleteket is e kisebb részmátrixokkal végezzük.

Ha egy mátrixot néhány vízszintes és függőleges vonallal részmátrixokra bontunk, azt mondjuk, hogy e mátrix a részmátrixokból – más néven blokkokból – alkotott *blokkmátrix*. Egy egyenletrendszer  $[\mathbf{A}|\mathbf{b}]$  bővített mátrixa egy két blokkból álló blokkmátrixnak is tekinthető.

Az alábbi példa egy 5-ismeretlenes, 5 egyenletből álló egyenletrendszer bővített mátrixának redukált lépcsős alakját mutatja:

$$\left[ \begin{array}{ccc|cc|c} 1 & 0 & 0 & 1 & 2 & 4 \\ 0 & 1 & 0 & 2 & 0 & 3 \\ 0 & 0 & 1 & 1 & 0 & 3 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right] = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \mathbf{B}_{13} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{B}_{23} \end{bmatrix}.$$

Az első blokkoszlop a kötött változóknak, a második a szabad változóknak, a harmadik az egyenletrendszer jobb oldalának felel meg. A második blokkosor a zérussorokat tartalmazza.

Egy blokkmátrix sorait és oszlopait szokás a mátrix *blokksorainak* és *blokkoszlopainak* nevezni a közönséges soroktól és oszlopoktól való

megkülönböztetés végett. Pl. a fenti blokkmátrix első blokksorának elemei  $\mathbf{B}_{11}$ ,  $\mathbf{B}_{12}$ ,  $\mathbf{B}_{13}$ .<sup>1</sup>

4.11. ÁLLÍTÁS (MŰVELETEK BLOKKMÁTRIXOKKAL). *Blokkmátrixok skálárral való szorzása és két azonos módon particionált blokkmátrix összeadása blokkonként is elvégezhető, azaz*

$$c[\mathbf{A}_{ij}] := [c\mathbf{A}_{ij}], \quad [\mathbf{A}_{ij}] + [\mathbf{B}_{ij}] := [\mathbf{A}_{ij} + \mathbf{B}_{ij}].$$

Ha  $\mathbf{A} = [\mathbf{A}_{ik}]_{m \times t}$ ,  $\mathbf{B} = [\mathbf{B}_{kj}]_{t \times n}$  két blokkmátrix, és minden  $k$ -ra az  $\mathbf{A}_{ik}$  blokk oszlopainak száma megegyezik  $\mathbf{B}_{kj}$  sorainak számával, akkor a  $\mathbf{C} = \mathbf{AB}$  szorzat kiszámítható a szorzási szabály blokkokra való alkalmazásával is, azaz  $\mathbf{C}$  olyan blokkmátrix, melynek  $i$ -edik blokksorában és  $j$ -edik blokkoszlopában álló blokk

$$\mathbf{C}_{ij} = \sum_{k=1}^t \mathbf{A}_{ik} \mathbf{B}_{kj}.$$

4.12. PÉLDA (MŰVELETEK BLOKKMÁTRIXOKKAL). *Végezzük el az alábbi műveletet blokkmátrixokkal számolva!*

$$\left[ \begin{array}{cc|c} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 0 & 3 & 1 \end{array} \right] \left[ \begin{array}{c|c} 1 & 1 \\ 1 & 2 \\ 0 & 1 \end{array} \right] + \left[ \begin{array}{c|c} 1 & 4 \\ 1 & 1 \\ 6 & 0 \end{array} \right].$$

**MEGOLDÁS.** Először ellenőrizzük, hogy a műveletek valóban elvégezhetők lesznek blokkmátrixokként számolva is. Valóban, az első mátrix blokkoszlopai 2- és 1-oszloposak, míg a második mátrix blokksorai 2- és 1-sorosak. A szorzat pedig azonos típusú lesz a harmadik mátrixszal, tehát összeadhatók. A blokkokat mátrixokkal helyettesítve, majd a mátrixműveleteket elvégezve kapjuk, hogy

$$\begin{aligned} & \left[ \begin{array}{cc|c} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 0 & 3 & 1 \end{array} \right] \left[ \begin{array}{c|c} 1 & 1 \\ 1 & 2 \\ 0 & 1 \end{array} \right] + \left[ \begin{array}{c|c} 1 & 4 \\ 1 & 1 \\ 6 & 0 \end{array} \right] \\ &= \left[ \begin{array}{cc|c} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 0 & 3 & 1 \end{array} \right] + \left[ \begin{array}{c|c} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{array} \right] \left[ \begin{array}{cc|c} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 0 & 3 & 2 \end{array} \right] + \left[ \begin{array}{c|c} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{array} \right] + \left[ \begin{array}{c|c} 1 & 4 \\ 1 & 1 \\ 6 & 0 \end{array} \right] \\ &= \left[ \begin{array}{cc|c} 1 & 0 & 2 \\ 3 & 1 & 2 \\ 3 & 3 & 2 \end{array} \right] + \left[ \begin{array}{cc|c} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \end{array} \right] = \left[ \begin{array}{cc|c} 2 & 0 & 3 \\ 4 & 2 & 4 \\ 4 & 4 & 3 \end{array} \right] = \left[ \begin{array}{cc|c} 2 & 6 \\ 4 & 6 \\ 9 & 7 \end{array} \right] \end{aligned}$$

Ellenőrizzük a számítást közönséges mátrixműveletekkel!

<sup>1</sup> A BLOKKMÁTRIXOKRA a szakirodalomban a *hipermátrix* elnevezés is el van terjedve. Ekkor a blokksorokat hipersoroknak, a blokkoszlopokat hiperoszlopoknak nevezzük. Mi kerüljük e szóhasználatot a hipermátrix másik – többdimenziós tömb értelmű – jelentése miatt.

□

*Kronecker-szorzat és a vec-függvény\** Blokkmátrixok összege, skalárszorosa és szorzata ugyanazt az mátrixot adja eredményül, mint amit egyszerű mátrixműveletekkel kapunk. Vannak azonban olyan blokkmátrixműveletek is, amelyek nem származtathatóak egyszerű mátrixműveletekből.

A *vec* függvény egy tetszőleges mátrixot vektorrá alakít a mátrix oszlopvektorainak egymás alá tételével. Tehát ha  $\mathbf{A} = [\mathbf{a}_1 | \mathbf{a}_2 | \dots | \mathbf{a}_n]$ , akkor

$$\text{vec}(\mathbf{A}) = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_n \end{bmatrix}.$$

Például, ha

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \text{ akkor } \text{vec}(\mathbf{A}) = \begin{bmatrix} 1 \\ 3 \\ 2 \\ 4 \end{bmatrix}.$$

Legyen  $\mathbf{A}$  egy  $m \times n$ -es,  $\mathbf{B}$  egy  $p \times q$ -as mátrix. *Kronecker-szorzatukon* (vagy más néven *tenzorszorzatukon*) azt az  $\mathbf{A} \otimes \mathbf{B}$ -vel jelölt  $mp \times nq$  méretű mátrixot értjük, melynek blokkmátrix alakja

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots & a_{1n}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \dots & a_{2n}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \dots & a_{mn}\mathbf{B} \end{bmatrix}$$

Például

$$\begin{bmatrix} -1 & 2 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 0 & 1 & 2 \\ 3 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 0 & -1 & -2 & 0 & 2 & 4 \\ -3 & -3 & -3 & 6 & 6 & 6 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 3 & 3 & 3 \end{bmatrix}.$$

A *vec* függvény és a Kronecker-szorzat tulajdonságai közül néhányat kiemelünk.

**4.13. TÉTEL (A KRONECKER-SZORZAT TULAJDONSÁGAI).** *Adva van az  $\mathbf{A}_{m \times n}$ ,  $\mathbf{B}_{m \times n}$ ,  $\mathbf{C}_{p \times q}$  és  $\mathbf{D}_{r \times s}$  mátrix. Ekkor*

- a)  $(\mathbf{A} + \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes \mathbf{C} + \mathbf{B} \otimes \mathbf{C}$ ,  $\mathbf{C} \otimes (\mathbf{A} + \mathbf{B}) = \mathbf{C} \otimes \mathbf{A} + \mathbf{C} \otimes \mathbf{B}$ ,
- b)  $(\mathbf{A} \otimes \mathbf{C}) \otimes \mathbf{D} = \mathbf{A} \otimes (\mathbf{C} \otimes \mathbf{D})$ ,
- c)  $(\mathbf{A} \otimes \mathbf{C})^T = \mathbf{C}^T \otimes \mathbf{A}^T$ ,

A *vec* függvény és a Kronecker-szorzat közös tulajdonságait a lineáris mátrixegyenletek megoldásánál fogjuk használni:

4.14. TÉTEL (A KRONECKER-SZORZAT ÉS A vec-FÜGGVÉNY TULAJDONSÁGAI). *Adva van az  $\mathbf{A}_{m \times n}$ ,  $\mathbf{B}_{p \times q}$  és  $\mathbf{X}_{n \times p}$  mátrix. Ekkor*

- a)  $\text{vec}(\mathbf{AX}) = (\mathbf{I}_p \otimes \mathbf{A}) \text{vec}(\mathbf{X})$ ,  $\text{vec}(\mathbf{XB}) = (\mathbf{B}^T \otimes \mathbf{I}_n) \text{vec}(\mathbf{X})$ ,  
 b)  $\text{vec}(\mathbf{AXB}) = (\mathbf{B}^T \otimes \mathbf{A}) \text{vec}(\mathbf{X})$ ,  
 c)  $\text{vec}(\mathbf{AX} + \mathbf{XB}) = (\mathbf{I}_p \otimes \mathbf{A} + \mathbf{B}^T \otimes \mathbf{I}_n) \text{vec}(\mathbf{X})$ .

**BIZONYÍTÁS.** A fenti állítások mindegyike közvetlenül bizonyítható a definíció alapján. Szemléltetésül megmutatjuk a b) bizonyítását.

$$\begin{aligned} [\mathbf{AXB}]_{*j} &= \mathbf{AXB}_{*j} = \sum_{i=1}^n b_{ij}(\mathbf{AX})_{*i} = \sum_{i=1}^n (b_{ij}\mathbf{A})\mathbf{X}_{*i} \\ &= [b_{1j}\mathbf{A} \mid \dots \mid b_{nj}\mathbf{A}] \text{vec}(\mathbf{X}) \\ &= [\mathbf{B}^T \otimes \mathbf{A}]_{*j} \text{vec}(\mathbf{X}) \end{aligned} \quad \square$$

**Műveletek hipermátrixokkal\*** Bizonyos adatok, természetüknél fogva, 2-nél magasabb dimenziós tömbben rendezhetők el jól. Kérdés, hogy hogyan tudjuk ezeket kezelni, tudunk-e ezekkel is számolni, és hogyan?

4.15. DEFINÍCIÓ (HIPERMÁTRIX). Legyen  $n_1, n_2, \dots, n_d \in \mathbb{N}^+$  és legyen  $S$  egy tetszőleges halmaz (pl.  $S = \mathbb{R}, \mathbb{Q}, \mathbb{N}, \mathbb{Z}, \dots$ ).  $d$ -edrendű (vagy  $d$ -dimenziós)  $n_1 \times n_2 \times \dots \times n_d$ -típusú hipermátrixnak nevezzük az

$$\mathbf{A} : \{1, \dots, n_1\} \times \{1, \dots, n_2\} \times \dots \times \{1, \dots, n_d\} \rightarrow S$$

alakú leképezést. Az  $\mathbf{A}(i_1, i_2, \dots, i_d)$  elemet  $a_{i_1 i_2 \dots i_d}$ -vel jelöljük, melyre úgy gondolhatunk, mint egy  $d$ -dimenziós táblázat egy elemére és a mátrixoknál megszokotthoz hasonlóan írhatjuk, hogy

$$\mathbf{A} = [a_{i_1 i_2 \dots i_d}]_{i_1, i_2, \dots, i_d=1}^{n_1, n_2, \dots, n_d} \text{ vagy egyszerűbben } \mathbf{A} = [a_{i_1 i_2 \dots i_d}].$$

Ha  $n_1 = n_2 = \dots = n_d = n$ , akkor a hiper-kockamátrixról beszélünk.

Az  $S$  elemeiből képzett összes  $n_1 \times n_2 \times \dots \times n_d$ -típusú hipermátrixok halmazát

$$\mathcal{S}^{n_1 \times n_2 \times \dots \times n_d}$$

jelöli.

A másodrendű hipermátrixok egybeesnek a mátrixokkal.

A 3-adrendű hipermátrixok elemeinek leírását papírra (tehát 2-dimenzióban) úgy oldhatjuk meg, hogy például a harmadik index szerint „szeletekre” vágjuk. E „szeletek” mindegyike egy mátrix, melyeket függőleges vonallal elválasztva egymás mellé írunk. Így például a

$4 \times 2 \times 3$ -típusú hipermátrixok általános alakja

$$\left[ \begin{array}{cc|cc|cc} a_{111} & a_{121} & a_{112} & a_{122} & a_{113} & a_{123} \\ a_{211} & a_{221} & a_{212} & a_{222} & a_{213} & a_{223} \\ a_{311} & a_{321} & a_{312} & a_{322} & a_{313} & a_{323} \\ a_{411} & a_{421} & a_{412} & a_{422} & a_{413} & a_{423} \end{array} \right]$$

Két azonos típusú hipermátrix összeadása és egy hipermátrix skálárral való szorzása a mátrixokhoz hasonlóan elemenként történik, azaz

$$\begin{aligned} [a_{i_1 i_2 \dots i_d}] + [b_{i_1 i_2 \dots i_d}] &:= [a_{i_1 i_2 \dots i_d} + b_{i_1 i_2 \dots i_d}], \\ c[a_{i_1 i_2 \dots i_d}] &:= [ca_{i_1 i_2 \dots i_d}], \end{aligned}$$

ahol  $S$  elemein az összeadás és a szorzás is értelmezve van.

**4.16. DEFINÍCIÓ (HIPERMÁTRIX TRANSZPONÁLTJA).** Legyen  $\pi$  az  $\{1, 2, \dots, d\}$  halmaz egy permutációja. A  $d$ -edrendű  $\mathbf{A} = [a_{i_1 i_2 \dots i_d}] \in S^{n_1 \times n_2 \times \dots \times n_d}$  hipermátrix  $\pi$ -transzponáltján az

$$\mathbf{A}^\pi = [a_{i_{\pi(1)} i_{\pi(2)} \dots i_{\pi(d)}}] \in S^{n_{\pi(1)} \times n_{\pi(2)} \times \dots \times n_{\pi(d)}}$$

hipermátrixot értjük. Egy  $\mathbf{A} \in S^{n \times n \times \dots \times n}$  hiper-kockamátrix szimmetrikus, ha minden  $\pi$  permutációra  $\mathbf{A}^\pi = \mathbf{A}$ , és ferdén szimmetrikus, ha  $\mathbf{A}^\pi = \text{sgn}(\pi)\mathbf{A}$ , ahol  $\text{sgn}(\pi) = -1$ , ha  $\pi$  páratlan permutáció, és  $1$ , ha páros.

Eszerint a  $2 \times 2 \times 2$ -es hipermátrixok és szimmetrikus hipermátrixok általános alakja

$$\left[ \begin{array}{cc|cc} a_{111} & a_{121} & a_{112} & a_{122} \\ a_{211} & a_{221} & a_{212} & a_{222} \end{array} \right], \quad \left[ \begin{array}{cc|cc} a & b & b & c \\ b & c & c & d \end{array} \right].$$

A  $3 \times 3 \times 3$ -as hipermátrixok, szimmetrikus és ferdén szimmetrikus hipermátrixok általános alakja

$$\begin{aligned} &\left[ \begin{array}{ccc|ccc|ccc} a_{111} & a_{121} & a_{131} & a_{112} & a_{122} & a_{132} & a_{113} & a_{123} & a_{133} \\ a_{211} & a_{221} & a_{231} & a_{212} & a_{222} & a_{232} & a_{213} & a_{223} & a_{233} \\ a_{311} & a_{321} & a_{331} & a_{312} & a_{322} & a_{332} & a_{313} & a_{323} & a_{333} \end{array} \right], \\ &\left[ \begin{array}{ccc|ccc|ccc} a & b & c & b & d & e & c & e & f \\ b & d & e & d & g & h & e & h & i \\ c & e & f & e & h & i & f & i & j \end{array} \right], \\ &\left[ \begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 0 & 0 & -a & 0 & a & 0 \\ 0 & 0 & a & 0 & 0 & 0 & -a & 0 & 0 \\ 0 & -a & 0 & a & 0 & 0 & 0 & 0 & 0 \end{array} \right], \end{aligned}$$

ahol  $a, b, c, d, e, f, g, h, i, j \in S$  nem feltétlenül különböző elemek.

A vektorok diadikus szorzatát általánosítja az alábbi definíció:

**4.17. DEFINÍCIÓ (HIPERMÁTRIXOK KÜLSŐ SZORZATA).** Legyen  $\mathbf{A} \in S^{n_1 \times \cdots \times n_d}$  egy  $d$ -edrendű és  $\mathbf{B} \in S^{m_1 \times \cdots \times m_e}$  egy  $e$ -edrendű hipermátrix. Külső szorzatukon azt a  $(d + e)$ -edrendű

$$\mathbf{C} = [c_{i_1 \dots i_d j_1 \dots j_e}]_{i_1, \dots, i_d, j_1, \dots, j_e=1}^{n_1, \dots, n_d, m_1, \dots, m_e} = \mathbf{A} \otimes \mathbf{B} \in S^{n_1 \times \cdots \times n_d \times m_1 \times \cdots \times m_e}$$

hipermátrixot értjük, melyre

$$c_{i_1 \dots i_d j_1 \dots j_e} = a_{i_1 \dots i_d} b_{j_1 \dots j_e}.$$

Például

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \otimes \begin{bmatrix} 0 & 1 \\ 2 & 3 \\ 4 & 0 \end{bmatrix} = \left[ \begin{array}{ccc|ccc} 0 & 2 & 4 & 1 & 3 & 0 \\ 0 & 4 & 8 & 2 & 6 & 0 \end{array} \right]$$

## Feladatok

### Táblázatok

4.1. Anti, Bori, Cili almát, banánt és citromot vesz a piacon, a hipermarketben vagy a csarnokban. Ha csak az ár számít, melyikük hol vásároljon?

|      | alma (kg) | banán (kg) | citrom (kg) |
|------|-----------|------------|-------------|
| Anti | 2         | 2          | 1           |
| Bori | 3         | 2          | 0.5         |
| Cili | 2         | 1          | 1           |

|        | csarnok<br>(Ft/kg) | hipermarket<br>(Ft/kg) | piac<br>(Ft/kg) |
|--------|--------------------|------------------------|-----------------|
| alma   | 180                | 100                    | 130             |
| banán  | 390                | 420                    | 360             |
| citrom | 210                | 210                    | 230             |

4.2. Egy  $f(x, y)$  kifejezésben elvégezzük az

$$x = 2a + b$$

$$y = 3a + b$$

helyettesítést, majd az így kapott  $f(2a + b, 3a + b)$  kifejezésben az

$$a = -3s + t$$

$$b = 4s - t$$

helyettesítést. Számítsuk ki a két helyettesítés kompozícióját a helyettesítések végrehajtásával, és a nekik megfelelő táblázatok szorzásával is, azaz írjuk fel azt a helyettesítést, mely e két helyettesítés kompozíciójával ekvivalens!

4.3. Tegyük fel, hogy egy kifejezésben elvégezzük a következő két helyettesítést:

$$\begin{aligned} x &= 2a + b + 6c & a &= -s & + & u \\ y &= 4a + b + 7c & b &= -3s - 6t + 10u \\ z &= 3a + b + 6c & c &= s + t - 2u \end{aligned}$$

Hogyan számíthatjuk ki a két helyettesítés kompozícióját? Írjuk fel azt a helyettesítést, mely e két helyettesítés kompozíciójával ekvivalens!

4.4. Két versengő kereskedelmi TV-csatorna valóságshow-műsora kezdetben fele-fele arányban vonzza a nézőket. Az első hét végére a tv1 nézőinek fele, míg a tv2 nézőnek negyede átpártol a másik csatornára.

1. Készítsük el az átpártolás  $2 \times 2$ -es táblázatát, és a
2. nézők megoszlásának  $2 \times 1$ -es vagy  $1 \times 2$ -es táblázatát!
3. Táblázatok szorzásának segítségével határozzuk meg, hogy mi a nézők megoszlása az első és a második hét végén, ha az átpártolók aránya az idővel nem változik.

4. Írjuk fel az átpártolók kéthetenkénti táblázatát, azaz azt, amelyből kiolvasható, hogy két hét elteltével az egyes csatornák nézőinek hányadrésze pártol át, és mennyi marad!

### Elemenkénti mátrixműveletek

4.5. Adva vannak az alábbi mátrixok!

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 1 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 1 & 0 \end{bmatrix}$$

Számítsuk ki a következő kifejezések közül azok értékét, amelyek értelmezve vannak! a)  $4\mathbf{A} - 3\mathbf{B}$ , b)  $2\mathbf{B} - \mathbf{C}$ , c)  $2\mathbf{B} - \mathbf{C}^T$ ,

4.6. **ELEMENKÉNTI MÁTRIXMŰVELET A DIGITÁLIS KÉPFELDOLGOZÁSBAN** Egy leegyszerűsített képformátummal dolgozunk: az egészelemű  $\mathbf{A}_{m \times n}$  mátrix reprezentáljon egy  $m \times n$  képpontból álló szürkeárnyaltos képet. Minden mátrixelem egy képpont árnyalatát adja meg a  $\{0, 1, \dots, k\}$  tartományból, ahol 0 a fekete,  $k - 1$  a fehér színnek felel meg és  $k$  az átlátszó pixeleket jelöli. Legyen egy képen a háttér átlátszó, és legyen  $\mathbf{B}_{m \times n}$  egy tetszőleges másik kép azonos módon reprezentált mátrixa. Konstruáljuk meg azt a  $\odot$  jellel jelölt műveletet, amellyel az elemenkénti

$$\mathbf{A} \odot \mathbf{B} := [a_{ij} \odot b_{ij}]$$

mátrixművelet az  $\mathbf{A}$  kép háttérébe másolja a  $\mathbf{B}$  képet. Képletben:

$$a_{ij} \odot b_{ij} = \begin{cases} b_{ij}, & \text{ha } a_{ij} = k, \\ a_{ij}, & \text{egyébként.} \end{cases}$$

A megoldásban használhatjuk a  $x \mapsto \lfloor x \rfloor$  függvényt, mely egy  $x$  számhoz annak alsó egész részét rendeli.

4.7.  $\mathbb{R}^{m \times n}$  BÁZISA Adjuk meg az  $\mathbb{R}^{m \times n}$  tér egy bázisát.

4.8. **MÁTRIXOK ÁLTAL KIFESZÍTETT ALTÉR** Jellemezzük az  $\mathbb{R}^{2 \times 2}$  térnek azt az alterét, melyet az alább megadott  $\mathbf{A}$ ,  $\mathbf{B}$  és  $\mathbf{C}$  mátrixok feszítenek ki! Másként fogalmazva: milyen összefüggések állnak fenn azon  $2 \times 2$ -es valós mátrixok elemei között, melyek az alábbi mátrixok lineáris kombinációiként állnak elő?

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

### Mátrixszorzás

4.9. Adva vannak az alábbi mátrixok!

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 1 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 1 & 0 \end{bmatrix} \quad \mathbf{D} = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix}$$

Számítsuk ki a következő kifejezések közül azok értékét, amelyek értelmezve vannak! a)  $\mathbf{AB}$ , b)  $\mathbf{AB}^T - \mathbf{D}$ , c)  $\mathbf{BC}$ , d)  $\mathbf{CB}$ , e)  $(\mathbf{DA})\mathbf{C}$ .



## 4.10. A SZORZÁS NEM FELCSERÉLHETŐ Legyen

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 2 & 3 \\ 2 & 1 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 2 & 1 \end{bmatrix}, \mathbf{C} = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix},$$

$$\mathbf{D} = \begin{bmatrix} 6 & 6 \\ -2 & -1 \end{bmatrix}, \mathbf{E} = \begin{bmatrix} -2 & -6 \\ 2 & 5 \end{bmatrix}.$$

Döntsük el, hogy fennállnak-e az  $\mathbf{AB} = \mathbf{BA}$ ,  $\mathbf{BC} = \mathbf{CB}$ ,  $\mathbf{CD} = \mathbf{DC}$  és  $\mathbf{DE} = \mathbf{ED}$  egyenlőségek.

## Blokkmátrixok

4.11.  $2 \times 2$ -ES BLOKKMÁTRIXOK SZORZÁSA Legyen  $\mathbf{A}$  és  $\mathbf{B}$  két  $2 \times 2$ -es blokkmátrix, azaz legyen

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}, \mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix}.$$

Írjuk fel szorzatukat a blokkok szorzatai segítségével.

4.12. Végezzük el az  $\mathbf{A} + 3\mathbf{C}$  és az  $\mathbf{AB}$  műveleteket közös mátrixműveletekkel és blokkmátrixként számolva is, ha

$$\mathbf{A} = \left[ \begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 2 \\ \hline 0 & 0 & 3 & 0 \end{array} \right], \mathbf{B} = \left[ \begin{array}{cc} 2 & 4 \\ 1 & 5 \\ \hline 2 & 2 \\ 0 & 1 \end{array} \right], \mathbf{C} = \left[ \begin{array}{cc|cc} 0 & 2 & 0 & 1 \\ 2 & 0 & 0 & 0 \\ \hline 1 & 1 & 1 & 1 \end{array} \right].$$

Mátrixműveletek  $\mathbb{Z}_m$ -ben\*

A mátrixműveletek minden további nélkül értelmezhetők  $\mathbb{Z}_m$  fölötti mátrixokra is (általában bármely gyűrű fölötti mátrixokra, ld. a 13. fejezetet a függelékben).

4.13. Egy lineáris kód  $\mathbf{G}$  generátormátrixa és  $\mathbf{H}$  ellenőrző mátrixa eleget tesz a  $\mathbf{GH}^T = \mathbf{O}$  összefüggésnek. Ellenőrizzük ezt a  $[4, 2, 3]_3$  Hamming kód esetén a következő mátrixokkal az  $\mathbb{F}_3$  testben számolva:

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \end{bmatrix} \quad \mathbf{H} = \begin{bmatrix} 2 & 2 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{bmatrix}$$

## Hiper mátrixok\*

4.14. MULTILINEÁRIS MÁTRIXSZORZÁS Defináljunk egy hiper mátrixműveletet a következőképp. Legyen  $\mathbf{X}_1 = [x_{ij}^{(1)}] \in S^{m_1 \times n_1}, \dots, \mathbf{X}_d = [x_{ij}^{(d)}] \in S^{m_d \times n_d}$  tetszőleges  $d$  mátrix, és legyen  $\mathbf{A} \in S^{n_1 \times \dots \times n_d}$  egy hiper mátrix. Ekkor a  $\mathbf{B} = (\mathbf{X}_1, \dots, \mathbf{X}_d) \cdot \mathbf{A}$  multilineáris mátrixszorzatot a

$$b_{i_1 \dots i_d} = \sum_{j_1, \dots, j_d=1}^{n_1, \dots, n_d} x_{i_1 j_1}^{(1)} \dots x_{i_d j_d}^{(d)} a_{j_1 \dots j_d},$$

képlet definiálja, ahol  $\mathbf{B} = [b_{i_1 \dots i_d}]_{i_1, \dots, i_d=1}^{m_1, \dots, m_d}$ .

Igazoljuk, hogy a) ha  $d = 1$ ,  $n_1 = n$  és  $m_1 = 1$ , akkor e szorzás megegyezik a skaláris szorzással; b) ha  $d = 2$ ,  $m_1 = m_2 = 1$  és  $\mathbf{X}_1 = \mathbf{X}_2$ , akkor e szorzás kvadratikussá alakot ad.

4.15. VEKTOROK SEGRE-FÉLE KÜLSŐ SZORZATA Legyen  $n_1, n_2, \dots, n_d \in \mathbb{N}^+$  és legyen  $\mathbf{a}_i = (a_{i1}, a_{i2}, \dots, a_{in_i}) \in S^{n_i}$  ( $i = 1, 2, \dots, d$ ). E vektorok Segre-féle külső szorzatán az

$$\mathbf{a}_1 \otimes \mathbf{a}_2 \otimes \dots \otimes \mathbf{a}_d = [a_{i_1 i_1} a_{i_2 i_2} \dots a_{i_d i_d}]_{i_1, i_2, \dots, i_d=1}^{n_1, n_2, \dots, n_d}$$

hipermátrixot értjük.

Számítsuk ki a következő Segre-féle külső szorzatot:

$$\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}.$$

## Projekt: a mátrixszorzás általánosítása hiper mátrixokra\*

A mátrixok szorzatát általánosíthatjuk az alábbi módon: legyen  $\mathbf{A} \in S^{n_1 \times \dots \times n_{d-1} \times k}$  egy  $d$ -edrendű és  $\mathbf{B} \in S^{k \times m_2 \times \dots \times m_e}$  egy  $e$ -edrendű hiper mátrix, ahol  $\mathbf{A}$  utolsó és  $\mathbf{B}$  első indexe ugyanazon az  $\{1, \dots, k\}$  tartományon fut végig.  $\mathbf{A}$  és  $\mathbf{B}$  szorzatán azt a  $(d + e - 2)$ -edrendű  $\mathbf{C} \in S^{n_1 \times \dots \times n_{d-1} \times m_2 \times \dots \times m_e}$  hiper mátrixot értjük, melyre

$$c_{i_1 \dots i_{d-1} j_2 \dots j_e} = \sum_{\ell=1}^k a_{i_1 \dots i_{d-1} \ell} b_{\ell j_2 \dots j_e}.$$

Megkülönböztetésül e szorzatot  $\mathbf{A}$  és  $\mathbf{B}$  kontrakciós szorzatának nevezzük. Ez azonban nem csak  $\mathbf{A}$  utolsó és  $\mathbf{B}$  első indexére, hanem bármely indexpárra értelmezhető, ha azokhoz azonos indextartomány tartozik.

4.16. KONTRAKCIÓS SZORZAT Defináljuk az  $\mathbf{A}$  és  $\mathbf{B}$  hiper mátrixok  $\langle \mathbf{A}, \mathbf{B} \rangle_{u,v}$ -val jelölt kontrakciós szorzatát, ahol az  $\mathbf{A}$  mátrix  $u$ -adik és a  $\mathbf{B}$  mátrix  $v$ -edik indextartománya azonos.

4.17. Számítsuk ki az

$$\left\langle \left[ \begin{array}{ccc|cc} 1 & 0 & 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1' \end{array} \right] \left[ \begin{array}{cc} 2 & 1 \\ 1 & 0 \\ 0 & 1 \end{array} \right] \right\rangle_{2,1}$$

kontrakciós szorzatot!

## A mátrixszorzás használata

Mátrixszorzással az eddig tanultak áttekinthetőbbé és könnyebben kezelhetővé válnak (pl. vektorok lineáris kombinációja, az egyenletrendszerek és megoldásuk felírása).

**Skaláris szorzat és diadikus szorzat mátrixszorzatos alakja** Két oszlopvektor nem szorozható össze, ha 1-nél nagyobb dimenziósak. Viszont az egyikük transzponálása után a szorzás elvégezhető.

Legyen  $\mathbf{a}$  és  $\mathbf{b}$  két  $\mathbb{R}^n$ -beli vektor. Az  $\mathbf{a}^T \mathbf{b}$  szorzat a két vektor skaláris szorzatát adja, azaz

$$\mathbf{a}^T \mathbf{b} = \mathbf{a} \cdot \mathbf{b},$$

ugyanis

$$\mathbf{a}^T \mathbf{b} = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = a_1 b_1 + a_2 b_2 + \dots + a_n b_n = \mathbf{a} \cdot \mathbf{b}.$$

Ha a második vektort transzponáljuk, a két vektor lehet különböző dimenziós is.

**4.18. DEFINÍCIÓ (DIADIKUS SZORZAT).** Legyen  $\mathbf{u} \in \mathbb{R}^m$ ,  $\mathbf{v} \in \mathbb{R}^n$ . Az  $\mathbf{u} \mathbf{v}^T$  szorzatot a két vektor diadikus szorzatának, röviden diádnak nevezzük. E szorzat egy  $m \times n$ -es mátrix:

$$\mathbf{u} \mathbf{v}^T = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix} \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} = \begin{bmatrix} u_1 v_1 & u_1 v_2 & \dots & u_1 v_n \\ u_2 v_1 & u_2 v_2 & \dots & u_2 v_n \\ \vdots & \vdots & \ddots & \vdots \\ u_m v_1 & u_m v_2 & \dots & u_m v_n \end{bmatrix}.$$

Két vektor diadikus szorzatát  $\mathbf{u} \otimes \mathbf{v}$  jelöli.

**4.19. PÉLDA (SKALÁRIS ÉS DIADIKUS SZORZAT).** Legyen  $\mathbf{u} = (1, 0, 2)$ ,  $\mathbf{v} = (3, 2, 1)$ . Írjuk fel mátrixszorzatos alakba skaláris és diadikus szorzatukat, és számítsuk ki!

**MEGOLDÁS.**

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = \begin{bmatrix} 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = 5,$$

$$\mathbf{u} \otimes \mathbf{v} = \mathbf{u} \mathbf{v}^T = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \begin{bmatrix} 3 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 1 \\ 0 & 0 & 0 \\ 6 & 4 & 2 \end{bmatrix} \quad \square$$

**Lineáris egyenletrendszer mátrixszorzatos alakja** A mátrixszorzást felhasználva a lineáris egyenletrendszerek egyszerű alakba írhatók.

**4.1 (LINEÁRIS EGYENLETRENDSZER MÁTRIXSZORZATOS ALAKJA).** Ha  $\mathbf{A}$  jelöli egy egyenletrendszer együtthatómátrixát, illetve  $\mathbf{b}$  a konstans tagok és  $\mathbf{x}$  az ismeretlenek oszlopvektorát, azaz

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \text{és} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

akkor az

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

egyenletrendszer

$$\mathbf{Ax} = \mathbf{b}$$

alakba írható.

Könnyen ellenőrizhető a mátrixszorzás elvégzésével, hogy a

$$\begin{array}{rclcl} ax & = & u & & x + 2y = 1 \\ 2x_1 + 3x_2 - x_3 = 5, & by & = & v & \text{és} & y = 1 \\ & cz & = & w & & 0 = 1 \end{array}$$

egyenletrendszerek mátrixszorzatos alakjai rendre:

$$\begin{bmatrix} 2 & 3 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 5, \quad \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} u \\ v \\ w \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

A szimultán egyenletrendszerek (ld. 97. oldal) ugyanígy fölírhatóak mátrixszorzatos alakba.

**4.20. PÉLDA (SZIMULTÁN EGYENLETRENDSZER MÁTRIXSZORZATOS ALAKJA).** Írjuk az alábbi két egyenletrendszert egyetlen mátrixszorzatos alakba!

$$\begin{array}{rcl} 2x_{11} + 3x_{21} & = & 7 \\ 3x_{11} - 4x_{21} & = & 2 \end{array} \quad \begin{array}{rcl} 2x_{12} + 3x_{22} & = & 9 \\ 3x_{12} - 4x_{22} & = & 5 \end{array}$$

**MEGOLDÁS.** A két egyenletrendszer mátrixszorzatos alakjai külön-külön

$$\begin{bmatrix} 2 & 3 \\ 3 & -4 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix} = \begin{bmatrix} 7 \\ 2 \end{bmatrix}, \quad \begin{bmatrix} 2 & 3 \\ 3 & -4 \end{bmatrix} \begin{bmatrix} x_{12} \\ x_{22} \end{bmatrix} = \begin{bmatrix} 9 \\ 5 \end{bmatrix}.$$

Ezek egyetlen mátrixszorzattá olvashatók:

$$\begin{bmatrix} 2 & 3 \\ 3 & -4 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 7 & 9 \\ 2 & 5 \end{bmatrix}.$$

Általánosan a szimultán egyenletrendszerek  $\mathbf{AX} = \mathbf{B}$  alakba írhatóak, ahol  $\mathbf{X}$  az ismeretlenekből,  $\mathbf{B}$  a jobb oldalakból alkotott mátrix.  $\square$

**Lineáris helyettesítés mátrixszorzatos alakja** Az egyenletrendszer mátrixszorzatos alakjához hasonlóan adódik a lineáris helyettesítés mátrixszorzatos alakja. Egyszerűen csak úgy kell tekintenünk az  $\mathbf{Ax} = \mathbf{b}$  egyenlőségre, hogy ott  $\mathbf{b}$  koordinátái a helyettesítendő változók, melyek helyébe az  $\mathbf{x}$  koordinátáinak egy lineáris kifejezését helyettesítjük. Ilyenkor inkább a  $\mathbf{b} = \mathbf{Ax}$  alakot használjuk, és az  $\mathbf{A}$  mátrixot a *lineáris helyettesítés mátrixának* nevezzük. Részletesebben lásd még a 4.37. feladatot.

Példaként íme egy lineáris helyettesítés és mátrixszorzatos alakja:

$$\begin{aligned} x &= 3a + 2b + 4c \\ y &= a - 3b + 2c \\ z &= 2a - b + 2c \end{aligned} \quad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 & 2 & 4 \\ 1 & -3 & 2 \\ 2 & -1 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

**Szorzás vektorral** Egy  $m \times n$ -es mátrix vektorral kétféleképp szorozható: jobbról egy  $n \times 1$ -es oszlopvektorral, balról egy  $1 \times m$ -es sorvektorral.

Az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszer oszlopmodelljéből láttuk, hogy az egyenletrendszer bal oldala az  $\mathbf{A}$  oszlopvektorainak az  $\mathbf{x}$  koordinátaival vett lineáris kombinációja. Hasonló állítás igaz a sorvektorral balról való szorzásra is.

**4.21. ÁLLÍTÁS (MÁTRIXSZORZÁS ÉS LINEÁRIS KOMBINÁCIÓ).** Legyen  $\mathbf{A}$   $m \times n$ -es mátrix,  $\mathbf{x}$   $n$ -dimenziós,  $\mathbf{y}$   $m$ -dimenziós vektor. Ekkor az  $\mathbf{Ax}$  szorzat az  $\mathbf{A}$  oszlopvektorainak

$$\mathbf{a}_{*1}x_1 + \mathbf{a}_{*2}x_2 + \cdots + \mathbf{a}_{*n}x_n$$

lineáris kombinációját, míg az  $\mathbf{y}^T \mathbf{A}$  szorzat az  $\mathbf{A}$  sorvektorainak

$$\mathbf{a}_{1*}y_1 + \mathbf{a}_{2*}y_2 + \cdots + \mathbf{a}_{m*}y_m$$

lineáris kombinációját adja.

**Szorzás standard egységvektorral** Könnyen igazolhatók azok az összefüggések, melyeket a standard egységvektorokkal való szorzással kapunk. Jelölje  $\mathbf{e}_i = (0, 0, \dots, 1, \dots, 0)$  azt a vektort, melynek  $i$ -edik koordinátája 1, a többi 0.

4.22. ÁLLÍTÁS (MÁTRIX ELEMEINEK, SOR- ÉS OSZLOPVEKTORAINAK ELŐÁLLÍTÁSA). Legyen  $\mathbf{A}$  egy  $m \times n$ -es mátrix,  $\mathbf{e}_i$   $m$ -dimenziós,  $\mathbf{e}_j$   $n$ -dimenziós standard egységvektor. Ekkor a standard  $\mathbf{e}_i$  sorvektorral balról való szorzás a mátrix  $i$ -edik sorvektorát, az  $\mathbf{e}_j$ -vel jobbról való szorzás a mátrix  $j$ -edik oszlopvektorát adja, azaz

$$\mathbf{e}_i^T \mathbf{A} = \mathbf{a}_{i*} \text{ és } \mathbf{A} \mathbf{e}_j = \mathbf{a}_{*j},$$

továbbá

$$\mathbf{e}_i^T (\mathbf{A} \mathbf{e}_j) = (\mathbf{e}_i^T \mathbf{A}) \mathbf{e}_j = a_{ij}.$$

Az  $\mathbf{e}_i \mathbf{e}_j^T$  diád egy olyan mátrix, melynek  $(i, j)$ -indexű eleme 1, az összes többi 0:

$$\mathbf{e}_i \mathbf{e}_j^T = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \begin{bmatrix} 0 & \dots & 1 & \dots & 0 \end{bmatrix} = \begin{bmatrix} 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{bmatrix}.$$

**A báziscsere mátrixszorzatos alakja** Ha egy vektornak két különböző bázisban is meg van adva a koordinátás alakja, akkor az egyikből a másikat egy egyszerű mátrixszorzással is megkaphatjuk.

4.23. PÉLDA (ÁTTÉRÉS STANDARD BÁZISRA). Az  $\mathbb{R}^3$  térnek  $\mathcal{B} = \{ (1, 2, 3), (0, 2, 3), (3, 5, 8) \}$  egy bázisa. Az  $\mathcal{B}$  bázisban megadott  $[\mathbf{v}]_{\mathcal{B}}$  vektornak írjuk fel a koordinátás alakját a standard bázisban egyetlen mátrixszorzással. Mi a  $\mathbf{v}$  vektor standard koordinátás alakja, ha  $[\mathbf{v}]_{\mathcal{B}} = (3, 2, -1)$ ?

**MEGOLDÁS.**  $[\mathbf{v}]_{\mathcal{B}} = (3, 2, -1)$  azt jelenti, hogy

$$\mathbf{v} = 3 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 3 \\ 5 \\ 8 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \\ 7 \end{bmatrix},$$

ami mátrixszorzatos alakban

$$\mathbf{v} = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 2 & 5 \\ 3 & 3 & 8 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \\ 7 \end{bmatrix}.$$

Legyen  $[\mathbf{v}]_{\mathcal{B}} = (x, y, z)$ . Ekkor

$$\mathbf{v} = x \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + y \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix} + z \begin{bmatrix} 3 \\ 5 \\ 8 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 2 & 5 \\ 3 & 3 & 8 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}. \quad \square$$

E példa a következő definícióhoz és állításhoz vezet:

**4.24. DEFINÍCIÓ (ÁTTÉRÉS MÁTRIXA).** Legyen  $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$  és  $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n\}$  az  $\mathcal{V}$  vektortér két bázisa. Írjuk fel  $\mathcal{B}$  vektorainak koordinátás alakját a  $\mathcal{C}$  bázisban, majd képezzünk ezekből mátrixot. Ezt nevezzük a  $\mathcal{B}$  bázisról a  $\mathcal{C}$ -re való áttérés mátrixának, melynek alakja tehát

$$\mathbf{A}_{\mathcal{C} \leftarrow \mathcal{B}} = [ [\mathbf{b}_1]_{\mathcal{C}} \mid [\mathbf{b}_2]_{\mathcal{C}} \mid \dots \mid [\mathbf{b}_n]_{\mathcal{C}} ]$$

(Érdemes a mátrix indexébe írni a két bázist  $\mathcal{C} \leftarrow \mathcal{B}$  alakban.)

**4.25. ÁLLÍTÁS (KOORDINÁTÁK VÁLTOZÁSA A BÁZIS CSERÉJÉNÉL).** Ha  $\mathcal{B}$  és  $\mathcal{C}$  a  $\mathcal{V}$  vektortér két (véges sok vektorból álló) bázisa és  $\mathbf{A}_{\mathcal{C} \leftarrow \mathcal{B}}$  az áttérés mátrixa, akkor bármely  $\mathbf{v}$  vektor  $\mathcal{B}$ -, illetve  $\mathcal{C}$ -beli koordinátás alakja közt fennáll a

$$[\mathbf{v}]_{\mathcal{C}} = \mathbf{A}_{\mathcal{C} \leftarrow \mathcal{B}} [\mathbf{v}]_{\mathcal{B}}$$

összefüggés.

**BIZONYÍTÁS.** Legyen  $[\mathbf{v}]_{\mathcal{B}} = (v_1, v_2, \dots, v_n)$ . A koordinátás alak jelentése szerint

$$\mathbf{v} = v_1 \mathbf{b}_1 + v_2 \mathbf{b}_2 + \dots + v_n \mathbf{b}_n.$$

Ennek koordinátás alakja a  $\mathcal{C}$  bázisban

$$\begin{aligned} [\mathbf{v}]_{\mathcal{C}} &= v_1 [\mathbf{b}_1]_{\mathcal{C}} + v_2 [\mathbf{b}_2]_{\mathcal{C}} + \dots + v_n [\mathbf{b}_n]_{\mathcal{C}} \\ &= [ [\mathbf{b}_1]_{\mathcal{C}} \mid [\mathbf{b}_2]_{\mathcal{C}} \mid \dots \mid [\mathbf{b}_n]_{\mathcal{C}} ] [\mathbf{v}]_{\mathcal{B}} \\ &= \mathbf{A}_{\mathcal{C} \leftarrow \mathcal{B}} [\mathbf{v}]_{\mathcal{B}}. \end{aligned} \quad \square$$

► A 4.23. példában a  $\mathcal{B}$  bázisról a standard  $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$  bázisra tértünk át, tehát az áttérés mátrixát jelölheti  $\mathbf{A}_{\mathcal{E} \leftarrow \mathcal{B}}$ .

► Az áttérés mátrixa felírható akkor is, ha  $\mathcal{C}$  a  $\mathcal{W}$  vektortérnek,  $\mathcal{B}$  a  $\mathcal{V} \leq \mathcal{W}$  altérnek egy bázisa. Erre mutat példát a 4.39. feladat.

**Bázisfelbontás\*** A 3.24. tétel második pontjának az előző feladatban is szemléltetett egyenes következménye, a következő állítás.

**4.26. ÁLLÍTÁS (BÁZISFELBONTÁS).** Jelölje egy  $m \times n$ -es  $\mathbf{A}$  mátrix redukált lépcsős alakjának nemzérus soraiból álló  $r \times n$ -es részmátrixát  $\mathbf{R}$ , az  $\mathbf{R}$  főoszlopainak megfelelő  $\mathbf{A}$ -beli oszlopok alkotta  $m \times r$ -es részmátrixot  $\mathbf{B}$ , ahol  $r = r(\mathbf{A})$ . Ekkor az  $\mathbf{R}$  mátrix  $j$ -edik oszlopa megegyezik az  $\mathbf{A}$  mátrix

$j$ -edik oszlopának a  $\mathbf{B}$  oszlopai alkotta bázisban felírt koordinátás alakjával. Képletben ez azt jelenti, hogy

$$\mathbf{A}_{*j} = \mathbf{B}\mathbf{R}_{*j}, \quad \text{azaz} \quad \mathbf{A} = \mathbf{B}\mathbf{R}.$$

Egy mátrix fenti  $\mathbf{A} = \mathbf{B}\mathbf{R}$  alakú felbontását *bázisfelbontásnak* nevezzük.

**4.27. PÉLDA (BÁZISFELBONTÁS).** Határozzuk meg az alábbi mátrix bázisfelbontását, és magyarázzuk meg a két mátrix oszlopainak jelentését!

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 8 & 6 & 2 \\ 1 & 2 & 7 & 0 & -11 \end{bmatrix}.$$

**MEGOLDÁS.** Az  $\mathbf{A}$  mátrix redukált lépcsős alakja:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 8 & 6 & 2 \\ 1 & 2 & 7 & 0 & -11 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 0 & 7 & 17 \\ 0 & 0 & 1 & -1 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

E mátrix első két sora alkotja az  $\mathbf{R}$  mátrixot, az  $\mathbf{A}$  mátrix első és harmadik oszlopa a  $\mathbf{B}$  mátrixot, így a felbontás

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 8 & 6 & 2 \\ 1 & 2 & 7 & 0 & -11 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 2 & 8 \\ 1 & 7 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 & 7 & 17 \\ 0 & 0 & 1 & -1 & -4 \end{bmatrix} = \mathbf{B}\mathbf{R}.$$

Ellenőrizzük, hogy az  $\mathbf{R}$  oszlopai az  $\mathbf{A}$  oszlopvektorainak koordinátás alakjai a  $\mathbf{B}$  oszlopai alkotta bázisban, azaz

$$[\mathbf{v}]_{\mathcal{E}} = \mathbf{B}[\mathbf{v}]_{\mathcal{B}},$$

ahol az  $\mathcal{E}$  indexszel a standard,  $\mathcal{B}$ -vel a  $\mathbf{B}$  mátrix oszlopai alkotta bázisbeli koordinátás alakját jelöltük ugyanannak a vektornak. Például

$$\begin{bmatrix} 4 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 2 & 8 \\ 1 & 7 \end{bmatrix} \begin{bmatrix} 7 \\ -1 \end{bmatrix}, \quad \text{azaz} \quad [\mathbf{a}_4]_{\mathcal{E}} = \begin{bmatrix} 4 \\ 6 \\ 0 \end{bmatrix}, \quad [\mathbf{a}_4]_{\mathcal{B}} = \begin{bmatrix} 7 \\ -1 \end{bmatrix}, \quad (4.3)$$

ahol  $\mathbf{a}_4$  az  $\mathbf{A}$  negyedik oszlopvektora. □

**Egységmátrix, elemi mátrixok** Egy adott  $\mathbf{B}$  mátrixhoz találhatunk olyan  $\mathbf{A}$ -t, hogy az 1-gyel való szorzáshoz hasonlóan  $\mathbf{A}\mathbf{B} = \mathbf{B}$  legyen. Például

$$\mathbf{A} = \begin{bmatrix} 3 & -4 \\ -2 & 5 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix}$$

esetén

$$\begin{bmatrix} 3 & -4 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix}.$$

Az azonban már nem igaz, hogy  $\mathbf{A}$ -t bármely  $2 \times 2$ -es  $\mathbf{B}$  mátrixszal szorozva  $\mathbf{B}$  lesz az eredmény. Ilyen mátrix is létezik, némi próbálkozás után bárki rátalálhat.

4.28. DEFINÍCIÓ (EGYSÉGMÁTRIX). Az  $n \times n$ -es

$$\mathbf{I}_n := \text{diag}(1, 1, \dots, 1) = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

mátrixot egységmátrixnak nevezzük.

► Az egységmátrix elnevezés onnan származik, hogy bármely  $m \times n$ -es  $\mathbf{A}$  mátrixra igaz, hogy

$$\mathbf{I}_m \mathbf{A}_{m \times n} = \mathbf{A}_{m \times n} \mathbf{I}_n = \mathbf{A}_{m \times n},$$

azaz e mátrix hasonló tulajdonsággal rendelkezik, mint a számok közt az egy.

► Az egységmátrixszal már találkoztunk: a Gauss–Jordan-módszernél egy  $n$ -ismeretlenes,  $n$  egyenletből álló egyértelműen megoldható egyenletrendszer együtthatómátrixa az elemi sorműveletek során egységmátrixszá transzformálódik!

Az egységmátrixon végrehajtott elemi sorműveletek olyan mátrixokat eredményeznek, melyek kapcsolatot létesítenek az elemi sorműveletek és a mátrixokkal való szorzás között.

4.29. DEFINÍCIÓ (ELEMÍ MÁTRIXOK). Az  $\mathbf{I}_n$  egységmátrixon végrehajtott egyetlen elemi sorművelettel kapott mátrixot elemi mátrixnak nevezzük.

Az alábbi mátrixok elemi mátrixok:

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Ezt igazolja, hogy mindegyikük  $\mathbf{I}_4$ -ből származik rendre a következő elemi sorműveletekkel:  $S_1 \leftrightarrow S_4$ ,  $5S_2$ ,  $S_1 + 2S_3$ ,  $1S_1$ . Az utolsó mátrix az egységmátrix, amely maga is elemi mátrix, mert például egy sorának 1-gyel való szorzásával megkapható.

Az egységmátrix jelölésére használt  $\mathbf{I}$  betű az angol *identity matrix* elnevezés első betűjéből származik. Az *azonosság* vagy *identitás* jelentésű *identity* szó az  $\mathbf{I}\mathbf{A} = \mathbf{A}$  összefüggésre utal (az  $x \mapsto x$  függvényt ugyanezen okból hívjuk identikus függvénynek). Ráadásul az  $\mathbf{I}$  betű hasonlít legjobban az 1-es számra.



4.30. PÉLDA (MÁTRIX BALRÓL SZORZÁSA ELEMI MÁTRIXSZAL). Vizsgáljuk meg mi történik, ha az előbbi mátrixokkal balról megszorozunk egy tetszőleges 4-soros mátrixot?

**MEGOLDÁS.** Legyen  $\mathbf{A}$  egy 4- sorból, és az egyszerűség kedvéért csak 2 oszlopból álló mátrix.

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} = \begin{bmatrix} a_{41} & a_{42} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{11} & a_{12} \end{bmatrix},$$

A szorzás eredményként fölcserélődött  $\mathbf{A}$  első és negyedik sora.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ 5a_{21} & 5a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix},$$

Itt az  $\mathbf{A}$  második sora be lett szorozva 5-tel.

$$\begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} = \begin{bmatrix} a_{11} + 2a_{31} & a_{12} + 2a_{32} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix},$$

A szorzás eredményként az  $\mathbf{A}$  első sorához hozzá lett adva harmadik sorának kétszerese.  $\square$

E példa eredménye kimondható tételként, melynek bizonyítása általánosan is úgy történik, mint az előző példában, ezért elhagyjuk:

4.31. TÉTEL (ELEMI SORMŰVELETEK MÁTRIXSZORZÁSSAL). Legyen  $\mathbf{E}$  az elemi mátrix, melyet  $\mathbf{I}_m$ -ből egy elemi sorművelettel kapunk. Ha ugyanezt a sorműveletet egy tetszőleges  $m \times n$ -es  $\mathbf{A}$  mátrixra alkalmazzuk, akkor eredményül az  $\mathbf{EA}$  mátrixot kapjuk.

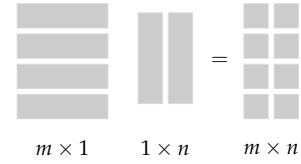
► Az elemi sorműveletek mátrixszorzással való elvégzésének nincs számítási praktikuma, annak célja az elemi sorműveletek algebraizálása, s így mélyebb megértése.

**Vektorokra particionált mátrixok** Fontosak azok a blokkmátrixok, amelyekben vagy oszlopvektor vagy sorvektor minden blokk. Két mátrix szorzatának felírásakor legalább az egyiküket vektorokra particionálására 4 eset lehetséges. Legyen a következőkben  $\mathbf{A}$  egy  $m \times t$ ,  $\mathbf{B}$  egy  $t \times n$  méretű mátrix.

1. [sorvektorok]  $\cdot$  [oszlopvektorok]: Bontsuk fel az  $\mathbf{A}_{m \times t}$  mátrixot sorvektoraira, és a  $\mathbf{B}_{t \times n}$  mátrixot oszlopvektoraira. Ekkor egy  $m \times 1$ -es

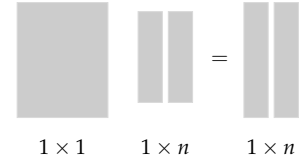
blokkmátrixot szorzunk egy  $1 \times n$ -essel, ami épp az  $\mathbf{AB}$  mátrixszorzat definícióját adja:

$$\mathbf{AB} = \begin{bmatrix} \mathbf{a}_{1*} \\ \mathbf{a}_{2*} \\ \vdots \\ \mathbf{a}_{m*} \end{bmatrix} \begin{bmatrix} \mathbf{b}_{*1} & \dots & \mathbf{b}_{*n} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_{1*}\mathbf{b}_{*1} & \mathbf{a}_{1*}\mathbf{b}_{*2} & \dots & \mathbf{a}_{1*}\mathbf{b}_{*n} \\ \mathbf{a}_{2*}\mathbf{b}_{*1} & \mathbf{a}_{2*}\mathbf{b}_{*2} & \dots & \mathbf{a}_{2*}\mathbf{b}_{*n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_{m*}\mathbf{b}_{*1} & \mathbf{a}_{m*}\mathbf{b}_{*2} & \dots & \mathbf{a}_{m*}\mathbf{b}_{*n} \end{bmatrix}.$$



2.  $[\text{mátrix}] \cdot [\text{oszlopvektorok}]$ : Ekkor egy  $1 \times 1$ -es blokkmátrixot szorzunk egy  $1 \times n$ -essel:

$$\mathbf{C} = \mathbf{AB} = \mathbf{A} \begin{bmatrix} \mathbf{b}_{*1} & \mathbf{b}_{*2} & \dots & \mathbf{b}_{*n} \end{bmatrix} = \begin{bmatrix} \mathbf{Ab}_{*1} & \mathbf{Ab}_{*2} & \dots & \mathbf{Ab}_{*n} \end{bmatrix}$$



Itt tehát a  $\mathbf{C}$  mátrix  $j$ -edik oszlopvektora az  $\mathbf{A}$  mátrix és a  $\mathbf{B}$   $j$ -edik oszlopának szorzata, vagyis  $\mathbf{c}_{*j} = \mathbf{Ab}_{*j}$ . Sematikusan ábrázolva:

$$\begin{bmatrix} \mathbf{A} \\ \vdots \\ \mathbf{a}_{*j} \end{bmatrix} \begin{bmatrix} \mathbf{B} \\ \vdots \\ \mathbf{b}_{*j} \end{bmatrix} = \begin{bmatrix} \mathbf{C} \\ \vdots \\ \mathbf{c}_{*j} \end{bmatrix}$$

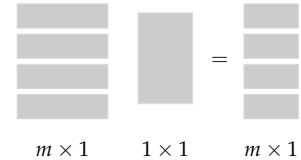
Ezzel az esettel már találkoztunk a szimultán egyenletrendszerek mátrixszorzatos alakjának fölírásánál (4.20. példa). Ha a fenti sematikus ábra egy szimultán egyenletrendszer mátrixszorzatos alakját reprezentálja, akkor a színesen kiemelt rész a szimultán egyenletrendszer egyetlen egyenletrendszerének felel meg.

3.  $[\text{sorvektorok}] \cdot [\text{mátrix}]$ : Ekkor egy  $m \times 1$ -es blokkmátrixot szorzunk egy  $1 \times 1$ -essel:

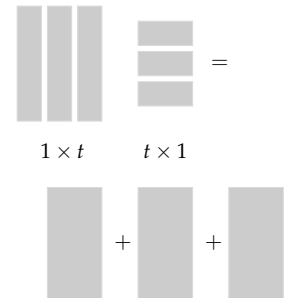
$$\mathbf{C} = \mathbf{AB} = \begin{bmatrix} \mathbf{a}_{1*} \\ \mathbf{a}_{2*} \\ \vdots \\ \mathbf{a}_{m*} \end{bmatrix} \mathbf{B} = \begin{bmatrix} \mathbf{a}_{1*}\mathbf{B} \\ \mathbf{a}_{2*}\mathbf{B} \\ \vdots \\ \mathbf{a}_{m*}\mathbf{B} \end{bmatrix}$$

Azaz itt a  $\mathbf{C} = \mathbf{AB}$  mátrix  $i$ -edik sora az  $\mathbf{A}$  mátrix  $i$ -edik sorának  $\mathbf{B}$ -szerese. Másként írva  $\mathbf{c}_{i*} = \mathbf{a}_{i*}\mathbf{B}$ , sematikusan ábrázolva:

$$\begin{bmatrix} \mathbf{A} \\ \vdots \\ \mathbf{a}_{i*} \end{bmatrix} \begin{bmatrix} \mathbf{B} \\ \vdots \\ \mathbf{b}_{*j} \end{bmatrix} = \begin{bmatrix} \mathbf{C} \\ \vdots \\ \mathbf{c}_{i*} \end{bmatrix}$$



4.  $[\text{oszlopvektorok}] \cdot [\text{sorvektorok}]$ : Ekkor egyetlen blokk sorból álló mátrixot szorzunk egy blokk oszlopból állóval, azaz egy  $1 \times t$ -es blokkmátrixot egy  $t \times 1$ -essel. Így egy skaláris szorzatra emlékeztető összeget kapunk:



$$\mathbf{AB} = \left[ \begin{array}{c|c|c} \mathbf{a}_{*1} & \dots & \mathbf{a}_{*t} \end{array} \right] \left[ \begin{array}{c} \mathbf{b}_{1*} \\ \mathbf{b}_{2*} \\ \dots \\ \mathbf{b}_{t*} \end{array} \right] = \mathbf{a}_{*1}\mathbf{b}_{1*} + \mathbf{a}_{*2}\mathbf{b}_{2*} + \dots + \mathbf{a}_{*t}\mathbf{b}_{t*}.$$

E felbontásban az  $\mathbf{AB}$  mátrixot *diádok összegére* bontottuk! Például az alábbi szorzat diádokra való bontása a következő:

$$\begin{aligned} \begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -2 & 0 \\ 1 & 1 \end{bmatrix} &= \begin{bmatrix} 0 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 4 \end{bmatrix} \begin{bmatrix} -2 & 0 \end{bmatrix} + \begin{bmatrix} 2 \\ 5 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 3 & 3 \end{bmatrix} + \begin{bmatrix} -2 & 0 \\ -8 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 2 \\ 5 & 5 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 0 & 8 \end{bmatrix}. \end{aligned}$$

Tehát a szorzatot három diád összegére bontottuk. Megjegyezzük, az eredmény maga is egy diád, hisz  $\begin{bmatrix} 0 & 2 \\ 0 & 8 \end{bmatrix} = \begin{bmatrix} 2 \\ 8 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix}$ .

E felbontásnak fontos speciális esete az, amikor  $\mathbf{A}$  egyetlen sorból, vagy  $\mathbf{B}$  egyetlen oszlopból áll. Tekintsük az előző példában szereplő  $\mathbf{A}$  mátrix első sorát, szorozzuk meg  $\mathbf{B}$ -vel, majd az  $\mathbf{A}$ -t a  $\mathbf{B}$  második oszlopával!

$$\begin{aligned} \begin{bmatrix} 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -2 & 0 \\ 1 & 1 \end{bmatrix} &= 0 \begin{bmatrix} 1 & 1 \end{bmatrix} + 1 \begin{bmatrix} -2 & 0 \end{bmatrix} + 2 \begin{bmatrix} 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 2 \end{bmatrix}, \\ \begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} &= 1 \begin{bmatrix} 0 \\ 3 \end{bmatrix} + 0 \begin{bmatrix} 1 \\ 4 \end{bmatrix} + 1 \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 2 \\ 8 \end{bmatrix}. \end{aligned}$$

Az első szorzat eredménye  $\mathbf{B}$  sorvektorainak, a második szorzat  $\mathbf{A}$  oszlopvektorainak lineáris kombinációja! Ez általánosan is kimondható:

**4.32. ÁLLÍTÁS (A SZORZAT OSZLOPAI ÉS SORAI).** Az  $\mathbf{AB}$  mátrix minden oszlopa az  $\mathbf{A}$  oszlopainak lineáris kombinációja, és minden sora a  $\mathbf{B}$  sorainak lineáris kombinációja.

**BIZONYÍTÁS.** Az  $\mathbf{AB}$  mátrix  $j$ -edik oszlopa

$$(\mathbf{AB})_{*j} = \mathbf{A}\mathbf{b}_{*j} = \mathbf{a}_{*1}b_{1j} + \mathbf{a}_{*2}b_{2j} + \dots + \mathbf{a}_{*t}b_{tj}$$

az  $i$ -edik sora pedig

$$(\mathbf{AB})_{i*} = \mathbf{a}_{i*}\mathbf{B} = a_{i1}\mathbf{b}_{1*} + a_{i2}\mathbf{b}_{2*} + \dots + a_{it}\mathbf{b}_{t*},$$

ami bizonyítja az állítást. □

**4.33. KÖVETKEZMÉNY (SZORZAT RANGJA).**  $r(\mathbf{AB}) \leq r(\mathbf{A})$  és  $r(\mathbf{AB}) \leq r(\mathbf{B})$ , így

$$r(\mathbf{AB}) \leq \min(r(\mathbf{A}), r(\mathbf{B})). \quad (4.4)$$

**BIZONYÍTÁS.** Az  $\mathbf{AB}$  független oszlopainak száma legföljebb annyi, mint  $\mathbf{A}$  független oszlopainak száma, hisz oszloptere az  $\mathbf{A}$  oszlopterének altere, és  $\mathbf{AB}$  független sorainak száma legföljebb annyi, mint  $\mathbf{B}$  független sorainak száma, hisz sortere  $\mathbf{B}$  sortérének altere.  $\square$

**Altér felírása mátrixszorzattal** Egy altér vektorai is fölírhatók egyetlen mátrixszorzással!

4.34. PÉLDA (NULLTÉR). Írjuk fel az

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 5 & 7 & 11 \\ 0 & 1 & 1 & 1 & -1 \end{bmatrix}$$

mátrix nullterének vektorait egy mátrix és egy vektor szorzataként!

**MEGOLDÁS.** A nulltér, azaz a mátrixhoz tartozó homogén lineáris egyenletrendszer megoldásainak tere könnyen leolvasható a redukált lépcsős alakból.

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 5 & 7 & 11 \\ 0 & 1 & 1 & 1 & -1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 1 & 2 & 7 \\ 0 & 1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

A szabad változókhoz rendelt paraméterek legyenek  $x_3 = t_1$ ,  $x_4 = t_2$ ,  $x_5 = t_3$ , amiből  $x_1 = -t_1 - 2t_2 - 7t_3$  és  $x_2 = -t_1 - t_2 + t_3$ . Innen

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = t_1 \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2 \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + t_3 \begin{bmatrix} -7 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 & -2 & -7 \\ -1 & -1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}. \quad \square$$

► Vegyük észre, hogy a redukált lépcsős alak és a megoldás blokk-szerkezete egyszerű kapcsolatot mutat:

$$\begin{bmatrix} \mathbf{I}_2 & \mathbf{S} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} -\mathbf{S} \\ \mathbf{I}_3 \end{bmatrix} \mathbf{t},$$

ahol  $\mathbf{t}$  a paraméterek vektora. Ez általánosítható tetszőleges homogén, és inhomogén lineáris egyenletrendszerekre (ld. a 4.49. feladatot).

## Feladatok

### Mátrixműveletek

4.18. **IGAZ – HAMIS** Döntsük el, igazak-e az alábbi állítások? Válaszunkat indokoljuk!

- Ha az  $\mathbf{AB}$  és a  $\mathbf{BA}$  szorzat is értelmezve van, akkor mindkét mátrix négyzetes.
- Ha az  $\mathbf{AB}$  és a  $\mathbf{BA}$  szorzat is értelmezve van, akkor mindkét szorzat négyzetes.
- Ha az  $(\mathbf{AB})\mathbf{C}$  szorzat értelmezve van, akkor biztosan értelmezve van az  $\mathbf{A}(\mathbf{BC})$  szorzat is.

A következőkben legyen

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 4 & 2 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 4 & 2 \\ 4 & 5 \end{bmatrix}, \mathbf{C} = \begin{bmatrix} -3 & 2 \\ 2 & -1 \end{bmatrix}, \mathbf{D} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}.$$

Végezzük el az alábbi műveleteket!

4.19.  $2\mathbf{A} - 3\mathbf{B}^T$

4.20.  $\mathbf{AB} - \mathbf{BA} + \mathbf{AC} - \mathbf{CA}$

4.21.  $(\mathbf{CD} - \mathbf{DC})(\mathbf{ABC})$

4.22.  $\mathbf{A}^2 - \mathbf{C}^2$

4.23.  $(\mathbf{C})_{2*}(\mathbf{D})_{*2}$

4.24.  $(\mathbf{A})_{*1}(\mathbf{B})_{2*}$

4.25. A fenti jelölések mellett igazak-e a következő egyenlőségek?

$$(\mathbf{A} + \mathbf{B})^2 = \mathbf{A}^2 + 2\mathbf{AB} + \mathbf{B}^2, \quad (\mathbf{A} + \mathbf{C})^2 = \mathbf{A}^2 + 2\mathbf{AC} + \mathbf{C}^2.$$

4.26. A fenti jelölések mellett igazak-e a következő egyenlőségek?

$$(\mathbf{C} + \mathbf{D})(\mathbf{C} - \mathbf{D}) = \mathbf{C}^2 - \mathbf{D}^2, \quad (\mathbf{A} + \mathbf{D})(\mathbf{A} - \mathbf{D}) = \mathbf{A}^2 - \mathbf{D}^2.$$

Számítsuk ki az alábbi vektorok skaláris és diadikus szorzatát! Írjuk fel mindkét műveletet mátrixszorzatos alakban!

4.27.  $\mathbf{a} = (1, 2)$ ,  $\mathbf{b} = (0, 1)$

4.28.  $\mathbf{u} = (1, 2, 0, 1)$ ,  $\mathbf{v} = (0, 1, 2, 3)$

4.29.  $\mathbf{a} = (1, 2, 0)$ ,  $\mathbf{b} = (0, 1, 2, 3)$

### Mátrixszorzatos alakok

Írjuk fel az alábbi egyenletrendszerek mátrixszorzatos alakját!

4.30.  $x + y = 1$

$$x - z = 2$$

$$z = 3$$

4.31.  $3x - 2y + 4z = 5$

4.32.  $2x + z = 1$

$$x - y - w = 2$$

$$y + z + w = 2$$

$$0 = 3$$

Írjuk fel az alábbi lineáris helyettesítések mátrixszorzatos alakját!

4.33.  $u = 2x - 4y$

$$v = x + 2y$$

4.34.  $x = 3a - 2b + c$

$$y = 2a - c$$

$$z = b + 2c$$

4.35.  $x = 3a + b$

$$y = 2a - b$$

$$z = b$$

4.36.  $x = 3a - 2b + c$

$$y = 2a - c$$

4.37. **LINEÁRIS HELYETTESÍTÉS MÁTRIXSZORZATOS ALAKJA** Írjuk fel az  $x_1, x_2, \dots, x_n$  változók lineáris kifejezéseinek az  $y_1, y_2, \dots, y_m$  változók helyébe való helyettesítését általánosan leíró

$$y_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n$$

$$y_2 = a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n$$

$$\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots$$

$$y_m = a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n$$

helyettesítés mátrixszorzatos alakját!

### Áttérés mátrixa

Írjuk fel az alábbi  $\mathcal{B}$  bázisról  $\mathcal{C}$ -re való áttérés mátrixát, ha  $\mathcal{B}$  vektorainak  $\mathcal{C}$ -beli koordinátás alakját ismerjük. Írjuk fel a megadott  $\mathbf{u}$  és  $\mathbf{v}$  vektorok  $\mathcal{C}$ -beli koordinátás alakját!

4.38.  $\mathcal{B} = \{(1, 1, 1), (0, 2, 2), (0, 0, 3)\}$ ,  $\mathcal{C}$  az  $\mathbb{R}^3$  standard bázisa,  $(\mathbf{u})_{\mathcal{B}} = (-1, 1, 1)$ ,  $(\mathbf{v})_{\mathcal{B}} = (-3, 2, 0)$ .

4.39. **ÁTTÉRÉS ALTÉRKRŐL**  $\mathcal{B} = \{(1, 1, 0, -2), (2, 3, 3, -2)\}$ ,  $\mathcal{C}$  az  $\mathbb{R}^4$  standard bázisa,  $(\mathbf{u})_{\mathcal{B}} = (-1, 1)$ ,  $(\mathbf{v})_{\mathcal{B}} = (-3, 2)$ .

Vessük össze e feladatot a 3.26., a 3.27. és a 3.29. példakkal!

### Elemi mátrixok

Keressük meg azt az  $\mathbf{E}$  mátrixot, mely megoldása az alábbi mátrixegyenletnek!

4.40.  $\mathbf{E} \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} = \begin{bmatrix} a & b \\ e & f \\ c & d \end{bmatrix}$

4.41.  $\mathbf{E} \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} = \begin{bmatrix} a & b \\ 3c & 3d \\ e & f \end{bmatrix}$

4.42.  $\mathbf{E} \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} = \begin{bmatrix} a & b \\ c + 2e & d + 2f \\ e & f \end{bmatrix}$

$$4.43. \mathbf{E} \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} = \begin{bmatrix} a-c & b-d \\ c & d \\ e & f \end{bmatrix}$$

Elemi sorműveletekkel, mátrixszorzás nélkül határozzuk meg az alábbi mátrixszorzatok értékét!

$$4.44. \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 & 2 \\ 3 & 3 & 3 \\ 4 & 4 & 4 \end{bmatrix}$$

$$4.45. \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

### Blokkmátrixok

Számítsuk ki az alábbi feladatokban megadott mátrixszorzatokat a kijelölt blokkmátrixokat használva!

$$4.46. \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 3 & 3 \end{bmatrix} \begin{bmatrix} 2 & 3 & 1 & 1 \\ 4 & 5 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$4.47. \begin{bmatrix} 2 & 3 & 1 & 1 \\ 4 & 5 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 3 \end{bmatrix}$$

$$4.48. \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ 3 & 3 & 3 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 3 & 3 \\ 2 & 3 \\ 3 & 4 \end{bmatrix}$$

- \* 4.49. LINEÁRIS EGYENLETRENDSZER MEGOLDÁSÁNAK BLOKK-MÁTRIX ALAKJA Tegyük fel, hogy az  $r$  rangú  $\mathbf{A}$  mátrix első  $r$  oszlopa lineárisan független – ez oszlopserékkel mindig

elérhető. Jelölje  $\mathbf{B}_r$  az  $\mathbf{A}$  első  $r$  oszlopából álló mátrixot, és legyen  $\mathbf{A}$ , illetve az  $[\mathbf{A}|\mathbf{b}]$  bővített mátrix redukált lépcsős alakja

$$\begin{bmatrix} \mathbf{I}_r & \mathbf{S} \\ \mathbf{O} & \mathbf{O} \end{bmatrix} \text{ illetve } \begin{bmatrix} \mathbf{I}_r & \mathbf{S} & \mathbf{d}_r \\ \mathbf{O} & \mathbf{O} & \mathbf{0} \end{bmatrix},$$

ahol  $\mathbf{d}_r$  egy  $r$ -dimenziós vektor. Ekkor

1. az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszer megoldható, és megoldása

$$\mathbf{x} = \begin{bmatrix} \mathbf{d}_r \\ \mathbf{0}_s \end{bmatrix} + \begin{bmatrix} -\mathbf{S} \\ \mathbf{I}_s \end{bmatrix} \mathbf{t}_s,$$

ahol  $s$  a szabad változók száma, azaz  $s = n - r$ , és  $\mathbf{t}_s$  a szabad paraméterek vektora, ráadásul  $\mathbf{A} = \mathbf{B}_r[\mathbf{I}_r|\mathbf{S}]$  és  $\mathbf{b} = \mathbf{B}_r\mathbf{d}_r$ , továbbá

2. az  $\mathbf{Ax} = \mathbf{0}$  homogén lineáris egyenletrendszer megoldása

$$\mathbf{x} = \begin{bmatrix} -\mathbf{S} \\ \mathbf{I}_s \end{bmatrix} \mathbf{t}_s,$$

ahol a  $[\begin{smallmatrix} -\mathbf{S} \\ \mathbf{I}_s \end{smallmatrix}]$  mátrix oszlopvektorai a nulltér bázisát alkotják.

### Vegyes feladatok

4.50. A *sudoku* egy olyan logikai játék, melyben egy olyan  $9 \times 9$ -es mátrixot kell megadni, melynek ismerjük néhány, de nem minden elemét. A feladat a nem ismert elemek meghatározása. A mátrix 9 darab  $3 \times 3$ -as blokkra van partícionálva és elegendő teszt annak a feltételnek, hogy minden sorában, minden oszlopában és minden blokkjában az 1-től 9-ig terjedő egészek mindegyike egyszer szerepel. Ez azt jelenti, hogy az egy sorban, egy oszlopban és egy blokkban lévő számok összege mindig 45. Fejezzük ki ezt mátrixműveletekkel, azaz írjunk fel a sudoku tábla  $\mathbf{A}$  mátrixát is tartalmazó olyan mátrixegyenleteket, melyeket minden helyesen kitöltött sudoku tábla mátrixa kielégít!

4.51. Hány eleme van a  $\mathbb{Z}_2^{2 \times 2}$ -nek, azaz a kételemű test fölötti  $2 \times 2$ -es mátrixok terének?

## Megoldások

4.1. A két táblázat szorzata:

|      | csarnok | hipermarket | piac |
|------|---------|-------------|------|
| Anti | 1350    | 1250        | 1210 |
| Bori | 1425    | 1245        | 1225 |
| Cili | 960     | 830         | 850  |

Tehát Antinak és Borinak a piacon, Cilinek a hipermarketben érdemes vásárolnia.

4.2. A két helyettesítést elvégezve:

$$\begin{aligned}x &= 2a + b = 2(-3s + t) + (4s - t) = -2s + t, \\y &= 3a + b = 3(-3s + t) + (4s - t) = -5s + 2t.\end{aligned}$$

A két helyettesítés kompozíciója a két helyettesítés táblázatának szorzatával megkapható:

$$\begin{array}{c|cc} & a & b \\ \hline x & 2 & 1 \\ y & 3 & 1 \end{array} \times \begin{array}{c|cc} & s & t \\ \hline a & -3 & 1 \\ b & 4 & -1 \end{array} = \begin{array}{c|cc} & s & t \\ \hline x & -2 & 1 \\ y & -5 & 2 \end{array}$$

4.3. A két helyettesítés kompozíciója a két helyettesítés táblázatának szorzatával megkapható. E szorzatból az olvasható le, hogy a kompozícióval kapott helyettesítés:  $x = s$ ,  $y = t$ ,  $z = u$ . Ez azt jelenti, hogy a két helyettesítés valamilyen értelemben egymás inverze.

4.4.

1. Kétféleképp adhatjuk meg a táblázatot, ha az első sor és oszlop a  $tv_1$ -é:

$$\begin{array}{c|cc} -re & tv_1 & tv_2 \\ \hline tv_1 & 1/2 & 1/4 \\ tv_2 & 1/2 & 3/4 \end{array} \quad \begin{array}{c|cc} -ról & tv_1 & tv_2 \\ \hline tv_1 & 1/2 & 1/2 \\ tv_2 & 1/4 & 3/4 \end{array}$$

2. A nézők kezdeti eloszlásának táblázatára a két lehetőség:

$$\begin{array}{c|c} & arány \\ \hline tv_1 & 1/2 \\ tv_2 & 1/2 \end{array} \quad \begin{array}{c|cc} & tv_1 & tv_2 \\ \hline arány & 1/2 & 1/2 \end{array}$$

3. Először válaszoljuk meg a kérdést a  $tv_1$ -re: a saját nézőinek fele marad ( $\frac{1}{2} \cdot \frac{1}{2}$ ), ehhez jön a  $tv_2$  nézőinek negyede ( $\frac{1}{2} \cdot \frac{1}{4}$ ), ez összesen  $\frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{4} = \frac{3}{8}$ . A  $tv_2$ -re a számítás:  $\frac{1}{2} \cdot \frac{3}{4} + \frac{1}{2} \cdot \frac{1}{2} = \frac{5}{8}$ . Ez táblázatok szorzásával az előző 2-2

felírást használva:

$$\begin{array}{c|cc} -re & tv_1 & tv_2 \\ \hline tv_1 & 1/2 & 1/4 \\ tv_2 & 1/2 & 3/4 \end{array} \times \begin{array}{c|cc} arány & & \\ \hline tv_1 & 1/2 & \\ tv_2 & 1/2 & \end{array} = \begin{array}{c|cc} arány & & \\ \hline tv_1 & 3/8 & \\ tv_2 & 5/8 & \end{array}$$

$$\begin{array}{c|cc} & tv_1 & tv_2 \\ \hline arány & 1/2 & 1/2 \end{array} \times \begin{array}{c|cc} -ról & tv_1 & tv_2 \\ \hline tv_1 & 1/2 & 1/2 \\ tv_2 & 1/4 & 3/4 \end{array} = \begin{array}{c|cc} & tv_1 & tv_2 \\ \hline arány & 3/8 & 5/8 \end{array}$$

4. Csak az átpártolás táblázatát nézve, a második hét végére a  $tv_1$  nézői első héten megmaradt felének csak a fele marad meg, míg a  $tv_2$ -től átpártolt negyednyi közönségnek is a fele, tehát a  $tv_1 \rightarrow tv_1$  „mozgás” a nézők  $3/8$ -adát érinti, mert  $\frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{4} = \frac{3}{8}$ . Hasonló számításokkal a többi érték is megkapható, melyet az alábbi, két, táblázatok közti szorzással is meg lehet adni:

$$\begin{array}{c|cc} -re & tv_1 & tv_2 \\ \hline tv_1 & 1/2 & 1/4 \\ tv_2 & 1/2 & 3/4 \end{array} \times \begin{array}{c|cc} -re & tv_1 & tv_2 \\ \hline tv_1 & 1/2 & 1/4 \\ tv_2 & 1/2 & 3/4 \end{array} = \begin{array}{c|cc} -re & tv_1 & tv_2 \\ \hline tv_1 & 3/8 & 5/16 \\ tv_2 & 5/8 & 11/16 \end{array}$$

$$\begin{array}{c|cc} -ról & tv_1 & tv_2 \\ \hline tv_1 & 1/2 & 1/2 \\ tv_2 & 1/4 & 3/4 \end{array} \times \begin{array}{c|cc} -ról & tv_1 & tv_2 \\ \hline tv_1 & 1/2 & 1/2 \\ tv_2 & 1/4 & 3/4 \end{array} = \begin{array}{c|cc} -ról & tv_1 & tv_2 \\ \hline tv_1 & 3/8 & 5/8 \\ tv_2 & 5/16 & 11/16 \end{array}$$

4.5. a)  $4\mathbf{A} - 3\mathbf{B} = \begin{bmatrix} 1 & 5 & 12 \\ 2 & 1 & -3 \end{bmatrix}$ , b)  $2\mathbf{B} - \mathbf{C}$  nincs értelmezve. c)  $2\mathbf{B} - \mathbf{C}^T = \begin{bmatrix} 1 & 0 & -1 \\ 3 & 0 & 2 \end{bmatrix}$ .

4.6. Ha  $a$  a  $[0, k]$  intervallumba eső szám, akkor  $0 \leq a/k \leq 1$ , így  $a/k$  egész része 0 vagy 1. Részletezve  $[a/k]$  pontosan akkor 1, ha  $a = k$ , azaz ha a pixel átlátszó, egyébként 0. Másrészt  $1 - [a/k]$  pontosan akkor 0, ha  $a = k$ , egyébként 1. Ezt kihasználva könnyen definiálható a kívánt művelet:

$$a \odot b = \left\lfloor \frac{a}{k} \right\rfloor b + \left(1 - \left\lfloor \frac{a}{k} \right\rfloor\right) a.$$

Így e művelettel elemenként definiált  $\mathbf{A} \odot \mathbf{B}$  művelet a kívánt eredményt adja. A 4.2. ábrán három képet  $32 \times 24$ -es mátrixszal szemléltetünk, a férfiarc mátrixát is megadtuk, a másik a háttér. A művelet eredménye a harmadik kép.

4.7. A standard bázisba azon mátrixok tartoznak, amelyekben egyetlen elem 1, a többi 0.

4.8. E mátrixok összes lineáris kombinációja

$$a\mathbf{A} + b\mathbf{B} + c\mathbf{C} = \begin{bmatrix} a+b+c & a+c \\ a & b+c \end{bmatrix}$$

alakú. Ha egy tetszőleges  $\begin{bmatrix} u & v \\ w & z \end{bmatrix}$  mátrixról el akarjuk dönteni, hogy a fenti alakú-e, azaz fönnáll-e valamely  $a, b, c$

ismeretlenekre az

$$\begin{bmatrix} a+b+c & a+c \\ a & b+c \end{bmatrix} = \begin{bmatrix} u & v \\ w & z \end{bmatrix}$$

egyenlőség, akkor meg kell oldani a mátrixok négy elemére vonatkozó négy egyenletből álló 3-ismeretlenes egyenletrendszert:

$$\begin{aligned} a+b+c &= u \\ a &+ c = v \\ a &= w \\ b+c &= z \end{aligned}$$

Ha ennek van megoldása, akkor létezik a megfelelő lineáris kombináció, tehát az adott  $\begin{bmatrix} u & v \\ w & z \end{bmatrix}$  mátrix a kifeszített térbe esik. Ennek az egyenletrendszernek a bővített mátrixát fölírva, majd elemi sorműveletekkel megoldva a következőt kapjuk:

$$\left[ \begin{array}{ccc|c} 1 & 1 & 1 & u \\ 1 & 0 & 1 & v \\ 1 & 0 & 0 & w \\ 0 & 1 & 1 & z \end{array} \right] \Rightarrow \left[ \begin{array}{ccc|c} 1 & 0 & 0 & w \\ 0 & 1 & 1 & u-w \\ 0 & 0 & 1 & v-w \\ 0 & 0 & 0 & w+z-u \end{array} \right].$$

A lépcsős alakból leolvasható, hogy ez az egyenletrendszer pontosan akkor oldható meg, ha  $w+z-u=0$ . Például az  $\begin{bmatrix} 5 & 4 \\ 3 & 2 \end{bmatrix}$  mátrix ebbe az altérbe esik. A fenti egyenletrendszer megoldásával az is megkapható, hogy mik a lineáris kombináció együtthatói. Azt kapjuk, hogy  $a=3$ ,  $b=1$  és  $c=1$ .

4.9. a)  $\mathbf{AB}$  nincs értelmezve. b)  $\mathbf{AB}^T - \mathbf{D} = \begin{bmatrix} 0 & 5 \\ 2 & 3 \end{bmatrix}$ ,

c)  $\mathbf{BC} = \begin{bmatrix} 3 & 3 \\ 5 & 4 \end{bmatrix}$ , d)  $\mathbf{CB} = \begin{bmatrix} 3 & 2 & 1 \\ 6 & 4 & 2 \\ 1 & 1 & 0 \end{bmatrix}$ , e)  $(\mathbf{DA})\mathbf{C} = \begin{bmatrix} 32 & 23 \\ 16 & 13 \end{bmatrix}$ .

4.10. A méretek alapján a  $\mathbf{BC}$  szorzat nincs értelmezve, a többi:

$$\mathbf{AB} = \begin{bmatrix} 3 & 2 & 1 \\ 9 & 8 & 7 \\ 3 & 4 & 5 \end{bmatrix}, \mathbf{BA} = \begin{bmatrix} 6 & 5 \\ 6 & 10 \end{bmatrix}, \mathbf{CB} = \begin{bmatrix} 3 & 2 & 1 \\ 3 & 4 & 5 \end{bmatrix},$$

$$\mathbf{CD} = \begin{bmatrix} -2 & -1 \\ 10 & 11 \end{bmatrix}, \mathbf{DC} = \begin{bmatrix} 12 & 12 \\ -2 & -3 \end{bmatrix}, \mathbf{DE} = \mathbf{ED} = \begin{bmatrix} 0 & -6 \\ 2 & 7 \end{bmatrix}$$

Összefoglalva:  $\mathbf{AB} \neq \mathbf{BA}$ , mert különböző típusúak,  $\mathbf{BC} \neq \mathbf{CB}$ , mert az egyik oldal nincs értelmezve,  $\mathbf{CD} \neq \mathbf{DC}$ , bár mindkét oldal értelmezve van és azonos típusú. Az előzőekkel ellentétben viszont fennáll a  $\mathbf{DE} = \mathbf{ED}$  egyenlőség.

Azaz vannak felcserélhető mátrixok, de a mátrixszorzás nem felcserélhető művelet, tehát nem kommutatív!

4.11. Az  $\mathbf{AB}$  szorzat felírható

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} \end{bmatrix}$$

alakban. A  $\mathbf{BA}$  hasonlóképp írható fel! Ellenőrizzük, hogy a 4.11. állítás feltétele (minden  $k$ -ra az  $\mathbf{A}_{ik}$  blokk oszlopainak száma megegyezik  $\mathbf{B}_{kj}$  sorainak számával) valóban szükséges, de elégséges is.

4.12. Számoljunk blokkmátrixként kezelve a mátrixokat:

$$\begin{aligned} \mathbf{A} + 3\mathbf{C} &= \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 3 & 0 \end{bmatrix} + 3 \begin{bmatrix} 0 & 2 & 0 & 1 \\ 2 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} + 3 \begin{bmatrix} 0 & 2 \\ 2 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} + 3 \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} + 3 \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 6 \\ 6 & 1 \\ 3 & 3 \end{bmatrix} \end{aligned}$$

Ellenőrizzük a számítást közönséges mátrixműveletekkel! Ezután tekintsük a blokkmátrixok szorzását!

$$\begin{aligned} \mathbf{AB} &= \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 1 & 5 \\ 2 & 2 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 1 & 5 \\ 2 & 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 4 \\ 1 & 5 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 2 \\ 2 & 2 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ 3 & 9 \\ 6 & 6 \end{bmatrix} \end{aligned}$$

Ugyanezt az eredményt kapjuk, ha ellenőrzésül egyszerű mátrixszorzással is elvégezzük a műveletet!

4.13. Valóban

$$\begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 2 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$



4.15. A Segre-féle külső szorzat:

$$\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 2 & 2 & 0 & 4 & 0 & 0 & 0 \\ 2 & 0 & 4 & 4 & 0 & 8 & 0 & 0 & 0 \end{bmatrix}.$$

4.16. Legyen  $\mathbf{A} \in S^{n_1 \times \dots \times n_d}$  egy  $d$ -edrendű és  $\mathbf{B} \in S^{m_1 \times \dots \times m_e}$  egy  $e$ -edrendű hipermátrix, és tegyük fel, hogy valamely  $1 \leq u \leq d$  és  $1 \leq v \leq e$  indexre  $n_u = m_v = k$ .  $\mathbf{A}$  és  $\mathbf{B}$  kontrakciós szorzatán azt a  $(d+e-2)$ -edrendű

$\mathbf{C} = [c_{i_1 \dots i_u \dots i_d j_1 \dots j_v \dots j_e}] = \langle \mathbf{A}, \mathbf{B} \rangle_{u:v} \in S^{n_1 \times \dots \times \overline{n_u} \dots n_d \times m_1 \times \dots \times \overline{m_v} \dots m_e}$  hipermátrixot értjük, melyre

$$c_{i_1 \dots \overline{i_u} \dots i_d j_1 \dots \overline{j_v} \dots j_e} = \sum_{\ell=1}^k a_{i_1 \dots \ell \dots i_d} b_{j_1 \dots \ell \dots j_e},$$

ahol a felülvonás az index törlését jelöli.

4.17. A kontrakciós szorzat:

$$\left\langle \begin{bmatrix} 1 & 0 & 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1' \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \right\rangle_{2:1} = \begin{bmatrix} 2 & 6 & 2 & 2 \\ 1 & 0 & 1 & 1 \end{bmatrix}$$

4.18. a) Hamis, b) igaz, c) igaz.

$$4.27. \mathbf{a} \cdot \mathbf{b} = \mathbf{a}^T \mathbf{b} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 2, \mathbf{a} \otimes \mathbf{b} = \mathbf{a} \mathbf{b}^T = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}.$$

$$4.28. \mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = \begin{bmatrix} 1 & 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix} = 5,$$

$$\mathbf{u} \otimes \mathbf{v} = \mathbf{u} \mathbf{v}^T = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 0 & 2 & 4 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 \end{bmatrix}.$$

4.29. A skaláris szorzat nem végezhető el, a diadikus szorzat

$$\mathbf{a} \otimes \mathbf{b} = \mathbf{a} \mathbf{b}^T = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 0 & 2 & 4 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

$$4.30. \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$$4.36. \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 & -2 & 1 \\ 2 & 0 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

4.37. A lineáris helyettesítés mátrixszorzatos alakja

$$\mathbf{y} = \mathbf{A} \mathbf{x},$$

ahol

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}, \quad \text{és} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

4.38. Az áttérés mátrixa

$$\mathbf{A}_{C \leftarrow B} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 2 & 3 \end{bmatrix},$$

Ezt fölhasználva kapjuk, hogy

$$[\mathbf{u}]_C = \mathbf{A}_{C \leftarrow B} [\mathbf{u}]_B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 4 \end{bmatrix},$$

továbbá

$$[\mathbf{v}]_C = \mathbf{A}_{C \leftarrow B} [\mathbf{v}]_B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} -3 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix}.$$

4.39. Az áttérés mátrixa

$$\mathbf{A}_{C \leftarrow B} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 3 \\ -2 & -2 \end{bmatrix}.$$

Ezt fölhasználva kapjuk, hogy

$$[\mathbf{u}]_C = \mathbf{A}_{C \leftarrow B} [\mathbf{u}]_B = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 3 \\ -2 & -2 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix},$$

továbbá

$$[\mathbf{v}]_C = \mathbf{A}_{C \leftarrow B} [\mathbf{v}]_B = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 3 \\ -2 & -2 \end{bmatrix} \begin{bmatrix} -3 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 6 \\ 2 \end{bmatrix}.$$

$$4.40. \mathbf{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$4.41. \quad \mathbf{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$4.42. \quad \mathbf{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

$$4.43. \quad \mathbf{E} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$4.44. \quad \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 8 & 8 & 8 \end{bmatrix}$$

$$4.45. \quad \begin{bmatrix} a \\ 2d - 2b \\ 3a + c \\ d \end{bmatrix}$$

$$4.46. \quad \left[ \begin{array}{cc|cc} 2 & 3 & 2 & 2 \\ 4 & 5 & 2 & 2 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 3 & 3 \end{array} \right]$$

$$4.47. \quad \left[ \begin{array}{cc|c} 2 & 3 & 10 \\ 4 & 5 & 14 \\ 0 & 0 & 5 \end{array} \right]$$

$$4.48. \quad \begin{bmatrix} 5 & 6 \\ 6 & 7 \\ 12 & 12 \end{bmatrix}$$

4.49. Mivel  $[\mathbf{I}_r|\mathbf{S}]$  az  $\mathbf{A}$  redukált lépcsős alakja, ezért ennek

bármely oszlopa az  $\mathbf{A}$  mátrix azonos sorszámu oszlopának koordinátás alakja az  $\mathbf{B}_r$  oszlopvektoraiban, mint bázisban felírva. Ez épp azt jelenti, hogy  $\mathbf{A} = \mathbf{B}_r[\mathbf{I}_r|\mathbf{S}]$ . Ez az oszlop-tér bármely oszlopára, így  $\mathbf{b}$ -re is igaz, hisz  $[\mathbf{A}|\mathbf{b}]$  redukált lépcsős alakja szerint az egyenletrendszer megoldható, így  $\mathbf{b}$  eleme az oszloptérnek. Eszerint tehát  $\mathbf{b} = \mathbf{B}_r\mathbf{d}_r$ .

Az, hogy minden megoldás fölírható ilyen alakba, a Gauss–Jordan-módszerből következik. Meg kell még mutatni, hogy a tételben felírt  $\mathbf{x}$  vektor valóban megoldás.

$$\begin{aligned} \mathbf{Ax} &= \mathbf{B}_r \begin{bmatrix} \mathbf{I}_r & \mathbf{S} \end{bmatrix} \left( \begin{bmatrix} \mathbf{d}_r \\ \mathbf{0}_s \end{bmatrix} + \begin{bmatrix} -\mathbf{S} \\ \mathbf{I}_s \end{bmatrix} \mathbf{t}_s \right) \\ &= \mathbf{B}_r (\mathbf{d}_r - \mathbf{S}\mathbf{t}_s + \mathbf{S}\mathbf{t}_s) = \mathbf{B}_r\mathbf{d}_r \\ &= \mathbf{b}. \end{aligned}$$

Ez bizonyítja az állítás első felét. A második felének bizonyításához csak azt kell látni, hogy  $[\begin{smallmatrix} -\mathbf{S} \\ \mathbf{I}_s \end{smallmatrix}]$  oszlopvektora a nulltér bázisát alkotják. Ez abból következik, hogy egyrészt kifeszítik a nullteret, másrészt lineárisan függetlenek, hisz az alsó blokkban lévő  $\mathbf{I}_s$  mátrix oszlopai lineárisan függetlenek.

4.50. Jelölje  $\mathbf{j}$  a csupa 1-esből álló 9-dimenziós vektort,  $\mathbf{j}_{456}$  azt, amelynek 4, 5, 6 indexű eleme 1-es, a többi 0. Ekkor a „minden sorösszeg 45” és a „minden oszlopösszeg 45” feltételek ekvivalensek az  $\mathbf{Aj} = 45\mathbf{j}$ ,  $\mathbf{j}^T\mathbf{A} = 45\mathbf{j}^T$  egyenletekkel, míg pl. az „első blokkoszlop, második blokkor metszetében álló blokk elemeinek összege 45” feltételnek a  $\mathbf{j}_{456}^T\mathbf{A}\mathbf{j}_{123} = 45$  egyenlet felel meg.

4.51.  $\mathbb{Z}_2^{2 \times 2}$ -be  $2^4 = 16$  mátrix tartozik:

$$\mathbb{Z}_2^{2 \times 2} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \dots, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\}.$$

## 5

# Mátrixműveletek algebrája

Áttekintjük a mátrixműveletek legfontosabb algebrai tulajdonságait. Ezek nem csak a mátrixokkal való számolás közben követendő szabályokról szólnak, de hozzásegítenek az lineáris egyenletrendszerek mélyebb megértéséhez, és olyan eszközöket adnak a kezünkbe, például a mátrixfelbontásokkal, melyek a lineáris algebra alkalmazásaiban is fontosak.

### Az alapműveletek tulajdonságai

---

*Az összeadás és a skalárral szorzás őrzi a valósok műveleti tulajdonságait, de a mátrixszorzás nem.*

---

**Az összeadás és a skalárral való szorzás tulajdonságai** Mivel a mátrixok összeadása és skalárral való szorzása elemenként végrehajtható műveletek, ezért műveleti tulajdonságaik természetes módon öröklik meg a számok műveleti tulajdonságait. Például azonos típusú mátrixok összeadása felcserélhető (kommutatív) és csoportosítható (asszociatív) művelet, míg összeg skalárral való szorzása disztributív. Tehát

$$\begin{aligned} \mathbf{A} + \mathbf{B} &= \mathbf{B} + \mathbf{A}, & \mathbf{A} + (\mathbf{B} + \mathbf{C}) &= (\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + \mathbf{B} + \mathbf{C}, \\ c(\mathbf{A} + \mathbf{B}) &= c\mathbf{A} + c\mathbf{B}, & (c + d)\mathbf{A} &= c\mathbf{A} + d\mathbf{A}. \end{aligned}$$

E tulajdonságok igazolását az Olvasóra hagyjuk (ld. 5.4. feladat).

**A szorzás tulajdonságai** A számok szorzásának algebrai tulajdonságai nem öröklődnek automatikusan a mátrixműveletekre, mint az összeadásnál. Nem is teljesülnek mind, pl. a mátrixszorzás *nem kommutatív*.

A mátrixokkal való számolás közben nem csak arra kell ügyelnünk, hogy bizonyos azonosságok nem teljesülnek, de arra is, hogy bizonyos elemi eljárások nem végezhetők el olyan tág körben, mint azt a valós számoknál megszoktuk.

## 5.1. ÁLLÍTÁS (MIRE VIGYÁZZUNK A MÁTRIXSZORZÁSNÁL?).

- a) A mátrixszorzás nem kommutatív, azaz  $\mathbf{AB} = \mathbf{BA}$  nem áll fenn bármely két összeszorozható mátrixra.
- b) Ha  $\mathbf{AB} = \mathbf{AC}$ , akkor az  $\mathbf{A} \neq \mathbf{O}$  feltétel kevés ahhoz, hogy a  $\mathbf{B} = \mathbf{C}$  következtetésre jussunk.
- c) Az  $\mathbf{AB} = \mathbf{O}$  egyenlőségből nem következik, hogy  $\mathbf{A}$  vagy  $\mathbf{B}$  a nullmátrix.

► A mátrixszorzás kommutativitásának cáfolására az egyik legegyszerűbb példa:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \text{ de } \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

A 4.10. feladatban további példákat mutatunk.

► A valós számok közt igaz, hogy ha  $a \neq 0$  és  $ab = ac$ , akkor  $a$ -val egyszerűsíthetünk, azaz akkor  $b = c$ . Mátrixokra egy ellenpélda:

$$\begin{bmatrix} 1 & -2 \\ 2 & -4 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & -3 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 2 & -4 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 1 & 2 \end{bmatrix}, \text{ de } \begin{bmatrix} 1 & -3 \\ 2 & -1 \end{bmatrix} \neq \begin{bmatrix} -1 & 3 \\ 1 & 2 \end{bmatrix}.$$

► Nullosztónak nevezzük egy algebrai struktúra olyan nemzérus elemét, melyhez található olyan nemzérus elem, mellyel vett szorzata zérus. Valósok közt ilyenek nincsenek, de a mátrixok közt igen, például

$$\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 2 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

Nullosztóval találkozhatunk a  $\mathbb{Z}_m$ -ben való számolásnál is, ha  $m$  összetett. Például  $\mathbb{Z}_6$ -ban  $2 \cdot 3 = 0$ . Összetett  $m$  esetén egyszerűsíteni sem lehet mindig  $\mathbb{Z}_m$ -ben, például  $\mathbb{Z}_{12}$ -ben  $9 \cdot 2 = 3 \cdot 2 = 6$ , de  $9 \neq 3$ .

5.2. TÉTEL (MÁTRIXSZORZÁS ALGEBRAI TULAJDONSÁGAI). Legyen  $\mathbf{A}$ ,  $\mathbf{B}$  és  $\mathbf{C}$  olyan, hogy a kijelölt műveletek elvégezhetők legyenek, legyen továbbá  $c$  tetszőleges skalár. Ekkor

- |                                                                                                            |                                    |
|------------------------------------------------------------------------------------------------------------|------------------------------------|
| a) $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$                                                     | csoporthozhatóság, asszociativitás |
| b) $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$                                       | disztributivitás                   |
| c) $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$                                       | disztributivitás                   |
| d) $(c\mathbf{A})\mathbf{B} = c(\mathbf{AB}) = \mathbf{A}(c\mathbf{B})$                                    |                                    |
| e) $\mathbf{A}_{m \times n} \mathbf{O}_{n \times t} = \mathbf{O}_{m \times t}$                             | szorzás nullmátrixszal             |
| f) $\mathbf{I}_m \mathbf{A}_{m \times n} = \mathbf{A}_{m \times n} \mathbf{I}_n = \mathbf{A}_{m \times n}$ | szorzás egységmátrixszal           |

**BIZONYÍTÁS.** A fenti tulajdonságok közül csak az elsőt bizonyítjuk, a többi hasonlóan, vagy még egyszerűbben bizonyítható.

a) Valójában többet bizonyítunk. Megmutatjuk, hogy ha az egyenlőség egyik oldalán kijelölt szorzások elvégezhetők, akkor a másik oldalon kijelöltek is. Legyen  $\mathbf{A}_{m \times s}$ ,  $\mathbf{B}_{s \times v}$  és  $\mathbf{C}_{v \times n}$  három tetszőleges mátrix. Az  $(\mathbf{AB})\mathbf{C}$  szorzatban  $\mathbf{AB}$  csak  $s = v$  esetén végezhető el, a szorzat típusa  $m \times v$ , ami  $\mathbf{C}$ -vel csak  $v = t$  esetén szorozható meg, és a szorzat  $m \times n$ -es. Tehát e szorzat csak akkor van értelmezve, ha  $\mathbf{B}$  típusa  $s \times t$ . Hasonló érveléssel ugyanezt kapjuk az  $\mathbf{A}(\mathbf{BC})$  szorzatról is.

Az indexek kezelésének könnyítésére elég lesz a bizonyítást sorvektor alakú  $\mathbf{A}$  és oszlopvektor alakú  $\mathbf{C}$  mátrixokra elvégezni, ugyanis az  $(\mathbf{AB})\mathbf{C}$  szorzat  $i$ -edik sorában és  $j$ -edik oszlopában álló elem az  $\mathbf{AB}$   $i$ -edik sorának, azaz az  $\mathbf{a}_{i*}\mathbf{B}$  sorvektornak és  $\mathbf{C}$   $j$ -edik oszlopának szorzata, azaz  $(\mathbf{a}_{i*}\mathbf{B})\mathbf{c}_{*j}$ . Hasonlóképp az  $\mathbf{A}(\mathbf{BC})$  szorzat  $i$ -edik sorában és  $j$ -edik oszlopában álló elem  $\mathbf{a}_{i*}(\mathbf{BC}_{*j})$ . Legyen tehát

$$\mathbf{A} = \begin{bmatrix} a_1 & a_2 & \dots & a_m \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}.$$

Ekkor a szorzat  $1 \times 1$ -es. Először számoljuk ki az  $\mathbf{AB}$  mátrixot, ami  $1 \times n$ -es:  $\left[ \sum_{k=1}^m a_k b_{k1} \quad \sum_{k=1}^m a_k b_{k2} \quad \dots \quad \sum_{k=1}^m a_k b_{kn} \right]$ . Innen számolva  $(\mathbf{AB})\mathbf{C}$ -t:

$$\left[ \sum_{k=1}^m a_k b_{k1} \quad \sum_{k=1}^m a_k b_{k2} \quad \dots \quad \sum_{k=1}^m a_k b_{kn} \right] \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \sum_{l=1}^n \sum_{k=1}^m a_k b_{kl} c_l.$$

Hasonlóan, először  $\mathbf{BC}$ -t fölírva, az  $\mathbf{A}(\mathbf{BC})$  mátrixra kapjuk, hogy

$$\begin{bmatrix} a_1 & a_2 & \dots & a_m \end{bmatrix} \begin{bmatrix} \sum_{l=1}^n b_{1l} c_l \\ \sum_{l=1}^n b_{2l} c_l \\ \vdots \\ \sum_{l=1}^n b_{ml} c_l \end{bmatrix} = \sum_{k=1}^m a_k \left( \sum_{l=1}^n b_{kl} c_l \right) = \sum_{k=1}^m \sum_{l=1}^n a_k b_{kl} c_l.$$

Az utolsó lépésben a belső szumma minden tagját beszoroztuk  $a_k$ -val, a számok közti összeadás és szorzás közti disztributivitást használva. Vagyis mindkét oldalon olyan összeg áll, amely az összes  $a_k b_{kl} c_l$  alakú szorzat összege, csak a tagok csoportosítása más.  $\square$

► Az asszociativitás következménye, hogy a többtényezős mátrixszorzatokat nem kell zárójelezni, hisz bármelyik zárójelezés ugyanazt az eredményt adja. Így  $\mathbf{ABC} = (\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$ . Az állítás igaz többtényezős szorzatokra is, vagyis az  $\mathbf{A}_1 \mathbf{A}_2 \dots \mathbf{A}_k$  szorzat független a végrehajtás sorrendjétől, de a tényezők sorrendje nem változtatható!

► Megjegyezzük, hogy az asszociativitás imént leírt bizonyítása hasonlóan mondható el, ha az  $\mathbf{A} = [a_{ik}]$  mátrix nem csak 1 sorból, és a  $\mathbf{C} = [c_{lj}]$  mátrix nem csak egy oszlopból áll: ekkor a  $\mathbf{D} = \mathbf{ABC}$  szorzat  $i$ -edik sorának  $j$ -edik elemére azt kapjuk, hogy az az összes  $a_{ik} b_{kl} c_{lj}$  alakú szorzatok összege, azaz

$$d_{ij} = \sum_{k=1}^m \sum_{l=1}^n a_{ik} b_{kl} c_{lj}. \quad (5.1)$$

Az 5.1 egyenlőség, és az ehhez hasonló számtalan hasonló kifejezés vezetett Einsteint arra a felismerésre, hogy az indexelt változók szorzatainak összegében a szumma jelek feleslegesek, hisz azokra az indexekre kell összegezni, amelyek legalább kétszer szerepelnek, míg az egyszer szereplőkre nem. Tehát az előző kettős szumma helyett írhatnánk azt is, hogy

$$d_{ij} = a_{ik} b_{kl} c_{lj},$$

hisz a jobb oldalon  $i$  és  $j$  csak egyszer szerepel, így  $k$ -ra és  $l$ -re kell összegezni, azt pedig tudjuk, hogy  $k = 1, \dots, m$  és  $l = 1, \dots, n$ . Ezt a jelölésbeli egyszerűsítést *Einstein-konvenciónak* nevezik. Einstein ezt a relativitás általános elméletéről írt híres dolgozatában használta először 1916-ban. A konvenció használata főként a lineáris algebra fizikai alkalmazásaiban terjedt el, mi e könyvben nem fogjuk használni.

**Mátrix hatványozása** Csak a négyzetes mátrixok szorozhatók meg önmagukkal, hisz ha egy  $m \times n$ -es mátrix megszorozható egy  $m \times n$ -essel, akkor  $m = n$ . Ezt figyelembe véve természetes módon definiálható négyzetes mátrixok pozitív egész kitevős hatványa:

$$\mathbf{A}^k = \underbrace{\mathbf{A}\mathbf{A}\dots\mathbf{A}}_{k \text{ tényező}}$$

Kicsit elegánsabban – rekurzióval – is definiálhatjuk e fogalmat:  $\mathbf{A}^1 := \mathbf{A}$  és  $\mathbf{A}^{k+1} := \mathbf{A}^k \mathbf{A}$ .

Mivel a mátrixszorzás asszociatív, mindegy, hogy milyen sorrendben végezzük el a hatványozást. Ezzel igazolható a következő két összefüggés is:

5.3. ÁLLÍTÁS (HATVÁNYOZÁS AZONOSSÁGAI). *Legyen  $\mathbf{A}$  egy négyzetes mátrix! Ekkor*

- a)  $\mathbf{A}^k \mathbf{A}^m = \mathbf{A}^{k+m}$ ,  
b)  $(\mathbf{A}^k)^m = \mathbf{A}^{km}$ .

Ha ki akarjuk terjeszteni a hatványozást 0 kitevőre is, kövessük a precedencia-elvet<sup>1</sup>, azaz olyan értelmet adjunk  $\mathbf{A}^0$ -nak, hogy a fenti összefüggések érvényben maradjanak. Például tekintsük az a) azonosságot  $m = 0$  esetén:

$$\mathbf{A}^k \mathbf{A}^0 = \mathbf{A}^{k+0} = \mathbf{A}^k.$$

Ez minden  $\mathbf{A}$  mátrix esetén csak az egységmátrixra igaz, tehát

$$\mathbf{A}^0 = \mathbf{I}_n,$$

ahol  $n$  a négyzetes  $\mathbf{A}$  mérete.

► A valós számoknál tanult, különböző alapú hatványokra érvényes azonosság itt a kommutativitás hiánya miatt nem érvényes, azaz általában  $(\mathbf{A}\mathbf{B})^k \neq \mathbf{A}^k \mathbf{B}^k$ .

5.4. PÉLDA (MÁTRIX HATVÁNYOZÁSA). *Számítsuk ki az*

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \text{és a} \quad \mathbf{B} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

*mátrixok  $k$ -adik hatványait!*

**MEGOLDÁS.** Számoljuk ki  $\mathbf{A}$  hatványait!

$$\mathbf{A}^2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

azaz  $\mathbf{A}^2 = \mathbf{I}_2$ , ebből pedig látjuk, hogy  $\mathbf{A}^3 = \mathbf{I}_2 \mathbf{A} = \mathbf{A}$ ,  $\mathbf{A}^4 = \mathbf{A}^3 \mathbf{A} = \mathbf{A} \mathbf{A} = \mathbf{I}_2, \dots$ . Tehát általában  $\mathbf{A}^{2k} = \mathbf{I}_2$  és  $\mathbf{A}^{2k+1} = \mathbf{A}$ .

<sup>1</sup> A latin eredetű *precedencia* szó előzményt jelent (lásd még *precedens*). A *precedencia elv* a matematikában fogalmak jelentésének olyan kiterjesztését jelenti, melynek során a korábban megismert tulajdonságok, összefüggések érvényben maradnak.

A másik feladatot a hatványozás rekurzív definícióját használva indukcióval kényelmesen meg tudjuk oldani. Először számoljuk ki  $\mathbf{A}$  néhány hatványát:

$$\mathbf{A}^2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \mathbf{A}^3 = \mathbf{A}^2 \mathbf{A} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}.$$

Ebből azt sejtjük, hogy  $\mathbf{A}^k = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$ . Ha be tudjuk látni ennek az összefüggésnek az öröklődését  $k$ -ról  $k+1$ -re, akkor kész vagyunk. Más szóval meg kell mutatnunk, hogy ha  $\mathbf{A}^k = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$ , akkor  $\mathbf{A}^{k+1} = \begin{bmatrix} 1 & k+1 \\ 0 & 1 \end{bmatrix}$ . Ezt a következő szorzás elvégzése igazolja:

$$\mathbf{A}^{k+1} = \mathbf{A}^k \mathbf{A} = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & k+1 \\ 0 & 1 \end{bmatrix} \quad \square$$

Miután mátrixok lineáris kombinációja és négyzetes mátrixok egész kitevős hatványa értelmezve van, ezért négyzetes mátrixokra is definiálhatjuk skalár együtthatós polinom helyettesítési értékét. Legyen

$$p(x) = a_k x^k + a_{k-1} x^{k-1} + \dots + a_2 x^2 + a_1 x + a_0$$

egy skalár együtthatós polinom. A  $p$  polinom  $\mathbf{X} \in \mathbb{R}^{n \times n}$  helyen vett helyettesítési értékén a

$$p(\mathbf{X}) = a_k \mathbf{X}^k + \dots + a_2 \mathbf{X}^2 + a_1 \mathbf{X} + a_0 \mathbf{I}_n$$

mátrixot értjük.

5.5. PÉLDA (POLINOM HELYETTESÍTÉSI ÉRTÉKE). *Legyen*

$$\mathbf{C} = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 3 & -4 \\ 3 & 4 & -6 \end{bmatrix}.$$

*Mutassuk meg, hogy  $p(\mathbf{C}) = \mathbf{O}$ , ha  $p(x) = x^3 + 2x^2 - 1$ .*

**MEGOLDÁS.** A  $p(\mathbf{C}) = \mathbf{C}^3 + 2\mathbf{C}^2 - \mathbf{I}$  műveleteit elvégezve kapjuk, hogy

$$\begin{aligned} p(\mathbf{C}) &= \mathbf{C}^3 + 2\mathbf{C}^2 - \mathbf{I} \\ &= \begin{bmatrix} 9 & 8 & -14 \\ 8 & 7 & -12 \\ 14 & 12 & -21 \end{bmatrix} + 2 \begin{bmatrix} -4 & -4 & 7 \\ -4 & -3 & 6 \\ -7 & -6 & 11 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad \square \end{aligned}$$

**A transzponálás tulajdonságai** A következő tétel a transzponálás és a többi művelet kapcsolatáról szól:

**5.6. TÉTEL (TRANSPONÁLÁS TULAJDONSÁGAI).** Legyenek  $\mathbf{A}$  és  $\mathbf{C}$  azonos típusú mátrixok,  $\mathbf{B}$  sorainak száma egyezzen meg  $\mathbf{A}$  oszlopainak számával,  $c$  pedig legyen tetszőleges skalár. Ekkor

- a)  $(\mathbf{A}^\top)^\top = \mathbf{A}$ ,
- b)  $(\mathbf{A} + \mathbf{C})^\top = \mathbf{A}^\top + \mathbf{C}^\top$ ,
- c)  $(c\mathbf{A})^\top = c\mathbf{A}^\top$ ,
- d)  $(\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top$ .

**BIZONYÍTÁS.** Az első három összefüggés magától értetődő, csak az utolsót bizonyítjuk.

Először megmutatjuk, hogy ha  $(\mathbf{AB})^\top$  elvégezhető, akkor  $\mathbf{B}^\top \mathbf{A}^\top$  is. Az  $m \times t$  típusú  $\mathbf{A}$  és a  $t \times n$  típusú  $\mathbf{B}$  szorzata  $m \times n$ -es, transzponáltja  $n \times m$ -es, így az  $n \times t$  típusú  $\mathbf{B}^\top$  és a  $t \times m$ -es  $\mathbf{A}^\top$  összeszorozhatók, szorzatuk  $n \times m$ -es, így a tételbeli egyenlőség két oldalának típusa azonos.

A tétel azon alapul, hogy két tetszőleges  $\mathbf{u}, \mathbf{v}$  vektorra  $\mathbf{u}^\top \mathbf{v} = \mathbf{v}^\top \mathbf{u}$ . Ezt az összefüggést a  $*$ -gal jelölt egyenlőségnél fogjuk használni. Az  $(\mathbf{AB})^\top$   $i$ -edik sorának  $j$ -edik eleme

$$((\mathbf{AB})^\top)_{ij} = (\mathbf{AB})_{ji} = (\mathbf{A})_{j*} (\mathbf{B})_{*i}.$$

A  $\mathbf{B}^\top \mathbf{A}^\top$   $i$ -edik sorának  $j$ -edik eleme

$$(\mathbf{B}^\top \mathbf{A}^\top)_{ij} = (\mathbf{B}^\top)_{i*} (\mathbf{A}^\top)_{*j} = (\mathbf{A})_{j*} (\mathbf{B})_{*i}.$$

Tehát  $(\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top$ . □

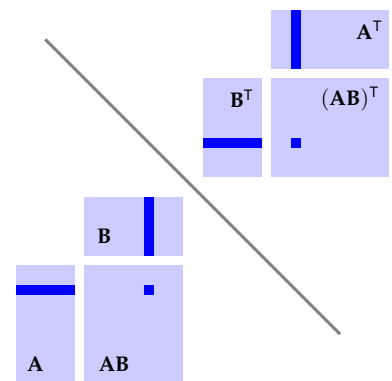
► A tétel b) pontjának indukcióval könnyen bizonyítható következménye, hogy többtagú összeg transzponáltja megegyezik a transzponáltak összegével. A c) pontot is figyelembe véve kapjuk, hogy mátrixok lineáris kombinációjának transzponáltja megegyezik a mátrixok transzponáltjainak azonos lineáris kombinációjával, azaz

$$(c_1 \mathbf{A}_1 + c_2 \mathbf{A}_2 + \dots + c_k \mathbf{A}_k)^\top = c_1 \mathbf{A}_1^\top + c_2 \mathbf{A}_2^\top + \dots + c_k \mathbf{A}_k^\top.$$

► A tétel d) pontjára „szemléletes igazolás” is adható, ami leolvasható az 5.1. ábráról.

► Indukcióval bizonyítható, hogy a d)-beli összefüggés többtényezős szorzatokra is fennáll, azaz

$$(\mathbf{A}_1 \mathbf{A}_2 \dots \mathbf{A}_k)^\top = \mathbf{A}_k^\top \dots \mathbf{A}_2^\top \mathbf{A}_1^\top.$$



5.1. ábra:  $(\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top$  szemléletes bizonyítása



**Mátrixszorzás inverze – mátrixok osztása\*** Lehet-e mátrixszal osztani, és ha igen, meg tudjuk-e vele oldani az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszert vagy az  $\mathbf{AX} = \mathbf{B}$  mátrixegyenletet úgy, ahogy az  $ax = b$  egyenletet megoldjuk az  $a$ -val való osztással?

Korábbi tanulmányainkban megtanultuk, hogy az összeadás és a szorzás invertálható műveletek, inverzeik a kivonás, ill. az osztás.

Azon, hogy az összeadás művelete invertálható, azt értjük, hogy bármely  $a$  és  $b$  valós esetén találunk olyan  $x$  valóst, hogy  $a + x = b$ , a megoldás  $x = b - a$ . A szorzás is invertálható, de csak a nemzérus valósok halmazán. Ez azt jelenti, hogy bármely  $a$  nemzérus valósához és  $b$  valósához található olyan  $x$  valós szám, hogy  $ax = b$ , a megoldás  $x = b/a$ .

Azonos típusú mátrixok közt az  $\mathbf{A} + \mathbf{X} = \mathbf{B}$  egyenlet megoldása ugyanolyan egyszerű, mint a számok közt:  $\mathbf{X} = \mathbf{B} - \mathbf{A}$ . A mátrixszorzás esete bonyolultabb.

► A mátrixszorzás nem kommutatív ezért az  $\mathbf{AX} = \mathbf{B}$  és az  $\mathbf{YA} = \mathbf{B}$  egyenletek megoldása különböző is lehet. Valójában a mátrixosztás művelete emiatt nem vezethető be, viszont be lehet vezetni egy balról és egy jobbról való osztást, az egyik jele  $\backslash$ , a másiké  $/$ . E jelöléssel a fenti két mátrixegyenlet és megoldása:

$$\begin{aligned}\mathbf{AX} = \mathbf{B} &\implies \mathbf{X} = \mathbf{A} \backslash \mathbf{B} && \mathbf{B} \text{ balról osztva } \mathbf{A}\text{-val,} \\ \mathbf{YA} = \mathbf{B} &\implies \mathbf{Y} = \mathbf{B} / \mathbf{A} && \mathbf{B} \text{ jobbról osztva } \mathbf{A}\text{-val.}\end{aligned}$$

Például

$$\begin{aligned}\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \backslash \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} &= \begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix}, \text{ mert } \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \text{ és} \\ \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} / \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}, \text{ mert } \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\end{aligned}$$

► A másik bonyodalmat az okozza, hogy míg az  $ax = b$  egyenlet a nullától különböző valósok közt mindig egyértelműen megoldható, addig pl. az  $\mathbf{Ax} = \mathbf{b}$  egyenletről tudjuk, hogy végtelen sok megoldása, de 0 megoldása is lehet. E nehézségekre a következő paragrafus egyszerű megoldást ad, melyet később a pszeudoinverz fogalma segítségével általánosítani fogunk (ld. a 7 fejezetet).

**Mátrix inverze** Tudjuk, hogy az  $ax = b$  egyenlet megoldásához elég ismerni a *reciprokát*, más néven a *multiplikatív inverzét*, és azzal szorozni  $b$ -t. Ez a gondolat átvihető a mátrixszorzásra is.

Egy nemnulla  $a$  szám reciproka az az  $a^{-1}$ -gyel jelölt szám, melyre  $aa^{-1} = a^{-1}a = 1$ . Az 1 szerepét mátrixszorzásnál az  $\mathbf{I}$  egységmátrix játssza. Világos, hogy adott  $\mathbf{A}$  mátrixhoz csak úgy létezhet olyan  $\mathbf{X}$ , melyre  $\mathbf{AX} = \mathbf{XA} = \mathbf{I}$ , ha  $\mathbf{A}$  négyzetes. Ez a következő definíciót adja:

Egy  $H$  halmazon értelmezett kétváltozós (más szóval bináris) *művelet* olyan függvényt értünk, mely  $H$ -beli elempárokhoz  $H$ -beli elemet rendel. Például a valós számok összeadása esetén a függvény valós számpárhoz valós számot rendel, mondjuk az  $(1.2, 0.4)$  számpárhoz a 1.6-ot. E függvényt a  $+$  jellel jelöljük, de a függvényeknél szokásos prefix „ $+(a, b)$ ” jelölés helyett műveleteknél az ún. infix jelölést használjuk, azaz  $a + b$ -t írunk (lásd erről még a következő szíljegyzetet).

A számítástechnikában gyakran találkozunk a műveletek *infix* jelölése mellett a *prefix* vagy *lengyel* és a *postfix* vagy *fordított lengyel* jelölésével is. A prefixnél a műveleti jel az argumentumai előtt, a postfixnél után van. Például a  $(3 + 4) \cdot 2$  kifejezést a prefix jelölést használó Lisp nyelvcsalád nyelveiben a

`(* (+ 3 4) 2)`

kód, míg például a postfix jelölést használó PostScript nyelvben a

`3 4 add 2 mul`

kód számítja ki. (A PostScript nyelvvel találkozhatunk a PDF formátumú fájlokban is.) Ugyanez a formula a komputer algebra nyelvek közül a Mapleben prefix módon

`'*('+(3,4),2)`

a Mathematicában

`Times[Plus[3,4],2]`

alakot ölt. A Sage két lehetőséget kínál:

`prod([sum([3,4]),2])`

`mul([add([3,4]),2])`

Általában egy algebrai struktúra egy elemének egy műveletre vonatkozó inverzéhez a művelet *semleges eleme* szükséges. Az összeadás semleges eleme a 0, mert bármely  $a$  elemhez adva  $a$ -t kapunk, hasonlóképp a szorzás semleges eleme az 1, mert bármely  $a$  elemet vele szorozva  $a$ -t kapunk. Összeadás esetén egy elem ellentettjét az  $a + x = 0$  egyenlet megoldásával kapjuk, szorzás esetén a reciprokot az  $ax = 1$  megoldásával. Az ellentettet, illetve a reciprokot additív, illetve multiplikatív *inverz*nek is nevezzük. Mátrixszorzás semleges eleme az egységmátrix.

5.7. DEFINÍCIÓ (MÁTRIX INVERZE). Legyen  $\mathbf{A}$  egy  $n \times n$ -es mátrix. Azt mondjuk, hogy  $\mathbf{A}$  invertálható, ha létezik olyan  $\mathbf{B}$  mátrix, melyre

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n.$$

A  $\mathbf{B}$  mátrixot  $\mathbf{A}$  inverzének nevezzük, és  $\mathbf{A}^{-1}$ -nel jelöljük. A nem invertálható mátrixot szingulárisnak nevezzük.

- Világos, hogy ha  $\mathbf{A}$  inverze  $\mathbf{B}$ , akkor  $\mathbf{B}$  inverze  $\mathbf{A}$ .
- Például az alábbi szorzatokban szereplő két mátrix egymás inverze:

$$\begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

- A definícióból nem derül ki, hogy egy mátrixnak lehet-e több inverze, de könnyen megmutatható, hogy nem. Ha ugyanis az  $\mathbf{A}$  mátrixnak  $\mathbf{B}$  és  $\mathbf{C}$  is inverze, azaz  $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$  és  $\mathbf{AC} = \mathbf{CA} = \mathbf{I}$ , akkor

$$\mathbf{C} = \mathbf{CI} = \mathbf{C}(\mathbf{AB}) = (\mathbf{CA})\mathbf{B} = \mathbf{IB} = \mathbf{B}.$$

- Az  $\mathbf{A}$  mátrix inverzére használhatjuk az  $\mathbf{A}^{-1}$  jelölést, mert megfelel a precedencia-elvnek. Például ha az 5.3. tétel érvényességét megtartva akarunk az  $\mathbf{A}^{-1}$  hatványnak értelmet adni, akkor fenn kell álljon rá az  $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}^{-1+1} = \mathbf{A}^0 = \mathbf{I}$  és a  $\mathbf{AA}^{-1} = \mathbf{A}^{1-1} = \mathbf{A}^0 = \mathbf{I}$  összefüggés.

- Minthogy a mátrixok közti műveleteket a számok közti műveletek táblázatokra való kiterjesztésén keresztül vezettük be, elvárjuk, hogy az  $1 \times 1$ -es mátrixok inverze essen egybe a számok multiplikatív inverzével (reciprokával), azaz ha  $\mathbf{A} = [a]$ , akkor  $\mathbf{A}^{-1} = [a^{-1}] = [1/a]$  legyen igaz. A fenti definíció ennek az elvárásunknak is megfelel.

Egy négyzetes  $\mathbf{A}$  mátrixot *nilpotensnek* nevezünk, ha van olyan  $k$  pozitív egész, hogy

$$\mathbf{A}^k = \mathbf{O}.$$

Például a  $\begin{bmatrix} -2 & 4 \\ -1 & 2 \end{bmatrix}$  mátrix nilpotens, mert  $\begin{bmatrix} -2 & 4 \\ -1 & 2 \end{bmatrix}^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ . Több alkalmazásban is fontos szerepet kap az alábbi példában szereplő inverz.

5.8. PÉLDA ( $\mathbf{I} - \mathbf{A}$  INVERZE NILPOTENS  $\mathbf{A}$  ESETÉN). Mutassuk meg, ha  $\mathbf{A}$  nilpotens, azaz valamely pozitív  $k$ -ra  $\mathbf{A}^k = \mathbf{O}$ , akkor  $\mathbf{I} - \mathbf{A}$  invertálható, és inverze  $\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \dots + \mathbf{A}^{k-1}$ .

**MEGOLDÁS.** Megmutatjuk, hogy  $(\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \dots + \mathbf{A}^{k-1}) = \mathbf{I}$ .

$$\begin{aligned} & (\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \dots + \mathbf{A}^{k-1}) \\ &= \mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \dots + \mathbf{A}^{k-1} - \mathbf{A} - \mathbf{A}^2 - \dots - \mathbf{A}^{k-1} - \mathbf{A}^k \\ &= \mathbf{I} - \mathbf{A}^k \\ &= \mathbf{I} \end{aligned}$$

Az  $(\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \dots + \mathbf{A}^{k-1})(\mathbf{I} - \mathbf{A}) = \mathbf{I}$  összefüggés ugyanígy bizonyítható. □

**Elemi mátrixok inverze** Minden  $R$  elemi sorművelethez van egy olyan  $R'$  sorművelet, hogy az  $R$  sorművelettel átalakított mátrixot az  $R'$  visszaalakítja (ld. 5.21. feladat). Nevezzük ezt az  $R'$  sorműveletet az  $R$  sorművelet inverzének. Könnyen ellenőrizhető, hogy az  $S_i \leftrightarrow S_j$  sorművelet inverze önmaga, a  $cS_i$  inverze  $\frac{1}{c}S_i$ , és  $S_i + cS_j$  inverze  $S_i - cS_j$ .

**5.9. ÁLLÍTÁS (SORMŰVELET INVERZÉNEK MÁTRIXA).** Minden elemi mátrix invertálható, nevezetesen egy sorművelet elemi mátrixának inverze meggyezik a sorművelet inverzének elemi mátrixával.

A bizonyításhoz elég belátni, hogy egy sorművelet és az inverz sorművelet mátrixainak szorzata az egységmátrix. Az általános bizonyítás végiggondolását az Olvasóra hagyjuk, itt csak egy-egy konkrét esetet mutatunk meg, nevezetesen  $3 \times 3$ -as mátrixokon az  $S_2 \leftrightarrow S_3$ , a  $3S_2$  és az  $S_1 + 4S_3$  sorműveletek és inverzeik mátrixának szorzatát:

$$\begin{aligned} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \\ \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} &= \begin{bmatrix} 1 & 0 & -4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

**Az inverz kiszámítása** A négyzetes  $\mathbf{A}$  mátrix inverzének kiszámításához meg kell oldani az  $\mathbf{AX} = \mathbf{I}$  mátrixegyenletet, ami egyúttal egy szimultán egyenletrendszer is, és az elemi sorműveletekkel megoldható. Előbb azonban egy kérdésre válaszolnunk kell: nem fordulhat-e elő, hogy az  $\mathbf{AX} = \mathbf{I}$  mátrixegyenlet megoldható, de a megoldás nem tesz eleget az  $\mathbf{XA} = \mathbf{I}$  mátrixegyenletnek? Négyzetes mátrixok esetén *nem* a válasz, ami azt jelenti, hogy mátrix invertálásához elég az  $\mathbf{AX} = \mathbf{I}$  mátrixegyenlet megoldása!

**5.10. TÉTEL (AZ INVERZ LÉTEZÉSÉHEZ ELÉG EGY FELTÉTEL).** A négyzetes  $\mathbf{A}$  mátrix pontosan akkor invertálható, ha létezik olyan  $\mathbf{B}$  mátrix, hogy az  $\mathbf{AB} = \mathbf{I}$  és a  $\mathbf{BA} = \mathbf{I}$  feltételek egyike teljesül. Ha ilyen  $\mathbf{B}$  mátrix létezik, az egyértelmű.

**BIZONYÍTÁS.** Az inverz mátrix egyértelműségét beláttuk az 5.7. definíció utáni megjegyzések közt. Így elég belátnunk, hogy négyzetes mátrixokra az  $\mathbf{AB} = \mathbf{I}$  és a  $\mathbf{BA} = \mathbf{I}$  feltételek bármelyikének teljesülése maga után vonja a másik teljesülését is! Sőt, elég e két állítás egyikét igazolni: megmutatjuk, hogy ha a négyzetes  $\mathbf{A}$  és  $\mathbf{B}$  mátrixok kielégítik

az  $\mathbf{AB} = \mathbf{I}$  egyenletet, akkor  $\mathbf{BA} = \mathbf{I}$  is fennáll, azaz  $\mathbf{A}$  és  $\mathbf{B}$  inverzei egymásnak.

Tekintsük az  $\mathbf{AX} = \mathbf{I}$  mátrixegyenletet. Ezt úgy oldjuk meg, hogy az  $[\mathbf{A}|\mathbf{I}]$  mátrixot redukált lépcsős alakra hozzuk. Ha ez  $[\mathbf{I}|\mathbf{B}]$  alakú, akkor  $\mathbf{B}$  az  $\mathbf{AX} = \mathbf{I}$  egyenlet megoldása, ezért  $\mathbf{AB} = \mathbf{I}$  fennáll. A redukált lépcsős alakban zérus sor nem keletkezhet, mert a mátrix jobb oldalát az  $\mathbf{I}$  mátrixból kaptuk, ami redukált lépcsős alak, s így egyértelmű. Ha elemi sorműveletekkel zérus sort kapnánk a jobb oldali félmátrixban, akkor volna olyan redukált lépcsős alakja is, mely zérus sort tartalmazna, ami ellentmondás. Ha csak a mátrix bal felén kapnánk zérus sort, akkor az  $\mathbf{AX} = \mathbf{I}$  egyenletnek nem lenne megoldása, vagyis nem állhatna fenn az  $\mathbf{AB} = \mathbf{I}$  egyenlőség sem.

Ezután megmutatjuk, hogy  $\mathbf{BA} = \mathbf{I}$ . Ehhez tekintsük a  $\mathbf{BY} = \mathbf{I}$  mátrixegyenletet. Ennek megoldásához a  $[\mathbf{B}|\mathbf{I}]$  mátrixot kell redukált lépcsős alakra hozni. A előzőekből tudjuk, hogy elemi sorműveletekkel az

$$[\mathbf{A}|\mathbf{I}] \Rightarrow [\mathbf{I}|\mathbf{B}]$$

átalakítás megvalósítható. Az átalakítás lépéseinek inverzeit fordított sorrendben elvégezve az

$$[\mathbf{I}|\mathbf{B}] \Rightarrow [\mathbf{A}|\mathbf{I}]$$

transzformációt kapjuk. Itt minden lépésben fölcserélve a két részmátrixot a kívánt

$$[\mathbf{B}|\mathbf{I}] \Rightarrow [\mathbf{I}|\mathbf{A}]$$

átalakítást kapjuk. Ez azt jelenti, hogy a  $\mathbf{BY} = \mathbf{I}$  mátrixegyenletnek, az  $\mathbf{Y} = \mathbf{A}$  megoldása, azaz  $\mathbf{BA} = \mathbf{I}$ . □

Összefoglalva:

**5.11. ÁLLÍTÁS (INVERZ KISZÁMÍTÁSA ELEMI SORMŰVELETEKKEL).** *A négyzetes  $\mathbf{A}$  mátrix invertálható, ha az  $[\mathbf{A}|\mathbf{I}]$  mátrix elemi sorműveletekkel  $[\mathbf{I}|\mathbf{B}]$  alakra hozható, ekkor  $\mathbf{A}$  inverze  $\mathbf{B}$ . Ha  $\mathbf{A}$  redukált lépcsős alakja nem az  $\mathbf{I}$  mátrix, akkor  $\mathbf{A}$  nem invertálható.*

**5.12. PÉLDA (AZ INVERZ KISZÁMÍTÁSA).** *Számítsuk ki az*

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 6 \end{bmatrix} \quad \text{és a} \quad \mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 \\ 4 & 3 & 2 & 1 \end{bmatrix}$$

*mátrixok inverzét!*

**MEGOLDÁS.** A kiküszöböléssel oszloponként haladva:

$$\begin{aligned} \left[ \begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 3 & 4 & 0 & 1 & 0 \\ 3 & 4 & 6 & 0 & 0 & 1 \end{array} \right] &\Rightarrow \left[ \begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & -2 & -2 & 1 & 0 \\ 0 & -2 & -3 & -3 & 0 & 1 \end{array} \right] \Rightarrow \\ \left[ \begin{array}{ccc|ccc} 1 & 0 & -1 & -3 & 2 & 0 \\ 0 & 1 & 2 & 2 & -1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array} \right] &\Rightarrow \left[ \begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & 0 & 1 \\ 0 & 1 & 0 & 0 & 3 & -2 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array} \right] \end{aligned}$$

Tehát

$$\mathbf{A}^{-1} = \begin{bmatrix} -2 & 0 & 1 \\ 0 & 3 & -2 \\ 1 & -2 & 1 \end{bmatrix}.$$

A  $\mathbf{B}$  inverzének kiszámítása hasonló lépésekkel:

$$\begin{aligned} \left[ \begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 & 0 & 0 & 1 & 0 \\ 4 & 3 & 2 & 1 & 0 & 0 & 0 & 1 \end{array} \right] &\Rightarrow \left[ \begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 & -3 & 0 & 1 & 0 \\ 0 & 3 & 2 & 1 & -4 & 0 & 0 & 1 \end{array} \right] \Rightarrow \\ \left[ \begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 2 & 1 & 2 & -3 & 0 & 1 \end{array} \right] &\Rightarrow \left[ \begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & -2 & 1 \end{array} \right] \end{aligned}$$

Tehát

$$\mathbf{B}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{bmatrix}.$$

□

**5.13. TÉTEL (2 × 2-ES MÁTRIX INVERZE).** Az  $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  mátrix pontosan akkor invertálható, ha  $ad - bc \neq 0$ , és ekkor

$$\mathbf{A}^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

**BIZONYÍTÁS.** Azt, hogy az  $\mathbf{A}$  mátrixnak valóban a fenti mátrix az inverze, egyszerű mátrixszorzással ellenőrizhetjük. Azt, hogy az  $ad - bc \neq 0$  feltétel az invertálhatóságnak elégséges feltétele, a képlet bizonyítja. A feltétel szükségességének belátásához vegyük észre, hogy  $ad - bc = 0$ , azaz  $ad = bc$  pontosan akkor áll fenn, ha  $\mathbf{A}$  egyik sora a másik skalárszorosa. Ekkor viszont az egyik sor kinullázható, vagyis az  $\mathbf{A}$  mátrix nem alakítható elemi sorműveletekkel egységmátrixszá. □

**Az inverz tulajdonságai** Megvizsgáljuk a mátrixinvertálás más műveletekkel való kapcsolatát.

**5.14. TÉTEL (AZ INVERZ ALAPTULAJDONSÁGAI).** Tegyük fel, hogy  $\mathbf{A}$  és  $\mathbf{B}$  egyaránt  $n \times n$ -es invertálható mátrixok,  $c \neq 0$  skalár és  $k$  pozitív egész. Ekkor igazak a következők:

- a)  $\mathbf{A}^{-1}$  invertálható, és inverze  $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$ ,
- b)  $c\mathbf{A}$  invertálható, és inverze  $\frac{1}{c}\mathbf{A}^{-1}$ ,
- c)  $\mathbf{AB}$  invertálható, és inverze  $\mathbf{B}^{-1}\mathbf{A}^{-1}$ ,
- d)  $\mathbf{A}^k$  invertálható, és inverze  $(\mathbf{A}^k)^{-1} = (\mathbf{A}^{-1})^k$ , definíció szerint ezt értjük  $\mathbf{A}^{-k}$ -n,
- e)  $\mathbf{A}^T$  invertálható, és  $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$ .

**BIZONYÍTÁS.** Az állítások közül a fontosabbakat bizonyítjuk, a többit feladatként az Olvasóra hagyjuk:

c) Az

$$(\mathbf{AB})(\mathbf{B}^{-1}\mathbf{A}^{-1}) = \mathbf{A}(\mathbf{BB}^{-1})\mathbf{A}^{-1} = \mathbf{AA}^{-1} = \mathbf{I}$$

szorzat bizonyítja, hogy  $\mathbf{AB}$  invertálható, és inverze  $\mathbf{B}^{-1}\mathbf{A}^{-1}$ .

d) Az  $(\mathbf{A}^k)^{-1} = (\mathbf{A}^{-1})^k$  egyenlőség igaz volta a

$$\underbrace{\mathbf{AA} \dots \mathbf{A}}_{k \text{ tényező}} \underbrace{\mathbf{A}^{-1}\mathbf{A}^{-1} \dots \mathbf{A}^{-1}}_{k \text{ tényező}} = \underbrace{\mathbf{A}^{-1}\mathbf{A}^{-1} \dots \mathbf{A}^{-1}}_{k \text{ tényező}} \underbrace{\mathbf{AA} \dots \mathbf{A}}_{k \text{ tényező}} = \mathbf{I}$$

felírásból leolvasható, mert a szorzatok közepén lévő két mátrix szorzata mindig  $\mathbf{I}$ , ami elhagyható, és e lépést  $k$ -szor ismételve végül a kívánt eredményt kapjuk:

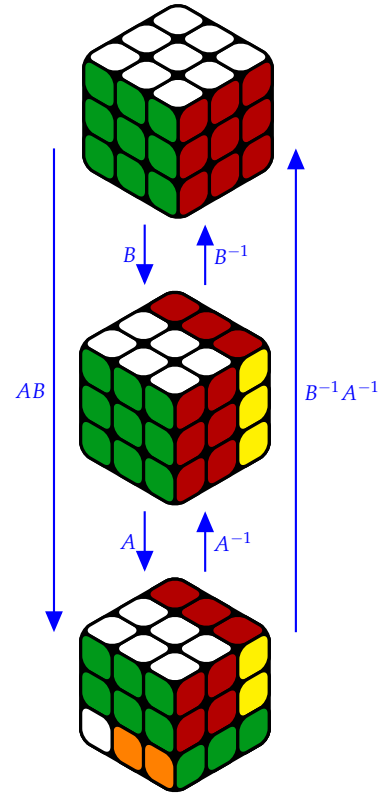
$$\underbrace{\mathbf{AA} \dots \mathbf{A}}_{k \text{ tényező}} \underbrace{(\mathbf{AA}^{-1})\mathbf{A}^{-1} \dots \mathbf{A}^{-1}}_{k \text{ tényező}} = \underbrace{\mathbf{AA} \dots \mathbf{A}}_{k-1 \text{ tényező}} \underbrace{\mathbf{A}^{-1} \dots \mathbf{A}^{-1}}_{k-1 \text{ tényező}} = \dots = \mathbf{I}.$$

► A (c) állítás indukcióval általánosítható véges sok mátrix szorzatára: ha az azonos méretű négyzetes  $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_m$  mátrixok mindegyike invertálható, akkor szorzatuk is, és

$$(\mathbf{A}_1\mathbf{A}_2 \dots \mathbf{A}_m)^{-1} = \mathbf{A}_m^{-1} \dots \mathbf{A}_2^{-1}\mathbf{A}_1^{-1}.$$

► A c) állításbeli összefüggéshez hasonlóval találkozhatunk a Rubik-kocka forgatása közben is. Egy forgatást jelöljön  $A$ , egy másikat  $B$ . A függvények kompozíciójához hasonlóan definiáljuk e két transzformáció szorzatát: a  $B$  majd az  $A$  forgatás egymás után való elvégzésével kapott transzformációt jelölje  $AB$  (ld. 5.2 ábra). E transzformáció inverze visszaállítja az eredeti állapotot, ehhez előbb az  $A$  transzformáció inverzét kell végrehajtani, majd a  $B$  inverzét, tehát  $(AB)^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ .

► Az  $\mathbf{A}^{-k}$  (d) pontbeli definíciója is megfelel a precedencia elvnek. Pl. az  $\mathbf{A}^m\mathbf{A}^n = \mathbf{A}^{m+n}$  összefüggés kiterjesztése negatív kitevőre az  $\mathbf{A}^k\mathbf{A}^{-k} = \mathbf{A}^0$  formulához vezet, amiből azt kapjuk, hogy  $\mathbf{A}^{-k} = (\mathbf{A}^k)^{-1}$ .



5.2. ábra: Jelölje  $A$  a bűvös kocka alsó,  $B$  a jobb hátsó oldalának elforgatását, és jelölje  $AB$  a  $B$ , majd az  $A$  egymás után való elvégzésével kapott transzformációt. (Ahogy a függvények összetételénél, előbb a jobb oldali, majd a bal oldali függvényt értékeljük ki, hajtjuk végre.) Ennek inverze  $(AB)^{-1}$  úgy kapható meg, ha előbb végrehajtjuk az  $A^{-1}$  majd a  $B^{-1}$  transzformációt. Ezek szorzata  $B^{-1}A^{-1}$ , tehát  $(AB)^{-1} = B^{-1}A^{-1}$ .

*Az invertálhatóság és az egyenletrendszerek megoldhatósága* A következő tétel a mátrixok invertálhatóságát, az egyenletrendszerek megoldásánál használt elemi sorműveleteket és az egyenletrendszerek megoldhatóságát kapcsolja össze.

**5.15. TÉTEL** (AZ INVERTÁLHATÓSÁG ÉS AZ EGYENLETRENDSZEREK). *Adva van egy  $n \times n$ -es  $\mathbf{A}$  mátrix. Az alábbi állítások ekvivalensek:*

- a)  $\mathbf{A}$  invertálható;
- b) az  $\mathbf{AX} = \mathbf{B}$  mátrixegyenlet bármely  $n \times t$ -es  $\mathbf{B}$  mátrixra egyértelműen megoldható;
- c) az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszer bármely  $n$  dimenziós  $\mathbf{b}$  vektorra egyértelműen megoldható;
- d) a homogén lineáris  $\mathbf{Ax} = \mathbf{0}$  egyenletrendszernek a triviális  $\mathbf{x} = \mathbf{0}$  az egyetlen megoldása;
- e)  $\mathbf{A}$  redukált lépcsős alakja  $\mathbf{I}$ ;
- f)  $\mathbf{A}$  előáll elemi mátrixok szorzataként.

**BIZONYÍTÁS.** Az állítások ekvivalenciáját az  $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (e) \Rightarrow (f) \Rightarrow (a)$ , implikációk igazolásával bizonyítjuk.

$(a) \Rightarrow (b)$ : Legyen tehát  $\mathbf{A}$  invertálható és legyen  $\mathbf{B}$  egy tetszőleges  $n \times t$  méretű mátrix. Ekkor az  $\mathbf{AX} = \mathbf{B}$  egyenlet mindkét oldalát  $\mathbf{A}^{-1}$ -gyel balról szorozva kapjuk, hogy  $\mathbf{A}^{-1}\mathbf{AX} = \mathbf{A}^{-1}\mathbf{B}$ , azaz  $\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$ . Ez azt mutatja, hogy egyrészt a mátrixegyenletnek van megoldása, másrészt hogy más megoldása nincs, mivel így minden megoldás megkapható, és  $\mathbf{A}$  inverze egyértelmű.

$(b) \Rightarrow (c)$ : Nyilvánvaló a  $\mathbf{B} = \mathbf{b}$  választással.

$(c) \Rightarrow (d)$ : Nyilvánvaló a  $\mathbf{b} = \mathbf{0}$  választással.

$(d) \Rightarrow (e)$ : Egy  $n$ -ismeretlenes,  $n$  egyenletből álló homogén lineáris egyenletrendszer pontosan akkor oldható meg egyértelműen, ha együtthatómátrixának redukált lépcsős alakja  $\mathbf{I}_n$ .

$(e) \Rightarrow (f)$ : Ha  $\mathbf{A}$  redukált lépcsős alakja  $\mathbf{I}_n$ , akkor létezik elemi sorműveletek olyan sorozata, mely az  $\mathbf{A} \Rightarrow \mathbf{I}_n$  transzformációt elvégzi. Jelölje az elemi sorműveletekhez tartozó elemi mátrixokat  $\mathbf{E}_1, \dots, \mathbf{E}_k$ . Ekkor tehát  $\mathbf{E}_1\mathbf{E}_2 \dots \mathbf{E}_k\mathbf{A} = \mathbf{I}_n$ . Innen  $\mathbf{A}$  kifejezhető az  $\mathbf{E}_1^{-1}$ -nel, ...,  $\mathbf{E}_k^{-1}$ -nel balról való beszorzás után:

$$\mathbf{A} = \mathbf{E}_k^{-1} \dots \mathbf{E}_2^{-1} \mathbf{E}_1^{-1}.$$

Elemi mátrixok inverze elemi mátrix, tehát  $\mathbf{A}$  előáll elemi mátrixok szorzataként.

$(f) \Rightarrow (a)$ : Az  $\mathbf{A} = \mathbf{E}_k^{-1} \dots \mathbf{E}_2^{-1} \mathbf{E}_1^{-1}$  mátrix minden tényezője invertálható, mivel mindegyik elemi mátrix, így szorzatuk is, és az inverz

$$\mathbf{A}^{-1} = \mathbf{E}_1\mathbf{E}_2 \dots \mathbf{E}_k.$$

□

► A tétel sok pontjának ekvivalenciája azt jelenti, hogy közülük bármely kettőre igaz, hogy „az egyik pontosan akkor igaz, ha a másik”.

Például „ $\mathbf{A}$  pontosan akkor invertálható, ha az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszer minden  $\mathbf{b}$  vektorra egyértelműen megoldható”.

► Később megmutatjuk azt is, hogy  $\mathbf{A}$  pontosan akkor invertálható, ha az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszer minden  $\mathbf{b}$  vektorra megoldható. Azaz az egyértelműség a feltételből kihagyható. Másként fogalmazva, ha  $\mathbf{Ax} = \mathbf{b}$  minden  $\mathbf{b}$  vektorra megoldható, akkor a megoldás minden  $\mathbf{b}$ -re egyértelmű.

5.16. PÉLDA (EGYENLETRENDSZER MEGOLDÁSA MÁTRIXINVERTÁLÁSSAL). Oldjuk meg az

$$2x + y = 2$$

$$5x + 3y = 3$$

egyenletrendszert mátrixinvertálással.

**MEGOLDÁS.** Az együtthatómátrix és inverze az 5.13. tétel szerint

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}, \quad \mathbf{A}^{-1} = \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix},$$

így az ismeretlenek  $(x, y)$  vektorára

$$\begin{bmatrix} x \\ y \end{bmatrix} = \mathbf{A}^{-1} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}. \quad \square$$

5.17. PÉLDA (MÁTRIXEGYENLET MEGOLDÁSA MÁTRIXINVERTÁLÁSSAL). Oldjuk meg az  $\mathbf{AX} = \mathbf{B}$  mátrixegyenletet, ahol

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}, \quad \text{és} \quad \mathbf{B} = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 3 & 1 \end{bmatrix}.$$

**MEGOLDÁS.** Az  $\mathbf{A}$  mátrix megegyezik az előző feladatbeli mátrixszal, így tudjuk, hogy invertálható, és ismerjük az inverzét. Az  $\mathbf{AX} = \mathbf{B}$  mátrixegyenlet megoldása:

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{B} = \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 1 & 3 & 2 \\ 4 & 3 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 6 & 5 \\ 3 & -9 & -8 \end{bmatrix}. \quad \square$$

► Megjegyezzük, hogy lineáris egyenletrendszert mátrixinvertálással ritkán oldunk meg, mert műveletigénye valamivel nagyobb, mint az egyszerű kiküszöbölésnek.

5.18. PÉLDA (MÁTRIX ELEMI MÁTRIXOK SZORZATÁRA BONTÁSA). Bontsuk fel az  $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$  mátrixot elemi mátrixok szorzatára!



**MEGOLDÁS.** Az 5.15. tétel bizonyításának (e)  $\Rightarrow$  (f) lépése szerint ha egy  $\mathbf{A}$  mátrixot elemi sorműveletekkel az egységmátrixba lehet transzformálni, akkor az elemi sorműveletek inverzei fordított sorrendben elvégezve az  $\mathbf{I}$ -t  $\mathbf{A}$ -ba transzformálják. Ez viszont azt jelenti, hogy a hozzájuk tartozó elemi mátrixok szorzata épp  $\mathbf{A}$ .

| Elemi sorműveletek                                                                                                                                                                                                                  | Elemi mátrixok                   | Elemi mátrixok inverzei                                                                                                                                                                                   |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| $\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$ $\Downarrow$ $\begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$ $\Downarrow$ $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ $\Downarrow$ $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ | $S_2 - 3S_1$ $-S_2$ $S_1 - 2S_2$ | $\mathbf{E}_1 = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}$ $\mathbf{E}_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ $\mathbf{E}_3 = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$              |
|                                                                                                                                                                                                                                     |                                  | $\mathbf{E}_1^{-1} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}$ $\mathbf{E}_2^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ $\mathbf{E}_3^{-1} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ |

A fenti átalakítás nyomán tehát  $\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{A} = \mathbf{I}$ , amiből  $\mathbf{A} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\mathbf{E}_3^{-1}$ , azaz

$$\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix},$$

így  $\mathbf{A}$ -t három elemi mátrix szorzatára bontottuk.  $\square$

**Invertálhatóság és bázis** Az 5.15. tétel szerint a négyzetes  $\mathbf{A}$  mátrix invertálhatósága azzal ekvivalens, hogy a homogén lineáris  $\mathbf{A}\mathbf{x} = \mathbf{0}$  egyenletrendszernek a triviális az egyetlen megoldása. Mivel  $\mathbf{A}\mathbf{x}$  az  $\mathbf{A}$  oszlopvektorainak egy lineáris kombinációja, ezért ez azt jelenti, hogy a nullvektor csak egyféleképp áll elő  $\mathbf{A}$  oszlopvektorainak lineáris kombinációjaként, a triviális módon. Tehát  $\mathbf{A}$  oszlopvektorai lineárisan függetlenek! Ez egyúttal azt is jelenti, hogy  $\mathbf{A}$  oszlopvektorai bázist alkotnak, és hogy  $r(\mathbf{A}) = n$ . Felhasználva a ?? tételt, mely szerint a sortér és az oszloptér dimenziója megegyezik a ranggal, a következő tételt kapjuk:

5.19. KÖVETKEZMÉNY (INVERTÁLHATÓSÁG ÉS BÁZIS). *Adva van egy valós  $n \times n$ -es  $\mathbf{A}$  mátrix. Az alábbi állítások ekvivalensek:*

- $\mathbf{A}$  invertálható;
- $\mathbf{A}$  oszlopvektorai lineárisan függetlenek;
- $\mathbf{A}$  oszlopvektorai bázist alkotnak  $\mathbb{R}^n$ -ben;
- $\mathbf{A}$  sorvektorai lineárisan függetlenek;
- $\mathbf{A}$  sorvektorai bázist alkotnak  $\mathbb{R}^n$ -ben;
- $r(\mathbf{A}) = n$ .

A fenti állításokat a tagadásukkal helyettesítjük és kiegészítjük azal, hogy ha egy mátrix sorvektorai közt lineáris kapcsolat van, akkor a redukált lépcsős alakban szükségképpen lesz zérusor:

5.20. KÖVETKEZMÉNY (SZINGULÁRIS MÁTRIXOK). *Adva van egy valós  $n \times n$ -es  $\mathbf{A}$  mátrix. Az alábbi állítások ekvivalensek:*

- a)  $\mathbf{A}$  szinguláris (azaz nem invertálható);
- b)  $\mathbf{A}$  oszlopvektorai lineárisan összefüggők;
- c) az  $\mathbf{A}$  oszlopvektorai által kifeszített altér dimenziója kisebb  $n$ -nél;
- d)  $\mathbf{A}$  sorvektorai lineárisan összefüggők;
- e) az  $\mathbf{A}$  sorvektorai által kifeszített altér dimenziója kisebb  $n$ -nél;
- f)  $\mathbf{A}$  bármely lépcsős alakjának (így redukált lépcsős alakjának is) van zérus sora;
- g)  $r(\mathbf{A}) < n$ .

**Báziscsere** Legyen  $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$  és  $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n\}$  az  $\mathbf{R}^n$  két bázisa, és jelölje  $\mathbf{X}_{\mathcal{C} \leftarrow \mathcal{B}}$  a  $\mathcal{B}$ -ről  $\mathcal{C}$ -re,  $\mathbf{Y}_{\mathcal{B} \leftarrow \mathcal{C}}$  a  $\mathcal{C}$ -ről  $\mathcal{B}$ -re való áttérés mátrixát. Legyen továbbá  $\mathbf{v}$  a tér egy tetszőleges vektora, a  $\mathcal{B}$  bázisbeli alakja  $[\mathbf{v}]_{\mathcal{B}}$ . A 4.25. tétel szerint

$$[\mathbf{v}]_{\mathcal{C}} = \mathbf{X}_{\mathcal{C} \leftarrow \mathcal{B}}[\mathbf{v}]_{\mathcal{B}}, \quad \text{és} \quad [\mathbf{v}]_{\mathcal{B}} = \mathbf{Y}_{\mathcal{B} \leftarrow \mathcal{C}}[\mathbf{v}]_{\mathcal{C}}.$$

A második egyenletbe helyettesítve az első kapjuk, hogy

$$[\mathbf{v}]_{\mathcal{B}} = \mathbf{Y}_{\mathcal{B} \leftarrow \mathcal{C}} \mathbf{X}_{\mathcal{C} \leftarrow \mathcal{B}} [\mathbf{v}]_{\mathcal{B}},$$

azaz  $\mathbf{Y}_{\mathcal{B} \leftarrow \mathcal{C}} \mathbf{X}_{\mathcal{C} \leftarrow \mathcal{B}}$  minden vektort önmagába visz, tehát egyenlő az egységmátrixszal, vagyis  $\mathbf{Y}_{\mathcal{B} \leftarrow \mathcal{C}}$  és  $\mathbf{X}_{\mathcal{C} \leftarrow \mathcal{B}}$  inverzei egymásnak.

5.21. TÉTEL (AZ ÁTTÉRÉS MÁTRIXÁNAK INVERZE). *Ha  $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$  és  $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n\}$  az  $\mathbf{R}^n$  két bázisa, akkor a  $\mathcal{B}$ -ről  $\mathcal{C}$ -re való áttérés  $\mathbf{X}_{\mathcal{C} \leftarrow \mathcal{B}}$  mátrixa, valamint a  $\mathcal{C}$ -ről  $\mathcal{B}$ -re való áttérés  $\mathbf{Y}_{\mathcal{B} \leftarrow \mathcal{C}}$  mátrixa is invertálható, és egymás inverzei, azaz  $\mathbf{X}_{\mathcal{C} \leftarrow \mathcal{B}}^{-1} = \mathbf{Y}_{\mathcal{B} \leftarrow \mathcal{C}}$  vagy más alakban  $\mathbf{X}_{\mathcal{C} \leftarrow \mathcal{B}} \mathbf{Y}_{\mathcal{B} \leftarrow \mathcal{C}} = \mathbf{I}_n$ .*

5.22. PÉLDA (AZ ÁTTÉRÉS MÁTRIXÁNAK INVERZE). Az  $\mathbf{R}^3$  egy  $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$  bázisában felírtuk a standard egységvektorokat:

$$\mathbf{i} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}_{\mathcal{B}}, \quad \mathbf{j} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}_{\mathcal{B}}, \quad \mathbf{k} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}_{\mathcal{B}}.$$

Írjuk fel  $\mathcal{B}$  bázisvektorainak standard bázisbeli koordinátás alakját!

**MEGOLDÁS.** Jelölje  $\mathcal{E}$  a standard bázist. Ennek vektorait kifejeztük a  $\mathcal{B}$  bázis elemeivel, az ezekből képzett mátrixszal tehát az  $\mathcal{E}$ -beli vektorok

$\mathcal{B}$ -beli koordinátás alakja fölírható, tehát ez a  $\mathcal{B} \leftarrow \mathcal{E}$  áttérés mátrixa, azaz

$$\mathbf{X}_{\mathcal{B} \leftarrow \mathcal{E}} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 2 & 4 \end{bmatrix}.$$

Ennek inverze a keresett mátrix:

$$\mathbf{Y}_{\mathcal{E} \leftarrow \mathcal{B}} = \mathbf{X}_{\mathcal{B} \leftarrow \mathcal{E}}^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 2 & 4 \end{bmatrix}^{-1} = \begin{bmatrix} 2 & -2 & 1 \\ -1 & 3 & -2 \\ 0 & -1 & 1 \end{bmatrix}.$$

Ennek oszlopvektorai adják a  $\mathcal{B}$  vektorainak  $\mathcal{E}$ -beli alakját. □

## Feladatok

### Igaz – hamis

5.1. A négyzetes  $\mathbf{A}$  mátrix pontosan akkor invertálható, ha elemi sorműveletekkel megkapható az  $\mathbf{I}$  mátrixból.

5.2. Ha elemi sorműveletek  $\mathbf{A}$ -t  $\mathbf{B}$ -be viszik, akkor az inverz sorműveletek  $\mathbf{B}$ -t  $\mathbf{A}$ -ba viszik.

5.3. Ha elemi sorműveletek  $\mathbf{A}$ -t  $\mathbf{B}$ -be viszik, akkor az inverz sorműveletek fordított sorrendben végrehajtva  $\mathbf{B}$ -t  $\mathbf{A}$ -ba viszik.

### Műveleti azonosságok

5.4. ÖSSZEADÁS ÉS SKALÁRRAL SZORZÁS TULAJDONSÁGAI Legyen  $\mathbf{A}$ ,  $\mathbf{B}$  és  $\mathbf{C}$  azonos típusú ( $m \times n$ -es) mátrix,  $c$  és  $d$  legyenek skalárok. Ekkor

- $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$  (felcserélhetőség, kommutativitás)
- $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$  (csoportosíthatóság, asszociativitás)
- $\mathbf{A} + \mathbf{O}_{m \times n} = \mathbf{A}$  (zérusmátrix)
- $\mathbf{A} + (-\mathbf{A}) = \mathbf{O}_{m \times n}$  (ellentett létezése)
- $c(d\mathbf{A}) = (cd)\mathbf{A}$  (csoportosíthatóság)
- $(c + d)\mathbf{A} = c\mathbf{A} + d\mathbf{A}$  (disztributivitás)
- $c(\mathbf{A} + \mathbf{B}) = c\mathbf{A} + c\mathbf{B}$  (disztributivitás)
- $0\mathbf{A} = \mathbf{O}_{m \times n}$ ,  $1\mathbf{A} = \mathbf{A}$ ,  $-1\mathbf{A} = -\mathbf{A}$

5.5. Egy algebrai kifejezésben végrehajtjuk az alábbi helyettesítést:

$$u = 3x_1 + 2x_2 + 4x_3$$

$$v = x_1 - 3x_2 + x_3$$

$$w = 2x_1 - x_2 - 3x_3$$

Írjuk fel e lineáris helyettesítést mátrixszorzatos alakban. Legyen  $(u^2 + v^2 + w^2)(2u - v - w)$  az a kifejezés, melyben a helyettesítést elvégezzük. Írjuk fel e kifejezést a helyettesítés előtt és után mátrixműveletek segítségével!

### Számítási feladatok

Bontsuk fel a következő mátrixokat elemi mátrixok szorzatára!

5.6.  $\begin{bmatrix} 1 & 3 \\ 2 & 8 \end{bmatrix}$

5.7.  $\begin{bmatrix} 1 & 2 \\ -2 & -1 \end{bmatrix}$

5.8.  $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

5.9.  $\begin{bmatrix} 2 & 4 \\ 3 & 8 \end{bmatrix}$

5.10.  $\begin{bmatrix} 2 & 0 & 4 \\ 0 & 2 & 0 \\ 3 & 2 & 7 \end{bmatrix}$

5.11.  $\begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 2 \\ 2 & 4 & 5 \end{bmatrix}$

5.12. Határozzuk meg az összes olyan  $2 \times 2$ -es  $\mathbf{A}$  mátrixot, melyre  $\mathbf{A}^2 = \mathbf{O}$ . Másként fogalmazva határozzuk meg a nullmátrix összes négyzetgyökét!

5.13. Számítsuk ki az

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

mátrix  $k$ -adik hatványait!

5.14. Írjuk fel a mátrixszorzás definícióját az Einstein-konvenciót használva.

### Blokkmátrixok

5.15. Mutassuk meg, hogy ha  $\mathbf{A}$  és  $\mathbf{D}$  invertálható mátrixok, akkor a következő ún. blokkdiagonális mátrix invertálható, és inverze

$$\begin{bmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} & \mathbf{O} \\ \mathbf{O} & \mathbf{D}^{-1} \end{bmatrix}.$$

továbbá tetszőleges, de megfelelő típusú  $\mathbf{B}$  mátrix esetén

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{D}^{-1} \\ \mathbf{O} & \mathbf{D}^{-1} \end{bmatrix}.$$

5.16. Mutassuk meg, hogy ha  $\mathbf{A}$  és  $\mathbf{D}$  négyzetes mátrixok, akkor

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{X} & -\mathbf{X}\mathbf{B}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}\mathbf{X} & \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{C}\mathbf{X}\mathbf{B}\mathbf{D}^{-1} \end{bmatrix},$$

ahol  $\mathbf{X} = (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}$ , és feltételezzük, hogy minden felírt mátrixinverz létezik.

Az előbbi két feladat valamelyikének felhasználásával számítsuk ki az alábbi mátrixok inverzét!

5.17.  $\begin{bmatrix} 2 & 3 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 7 & 3 & 3 \\ 0 & 0 & 8 & 1 & 2 \\ 0 & 0 & 4 & 4 & 3 \end{bmatrix}$

5.18.  $\begin{bmatrix} 2 & 3 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 & 1 \\ 0 & 0 & 7 & 3 & 3 \\ 0 & 0 & 8 & 1 & 2 \\ 0 & 0 & 4 & 4 & 3 \end{bmatrix}$

$$5.19. \begin{bmatrix} 2 & 3 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

### Bizonyítások

5.20. Bizonyítsuk be, hogy ha  $c\mathbf{A} = \mathbf{O}$ , akkor vagy  $c = 0$ , vagy  $\mathbf{A} = \mathbf{O}$ .

5.21. **SORMŰVELET ÉS ELEMI MÁTRIX INVERZE** Az  $S_i \leftrightarrow S_j$  sorművelethez tartozó elemi mátrixot jelölje  $\mathbf{E}_{ij}$ , a  $cS_i$ -hez tartozót  $\mathbf{E}_i(c)$  és a  $S_i + cS_j$  sorművelethez tartozót  $\mathbf{E}_{ij}(c)$ . Mutassuk meg, hogy  $\mathbf{E}_{ij}^{-1} = \mathbf{E}_{ij}$ ,  $\mathbf{E}_i(c)^{-1} = \mathbf{E}_i(\frac{1}{c})$  és

$$\mathbf{E}_{ij}(c)^{-1} = \mathbf{E}_{ij}(-c).$$

5.22. Mutassuk meg, hogy ha  $\mathbf{A}$  fölcserélhető  $\mathbf{B}$ -vel és  $\mathbf{B}$  invertálható, akkor  $\mathbf{A}$  fölcserélhető  $\mathbf{B}^{-1}$ -gyel is.

### Absztrakció

5.23. **INVERTÁLHATÓ MŰVELET** Legyen  $\odot$  egy  $H$ -n értelmezett kétváltozós művelet, azaz egy  $H^2 \rightarrow H$  függvény. Fogalmazzuk meg, mit értünk azon, hogy  $\odot$  invertálható egy  $R \subseteq H$  részhalmazán. Hogyan változik a definíció, ha a művelet kommutatív?

5.24. **ELEM INVERZE** Legyen  $\odot$  egy  $H$ -n értelmezett kétváltozós művelet.

1. Mit értünk azon, hogy  $e \in H$  e művelet semleges eleme?
2. Mit értünk azon, hogy  $b \in H$  az  $a \in H$  elem inverze?

### Műveletek speciális mátrixokkal

A gyakorlatban gyakran találkozunk olyan speciális mátrixokkal, melyekkel a műveletek egyszerűbben végezhetők el, és olyanokkal, melyek mátrixműveletek segítségével definiálhatók.

**Diagonális mátrixok** A mátrixműveletek definíciói alapján magától értetődő, hogy diagonális mátrixokkal hogyan végezhetők el a mátrixműveletek. Elsősorban egyszerű példákat mutatunk.

5.23. PÉLDA (MŰVELETEK DIAGONÁLIS MÁTRIXOKKAL). Legyen  $\mathbf{A} = \text{diag}(1, 2, 3)$ ,  $\mathbf{B} = \text{diag}(5, 4, 3)$ . Ellenőrizzük az alábbi számításokat, és fogalmazzunk meg általános összefüggéseket diagonális mátrixokkal végzett műveletekről.

$$\begin{aligned}\mathbf{AB} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 9 \end{bmatrix} \\ \mathbf{A}^2 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{bmatrix} \\ \mathbf{A}^{-1} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix} \\ \mathbf{A}^k &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}^k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2^k & 0 \\ 0 & 0 & 3^k \end{bmatrix}, \text{ ahol } k \text{ egész szám.}\end{aligned}$$

5.24. TÉTEL (MŰVELETEK DIAGONÁLIS MÁTRIXOKKAL). Legyen  $\mathbf{A} = \text{diag}(a_1, a_2, \dots, a_n)$ ,  $\mathbf{B} = \text{diag}(b_1, b_2, \dots, b_n)$ , és legyen  $k$  egész szám. Ekkor

- a)  $\mathbf{AB} = \text{diag}(a_1 b_1, a_2 b_2, \dots, a_n b_n)$ ,
- b)  $\mathbf{A}^k = \text{diag}(a_1^k, a_2^k, \dots, a_n^k)$ , speciálisan
- c)  $\mathbf{A}^{-1} = \text{diag}(a_1^{-1}, a_2^{-1}, \dots, a_n^{-1})$ .

**Permutáló mátrixok és kigyók** Könnyen kezelhetők a diagonális mátrixok sorainak permutációjával kapott mátrixok.

5.25. PÉLDA (SOROK PERMUTÁCIÓJA MÁTRIXSZORZÁSSAL). Alkalmazzunk több sorcserét az egységmátrixon. Az így kapott mátrixszal balról való szorzás milyen hatással van a beszorzott mátrixra? Szemléltessük ezt  $\mathbf{I}_4$ -en az  $S_1 \leftrightarrow S_2$ ,  $S_2 \leftrightarrow S_4$  sorcserékkel.

**MEGOLDÁS.** Legyen  $\mathbf{P}$  az  $\mathbf{I}_4$ -ből a fent megadott két sorcserével kapott mátrix, és legyen  $\mathbf{A}$  egy tetszőleges  $4 \times n$ -es mátrix. Ekkor

$$\mathbf{I}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{S_1 \leftrightarrow S_2} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{S_2 \leftrightarrow S_4} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} = \mathbf{P}$$

$$\mathbf{PA} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} = \begin{bmatrix} a_{21} & a_{22} \\ a_{41} & a_{42} \\ a_{31} & a_{32} \\ a_{11} & a_{12} \end{bmatrix}$$

Azt tapasztaljuk, hogy a  $\mathbf{PA}$  az  $\mathbf{A}$ -ból a soroknak épp azzal a permutációjával kapható, amely permutációval  $\mathbf{I}$ -ből a  $\mathbf{P}$ -t kaptuk.  $\square$

**5.26. DEFINÍCIÓ (PERMUTÁLÓ MÁTRIX, KÍGYÓ).** A *diagonális mátrixok sorainak permutációjával kapott mátrixot kígyónak (más néven transzverzálisnak) nevezzük, speciálisan az egységmátrixból ugyanígy kapott mátrixot permutáló mátrixnak (más néven permutációmátrixnak) hívjuk.*

► Könnyen látható, hogy a permutáló mátrix olyan négyzetes mátrix, melynek minden sorában és minden oszlopában *pontosan* egy 1-es van, az összes többi elem 0. A kígyó olyan négyzetes mátrix, melynek minden sorában és minden oszlopában *legfeljebb* egy nemnulla elem van.

► Minden kígyó megkapható egy diagonális mátrixból oszlopcserékkel is. Egy diagonális mátrixból akkor is kígyót kapunk, ha a sorok permutációja mellett az oszlopokat is permutáljuk.

**5.27. PÉLDA (KÍGYÓK).** Az alábbi mátrixok mindegyike kígyó, az utolsó kettő egyúttal permutáló mátrix is:

$$\begin{bmatrix} 0 & 5 & 0 \\ 0 & 0 & 9 \\ 3 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & \alpha & 0 & 0 \\ \gamma & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \beta \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

**5.28. TÉTEL (MŰVELETEK PERMUTÁLÓ MÁTRIXOKKAL).** Bármely két azonos méretű permutáló mátrix szorzata és egy permutáló mátrix bármely egész kitevős hatványa permutáló mátrix. Permutáló mátrix inverze meggyezik a transzponáltjával, azaz ha  $\mathbf{P}$  permutáló mátrix, akkor

$$\mathbf{P}^{-1} = \mathbf{P}^T.$$

**BIZONYÍTÁS.** Legyen  $\mathbf{A}$  és  $\mathbf{B}$  két permutáló mátrix. Szorzatuk sorvektorai  $\mathbf{a}_{i*}\mathbf{B}$  alakúak, ahol  $\mathbf{a}_{i*}$  megegyezik valamelyik standard egységvektorral, pl.  $\mathbf{a}_{i*} = \mathbf{e}_k$ . Ekkor a szorzatvektornak csak az az eleme 1, amelyik oszlop  $\mathbf{e}_k$ -vel megegyezik, és ilyen oszlop pontosan egy van. Tehát a szorzatmátrix minden sorában pontosan egy elem 1, a többi 0. Oszlopokra az állítás hasonlóan bizonyítható. A szorzatra vonatkozó állítás természetes következménye a pozitív egész kitevős hatványokra vonatkozó állítás. A negatív egész kitevőkre is igaz az állítás. Ennek igazolásához elég az inverzre belátni.

Tekintsük a  $\mathbf{PP}^T$  szorzatot. A  $(\mathbf{PP}^T)_{ij}$  elem a  $\mathbf{P}_{i*}$  vektornak és a  $(\mathbf{P}^T)_{*j} = \mathbf{P}_{i*}$  vektornak a szorzata, vagyis 1, míg

$$(\mathbf{PP}^T)_{ij} = (\mathbf{P})_{i*}(\mathbf{P}^T)_{*j} = (\mathbf{P})_{i*} \cdot (\mathbf{P})_{j*},$$

azaz a szorzat  $i$ -edik sorának  $j$ -edik eleme a  $\mathbf{P}$   $i$ -edik és  $j$ -edik sorvektorának skalárszorzata, ami 0, mivel két különböző sorban az 1-es különböző helyen van.  $\square$

► Az alábbi példa szemlélteti a tételben kimondott egyszerű állítást:

$$\mathbf{PP}^T = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

**Háromszögmátrixok** A Gauss-kiküszöbölés végrehajtásakor az együtthatómátrixot lépcsős alakra transzformáltuk, melyben a főátló alatt mindig csak nullák szerepelnek. Az ilyen mátrixok nem csak a Gauss-kiküszöbölésnél fontosak.

**5.29. DEFINÍCIÓ (HÁROMSZÖGMÁTRIX).** Azokat a mátrixokat, melyek főátlója alatt csak 0-elemek szerepelnek felső háromszögmátrixnak, azokat, melyek főátlója fölött csak 0-elemek vannak alsó háromszögmátrixnak nevezzük. Ha egy háromszögmátrix főátlójában csupa 1-es áll, egység háromszögmátrixról beszélünk.

A Gauss-kiküszöbölésnél kapott felső háromszögmátrixhoz hasonlóan azok az egyenletrendszerek is megoldhatók csak behelyettesítésekkel, amelyek együtthatómátrixa alsó háromszögmátrix. A különbség kizárólag annyi, hogy ekkor az első egyenlettel kezdjük, és az első változó értékét határozzuk meg először. Például az

$$\begin{aligned} x &= 3 \\ 2x + 3y &= 3 \\ 2x + y + 2z &= 3 \end{aligned}$$

egyenletrendszer első egyenletéből  $x = 3$ , a másodikba való behelyettesítés után  $y = -1$ , végül a harmadikba való behelyettesítés után  $z = -1$ .<sup>2</sup>

<sup>2</sup> Az angol nyelvű lineáris algebra tankönyvek különbséget tesznek a felső és az alsó háromszögmátrixú egyenletrendszerek megoldása között. *Forward substitution*, illetve *backward substitution* a neve a behelyettesítésnek ha alsó, illetve ha felső háromszögmátrix az együttható mátrix. Ez arra utal, hogy a változókat előre vagy hátra haladva számoljuk ki. Mi nem fogjuk használni e finom különbségtételt.



**5.30. TÉTEL (MŰVELETEK HÁROMSZÖGMÁTRIXOKKAL).** *Felső háromszög-mátrixok összege, szorzata, és invertálható felső háromszögmátrix inverze felső háromszögmátrix. Analóg tétel igaz az alsó háromszögmátrixokra is. Egy háromszögmátrix pontosan akkor invertálható, ha főátlóbeli elemeinek egyike sem zérus.*

A bizonyítást feladatként az Olvasóra hagyjuk.

**Szimmetrikus és ferdén szimmetrikus mátrixok** Gyakran használunk olyan mátrixokat, melyekben az elemek egyenlők vagy ellentettjei a főátlóra nézve szimmetrikusan elhelyezkedő párjuknak. E tulajdonság a transzponálttal könnyen kifejezhető.

**5.31. DEFINÍCIÓ (SZIMMETRIKUS ÉS FERDÉN SZIMMETRIKUS MÁTRIXOK).** *A négyzetes  $\mathbf{A}$  mátrixot szimmetrikusnak nevezzük, ha  $\mathbf{A}^T = \mathbf{A}$ , és ferdén szimmetrikusnak nevezzük, ha  $\mathbf{A}^T = -\mathbf{A}$ .*

**5.32. PÉLDA (SZIMMETRIKUS ÉS FERDÉN SZIMMETRIKUS MÁTRIXOK).** *Az alábbi mátrixok közül az  $\mathbf{A}$  szimmetrikus, a  $\mathbf{B}$  ferdén szimmetrikus, a  $\mathbf{C}$  egyik osztályba sem tartozik.*

$$\mathbf{A} = \begin{bmatrix} 5 & 6 & 1 \\ 6 & 2 & 0 \\ 1 & 0 & 3 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 1 & -2 \\ -1 & 0 & 3 \\ 2 & -3 & 0 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 9 & 9 \\ -9 & 2 & 9 \\ -9 & -9 & 3 \end{bmatrix}.$$

Ha  $\mathbf{A}$  ferdén szimmetrikus, akkor minden elemére  $a_{ij} = -a_{ji}$ , azaz  $i = j$  esetén  $a_{ii} = -a_{ii}$ . Ez csak  $a_{ii} = 0$  esetén áll fenn, azaz a ferdén szimmetrikus mátrixok főátlójában csupa 0 áll.

**5.33. ÁLLÍTÁS (MŰVELETEK (FERDÉN) SZIMMETRIKUS MÁTRIXOKKAL).** *Szimmetrikus mátrixok összege, skalárszorosa, inverze szimmetrikus. Ferdén szimmetrikus mátrixok összege, skalárszorosa, inverze ferdén szimmetrikus.*

Az állítás bizonyítását feladatként az olvasóra hagyjuk.

**5.34. TÉTEL (FELBONTÁS SZIMMETRIKUS ÉS FERDÉN SZIMMETRIKUS MÁTRIX ÖSSZEGÉRE).** *Minden négyzetes mátrix előáll egy szimmetrikus és egy ferdén szimmetrikus mátrix összegeként, nevezetesen minden  $\mathbf{A}$  négyzetes mátrixra*

$$\mathbf{A} = \underbrace{\frac{1}{2}(\mathbf{A} + \mathbf{A}^T)}_{\text{szimmetrikus}} + \underbrace{\frac{1}{2}(\mathbf{A} - \mathbf{A}^T)}_{\text{ferdén szimm.}}$$

**BIZONYÍTÁS.** Ha egy mátrix szimmetrikus, konstansszorosa is, így elég

megmutatni, hogy az  $\mathbf{A} + \mathbf{A}^\top$  mátrix szimmetrikus:

$$(\mathbf{A} + \mathbf{A}^\top)^\top = \mathbf{A}^\top + (\mathbf{A}^\top)^\top = \mathbf{A}^\top + \mathbf{A} = \mathbf{A} + \mathbf{A}^\top$$

Hasonlóképp  $\mathbf{A} - \mathbf{A}^\top$  ferdén szimmetrikus, hiszen

$$(\mathbf{A} - \mathbf{A}^\top)^\top = \mathbf{A}^\top - (\mathbf{A}^\top)^\top = \mathbf{A}^\top - \mathbf{A} = -(\mathbf{A} - \mathbf{A}^\top)$$

Végül a két mátrix összege valóban  $\mathbf{A}$ :

$$\frac{1}{2}(\mathbf{A} + \mathbf{A}^\top) + \frac{1}{2}(\mathbf{A} - \mathbf{A}^\top) = \frac{1}{2}\mathbf{A} + \frac{1}{2}\mathbf{A}^\top + \frac{1}{2}\mathbf{A} - \frac{1}{2}\mathbf{A}^\top = \mathbf{A}. \quad \square$$

Fontos következményei lesznek az alábbi egyszerű állításnak.

**5.35. TÉTEL** ( $\mathbf{A}^\top \mathbf{A}$  ÉS  $\mathbf{A} \mathbf{A}^\top$  SZIMMETRIKUS). Az  $\mathbf{A}^\top \mathbf{A}$  és az  $\mathbf{A} \mathbf{A}^\top$  mátrixok tetszőleges  $\mathbf{A}$  mátrix esetén szimmetrikusak.

**BIZONYÍTÁS.**  $(\mathbf{A} \mathbf{A}^\top)^\top = (\mathbf{A}^\top)^\top \mathbf{A}^\top = \mathbf{A} \mathbf{A}^\top$ . Az állítás másik fele ugyanígy bizonyítható.  $\square$

*Mátrix és diád összegének inverze\** Összegmátrix inverzére – a valósok összegének inverzéhez hasonlóan – nincs egyszerű képlet, de speciális mátrixokra nagyon hasznos eredmények vannak. Ilyen a Sherman–Morrison-formula is, melynek segítségével leírhatjuk, hogy hogyan változik egy mátrix inverze, ha egyetlen csak egyetlen elemén változtatunk.

**5.36. TÉTEL** (SHERMAN–MORRISON-FORMULA). Tegyük fel, hogy az  $\mathbf{A} \in \mathbb{R}^{n \times n}$  mátrix invertálható, és  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$  két olyan vektor, hogy  $1 + \mathbf{v}^\top \mathbf{A}^{-1} \mathbf{u} \neq 0$ . Ekkor  $\mathbf{A} + \mathbf{u} \mathbf{v}^\top$  invertálható, és

$$(\mathbf{A} + \mathbf{u} \mathbf{v}^\top)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1} \mathbf{u} \mathbf{v}^\top \mathbf{A}^{-1}}{1 + \mathbf{v}^\top \mathbf{A}^{-1} \mathbf{u}}.$$

**BIZONYÍTÁS.** Elég megmutatni, hogy

$$\left( \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1} \mathbf{u} \mathbf{v}^\top \mathbf{A}^{-1}}{1 + \mathbf{v}^\top \mathbf{A}^{-1} \mathbf{u}} \right) (\mathbf{A} + \mathbf{u} \mathbf{v}^\top) = \mathbf{I},$$

mert ez a formula igazolása mellett azt is bizonyítja, hogy  $\mathbf{A} + \mathbf{u} \mathbf{v}^\top$

invertálható.

$$\begin{aligned}
 & \left( \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1} \mathbf{u} \mathbf{v}^T \mathbf{A}^{-1}}{1 + \mathbf{v}^T \mathbf{A}^{-1} \mathbf{u}} \right) (\mathbf{A} + \mathbf{u} \mathbf{v}^T) \\
 &= \mathbf{A}^{-1} \mathbf{A} + \mathbf{A}^{-1} \mathbf{u} \mathbf{v}^T - \frac{\mathbf{A}^{-1} \mathbf{u} \mathbf{v}^T \mathbf{A}^{-1} \mathbf{A}}{1 + \mathbf{v}^T \mathbf{A}^{-1} \mathbf{u}} - \frac{\mathbf{A}^{-1} \mathbf{u} \mathbf{v}^T \mathbf{A}^{-1} \mathbf{u} \mathbf{v}^T}{1 + \mathbf{v}^T \mathbf{A}^{-1} \mathbf{u}} \\
 &= \mathbf{I} + \mathbf{A}^{-1} \mathbf{u} \mathbf{v}^T - \frac{\mathbf{A}^{-1} \mathbf{u} \mathbf{v}^T}{1 + \mathbf{v}^T \mathbf{A}^{-1} \mathbf{u}} - \frac{\mathbf{A}^{-1} \mathbf{u} (\mathbf{v}^T \mathbf{A}^{-1} \mathbf{u}) \mathbf{v}^T}{1 + \mathbf{v}^T \mathbf{A}^{-1} \mathbf{u}} \\
 &= \mathbf{I} + \mathbf{A}^{-1} \mathbf{u} \mathbf{v}^T - \frac{\mathbf{A}^{-1} \mathbf{u} (1 + \mathbf{v}^T \mathbf{A}^{-1} \mathbf{u}) \mathbf{v}^T}{1 + \mathbf{v}^T \mathbf{A}^{-1} \mathbf{u}} \\
 &\stackrel{*}{=} \mathbf{I} + \mathbf{A}^{-1} \mathbf{u} \mathbf{v}^T - \mathbf{A}^{-1} \mathbf{u} \mathbf{v}^T \\
 &= \mathbf{I}.
 \end{aligned}$$

A  $*$ -gal jelzett egyenlőségnél azt használtuk ki, hogy  $1 \times 1$ -es mátrixszal való szorzás egybeesik a skalárral való szorzással, a skalár tényező pedig egy mátrixszorzatban áttehető más helyre, így az adott törtkifejezésben egyszerűsíthettünk vele.  $\square$

A Sherman–Morrison-formula sokhelyütt használható, mi itt egyet emelünk ki: megvizsgáljuk, hogy hogyan változik egy mátrix inverze, ha a mátrixnak csak egyetlen elemén változtatunk.

**5.37. PÉLDA (INVERZ VÁLTOZÁSA).** Legyen  $\mathbf{A}$  invertálható mátrix, és változtassunk meg az  $a_{ij}$  elemet  $a_{ij} + \varepsilon$ -ra. Fejezzük ki az így kapott mátrix inverzét  $\mathbf{A}^{-1}$  segítségével.

**MEGOLDÁS.** Első lépésként kifejezzük az új mátrixot  $\mathbf{A}$ -ból mátrixműveletekkel. Legyen  $\mathbf{e}_i$  és  $\mathbf{e}_j$  az  $i$ -edik és  $j$ -edik standard egységvektor. Ekkor a módosított mátrix

$$\mathbf{B} = \mathbf{A} + \varepsilon \mathbf{e}_i \mathbf{e}_j^T.$$

Erre alkalmazható a Sherman–Morrison-formula az  $\mathbf{u} = \mathbf{e}_i$  és  $\mathbf{v} = \varepsilon \mathbf{e}_j$  választással.

$$\begin{aligned}
 \mathbf{B}^{-1} &= \left( \mathbf{A} + \varepsilon \mathbf{e}_i \mathbf{e}_j^T \right)^{-1} \\
 &= \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1} \mathbf{e}_i (\varepsilon \mathbf{e}_j^T) \mathbf{A}^{-1}}{1 + \varepsilon \mathbf{e}_j^T \mathbf{A}^{-1} \mathbf{e}_i} \\
 &= \mathbf{A}^{-1} - \varepsilon \frac{(\mathbf{A}^{-1})_{*i} (\mathbf{A}^{-1})_{j*}}{1 + \varepsilon (\mathbf{A}^{-1})_{ji}}
 \end{aligned}$$

**5.38. PÉLDA (INVERZ VÁLTOZÁSA SZÁMPÉLDÁN).** Adva van egy  $\mathbf{A}$  mátrix és annak inverze:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 0 & 0 & 3 \\ 3 & 0 & 0 & 2 \\ 4 & 3 & 2 & 1 \end{bmatrix} \quad \mathbf{A}^{-1} = \begin{bmatrix} 0 & -2/5 & 3/5 & 0 \\ -2/5 & 7/5 & -8/5 & 3/5 \\ 3/5 & -8/5 & 7/5 & -2/5 \\ 0 & 3/5 & -2/5 & 0 \end{bmatrix}.$$

Változtassuk meg  $a_{11}$  értékét 1-ről  $11/10$ -re. Az így kapott mátrixot jelölje **B**. Határozzuk meg inverzét!

**MEGOLDÁS.** Az előző példa alkalmazásával

$$\begin{aligned}
 \mathbf{B}^{-1} &= \mathbf{A}^{-1} - \frac{1}{10} \frac{(\mathbf{A}^{-1})_{*1}(\mathbf{A}^{-1})_{1*}}{1 + \frac{1}{10}(\mathbf{A}^{-1})_{11}} \\
 &= \begin{bmatrix} 0 & -2/5 & 3/5 & 0 \\ -2/5 & 7/5 & -8/5 & 3/5 \\ 3/5 & -8/5 & 7/5 & -2/5 \\ 0 & 3/5 & -2/5 & 0 \end{bmatrix} - \frac{1}{10} \frac{\begin{bmatrix} 0 \\ -2/5 \\ 3/5 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & -2/5 & 3/5 & 0 \end{bmatrix}}{1 + \frac{1}{10} \cdot 0} \\
 &= \begin{bmatrix} 0 & -2/5 & 3/5 & 0 \\ -2/5 & 7/5 & -8/5 & 3/5 \\ 3/5 & -8/5 & 7/5 & -2/5 \\ 0 & 3/5 & -2/5 & 0 \end{bmatrix} - \frac{1}{10} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 4/25 & -6/25 & 0 \\ 0 & -6/25 & 9/25 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\
 &= \begin{bmatrix} 0 & -2/5 & 3/5 & 0 \\ -2/5 & 173/125 & -197/125 & 3/5 \\ 3/5 & -197/125 & 341/250 & -2/5 \\ 0 & 3/5 & -2/5 & 0 \end{bmatrix}.
 \end{aligned}$$

Tizedestörtekkel számolva:

$$\begin{aligned}
 \mathbf{A}^{-1} &= \begin{bmatrix} 0.0 & -0.4 & 0.6 & 0.0 \\ -0.4 & 1.4 & -1.6 & 0.6 \\ 0.6 & -1.6 & 1.4 & -0.4 \\ 0.0 & 0.6 & -0.4 & 0.0 \end{bmatrix}, \\
 \mathbf{B}^{-1} &= \begin{bmatrix} 0.000 & -0.400 & 0.600 & 0.000 \\ -0.400 & 1.384 & -1.576 & 0.600 \\ 0.600 & -1.576 & 1.364 & -0.400 \\ 0.000 & 0.600 & -0.400 & 0.000 \end{bmatrix}. \quad \square
 \end{aligned}$$

**Gyorsszorzás\*** Két  $2 \times 2$ -es mátrix szokásos módon való összeszorzásához 8 szorzásra és 4 összeadásra van szükség. Strassen 1969-ben egy olyan módszert talált, mellyel e mátrixszorzást 7 szorzással is el lehet végezni, igaz azon az áron, hogy az összeadások száma 16-ra nő.

5.1 (STRASSEN-FORMULÁK). Legyen  $\mathbf{A}$ ,  $\mathbf{B}$  és  $\mathbf{C}$  is  $2 \times 2$ -es. A  $\mathbf{C} = \mathbf{AB}$  szorzás elvégezhető a következő formulákkal:

$$\begin{aligned} d_1 &= (a_{11} + a_{22})(b_{11} + b_{22}) & c_{11} &= d_1 + d_4 - d_5 + d_7 \\ d_2 &= (a_{21} + a_{22})b_{11} & c_{21} &= d_2 + d_4 \\ d_3 &= a_{11}(b_{12} - b_{22}) & c_{12} &= d_3 + d_5 \\ d_4 &= a_{22}(-b_{11} + b_{21}) & c_{22} &= d_1 + d_3 - d_2 + d_6 \\ d_5 &= (a_{11} + a_{12})b_{22} \\ d_6 &= (-a_{11} + a_{21})(b_{11} + b_{12}) \\ d_7 &= (a_{12} - a_{22})(b_{21} + b_{22}) \end{aligned}$$

Az ötlet nagyszerűsége abban van, hogy e módszer kiterjeszthető tetszőleges méretű négyzetes mátrixokra is, és elegendően nagy  $n$ -ekre az e módon elvégzett mátrixszorzás műveletigénye kisebb lesz a hagyományos módon elvégzetténél. A standard mátrixszorzás műveletigénye  $2n^3 - n^2$  (ebből  $n^3$  szorzás és  $n^3 - n^2$  összeadás – gondoljunk utána!), a Strassen-formulákkal való szorzásé  $n = 2^k$  esetén legföljebb  $7 \cdot 7^k - 6 \cdot 4^k$ . Ez  $n = 2^{10}$  esetén már kevesebb műveletet ad. Az általánosítás lényege, hogy a Strassen-formulák  $2 \times 2$ -es blokkmátrixokra is használhatók, mert a szorzás kommutativitását nem használják, így ha  $M(n)$  jelöli két  $n \times n$ -es mátrix összeszorzásához szükséges szorzások, és  $S(n)$  a szükséges összeadások számát, akkor  $M(2n) \leq 7M(n)$  és  $S(2n) \leq 18n^2 + 7S(n)$ . Az  $M(1) = 1$ ,  $S(1) = 0$  kezdeti feltételeket is használva megmutatható, hogy  $M(2^k) \leq 7^k$ ,  $S(2^k) \leq 6(7^k - 4^k)$ . E képletekből a felső egészrész jelét használva és a  $k = \lceil \log_2 n \rceil$  jelöléssel az műveletek összámára a  $cn^{\log_2 7} \leq cn^{2.81}$  felső becslést kapjuk, ami a  $2n^3 - n^2$  értéknél jobb, függetlenül a  $c$  konstans konkrét értékétől. Mivel a két összeszorzandó mátrix mindegyikének mind az  $n^2$  elemét használni kell, ezért a szükséges műveletek számának alsó becslése  $cn^2$ . A  $cn^{2.81}$  felső becslés mára  $cn^{2.376}$ -ra lett javítva (Coppersmith és Winograd, 1990), de az a sejtés, hogy a kitevő 2-re, de legalább  $2 + \varepsilon$ -ra lenyomható, ahol  $\varepsilon$  tetszőlegesen kis pozitív szám.

A módszer gyengéje numerikus instabilitása, így a gyakorlatban csak bizonyos mátrixokra érdemes használni, például nagyméretű egészelemű mátrixokra tetszőleges pontosságú aritmetika használata esetén.

## Feladatok

**5.25\* IGAZ – HAMIS** Döntsük el, igazak-e az alábbi állítások? Válaszunkat indokoljuk!

- Szimmetrikus mátrixok összege és skalárszorosa is szimmetrikus, így szimmetrikus mátrixok tetszőleges lineáris kombinációja is szimmetrikus.
- Ferdén szimmetrikus mátrixok összege és skalárszorosa is ferdén szimmetrikus, így ferdén szimmetrikus mátrixok tetszőleges lineáris kombinációja is ferdén szimmetrikus.
- Minden lépcsős alakú mátrix felső háromszögmátrix.
- Minden felső háromszögmátrix lépcsős alakú.

**5.26.** Számítsuk ki az alábbi mátrixok inverzeit, négyzetét és köbét!

$$\begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 4 \\ 3 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 2 \\ 0 & 4 & 0 \\ 3 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 3 & 0 \\ 4 & 0 & 0 & 0 \end{bmatrix}.$$

**5.27.** Hogyan oldanánk meg a következő egyenletrendszert a lehető legkevesebb lépésben?

$$\begin{aligned} x + 4y + 3z + 5w &= 3 \\ 6x + 3y &= 3 \\ 2x + 3y + 2z &= 3 \\ 2x + 4y + 3z + 5w &= 4 \end{aligned}$$

## Bizonyítások

**5.28.** Mutassuk meg, hogy minden permutáló mátrix oszlopcserékkel is megkapható az egységmátrixból, és hogy permutáló mátrixszal jobbról való szorzás a beszorzott mátrix oszlopain ugyanazt a permutációt hajtja végre, mint amellyel a permutáló mátrix az egységmátrixból megkapható.

**5.29.** Bizonyítsuk be, hogy bármely két azonos méretű kigyó szorzata és egy kigyó bármely pozitív egész kitevős hatványa kigyó.

**5.30\*** Mutassuk meg, hogy egy  $\mathbf{K}$  kigyó pontosan akkor invertálható, ha minden sorában pontosan egy elem nem 0, és ekkor inverze megkapható úgy, hogy minden nemnulla elem helyébe annak reciprokát írjuk, majd az így kapott mátrixot transzponáljuk.

**5.31\* GYORSINVERTÁLÁS** Legyen  $\mathbf{B} = \mathbf{A}^{-1}$ , mindketten  $2 \times 2$ -es mátrixok. Mutassuk meg, hogy az alábbi eljárással definiált mátrixinvertálás segítségével  $n \times n$ -es mátrixokra olyan algoritmus készíthető, melynek műveletigénye legfeljebb  $cn^{2.81}$ .

$$\begin{aligned} c_1 &= a_{11}^{-1} & b_{12} &= c_3 c_6 \\ c_2 &= a_{21} c_1 & b_{21} &= c_6 c_2 \\ c_3 &= c_1 a_{12} & c_7 &= c_3 b_{21} \\ c_4 &= a_{21} c_3 & b_{11} &= c_1 - c_7 \\ c_5 &= c_4 - a_{22} & b_{22} &= -c_6 \\ c_6 &= c_5^{-1} \end{aligned}$$

## Mátrixfelbontások

Mátrixfelbontáson egy mátrixnak adott tulajdonságú mátrixok szorzataként való fölírását értjük. Egy ilyen felbontással már találkoztunk, amikor invertálható mátrixot elemi mátrixok szorzatára bontottunk. E szakaszban a kiküszöbölési eljárásra épülő további felbontásokkal találkozunk. Ezek egyike, az LU-felbontás bizonyos lineáris algebrai feladatok számítógépes megoldásának gyakran használt eszköze.

**Az LU-felbontás** Tegyük fel, hogy egy  $\mathbf{A}$  mátrixból el lehet jutni egy  $\mathbf{U}$  felső háromszögalakhoz csak olyan sorműveletekkel, melyekben egy sor konstansszorosát valamely alatta lévő sorhoz adjuk. Minden ilyen elemi sorművelethez olyan elemi mátrix tartozik, mely alsó háromszög alakú. Ekkor tehát léteznek olyan  $\mathbf{E}_1, \dots, \mathbf{E}_k$  elemi alsó háromszögmátrixok, melyekre

$$\mathbf{E}_k \dots \mathbf{E}_1 \mathbf{A} = \mathbf{U}. \quad (5.2)$$

Innen

$$\mathbf{A} = (\mathbf{E}_k \dots \mathbf{E}_1)^{-1} \mathbf{U}, \quad (5.3)$$

ahol  $(\mathbf{E}_k \dots \mathbf{E}_1)^{-1}$  alsó háromszögmátrixok szorzatának inverze, tehát maga is alsó háromszögmátrix. Ráadásul mindegyik mátrixban, így szorzatukban, és annak inverzében is a főátló csupa 1-esből áll. Ez a következő definícióhoz vezet:

**5.39. DEFINÍCIÓ (LU-FELBONTÁS).** Azt mondjuk, hogy az  $m \times n$ -es  $\mathbf{A}$  mátrix egy  $\mathbf{A} = \mathbf{LU}$  alakú tényezőkre bontása LU-felbontás (LU-faktorizáció vagy LU-dekompozíció), ha  $\mathbf{L}$  alsó egység háromszögmátrix (tehát a főátlóban 1-ek, fölötte 0-k vannak),  $\mathbf{U}$  pedig felső háromszögmátrix.

Az LU-felbontásban az L és U betűk az alsó és felső jelentésű angol lower és upper szavak kezdőbetűi.

► Nincs minden mátrixnak LU-felbontása (ld. ?? feladat), például az

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix} \begin{bmatrix} b & c \\ 0 & d \end{bmatrix}$$

egyenlőség a paraméterek semmilyen értékére sem áll fönn.

► Az LU-felbontás nem egyértelmű, például

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & a & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

felbontás minden  $a$  paraméterértékre fönnáll. Megmutatható viszont, hogy ha  $\mathbf{A}$  invertálható, és létezik LU-felbontása, akkor az egyértelmű (ld. 5.42. tétel).

5.40. PÉLDA (AZ LU-FELBONTÁS KISZÁMÍTÁSA). *Elemi sorműveletekkel hozzuk felső háromszögalakra az*

$$\mathbf{A} = \begin{bmatrix} 4 & 8 & 4 & 8 \\ 2 & 6 & 4 & 4 \\ 1 & 3 & 2 & 4 \end{bmatrix} \text{ és a } \mathbf{B} = \begin{bmatrix} 4 & 8 & 8 \\ 2 & 6 & 4 \\ 1 & 3 & 4 \end{bmatrix}$$

*mátrixot, majd e lépéseket fölhasználva írjuk föl mindkét mátrix egy-egy LU-felbontását!*

**MEGOLDÁS.** Először nézzük az  $\mathbf{A}$  mátrixot! Oszloponként haladva végezzük el a Gauss-kiküszöbölést. Minden elemi sorművelet mellett (zárójelben) megadjuk a hozzá tartozó elemi mátrixot:

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} 4 & 8 & 4 & 8 \\ 2 & 6 & 4 & 4 \\ 1 & 3 & 2 & 4 \end{bmatrix} \xrightarrow{S_2 - 1/2 S_1} \left( \mathbf{E}_1 = \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \\ \mathbf{E}_1 \mathbf{A} &= \begin{bmatrix} 4 & 8 & 4 & 8 \\ 0 & 2 & 2 & 0 \\ 1 & 3 & 2 & 4 \end{bmatrix} \xrightarrow{S_3 - 1/4 S_1} \left( \mathbf{E}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1/4 & 0 & 1 \end{bmatrix} \right) \\ \mathbf{E}_2 \mathbf{E}_1 \mathbf{A} &= \begin{bmatrix} 4 & 8 & 4 & 8 \\ 0 & 2 & 2 & 0 \\ 0 & 1 & 1 & 2 \end{bmatrix} \xrightarrow{S_3 - 1/2 S_2} \left( \mathbf{E}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1/2 & 1 \end{bmatrix} \right) \\ \mathbf{E}_3 \mathbf{E}_2 \mathbf{E}_1 \mathbf{A} &= \begin{bmatrix} 4 & 8 & 4 & 8 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} = \mathbf{U}. \end{aligned}$$

Tehát  $\mathbf{E}_3 \mathbf{E}_2 \mathbf{E}_1 \mathbf{A} = \mathbf{U}$ , amiből az  $(\mathbf{E}_3 \mathbf{E}_2 \mathbf{E}_1)^{-1} = \mathbf{E}_1^{-1} \mathbf{E}_2^{-1} \mathbf{E}_3^{-1}$  mátrixszal való beszorzás után  $\mathbf{A} = (\mathbf{E}_1^{-1} \mathbf{E}_2^{-1} \mathbf{E}_3^{-1}) \mathbf{U}$ . Kiszámoljuk az elemi mátrixok inverzeinek szorzatát, azaz az  $\mathbf{L} = \mathbf{E}_1^{-1} \mathbf{E}_2^{-1} \mathbf{E}_3^{-1}$  mátrixot. Fölhasználjuk a 195. oldalon mondottakat, miszerint az  $S_i + c S_j$  sorművelet mátrixának inverze egyenlő az  $S_i - c S_j$  mátrixával:

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1/2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/4 & 1/2 & 1 \end{bmatrix}.$$

Meglepő (de általánosítható) módon ezeknek az elemi mátrixoknak a szorzata a főátló alatti számok átmásolásával megkapható. Az eredmény egy alsó egység háromszögmátrix. Így az  $\mathbf{A}$  mátrix LU-felbontása:

$$\mathbf{A} = \begin{bmatrix} 4 & 8 & 4 & 8 \\ 2 & 6 & 4 & 4 \\ 1 & 3 & 2 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/4 & 1/2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 8 & 4 & 8 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \quad (5.4)$$

Mivel az  $\mathbf{A}$  átalakítása közben az oszlopok közt nem végeztünk műveletet, és a  $\mathbf{B}$  mátrix az  $\mathbf{A}$ -ból a harmadik oszlopa elhagyásával kapható

OCTAVE: A

A =

```
4 8 4 8
2 6 4 4
1 3 2 4
```

OCTAVE: [L U]=lu(A)

L =

```
1.00 0.00 0.00
0.50 1.00 0.00
0.25 0.50 1.00
```

U =

```
4 8 4 8
0 2 2 0
0 0 0 2
```

OCTAVE: B

B =

```
4 8 8
2 6 4
1 3 4
```

OCTAVE: [L U]=lu(B)

L =

```
1.00 0.00 0.00
0.50 1.00 0.00
0.25 0.50 1.00
```

U =

```
4 8 8
0 2 0
0 0 2
```

5.3. kód: Egy mátrix LU-felbontásának kiszámítása mátrixalapú nyelvben



meg, ezért az előző felbontásból azonnal adódik a **B** felbontása is:

$$\mathbf{B} = \begin{bmatrix} 4 & 8 & 8 \\ 2 & 6 & 4 \\ 1 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/4 & 1/2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 8 & 8 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad (5.5) \quad \square$$

Az 5.40. példában követett eljárás egyszerűen általánosítható tetszőleges mátrixra.

Az alábbi algoritmus a Gauss-elimináció lépcsős alak helyett felső háromszögalakú mátrixot adó megváltoztatásával vagy talál egy  $m \times m$ -es **L** és egy  $m \times n$ -es **U** mátrixot, melyekre  $\mathbf{A} = \mathbf{LU}$ , vagy hibaüzenetet ad.

**5.41. ALGORITMUS (EGY LU-FELBONTÁS ELŐÁLLÍTÁSA).** Legyen **A** egy tetszőleges  $m \times n$ -es valós (vagy bármely más test feletti) mátrix.

Első lépésként tekintsük az **A** mátrix első sorának első elemét. Ha ez 0, de az első oszlopban alatta nemnulla elem is van, akkor „a mátrixnak nincs LU-felbontása” üzenettel az algoritmus leáll. Ha alatta minden elem 0, az algoritmust a második sor második elemével folytatjuk (Gauss-kiküszöbölés esetén az első sor második elemével folytatnánk). Végül, ha az első sor első eleme nem 0, akkor az első oszlop további elemei eliminálhatók az  $S_2 - l_{21}S_1$ ,  $S_3 - l_{31}S_1, \dots, S_n - l_{n1}S_1$  sorműveletekkel, ahol  $l_{k1} = a_{k1}/a_{11}$ .

Az algoritmust hasonlóan folytatjuk sorban haladva a főátló elemein. Ha valamelyikük 0, de alatta van a mátrixnak nemnulla eleme, leállunk, ha alatta már minden elem 0, folytatjuk a következő főátlóbeli elemmel, ha pedig nemnulla, akkor elimináljuk az alatta lévő elemeket. Az  $i$ -edik lépésben tehát az  $S_{i+1} - l_{i+1,i}S_i$ ,  $S_{i+2} - l_{i+2,i}S_i, \dots, S_n - l_{ni}S_i$  sorműveleteket hajtjuk végre.

Az elimináció végén megmaradt felső háromszögmátrix lesz **U**. A kiküszöbölés konstans  $l_{ij}$  elemeit írjuk az **I** <sub>$m$</sub>  egységmátrix  $i$ -edik sorának  $j$ -edik oszlopába. Ez lesz az **L** mátrix, azaz

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ l_{21} & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{m1} & l_{m2} & \dots & 1 \end{bmatrix}. \quad (5.6)$$

**5.42. TÉTEL (AZ LU-FELBONTÁS LÉTEZÉSE ÉS EGYÉRTELMŰSÉGE).** A fenti algoritmusra igaz, hogy

- a) pontosan akkor áll le hibaüzenettel, ha **A**-nak nincs LU-felbontása,
- b) a megkonstruált **L** és **U** mátrixok LU-felbontást adnak,
- c) ha **A** invertálható, akkor e felbontás egyértelmű.

**BIZONYÍTÁS.** Csak a b) állítást igazoljuk, a többit az Olvasóra hagyjuk (ld. 5.45. feladat). Jelölje az  $S_j - l_{ji}S_i$  sorművelet elemi mátrixát **L** <sub>$ji$</sub>  ( $1 \leq i < j \leq m$ ). Jelölje e mátrixoknak a végrehajtás sorrendjében jobbról balra vett szorzatát **E**, azaz legyen

$$\mathbf{E} = (\mathbf{L}_{m-1,m})(\mathbf{L}_{m-2,m}\mathbf{L}_{m-2,m-1}) \dots (\mathbf{L}_{m2} \dots \mathbf{L}_{42}\mathbf{L}_{32})(\mathbf{L}_{m1} \dots \mathbf{L}_{31}\mathbf{L}_{21}).$$

Az algoritmus szerint ekkor  $\mathbf{EA} = \mathbf{U}$ . Vizsgáljuk meg az  $\mathbf{EL}$  szorzatot az algoritmusbeli  $\mathbf{L}$  mátrixszal. Mivel e mátrix főátlójában csupa 1,  $ji$ -edik helyén  $l_{ji}$  áll, ezért az elemi  $\mathbf{L}_{ji}$  mátrix épp ezt az elemet fogja eliminálni, és így  $\mathbf{E}$  minden főátló alatti elemet eliminál, azaz  $\mathbf{EL} = \mathbf{I}$ . Eszerint  $\mathbf{E}^{-1} = \mathbf{L}$ , tehát  $\mathbf{A} = \mathbf{E}^{-1}\mathbf{U} = \mathbf{LU}$ .  $\square$

**Egyenletrendszer megoldása LU-felbontással** Ha már ismerjük egy  $\mathbf{A}$  mátrix LU-felbontását, akkor az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszer könnyen megoldható. Az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszer megoldása az  $\mathbf{Ly} = \mathbf{b}$ ,  $\mathbf{Ux} = \mathbf{y}$  egyenletrendszerek megoldásával ekvivalens. Ha ugyanis  $\mathbf{x}$  megoldása az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszernek, akkor  $\mathbf{LUx} = \mathbf{b}$ , és az  $\mathbf{y} = \mathbf{Ux}$  jelöléssel  $\mathbf{Ly} = \mathbf{b}$ . Másrészt, ha  $\mathbf{y}$  megoldása az  $\mathbf{Ly} = \mathbf{b}$  egyenletrendszernek, és  $\mathbf{x}$  az  $\mathbf{Ux} = \mathbf{y}$  egyenletrendszernek, akkor  $\mathbf{y}$ -t behelyettesítve  $\mathbf{L}(\mathbf{Ux}) = \mathbf{b}$ , azaz  $\mathbf{Ax} = \mathbf{b}$ . Tömören:

$$\mathbf{Ax} = \mathbf{b} \text{ megoldható} \iff \mathbf{Ly} = \mathbf{b}, \mathbf{Ux} = \mathbf{y} \text{ megoldható.}$$

Az  $\mathbf{L}$  és  $\mathbf{U}$  alakjából következik, hogy az  $\mathbf{Ly} = \mathbf{b}$ , és az  $\mathbf{Ux} = \mathbf{y}$  egyenletrendszerek egyszerű visszahelyettesítésekkel megoldhatók.

**5.43. PÉLDA (EGYENLETRENDSZER MEGOLDÁSA LU-FELBONTÁSSAL).**  
Oldjuk meg a következő egyenletrendszert!

$$4x_1 + 8x_2 + 8x_3 = 8$$

$$2x_1 + 6x_2 + 4x_3 = 4$$

$$x_1 + 3x_2 + 4x_3 = 4$$

**MEGOLDÁS.** Mivel ismerjük az együtthatómátrix LU-felbontását – az épp az (5.5)-beli felbontás –, ezért ezt használjuk, és először megoldjuk az  $\mathbf{Ly} = \mathbf{b}$  egyenletrendszert:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/4 & 1/2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \\ 4 \end{bmatrix}$$

Ebből  $y_1 = 8$ , ezt a második egyenletbe helyettesítve kapjuk, hogy  $y_2 = 0$ , majd ezeket a harmadikba helyettesítve kapjuk, hogy  $y_3 = 2$ . Ezután megoldjuk az  $\mathbf{Ux} = \mathbf{y}$  egyenletrendszert, aminek alakja

$$\begin{bmatrix} 4 & 8 & 8 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 8 \\ 0 \\ 2 \end{bmatrix}.$$

Ismét egyszerű visszahelyettesítésekkel kapjuk, hogy  $x_3 = 1$ ,  $x_2 = 0$  és  $x_1 = 1$ . A megoldás  $\mathbf{x} = (0, 0, 1)$ .  $\square$

**Mátrix invertálása LU-felbontással** Mátrix invertálásához elég megoldanunk az  $\mathbf{AX} = \mathbf{I}$  egyenletrendszer. Ha  $\mathbf{A} = \mathbf{LU}$  egy LU-felbontása  $\mathbf{A}$ -nak, akkor az  $\mathbf{LUX} = \mathbf{I}$  megoldása a vele ekvivalens két mátrixegyenlet megoldásával megkapható:

$$\mathbf{AX} = \mathbf{I} \iff \mathbf{LY} = \mathbf{I}, \mathbf{UX} = \mathbf{Y}.$$

E két utóbbi egyenletrendszer viszont megoldható kizárólag visszahelyettesítésekkel is!

5.44. PÉLDA (MÁTRIX INVERTÁLÁSA LU-FELBONTÁSSAL). *Invertáljuk az 5.40. példában megadott*

$$\mathbf{B} = \begin{bmatrix} 4 & 8 & 8 \\ 2 & 6 & 4 \\ 1 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/4 & 1/2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 8 & 8 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

*mátrixot az LU-felbontása segítségével!*

**MEGOLDÁS.** A  $\mathbf{B}$  mátrix LU-felbontását használva először megoldjuk az  $\mathbf{LY} = \mathbf{I}$  mátrixegyenletet:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/4 & 1/2 & 1 \end{bmatrix} \begin{bmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Az  $\mathbf{L}$  első sorával való szorzásból:  $[y_{11} \ y_{12} \ y_{13}] = [1 \ 0 \ 0]$ . A második sorral való szorzásból  $\frac{1}{2}[y_{11} \ y_{12} \ y_{13}] + [y_{21} \ y_{22} \ y_{23}] = [0 \ 1 \ 0]$ . Behelyettesítés után  $[y_{21} \ y_{22} \ y_{23}] = [-\frac{1}{2} \ 1 \ 0]$ . Végül a harmadik sorral való szorzásból:

$$\frac{1}{4}[y_{11} \ y_{12} \ y_{13}] + \frac{1}{2}[y_{21} \ y_{22} \ y_{23}] + [y_{31} \ y_{32} \ y_{33}] = [0 \ 0 \ 1],$$

amiből behelyettesítés után kifejezve  $\mathbf{Y}$  harmadik sorát kapjuk, így  $[y_{31} \ y_{32} \ y_{33}] = [0 \ -\frac{1}{2} \ 1]$ . Azaz

$$\mathbf{Y} = \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & -1/2 & 1 \end{bmatrix}.$$

Ezután ugyanígy, egyszerű helyettesítésekkel megoldható az  $\mathbf{UX} = \mathbf{Y}$ , azaz a

$$\begin{bmatrix} 4 & 8 & 8 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & -1/2 & 1 \end{bmatrix}$$

mátrixegyenlet is, melynek megoldása

$$\mathbf{X} = \begin{bmatrix} 3/4 & -1/2 & -1 \\ -1/4 & 1/2 & 0 \\ 0 & -1/4 & 1/2 \end{bmatrix}.$$

□

*Az LU-felbontás a gyakorlatban* Ismét végigszámoljuk az 5.40. példabeli mátrix felbontását. Először írjunk le egy egységmátrixot, de a főátló alatti helyeket üresen hagyva, ebből lesz  $L$ . Írjuk mellé az  $A$  mátrixot, és amikor elvégzünk egy  $S_i - l_{ji}S_j$  sorműveletet rajta, akkor az  $l_{ji}$  értéket bejegyezzük az  $L$  mátrix  $j$ -edik sorának  $i$ -edik oszlopába. Az alábbi számítások bal hasábjában látjuk a fentiek szerinti lépéseket.

$$\begin{array}{ccc}
 \begin{bmatrix} 1 & 0 & 0 \\ & 1 & 0 \\ & & 1 \end{bmatrix} & \begin{bmatrix} 4 & 1 & 2 \\ 2 & 4 & 1 \\ 1 & 2 & 4 \end{bmatrix} & \begin{bmatrix} 4.00 & 1.00 & 2.00 \\ 2.00 & 4.00 & 1.00 \\ 1.00 & 2.00 & 4.00 \end{bmatrix} \\
 \Downarrow & & \Downarrow \\
 \begin{bmatrix} 1 & 0 & 0 \\ \textcolor{blue}{1/2} & 1 & 0 \\ & & 1 \end{bmatrix} & \begin{bmatrix} 4 & 1 & 2 \\ 0 & 7/2 & 0 \\ 1 & 2 & 4 \end{bmatrix} & \begin{bmatrix} 4.00 & 1.00 & 2.00 \\ \textcolor{blue}{0.50} & 3.50 & 0.00 \\ 1.00 & 2.00 & 4.00 \end{bmatrix} \\
 \Downarrow & & \Downarrow \\
 \begin{bmatrix} 1 & 0 & 0 \\ \textcolor{blue}{1/2} & 1 & 0 \\ \textcolor{blue}{1/4} & & 1 \end{bmatrix} & \begin{bmatrix} 4 & 1 & 2 \\ 0 & 7/2 & 0 \\ 0 & 7/4 & 7/2 \end{bmatrix} & \begin{bmatrix} 4.00 & 1.00 & 2.00 \\ \textcolor{blue}{0.50} & 3.50 & 0.00 \\ \textcolor{blue}{0.25} & 1.75 & 3.50 \end{bmatrix} \\
 \Downarrow & & \Downarrow \\
 \begin{bmatrix} 1 & 0 & 0 \\ \textcolor{blue}{1/2} & 1 & 0 \\ \textcolor{blue}{1/4} & \textcolor{blue}{1/2} & 1 \end{bmatrix} & \begin{bmatrix} 4 & 1 & 2 \\ 0 & 7/2 & 0 \\ 0 & 0 & 7/2 \end{bmatrix} & \begin{bmatrix} 4.00 & 1.00 & 2.00 \\ \textcolor{blue}{0.50} & 3.50 & 0.00 \\ \textcolor{blue}{0.25} & \textcolor{blue}{0.50} & 3.50 \end{bmatrix}
 \end{array}$$

Vegyük észre, hogy az  $A$  mátrixon folytatott elemi átalakítások eredménye és az  $L$  már kiszámolt elemei egyetlen mátrixban is „elférnek”, ugyanis  $L$ -ben épp akkor és oda kerül egy elem, amikor és ahova  $A$ -ban 0. Ezt a számítógépprogramok kihasználják, ha igen nagy méretű  $A$  mátrixot kell felbontani, és az  $L$  és  $U$  mátrixot is az  $A$  helyében konstruálják meg. A fenti számítások jobb hasábjában ezt a számítógépes technikát alkalmazzuk. Színes háttérrel jelöljük az  $L$ -beli elemeket.

Az LU-felbontás műveletigénye megegyezik a Gauss-kiküszöbölésével, azaz egy  $n$ -edrendű mátrixra nagyságrendileg  $2n^3/3$ . Egyenletrendszer megoldásánál is azonos a lépésszám, hisz a Gauss-módszernél a kiküszöbölést a jobb oldallal is meg kell csinálni, az LU-felbontásnál viszont az alsó háromszögmátrixhoz tartozó egyenletrendszert is meg kell oldani: mindkettő  $n(n-1)/2$  összeadás/kivonás és ugyanennyi szorzás/osztás. Az LU-felbontásnak viszont több olyan előnyös tulajdonsága van, ami miatt használata meghatározó az egyenletrendszerek megoldásában és amellet több más feladatban is. Néhány a legfontosabbak közül:

1. Mivel az egyenletrendszer együtthatómátrixának LU-felbontásához nincs szükség az egyenletrendszer jobb oldalára, ezért használható olyan esetekben, amikor a jobb oldal még nem ismeretes, vagy több

különböző jobb oldallal is dolgozni kell.

2. Az LU-felbontás ismeretében több mátrixokkal kapcsolatos számítás gyorsabban elvégezhető mint egyébként, pl. ilyen a mátrix inverzének, vagy a később tanulandó determinánsának meghatározása.

3. Korábban említettük, hogy az LU-felbontás igen memóriatakarékos, ráadásul vannak olyan speciális mátrixosztályok (pl. a szalagmátrixok, vagy a ritka mátrixok), melyekre létezik a kiküszöbölésnél gyorsabb algoritmus az LU-felbontásra.

4. A komputer algebra programok úgy működnek, hogy ha egy mátrixon valamilyen számítást kell elvégezni, ami megoldható az LU-felbontással (vagy a következő pontban tárgyalandó PLU-felbontással), akkor azzal oldják meg. Így ha később egy másik számítást is el kell e mátrixszal végezni, e felbontás ismeretében az már sokkal gyorsabb lehet.

**PLU-felbontás\*** Nincs minden  $A$  mátrixnak LU-felbontása, de sorcserékkel – azaz egy permutáló mátrixszal való balról szorzással – olyan alakra hozható, melynek van LU-felbontása. Létezik tehát olyan  $P$  permutáló mátrix, hogy

$$PA = LU, \text{ azaz } A = P^T LU.$$

(Itt kihasználtuk, hogy permutáló mátrix inverze megegyezik transzponáltjával.)

**5.45. DEFINÍCIÓ (PLU-FELBONTÁS).** Egy tetszőleges  $m \times n$ -es  $A$  mátrixnak egy permutáló, egy egység főátlójú négyzetes alsó háromszög és egy  $m \times n$ -es felső háromszögmátrix szorzatára való bontását PLU-felbontásnak nevezzük.

► Szokás – kissé bonyolultabb, de kisebb méretű mátrixokat tartalmazó módon – úgy is definiálni a PLU-felbontást, hogy ha  $r = \min(m, n)$ , akkor  $P$   $m \times m$ -es permutáló mátrix,  $L$  1-esekből álló főátlójú  $m \times r$ -es alsó, míg az  $U$   $r \times n$ -es felső háromszögmátrix. E definíció az  $m > n$  esetben ad más felbontást az általunk adottnál. Például a következő első felbontás a definíciót, a második e megjegyzés szerinti felbontást adja:

$$\begin{aligned} \begin{bmatrix} 0 & 1 \\ 1 & 2 \\ 2 & 3 \end{bmatrix} &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/2 & 1/2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix} \end{aligned}$$

A PLU-felbontást megadó algoritmus minimális változtatással megkapható az LU-felbontáséból. Az algoritmust – épp ahogy a számítógépek is számolnak – az  $L$  és  $U$  mátrix elemeit egyetlen mátrixban tárolva fogjuk végrehajtani. Az LU-hoz képest csak annyi a változás, hogy sorcserék elvégzését is megengedjük. Az LU-felbontás algoritmus akkor akad el, amikor egy főátlóbeli elem 0, de van alatta nem nulla elem az oszlopban. Most ilyen esetben e két sort kicseréljük. Sőt, olyankor is kicserélhetünk egy sort egy alatta lévővel, ha főátlóbeli elem nem 0. Egy ilyen cserét a kerekítési hibák csökkentése érdekében lehet érdemes megtenni. A korábban már említett részleges főelemkiválasztás szabálya szerint mindig a legnagyobb abszolút értékű elemet érdemes főelemnek választani. Hogy az algoritmus végén tudjuk, hogyan változott a vektorok sorrendje, a sorindexek változását folyamatosan följegyezzük – praktikusán a mátrix sorvektorai elé írva. Lássunk egy példát. Azt, hogy az  $L$  és  $U$  mátrixok összeolvasztásából kapott mátrixon is elvégezhetők a sorcserék, később igazoljuk!

5.46. PÉLDA (PLU-FELBONTÁS). *Határozzuk meg az*

$$A = \begin{bmatrix} -1 & 6 & 1 & -7 & 4 \\ 1 & 4 & 4 & -7 & 5 \\ 4 & -8 & 4 & 8 & -4 \\ 3 & -6 & 8 & 6 & -8 \end{bmatrix} \quad (5.7)$$

*mátrix PLU-felbontását úgy, hogy minden lépésben részleges főelemkiválasztással a főátlóbeli elem alatti legnagyobb abszolút értékű elemet választjuk ki.*

**MEGOLDÁS.** A mátrix sorindexeit a következőképp fogjuk jelölni:

$$\begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} \begin{bmatrix} -1 & 6 & 1 & -7 & 4 \\ 1 & 4 & 4 & -7 & 5 \\ 4 & -8 & 4 & 8 & -4 \\ 3 & -6 & 8 & 6 & -8 \end{bmatrix} \quad (5.8)$$

Mivel az első oszlopban 4 a legnagyobb abszolút értékű szám, végrehajtunk egy  $S_{i \leftrightarrow j}$  sorcserét, majd elimináljuk az első oszlop összes többi elemét:

$$\begin{matrix} 3 \\ 2 \\ 1 \\ 4 \end{matrix} \begin{bmatrix} 4 & -8 & 4 & 8 & -4 \\ 1 & 4 & 4 & -7 & 5 \\ -1 & 6 & 1 & -7 & 4 \\ 3 & -6 & 8 & 6 & -8 \end{bmatrix} \rightarrow \begin{matrix} 3 \\ 2 \\ 1 \\ 4 \end{matrix} \begin{bmatrix} 4 & -8 & 4 & 8 & -4 \\ 1/4 & 6 & 3 & -9 & 6 \\ -1/4 & 4 & 2 & -5 & 3 \\ 3/4 & 0 & 5 & 0 & -5 \end{bmatrix}$$

A második oszlop második eleme alatt nincs nagyobb abszolút értékű szám, most nem kell sort cserélni, a negyedik sort eliminálni sem kell,

```
OCTAVE: A = [
> -1 6 1 -7 4
> 1 4 4 -7 5
> 4 -8 4 8 -4
> 3 -6 8 6 -8]
A =
```

```
-1 6 1 -7 4
1 4 4 -7 5
4 -8 4 8 -4
3 -6 8 6 -8
```

```
OCTAVE: [L U P] = lu(a)
L =
```

```
1.00 0.00 0.00 0.00
0.25 1.00 0.00 0.00
0.75 0.00 1.00 0.00
-0.25 0.67 0.00 1.00
```

```
U =
```

```
4 -8 4 8 -4
0 6 3 -9 6
0 0 5 0 -5
0 0 0 1 -1
```

```
P =
```

```
Permutation Matrix
```

```
0 0 1 0
0 1 0 0
0 0 0 1
1 0 0 0
```

```
OCTAVE: transpose(P)*L*U
ans =
```

```
-1 6 1 -7 4
1 4 4 -7 5
4 -8 4 8 -4
3 -6 8 6 -8
```

5.4. kód: Egy mátrix PLU-felbontásának kiszámítása mátrixalapú nyelvben

(azaz kivonhatjuk belőle a második sor 0-szorosát):

$$\rightarrow \begin{matrix} 3 \\ 2 \\ 1 \\ 4 \end{matrix} \begin{bmatrix} 4 & -8 & 4 & 8 & -4 \\ 1/4 & & 6 & 3 & -9 & 6 \\ -1/4 & 2/3 & 0 & 1 & -1 \\ 3/4 & 0 & 5 & 0 & -5 \end{bmatrix}$$

A harmadik oszlopban 0 áll a főátlón, kicseréljük a harmadik és negyedik sort. Az eddig még nem indokolt mozzanat: az **L** és az **U** mátrixba első részén egyaránt végrehajtható e művelet, és épp ezt tesszük a sor-indexekkel is:

$$\rightarrow \begin{matrix} 3 \\ 2 \\ 4 \\ 1 \end{matrix} \begin{bmatrix} 4 & -8 & 4 & 8 & -4 \\ 1/4 & & 6 & 3 & -9 & 6 \\ 3/4 & 0 & 5 & 0 & -5 \\ -1/4 & 2/3 & 0 & 1 & -1 \end{bmatrix} \rightarrow \begin{matrix} 3 \\ 2 \\ 4 \\ 1 \end{matrix} \begin{bmatrix} 4 & -8 & 4 & 8 & -4 \\ 1/4 & & 6 & 3 & -9 & 6 \\ 3/4 & 0 & 5 & 0 & -5 \\ -1/4 & 2/3 & 0 & 1 & -1 \end{bmatrix}$$

Az utolsó lépésben nem is volt tennivalónk, mivel a főátló alatt 0 volt, így csak jeleztük, hogy mi kerül az **L** mátrixba e helyen. (Az a két nulla nem ugyanaz a nulla! A második már az **L** eleme!) Végül ebből az alakból leolvasható az **L**, **U** és az indexekből a **P** mátrix. Ezeket egyből a **PA = LU** egyenlőségben adjuk meg:

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -1 & 6 & 1 & -7 & 4 \\ 1 & 4 & 4 & -7 & 5 \\ 4 & -8 & 4 & 8 & -4 \\ 3 & -6 & 8 & 6 & -8 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/4 & 1 & 0 & 0 \\ 3/4 & 0 & 1 & 0 \\ -1/4 & 2/3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & -8 & 4 & 8 & -4 \\ 0 & 6 & 3 & -9 & 6 \\ 0 & 0 & 5 & 0 & -5 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}.$$

Mindkét oldalt **P<sup>T</sup>**-tal szorozva megkapjuk a PLU-felbontást, azaz az **A = P<sup>T</sup>LU** egyenlőséget:

$$\begin{bmatrix} -1 & 6 & 1 & -7 & 4 \\ 1 & 4 & 4 & -7 & 5 \\ 4 & -8 & 4 & 8 & -4 \\ 3 & -6 & 8 & 6 & -8 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/4 & 1 & 0 & 0 \\ 3/4 & 0 & 1 & 0 \\ -1/4 & 2/3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & -8 & 4 & 8 & -4 \\ 0 & 6 & 3 & -9 & 6 \\ 0 & 0 & 5 & 0 & -5 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}.$$

Ezzel megoldottuk a feladatot! □

E felbontás egy lehetséges használatára mutatunk példát.

5.47. PÉLDA. Oldjuk meg az

$$\begin{bmatrix} -1 & 6 & 1 & -7 & 4 \\ 1 & 4 & 4 & -7 & 5 \\ 4 & -8 & 4 & 8 & -4 \\ 3 & -6 & 8 & 6 & -8 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 \\ 4 \\ 4 \\ 8 \end{bmatrix}$$

egyenletrendszert az együtthatómátrix PLU-felbontását használva!

**MEGOLDÁS.** Az **Ax = b** egyenletet **P**-vel szorozva és **PA** helyébe **LU**-t írva kapjuk, hogy az **LUx = Pb** egyenletrendszert kell megoldani,

ahol  $\mathbf{b} = (1, 4, 4, 8)$ . Ez – hasonlóan az LU-felbontásnál tanultakhoz – az  $\mathbf{Ly} = \mathbf{Pb}$  és az  $\mathbf{Ux} = \mathbf{y}$  egyenletrendszerek megoldásával ekvivalens.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/4 & 1 & 0 & 0 \\ 3/4 & 0 & 1 & 0 \\ -1/4 & 2/3 & 0 & 1 \end{bmatrix} \mathbf{y} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 4 \\ 8 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 8 \\ 1 \end{bmatrix}$$

Ennek az egyenletrendszernek a megoldása fejben számolva is leolvasható:  $\mathbf{y} = (4, 3, 5, 0)$ . Az  $\mathbf{Ux} = \mathbf{y}$  egyenletrendszer bővített mátrixa, és annak redukált lépcsős alakja:

$$\begin{bmatrix} 4 & -8 & 4 & 8 & -4 & 4 \\ 0 & 6 & 3 & -9 & 6 & 3 \\ 0 & 0 & 5 & 0 & -5 & 5 \\ 0 & 0 & 0 & 1 & -1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 1 & -1 & 0 \end{bmatrix}$$

Innen a megoldás  $\mathbf{x} = (-2t, 0, 1 + t, t, t)$ . □

5.48. ALGORITMUS (EGY PLU-FELBONTÁS ELŐÁLLÍTÁSA). Legyen  $\mathbf{A}$  egy (test fölött értelmezett) tetszőleges  $m \times n$ -es mátrix, legyen  $r = \min(m, n)$ , és képezzük az  $\mathbf{A}_k, \mathbf{P}_k, \mathbf{L}_k, \mathbf{U}_k$  ( $k = 0, 1, \dots, r$ ) mátrixok sorozatát a következő eljárás szerint:

- a)  $\mathbf{A}_0 = \mathbf{U}_0 = \mathbf{A}$ ,  $\mathbf{L}_0 = \mathbf{I}$ , így  $\mathbf{A}_0 = \mathbf{L}_0 \mathbf{U}_0$ ,
- b) a  $k$ -edik lépésben az  $\mathbf{A}_{k-1} = \mathbf{L}_{k-1} \mathbf{U}_{k-1}$  összefüggés mátrixaiból megkonstruáljuk az  $\mathbf{A}_k = \mathbf{L}_k \mathbf{U}_k$  egyenlőségbeli mátrixokat:
  1. ha  $\mathbf{U}_{k-1}$  főátlóján a  $k$ -edik elem és alatta minden elem 0, akkor legyen  $\mathbf{A}_k = \mathbf{A}_{k-1}$ ,  $\mathbf{L}_k = \mathbf{L}_{k-1}$ ,  $\mathbf{U}_k = \mathbf{U}_{k-1}$ ,  $\mathbf{P}_k = \mathbf{I}$ , megnöveljük  $k$  értékét 1-gyel és visszatérünk e pontra, egyébként a következővel folytatjuk.
  2. ha  $\mathbf{U}_{k-1}$  főátlóján a  $k$ -edik elem 0, és van olyan  $i > k$ , hogy az  $i$ -edik sorban alatta nem nulla elem van, akkor az  $S_k \leftrightarrow S_i$  sorcserét végző  $\mathbf{P}_k$  elemi mátrixszal kicseréljük e két sort, kapjuk az  $\mathbf{U}'_{k-1} = \mathbf{P}_k \mathbf{U}_{k-1}$ . E sorcserét végrehajtjuk az  $\mathbf{A}_{k-1}$  mátrixon is, Ez lesz  $\mathbf{A}_k$ . Kihasználva, hogy  $\mathbf{P}_k \mathbf{P}_k = \mathbf{I}$ , az  $\mathbf{L}_{k-1}$ -en elvégzendő transzformáció is adódik:

$$\begin{aligned} \mathbf{A}_k &= \mathbf{P}_k \mathbf{A}_{k-1} = \mathbf{P}_k \mathbf{L}_{k-1} \mathbf{U}_{k-1} = \mathbf{P}_k \mathbf{L}_{k-1} \mathbf{I} \mathbf{U}_{k-1} \\ &= \mathbf{P}_k \mathbf{L}_{k-1} \mathbf{P}_k \mathbf{P}_k \mathbf{U}_{k-1} = (\mathbf{P}_k \mathbf{L}_{k-1} \mathbf{P}_k) (\mathbf{P}_k \mathbf{U}_{k-1}) \\ &= \mathbf{L}'_k \mathbf{U}'_k. \end{aligned}$$

Itt tehát  $\mathbf{L}'_k = \mathbf{P}_k \mathbf{L}_{k-1} \mathbf{P}_k$ , azaz az  $\mathbf{L}_{k-1}$  mátrix  $k$ -edik, és egy  $i \geq k$  indexre az  $i$ -edik sorának főátló alatti részei felcserélődnek. A  $\mathbf{P}_k \mathbf{L}_{k-1}$  szorzat megcseréli az  $\mathbf{L}_{k-1}$  mátrix  $k$ -edik és  $i$ -edik sorát, megcserélve a főátlóbeli 1-eseket is. A  $\mathbf{P}_k$ -val való jobbról szorzás eredményeként a  $k$ -edik és  $i$ -edik oszlopok helyet cserélnek, mivel azonban  $\mathbf{L}_{k-1}$ -ben ezekben az oszlopokban csak zérusok vannak a főátló elemein kívül, ezért e két 1-es visszakerül a főátlóra. Szemléltetésként legyen  $m = 5$ ,  $k = 3$ ,



$i = 5$ :

$$\mathbf{P}_k = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \quad \mathbf{L}_{k-1} = \left[ \begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ * & 1 & 0 & 0 & 0 \\ \hline k_1 & k_2 & 1 & 0 & 0 \\ * & * & 0 & 1 & 0 \\ i_1 & i_2 & 0 & 0 & 1 \end{array} \right]$$

$$\mathbf{P}_k \mathbf{L}_{k-1} = \left[ \begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ * & 1 & 0 & 0 & 0 \\ \hline i_1 & i_2 & 0 & 0 & 1 \\ * & * & 0 & 1 & 0 \\ k_1 & k_2 & 1 & 0 & 0 \end{array} \right] \quad \mathbf{P}_k \mathbf{L}_{k-1} \mathbf{P}_k = \left[ \begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ * & 1 & 0 & 0 & 0 \\ \hline i_1 & i_2 & 1 & 0 & 0 \\ * & * & 0 & 1 & 0 \\ k_1 & k_2 & 0 & 0 & 1 \end{array} \right]$$

3. Ezután az  $\mathbf{U}'_{k-1}$  mátrix  $k$ -adik oszlopának főátló alatti elemeit egymás után elimináljuk, mindegyiket egy  $S_i \leftarrow S_i - cS_k$  elemi sorműveletet végző  $\mathbf{E}$  elemi mátrixszal. Ennek inverze a vele való jobbról szorzás esetén az  $O_i \leftarrow O_i + cO_k$  oszlopműveletet végzi, így az

$$\mathbf{L}'_k \mathbf{U}'_k = \mathbf{L}'_k \mathbf{I} \mathbf{U}'_k = \mathbf{L}'_k \mathbf{E}^{-1} \mathbf{E} \mathbf{U}'_k$$

egyenlőségnek megfelelően az  $\mathbf{U}'_k$ -n végrehajtott sorművelet mellett az  $\mathbf{L}'_k$  mátrix  $i$ -edik oszlopának  $c$ -szeresét kell a  $k$ -adik oszlophoz adni, azaz a  $c$  számot beírni az  $i$ -edik sor  $k$ -adik oszlopába. Az eliminációt a kapott mátrixokon tovább folytatjuk, míg az összes elemet nem elimináltuk a  $k$ -adik oszlopban a főátló alatt.  $\mathbf{L}_k$  és  $\mathbf{U}_k$  jelöli e procedura végén kapott mátrixokat, tehát  $\mathbf{U}_k$ -ban már a főátló alatt minden elem 0, és  $\mathbf{L}_k$  a párja, melyekre  $\mathbf{A}_k = \mathbf{L}_k \mathbf{U}_k$ . Megnöveljük  $k$  értékét, és ha  $k < r$ , visszatérünk az 1. pontra, egyébként a következő pontra lépünk.

- c) Legyen  $\mathbf{P} = \mathbf{P}_r \mathbf{P}_{r-1} \dots \mathbf{P}_1$ ,  $\mathbf{L} = \mathbf{L}_r$ ,  $\mathbf{U} = \mathbf{U}_r$ . Ekkor a fentiek szerint  $\mathbf{PA} = \mathbf{LU}$  az  $\mathbf{A}$  egy PLU-felbontása.

Az, hogy ez az algoritmus valóban PLU-felbontást ad, az algoritmus leírásában bizonyítottuk, illetve onnan kiolvasható.

## Feladatok

Adjuk meg az alábbi mátrixok egy LU-felbontását!

$$5.32. \begin{bmatrix} 4 & 4 & 4 \\ 2 & 5 & 5 \\ 1 & 2 & 4 \end{bmatrix}$$

$$5.33. \begin{bmatrix} 4 & 8 & 4 \\ 2 & 7 & 8 \\ 1 & 3 & 4 \end{bmatrix}$$

$$5.34. \begin{bmatrix} 5 & -4 & -2 \\ 4 & -5 & -5 \\ -3 & 1 & -4 \end{bmatrix}$$

$$5.35. \begin{bmatrix} -3 & 1 & -3 & 0 \\ -2 & 4 & 3 & -4 \\ 1 & 3 & 3 & 0 \\ -3 & 0 & -3 & -1 \end{bmatrix}$$

$$5.36. \begin{bmatrix} -2 & -2 & 0 & 3 \\ 0 & 2 & -2 & -1 \\ -1 & 0 & 2 & -2 \\ 1 & 0 & 2 & 1 \end{bmatrix}$$

$$5.37. \begin{bmatrix} 2.0 & 2.0 & -2.0 \\ -0.5 & 0.0 & -1.0 \\ 1.0 & 1.5 & 1.0 \end{bmatrix}$$

Az előző feladatokban megkonstruált LU-felbontásokat használva oldjuk meg az alábbi egyenletrendszereket, azaz oldjuk meg előbb az  $\mathbf{Ly} = \mathbf{b}$ , majd az  $\mathbf{Ux} = \mathbf{y}$  egyenletrendszereket!

$$5.38. \begin{bmatrix} 4 & 4 & 4 \\ 2 & 5 & 5 \\ 1 & 2 & 4 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 0 \\ 3 \\ 5 \end{bmatrix}$$

$$5.39. \begin{bmatrix} 4 & 8 & 4 \\ 2 & 7 & 8 \\ 1 & 3 & 4 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix}$$

$$5.40. \begin{bmatrix} 5 & -4 & -2 \\ 4 & -5 & -5 \\ -3 & 1 & -4 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 3 \\ -1 \\ -7 \end{bmatrix}$$

$$5.41. \begin{bmatrix} 2.0 & 2.0 & -2.0 \\ -0.5 & 0.0 & -1.0 \\ 1.0 & 1.5 & 1.0 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 5.6 \\ -1.0 \\ 4.6 \end{bmatrix}$$

## 5.42. VÉGTELEN SOK MEGOLDÁS

$$4x_1 + 8x_2 + 4x_3 + 8x_4 = 8$$

$$2x_1 + 6x_2 + 4x_3 + 4x_4 = 4$$

$$x_1 + 3x_2 + 2x_3 + 4x_4 = 4$$

Határozzuk meg az alábbi mátrixok inverzét az LU-felbontásuk ismeretében, azaz oldjuk meg az  $\mathbf{LY} = \mathbf{I}$  és az  $\mathbf{UX} = \mathbf{Y}$  mátrix-egyenleteket!

$$5.43. \begin{bmatrix} 4 & 4 & 4 \\ 2 & 5 & 5 \\ 1 & 2 & 4 \end{bmatrix}$$

$$5.44. \begin{bmatrix} 4 & 8 & 4 \\ 2 & 7 & 8 \\ 1 & 3 & 4 \end{bmatrix}$$

5.45. Bizonyítsuk be, hogy ha  $\mathbf{A} = \mathbf{LU}$  az  $\mathbf{A}$  egy LU-felbontása, és  $\mathbf{A}$  invertálható, akkor e felbontás egyértelmű!

Adjuk meg az alábbi mátrixok egy PLU-felbontását! Alkalmazzunk részleges főelemkiválasztást!

$$5.46. \begin{bmatrix} 1 & 1 & 2 \\ 3 & 3 & 3 \\ 2 & 2 & 3 \end{bmatrix}$$

$$5.47. \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 2 & 3 & 3 & 4 \\ 3 & 4 & 6 & 7 & 9 \end{bmatrix}$$

$$5.48. \begin{bmatrix} 0.0 & -1.0 & 1.5 \\ 0.5 & -2.0 & 2.0 \\ 0.0 & 2.0 & 2.0 \end{bmatrix}$$

5.49. Igazoljuk, hogy az

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

mátrixnak nincs LU-felbontása, és annak a mátrixnak sincs, melyet az  $\mathbf{A}$  első két sorának felcserélésével kapunk.

## Megoldások

5.4. A bizonyítások közvetlenül következnek a valós számok közti műveletek tulajdonságaiból. Mintaként bebizonyítjuk az (a) állítást.

$$\begin{aligned}\mathbf{A} + \mathbf{B} &= [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}] \\ &\stackrel{*}{=} [b_{ij} + a_{ij}] = [b_{ij}] + [a_{ij}] = \mathbf{B} + \mathbf{A}.\end{aligned}$$

A  $*$ -gal jelzett egyenlőségnél használjuk a számok összeadásának kommutativitását. A többi állítás hasonlóan bizonyítható.

5.5. Helyettesítés előtt:

$$\mathbf{u}^T \mathbf{u} \begin{bmatrix} 2 & -1 & -1 \end{bmatrix} \mathbf{u}.$$

Az  $\mathbf{u} = \mathbf{A}\mathbf{x}$  helyettesítés elvégzése után

$$\mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} \begin{bmatrix} 2 & -1 & -1 \end{bmatrix} \mathbf{A} \mathbf{x},$$

ahol

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}, \mathbf{A} = \begin{bmatrix} 3 & 2 & 4 \\ 1 & -3 & 1 \\ 2 & -1 & -3 \end{bmatrix}.$$

$$5.6. \begin{bmatrix} 1 & 3 \\ 2 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$$

$$5.7. \begin{bmatrix} 1 & 2 \\ -2 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

$$5.8. \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

$$5.9. \begin{bmatrix} 2 & 4 \\ 3 & 8 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

$$5.10. \begin{bmatrix} 2 & 0 & 4 \\ 0 & 2 & 0 \\ 3 & 2 & 7 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$5.11. \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 2 \\ 2 & 4 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

5.12. Legyen  $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ . Négyzete a zérusmátrix, azaz

$$\mathbf{A}^2 = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a^2 + bc & b(a+d) \\ c(a+d) & bc + d^2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Innen vagy  $a = d = 0$  és  $c$  vagy  $d$  legalább egyike 0, vagy  $a \neq 0, c \neq 0$  és  $b = -a^2/c, d = -a$ .

5.13. A feladat érdekes, abban a Fibonacci sorozat elemei bukkannak föl. Ez az  $f_0 = 0, f_1 = 1, f_{k+1} = f_k + f_{k-1}$  egyenlőségekkel definiált sorozat, melynek első néhány tagja: 0, 1, 1, 2, 3, 5, 8, 13, 21, ... Tekintsük  $\mathbf{B}$  néhány hatványát:

$$\begin{aligned}\mathbf{A}^2 &= \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} f_1 & f_2 \\ f_2 & f_3 \end{bmatrix}, \\ \mathbf{A}^3 &= \mathbf{A}^2 \mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} f_2 & f_3 \\ f_3 & f_4 \end{bmatrix}.\end{aligned}$$

Ennek alapján azt sejtjük, hogy

$$\mathbf{A}^n = \begin{bmatrix} f_{n-1} & f_n \\ f_n & f_{n+1} \end{bmatrix}.$$

Az állítás  $n = 1, 2, 3$  esetén igaz, és  $n$ -ről öröklődik  $n+1$ -re, ugyanis

$$\begin{aligned}\mathbf{A}^{n+1} &= \mathbf{A}^n \mathbf{A} = \begin{bmatrix} f_{n-1} & f_n \\ f_n & f_{n+1} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} f_n & f_{n+1} \\ f_{n-1} + f_n & f_n + f_{n+1} \end{bmatrix} = \begin{bmatrix} f_n & f_{n+1} \\ f_{n+1} & f_{n+2} \end{bmatrix}.\end{aligned}$$

5.14.  $\mathbf{C} = \mathbf{A}\mathbf{B}$  Einstein-konvencióval:  $c_{ij} = a_{ik}b_{kj}$ .

$$5.17. \begin{bmatrix} 2 & -3 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 & 0 \\ 0 & 0 & -5 & 3 & 3 \\ 0 & 0 & -16 & 9 & 10 \\ 0 & 0 & 28 & -16 & -17 \end{bmatrix}$$

$$5.18. \begin{bmatrix} 2 & -3 & 7 & -4 & -4 \\ -1 & 2 & -7 & 4 & 4 \\ 0 & 0 & -5 & 3 & 3 \\ 0 & 0 & -16 & 9 & 10 \\ 0 & 0 & 28 & -16 & -17 \end{bmatrix}$$

$$5.19. \begin{bmatrix} -1 & 0 & 1 & 1 & 1 \\ 2 & -1 & -1 & -1 & -1 \\ -1 & 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

5.22. A fölcserélhetőségre vonatkozó  $\mathbf{AB} = \mathbf{BA}$  egyenletet szorozzuk meg mindkét oldalról  $\mathbf{B}^{-1}$ -gyel:

$$\mathbf{B}^{-1}(\mathbf{AB})\mathbf{B}^{-1} = \mathbf{B}^{-1}(\mathbf{BA})\mathbf{B}^{-1}.$$

Az asszociativitást használva

$$(\mathbf{B}^{-1}\mathbf{A})(\mathbf{BB}^{-1}) = (\mathbf{B}^{-1}\mathbf{B})(\mathbf{AB}^{-1}),$$

amiből a  $\mathbf{BB}^{-1} = \mathbf{I}$  azonosság fölhasználásával kapjuk, hogy

$$\mathbf{B}^{-1}\mathbf{A} = \mathbf{AB}^{-1}.$$

5.23. Azt mondjuk, hogy a  $\odot$  művelet invertálható a  $H$  egy  $R$  részhalmazán, ha bármely  $a, b, c \in R$  elem esetén az

$$a \odot x = b, \quad y \odot a = c$$

egyenletek mindegyike megoldható, azaz vannak olyan  $x, y \in H$  elemek, melyek kielégítik a fenti egyenleteket. Ha a definícióbeli  $\odot$  kommutatív művelet, akkor elég a fenti két egyenlet egyikét tekinteni.

5.24. a) Az  $e \in H$  semleges elem, ha minden  $a \in H$  elemre  $a \odot e = e \odot a = a$ . b) Azt mondjuk, hogy a  $\odot$  műveletre nézve  $a$  inverze  $b$ , ha  $a \odot b = b \odot a = e$ .

5.27. Az első egyenletet kivonjuk az utolsóból, innen  $x = 1$ , ezután visszahelyettesítés a második, harmadik, majd az első egyenletbe.

$$5.32. \mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/4 & 1/3 & 1 \end{bmatrix}, \mathbf{U} = \begin{bmatrix} 4 & 4 & 4 \\ 0 & 3 & 3 \\ 0 & 0 & 2 \end{bmatrix}.$$

$$5.33. \mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/4 & 1/3 & 1 \end{bmatrix}, \mathbf{U} = \begin{bmatrix} 4 & 8 & 4 \\ 0 & 3 & 6 \\ 0 & 0 & 1 \end{bmatrix}$$

$$5.34. \mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ 4/5 & 1 & 0 \\ -3/5 & 7/9 & 1 \end{bmatrix}, \mathbf{U} = \begin{bmatrix} 5 & -4 & -2 \\ 0 & -9/5 & -17/5 \\ 0 & 0 & -23/9 \end{bmatrix}.$$

$$5.35. \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2/3 & 1 & 0 & 0 \\ -1/3 & 1 & 1 & 0 \\ 1 & -3/10 & -1/2 & 1 \end{bmatrix} \begin{bmatrix} -3 & 1 & -3 & 0 \\ 0 & 10/3 & 5 & -4 \\ 0 & 0 & -3 & 4 \\ 0 & 0 & 0 & -2/10 \end{bmatrix}$$

$$5.36. \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1/2 & 1/2 & 1 & 0 \\ -1/2 & -1/2 & 1/3 & 1 \end{bmatrix} \begin{bmatrix} -2 & -2 & 0 & 3 \\ 0 & 2 & -2 & -1 \\ 0 & 0 & 3 & -3 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

$$5.37. \begin{bmatrix} 1.00 & 0.00 & 0.00 \\ 0.25 & 1.00 & 0.00 \\ 0.50 & 1.00 & 1.00 \end{bmatrix} \begin{bmatrix} 2.00 & 2.00 & -2.00 \\ 0.00 & 0.50 & -1.50 \\ 0.00 & 0.00 & 3.50 \end{bmatrix}.$$

$$5.38. \mathbf{y} = (0, 3, 4), \mathbf{x} = (-1, -1, 2).$$

$$5.39. \mathbf{y} = (0, 3, 1), \mathbf{x} = (1, -1, 1).$$

$$5.40. \mathbf{y} = (3, -17/5, -23/9), \mathbf{x} = (1, 0, 1).$$

$$5.41. \mathbf{y} = (5.6, 0.4, 1.4), \mathbf{x} = (1.2, 2.0, 0.4).$$

5.42. Mivel ismerjük az együtthatómátrix LU-felbontását – az épp az (5.4)-beli felbontás –, ezért ezt használjuk, és először megoldjuk az  $\mathbf{Ly} = \mathbf{b}$  egyenletrendszert:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/4 & 1/2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \\ 4 \end{bmatrix}$$

Ebből  $y_1 = 8$ , ezt a második egyenletbe helyettesítve kapjuk, hogy  $y_2 = 0$ , majd ezeket a harmadikba helyettesítve kapjuk, hogy  $y_3 = 2$ . Ezután megoldjuk az  $\mathbf{Ux} = \mathbf{y}$  egyenletrendszert, aminek alakja

$$\begin{bmatrix} 4 & 8 & 4 & 8 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 8 \\ 0 \\ 2 \end{bmatrix}.$$

Ismét egyszerű visszahelyettesítésekkel kapjuk, hogy  $x_4 = 1$ ,  $x_3 = s$  a szabad változó,  $x_2 = -s$  és  $x_1 = s$ . A megoldás  $\mathbf{x} = (s, -s, s, 1) = (0, 0, 0, 1) + s(1, -1, 1, 0)$ .

5.43. Az  $\mathbf{LY} = \mathbf{I}$  egyenletből

$$\mathbf{Y} = \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ -1/12 & -1/3 & 1 \end{bmatrix},$$

míg az  $\mathbf{UX} = \mathbf{Y}$  egyenlet megoldása, egyúttal  $\mathbf{A}$  inverze

$$\mathbf{X} = \begin{bmatrix} 5/12 & -1/3 & 0 \\ -1/8 & 1/2 & -1/2 \\ -1/24 & -1/6 & 1/2 \end{bmatrix}.$$

$$5.44. \mathbf{Y} = \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ -1/12 & -1/3 & 1 \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 1/3 & -5/3 & 3 \\ 0 & 1 & -2 \\ -1/12 & -1/3 & 1 \end{bmatrix}.$$

5.45. Tegyük fel, hogy létezik az  $n$ -edrendű  $\mathbf{A}$  mátrixnak két LU-felbontása is, azaz  $\mathbf{A} = \mathbf{L}_1\mathbf{U}_1 = \mathbf{L}_2\mathbf{U}_2$ . Mivel  $\mathbf{A}$  invertálható, ezért  $\mathbf{L}_1$ ,  $\mathbf{U}_1$ ,  $\mathbf{L}_2$  és  $\mathbf{U}_2$  is. Ugyanis ha pl.  $\mathbf{L}_1$  nem volna invertálható, akkor az oszloptérének dimenziója kisebb lenne  $n$ -nél, és mivel  $\mathbf{L}_1\mathbf{U}_1$  oszlopvektorai az  $\mathbf{L}_1$  oszlopvektorainak lineáris kombinációi, ezért e szorzat oszloptérének dimenziója is kisebb lenne  $n$ -nél, azaz  $\mathbf{A}$  nem lenne invertálható. A többi mátrix invertálhatósága hasonlóan igazolható. Balról  $\mathbf{L}_1$ , jobbról  $\mathbf{U}_2$  inverzével szorozva kapjuk, hogy

$$\mathbf{U}_1\mathbf{U}_2^{-1} = \mathbf{L}_1^{-1}\mathbf{L}_2.$$

A bal oldalon két felső háromszögmátrix szorzataként egy felső háromszögmátrix van, míg a jobb oldalon két alsó háromszögmátrix szorzata, ami alsó háromszögmátrix (?? feladat). Ráadásul a jobb oldal egység főátlójú (?? feladat). Ez csak akkor állhat fenn, ha  $\mathbf{U}_1\mathbf{U}_2^{-1} = \mathbf{L}_1^{-1}\mathbf{L}_2 = \mathbf{I}$ , azaz ha  $\mathbf{L}_1 = \mathbf{L}_2$  és  $\mathbf{U}_1 = \mathbf{U}_2$ .

$$5.46. \quad \mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{2}{3} & 1 & 0 \\ \frac{1}{3} & \frac{1}{2} & 1 \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} 3 & 3 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

$$5.47. \quad \mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{2}{3} & -1 & 1 \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} 3 & 4 & 6 & 7 & 9 \\ 0 & \frac{2}{3} & 1 & \frac{5}{3} & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{P} =$$

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

$$5.48. \quad \mathbf{L} = \begin{bmatrix} 1.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 \\ 0.0 & -0.5 & 1.0 \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} 0.5 & -2.0 & 2.0 \\ 0.0 & 2.0 & 2.0 \\ 0.0 & 0.0 & 2.5 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

5.49. Egyrészt  $\det \mathbf{A} = 1 \neq 0$ , és így  $\det \mathbf{U} \neq 0$ , tehát  $u_{11} \neq 0$ , másrészt  $\mathbf{A}$  és egy LU-felbontás bal felső elemére  $0 = a_{11} = (\mathbf{LU})_{11} = l_{11}u_{11} \neq 0$ , ami ellentmondás. Az első két sor felcserélése után kapott mátrixnál hasonló ellentmondásra jutunk az  $a_{22}$  elemmel.



## 6

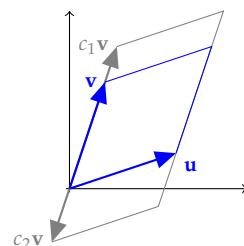
# Determináns

Egy valós négyzetes mátrix sorvektorai által kifeszített paralelepipedon térfogata jó jellemezője lehet a mátrixnak. Ehhez közel áll a determináns fogalma, melyet egy négyzetes mátrixokon értelmezett skálárértékű függvényként definiálunk. E skálár könnyen kiszámolható az elemi sorműveletekkel, sőt definiálni is segítségükkel fogjuk. A determináns fontos tulajdonságainak egyike, hogy pontosan akkor nulla, ha a sorvektorok lineárisan összefüggők.

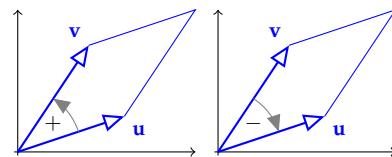
**Parallelogramma előjeles területe** A **parallelogramma területéről** szóló 1.42. példában láttuk, hogy az  $(a, b)$  és a  $(c, d)$  vektorok által kifeszített parallelogramma területe  $|ad - bc|$ , és hogy  $ad - bc$  pontosan akkor pozitív, ha az  $(a, b)$  és a  $(c, d)$  vektorok jobbrendszert alkotnak, és pontosan akkor negatív, ha az  $(a, b)$  és a  $(c, d)$  vektorok balrendszert alkotnak. Ez vezet a következő definícióhoz: Két síkbeli vektor által kifeszített parallelogramma **előjeles területe** megegyezik területével, ha a két vektor jobbrendszert alkot, és a terület  $-1$ -szeresével, ha a két vektor balrendszert alkot.

Az előzőek szerint az  $\mathbf{u} = (a, b)$  és a  $\mathbf{v} = (c, d)$  vektorok által kifeszített parallelogramma előjeles területe  $ad - bc$ . Az előjeles terület tehát egy vektorpárokon értelmezett valós értékű függvény – jelölje most  $f$  –, mely eleget tesz a következő tulajdonságoknak.

1.  $f(c\mathbf{u}, \mathbf{v}) = cf(\mathbf{u}, \mathbf{v})$ , és  $f(\mathbf{u}, c\mathbf{v}) = cf(\mathbf{u}, \mathbf{v})$ , azaz ha  $f$  egyik argumentumát  $c$ -vel szorozzuk, a függvényérték is  $c$ -szeresére változik. (E tulajdonságot úgy fogjuk kifejezni, hogy  $f$  homogén mindkét változójában.) Ez igaz az előjeles területre, hisz egy parallelogramma egyik oldalának  $c$ -szeresére növelése  $c$ -szerezi a területét is. Ha  $c$  negatív, akkor a vektorok körüljárása is változik összhangban azazal, hogy az előjeles területének is megváltozik az előjele (6.1. ábra).
2.  $f(\mathbf{u}, \mathbf{v}) = -f(\mathbf{v}, \mathbf{u})$ . Ez nyilvánvaló, hisz a két vektor sorrendjét megcserélve megváltozik orientációjuk (jobbrendszerből balrendszerbe és viszont, ld. 6.2. ábra).



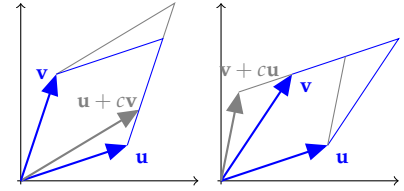
6.1. ábra: Az  $f(\mathbf{u}, c_1\mathbf{v}) = c_1f(\mathbf{u}, \mathbf{v})$  ( $c_1 = 4/3 > 0$ ) és az  $f(\mathbf{u}, c_2\mathbf{v}) = c_2f(\mathbf{u}, \mathbf{v})$  ( $c_2 = -1/2 < 0$ ) összefüggések szemléltetése.



6.2. ábra: Két vektor sorrendjének cseréje megváltoztatja a orientációjukat.

3.  $f(\mathbf{u}, \mathbf{v}) = f(\mathbf{u} + c\mathbf{v}, \mathbf{v}) = f(\mathbf{u}, \mathbf{v} + c\mathbf{u})$ , azaz az  $\begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  mátrix sorvektorai által kifeszített paralelogramma területe megegyezik a hozzáadás sorművelete után kapott mátrix sorvektoraihoz tartozó paralelogramma területével (6.3. ábra).
4.  $f(\mathbf{u}, \mathbf{u}) = 0$ , és hasonlóképp  $f(\mathbf{u}, \mathbf{0}) = f(\mathbf{0}, \mathbf{u}) = 0$  tetszőleges  $\mathbf{u}$  vektorra, ugyanis az elfajuló paralelogramma területe 0.
5.  $f(\mathbf{i}, \mathbf{j}) = 1$ , azaz a standard bázis által kifeszített egységnégyzet területe 1.

Az állítások a fenti ábrákkal szemléltetett egyszerű geometriai érvelések mellett az  $f((a, b), (c, d)) = ad - bc$  formulával is bizonyíthatóak. E tulajdonságok segítségével általánosítani tudjuk az előjeles terület fogalmát, és bevezethetjük az előjeles térfogat fogalmát az  $n$ -dimenziós valós tér paralelepipedonjaira.



6.3. ábra: Az  $f(\mathbf{u}, \mathbf{v}) = f(\mathbf{u} + c\mathbf{v}, \mathbf{v})$  és az  $f(\mathbf{u}, \mathbf{v}) = f(\mathbf{u}, \mathbf{v} + c\mathbf{u})$  összefüggések szemléltetése.

**Paralelepipedon előjeles térfogata** A vegyes szorzat tárgyalásakor láttuk (1.33. definíció), hogy a valós háromdimenziós térben három vektor vegyes szorzata a vektorok által kifeszített paralelepipedon előjeles térfogatát adja, ahol az előjel aszerint pozitív vagy negatív, hogy a három vektor jobb- vagy balrendszert alkot. A háromdimenziós tér paralelepipedonjainak  $f$  előjeles térfogatára a paralelogrammánál láttakhoz hasonló tulajdonságok igazolhatók.

1.  $f$  homogén mindhárom argumentumában, azaz egy konstans tényező bármelyik argumentumból kiemelhető, pl.  $f(c\mathbf{u}, \mathbf{v}, \mathbf{w}) = cf(\mathbf{u}, \mathbf{v}, \mathbf{w})$ .
2. Bármely két argumentum felcserélése megváltoztatja a függvényérték előjelét, pl.  $f(\mathbf{u}, \mathbf{v}, \mathbf{w}) = -f(\mathbf{w}, \mathbf{v}, \mathbf{u})$ ,  $f(\mathbf{u}, \mathbf{v}, \mathbf{w}) = -f(\mathbf{u}, \mathbf{w}, \mathbf{v})$ .
3.  $f$  bármely argumentumához hozzáadva egy másik konstansszorzót, a függvényérték nem változik, pl.  $f(\mathbf{u}, \mathbf{v}, \mathbf{w}) = f(\mathbf{u} + c\mathbf{w}, \mathbf{v}, \mathbf{w})$ .
4. Ha  $f$  bármely két argumentuma megegyezik, a függvényérték 0, pl.  $f(\mathbf{u}, \mathbf{v}, \mathbf{u}) = 0$ . Ugyancsak 0 értéket kapunk, ha  $f$  bármelyik argumentuma a  $\mathbf{0}$ -vektor, pl.  $f(\mathbf{u}, \mathbf{0}, \mathbf{w}) = 0$ .
5.  $f(\mathbf{i}, \mathbf{j}, \mathbf{k}) = 1$ , azaz az egységkocka térfogata 1.

Látjuk, hogy  $f$  eddig megismert tulajdonságai vagy azonnal megadják  $f$  értékét (ha az 0 vagy 1), vagy az argumentumok olyan megváltoztatását tartalmazzák, amelyekhez hasonlókat a mátrixok elemi sorműveleteinél láttunk. Valóban, e tulajdonságok nem csak egy új fogalom – az előjeles térfogat általánosítását –, de egyúttal annak egyszerű kiszámítási módját is lehetővé teszik.



## A determináns mint sorvektorainak függvénye

**A determináns definíciója** Az  $n$  darab  $n$ -dimenziós vektor által kifeszített paralelepipedon előjeles térfogata helyett olyan fogalmat fogunk definiálni, mely speciális esetként ezt is tartalmazza. Ez lesz a determináns. A determináns tehát olyan függvény, mely  $n$  darab  $n$ -dimenziós vektorhoz – vagy ami ezzel ekvivalens, a belőlük képzett  $n \times n$ -es mátrixhoz – egy skalárt rendel. A definícióban csak az előjeles térfogat vizsgálatában megismert függvénytulajdonságokat használjuk.

Az  $\mathbf{A} = [a_{ij}]_n$  mátrixhoz rendelt skalárt, azaz determinánsának értékét  $\det(\mathbf{A})$ ,  $|\mathbf{A}|$  vagy  $|a_{ij}|_n$  jelöli. Részletezve az általános jelölést az  $\mathbf{A}$  mátrixra és determinánsára:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, \quad \det(\mathbf{A}) = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

E jelölésnek megfelelően a  $\det(\mathbf{A})$  determináns sorain, oszlopain, elemein az  $\mathbf{A}$  mátrix sorait, oszlopait, elemeit értjük.

**6.1. DEFINÍCIÓ (DETERMINÁNS).** Determinánson azt a négyzetes mátrixokon értelmezett és  $\det$ -tel jelölt skalár értékű függvényt értjük, mely eleget tesz a következő feltételeknek:

- D1. értéke  $c$ -szeresére változik, ha egy sorát  $c$ -vel szorozzuk,
- D2. értéke  $-1$ -szeresére változik, ha két különböző sorát fölcseréljük,
- D3. értéke nem változik a hozzáadás elemi sorművelete közben,
- D4. az egységmátrixhoz 1-et rendel.

► Mint látjuk, a definíció első három feltétele azt mondja ki, hogy hogyan változik a determináns értéke elemi sorműveletek közben. Az egyetlen változás, hogy itt a skalárral való szorzásnál nem kötöttük ki, hogy  $c$  nem lehet 0. Látni fogjuk, hogy e kikötés elhagyása nem fog gondot okozni az elemi sorműveletek determinánsokra való alkalmazásában.

► A definícióból nem látszik, hogy e feltételeket kielégítő függvény létezik-e és ha igen, egyértelmű-e. Ezeket később igazolni fogjuk.

► A definícióban nem törekedtünk a feltételek minimalizálására, inkább a természetesség és egyszerűség volt a fontosabb szempont. Például a D2. feltétel elhagyható, hisz a sorcsere előállítható a két másik sorművelet segítségével, amint azt a 2.20. feladatból is láthattuk.

► Az  $1 \times 1$ -es  $[a]$  mátrix determinánsa  $\det([a]) = a$ , ugyanis a determináns definíciója szerint  $\det([1]) = 1$ , és  $\det([a]) = \det([a \cdot 1]) = a \det([1]) = a$ . A D2. és D3. feltételek teljesülnek, hisz e determinánsnak csak egy sora van. A jelölésbeli zavarok elkerülésére az  $1 \times 1$ -es  $[a]$

A determinánsokat először SEKI Takakazu (関 孝和, 1642–1708) vizsgálta, eredményei 1683-ban jelentek meg. Seki 2-től 5-öndrendűek értékét tudta kiszámolni. Európában is 1683-ban jelenik meg e fogalom először Leibniz egy l'Hôpitalnak írt levelében, melyet később rezultánsnak hívtak. A determináns név Gauss-tól származik. A mátrixok és determinánsok történetének szép összefoglalója olvasható a [MacTutor History of Mathematics](#) weboldalon.

mátrix determinánsára csak a  $\det([a])$  vagy  $\det(a)$  jelölést használjuk, mert  $|a|$  az  $a$  abszolút értékét jelöli!

► Hamarosan igazolni fogjuk, hogy a fejezet elején említett

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

képlet illeszkedik e definícióhoz.

► A determináns tekinthető olyan  $n$ -változós függvénynek, melynek  $n$  argumentumába a mátrix  $n$  sorvektora kerül. Nem okoz félreértést, ha ezt a függvényt is  $\det$  jelöli. Ha tehát  $\mathbf{A}$  sorvektorai  $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n$ , akkor  $\det(\mathbf{A})$  megegyezik a  $\det(\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n)$  függvényértékkel. Például a  $3 \times 3$ -as egységmátrix determinánsa az alábbi alakokba írható:

$$\det(\mathbf{I}_3) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = \det((1,0,0), (0,1,0), (0,0,1)) = \det(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3),$$

ahol  $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$  a standard egységvektorokat jelöli. A determináns fenti definíciója könnyen fölírható e jelöléssel is (ld. 6.11. feladat).

**Mikor 0 a determináns értéke** Gyakran vízvázasztó, hogy egy determináns értéke zérus-e.

**6.2. TÉTEL (RÁNÉZÉSRE 0 DETERMINÁNSOK).** *Ha egy mátrixnak van egy zérussora, akkor determinánsa 0. Ha egy mátrixnak van két azonos sora, akkor determinánsa 0.*

**BIZONYÍTÁS.** Ha egy mátrixnak van egy zérussora, akkor e sort bármely  $c$  számmal beszorozva e sor nem változik, így a determináns értéke sem. Másrészt a determináns definíciójának  $D1$ . pontja szerint a determináns értéke  $c$ -szeresére változik. E két feltétel csak úgy állhat fenn minden  $c$  skalárra, ha  $\det(\mathbf{A}) = 0$ . (Ennek következményeként a definíció  $D1$ . pontjában nem kell a  $c = 0$  lehetőséget kizárni.)

Ha egy determinánsnak két azonos sora van, akkor  $D3$ . szerint értéke nem változik, ha az egyik sort a másiktól kivonjuk, így egy zérussort kapunk, akkor pedig a determináns értéke 0.  $\square$

**6.3. TÉTEL (ZÉRUS ÉRTÉKŰ DETERMINÁNS).** *Legyen  $\mathbf{A}$  négyzetes mátrix. A következő állítások ekvivalensek:*

1.  $\det(\mathbf{A}) = 0$ ,
2.  $\mathbf{A}$  sorvektorai lineárisan összefüggők,
3.  $\mathbf{A}$  szinguláris,
4. a homogén lineáris  $\mathbf{Ax} = \mathbf{0}$  egyenletrendszernek van nemtriviális megoldása.

**BIZONYÍTÁS.** Az 5.20. tételben láttuk, hogy négyzetes mátrix sorvektorai pontosan akkor lineárisan összefüggők, ha a mátrix szinguláris,

azaz ha a lépcsős alakra hozás során keletkezik egy 0-sor, ez pedig azzal ekvivalens, hogy a determináns értéke 0. Az utolsó állítás ekvivalenciája a mátrix invertálhatóságáról szóló 5.15. tétel közvetlen következménye.  $\square$

6.4. PÉLDA (ZÉRUS ÉRTÉKŰ DETERMINÁNSOK). *A sorvektorok lineáris összefüggőségének igazolásával mutassuk meg, hogy*

$$\begin{vmatrix} 5 & 6 & 8 \\ 2 & 1 & 2 \\ 3 & 5 & 6 \end{vmatrix} = 0, \quad \begin{vmatrix} 2 & -1 & 0 & -1 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ -1 & 0 & -1 & 2 \end{vmatrix} = 0.$$

**MEGOLDÁS.** Az első determináns első sora a második és a harmadik összege. De fogalmazhatunk úgy is, hogy az első sorból kivonva a másodikat és a harmadikat, a nullvektort kapjuk. Tehát az első mátrix sorvektorai lineárisan összefüggők, így determinánsa 0.

A második determináns sorvektorainak összege a nullvektor, tehát ezek is lineárisan összefüggők, így ez a determináns is 0.  $\square$

Az előző 6.3. tétel, valamint az 5.15. tétel fontos következménye a determinánsnak az egyenletrendszerek megoldhatóságával való kapcsolatáról szól:

6.5. TÉTEL (EGYENLETRENDSZER MEGOLDHATÓSÁGA ÉS A DETERMINÁNS). *Legyen  $\mathbf{A}$  négyzetes mátrix. Ekkor az alábbi állítások ekvivalensek:*

1.  $\det \mathbf{A} \neq 0$ ,
2. az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszer tetszőleges  $\mathbf{b}$ -re egyértelműen megoldható,
3. az  $\mathbf{Ax} = \mathbf{0}$  egyenletrendszernek csak triviális megoldása van.

► A legegyszerűbb eseteket leszámítva a sorvektorok lineáris összefüggősége „ránézésre” nem látható, de az összefüggőséget bizonyító skalárok – ha szükségünk van rá – megkaphatók az  $\mathbf{A}^T \mathbf{x} = \mathbf{0}$  egyenletrendszer nemtriviális megoldásaiból.

► A gyakorlatban – például mért vagy közelítő számítással kapott adatok esetén – annak eldöntése, hogy egy determináns nulla-e, nagy óvatosságot igényel! Az, hogy egy mátrix „közel szinguláris”, nem feltétlenül olvasható le abból, hogy a determináns értéke „közel van a nullához”. Például az

$$\begin{vmatrix} \frac{1}{n} & 0 \\ 0 & n \end{vmatrix} = 1, \quad \text{és az} \quad \begin{vmatrix} \frac{1}{2} & 0 & \dots & 0 \\ 0 & \frac{1}{2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{2} \end{vmatrix} = \frac{1}{2^n}$$

determinánsok közül az első értéke tetszőlegesen nagy  $n$ -re is 1, pedig  $\frac{1}{n}$  tetszőlegesen közel lehet 0-hoz, és az  $\begin{bmatrix} 0 & 0 \\ 0 & n \end{bmatrix}$  mátrix már szinguláris. A

második determinánsbeli  $\frac{1}{2}\mathbf{I}_n$  mátrix nem szinguláris, pedig determinánsának értéke tetszőlegesen közel lehet 0-hoz, igaz, csak elegendően nagy  $n$  esetén.

► A véletlen valós mátrixok determinánsa 1 valószínűséggel nem 0, ha a mátrix elemeit valamely folytonos valószínűségeloszlás szerint választjuk. Másként fogalmazva, ha egy valós elemű mátrix determinánsa 0, akkor annak különleges oka van! Ez az ok, a sorvektorok közti lineáris kapcsolat, ami „igen ritkán” esik meg „véletlenül”.

**A determináns értékének kiszámítása** A determináns kiszámításához az elemi sorműveleteket fogjuk használni. A 6.1. definíció pontosan megmondja, hogyan változik a determináns értéke az elemi sorműveletek közben. Ha a lépcsős alakra hozás közben nem keletkezik zérussor, akkor a lépcsős alak háromszög alakú, illetve a redukált lépcsős alak diagonális. Ezek értékéről szól a következő tétel:

**6.6. TÉTEL (HÁROMSZÖGMÁTRIX DETERMINÁNSA).** *Az alsó vagy felső háromszögmátrix, s így a diagonális mátrix determinánsa megegyezik a főátlóbeli elemek szorzatával.*

**BIZONYÍTÁS.** Ha egy háromszögmátrix főátlójában van 0, akkor a redukált lépcsős alakra hozás után a főelemek száma kevesebb lesz, mint a sorok száma, azaz a mátrixban lesz egy zérussor, így determinánsának értéke 0. Ha nincs 0-elem a főátlóban, mind az alsó, mind a felső háromszögmátrix csak a hozzáadás sorműveletével – azaz a determináns értékének megváltoztatása nélkül – diagonálissá alakítható a főátlón kívüli elemek kiküszöbölésével, azaz

$$\begin{vmatrix} a_{11} & 0 & \dots & 0 \\ ? & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ ? & ? & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & ? & \dots & ? \\ 0 & a_{22} & \dots & ? \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix}$$

Egy diagonális mátrix determinánsában minden sorból kiemelve a főátlóban szereplő számot kapjuk, hogy

$$\begin{vmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix} \stackrel{D_I}{=} a_{11}a_{22}\dots a_{nn} \begin{vmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{vmatrix} = a_{11}a_{22}\dots a_{nn},$$

tehát a determináns értéke valóban a főátlóbeli elemek szorzata. □

Például az alábbi determináns értéke egyetlen sorcsere után azonnal leolvasható:

$$\begin{vmatrix} 3 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 2 & 0 \\ 3 & 0 & 2 & 0 & 0 \\ 3 & 2 & 0 & 0 & 0 \\ 3 & 3 & 3 & 3 & 3 \end{vmatrix} = - \begin{vmatrix} 3 & 0 & 0 & 0 & 0 \\ 3 & 2 & 0 & 0 & 0 \\ 3 & 0 & 2 & 0 & 0 \\ 3 & 0 & 0 & 2 & 0 \\ 3 & 3 & 3 & 3 & 3 \end{vmatrix} = -3 \cdot 2 \cdot 2 \cdot 2 \cdot 3 = -72$$

A determináns kézzel való kiszámításának módja tehát a következő: elemi sorműveletekkel hozzuk a determinánst olyan alakra, melynek vagy van egy zérussora, vagy háromszög alakú. Az elemi sorműveletek közben pedig gondosan adminisztráljuk hatásukat, azaz

- két sor cseréjekor szorozzuk meg a determinánst  $-1$ -gyel,
- egy sorának  $c$ -vel való szorzásakor pedig szorozzuk meg a determinánst  $1/c$ -vel.

A háttérben lényegében ezt teszik a számítógépek is (ld. a 6.4. kódot).

**6.7. PÉLDA (DETERMINÁNS KISZÁMÍTÁSA HÁROMSZÖG ALAKRA HOZÁS-SAL).** Számítsuk ki  $a$

$$\begin{vmatrix} 2 & 2 & -3 \\ 2 & 2 & -4 \\ 4 & 5 & -6 \end{vmatrix} \quad \text{és } a \quad \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{vmatrix}$$

*determinánsok értékét!*

**MEGOLDÁS.** Elemi sorműveletekkel kapjuk, hogy

$$\begin{vmatrix} 2 & 2 & -3 \\ 2 & 2 & -4 \\ 4 & 5 & -6 \end{vmatrix} \xrightarrow{\substack{S_2 - S_1 \\ S_3 - 2S_1}} \begin{vmatrix} 2 & 2 & -3 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{vmatrix} \xrightarrow{S_2 \leftrightarrow S_3} - \begin{vmatrix} 2 & 2 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{vmatrix} = -(-2) = 2.$$

A következő determinánsnál sorcsere nélkül eliminálhatók a főátló alatti elemek, ezért a sorműveleteket nem is jelezzük.

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 2 & 5 & 9 \\ 0 & 3 & 9 & 19 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 3 & 10 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 1.$$

Egy érdekes észrevétel: a fenti determinánsban és sorlécsős alakjában is a Pascal-háromszög számai találhatók. Ez nem véletlen, erről szólnak a 6.15. és a 6.16. feladatok.  $\square$

**Elemi mátrixok determinánsa** Az elemi mátrixok egyetlen sorművelettel kaphatók az egység mátrixból, így ezek determinánsa könnyen számolható.

```
sage: M = matrix(3,range(9))
sage: M[2,2]=9
sage: M
[0 1 2]
[3 4 5]
[6 7 9]
sage: M.det()
-3
sage: det(M)
-3
```

6.4. kód: Determináns kiszámítása

**6.8. KÖVETKEZMÉNY (ELEMI MÁTRIXOK DETERMINÁNSA).** *A hozzáadás sorműveletével kapott elemi mátrix determinánusa 1, a sorcserével kapotté  $-1$ , egy sor  $c$ -vel való szorzásával kapotté  $c$ .*

**BIZONYÍTÁS.** Az állítás abból következik, hogy az elemi mátrixok az 1 determinánsú egységmátrixból kaphatók egyetlen sorművelettel.  $\square$

Például:

$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 4 & 0 & 1 \end{vmatrix} = 1, \quad \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{vmatrix} = -1, \quad \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{vmatrix} = 3.$$

**Permutáló mátrix determinánusa\*** A permutáló mátrix minden sorában és oszlopában egyetlen 1-es van, így csak elemi sorcserékkel megkapható az egységmátrixból. A sorcsere csak a determináns előjelét változtatja meg, ezért permutáló mátrix determinánusa 1, ha páros sok sorcserére volt szükség,  $-1$ , ha páratlan sokra. Például az alábbi determinánsok közül az első determináns két sorcserével, a második három sorcserével kapható meg az egységmátrixból, tehát

$$\begin{vmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 1, \quad \begin{vmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{vmatrix} = -1.$$

Azt mondjuk, hogy egy permutáló mátrix két sora *inverzióban* áll, ha az előbb álló sorbeli 1-es hátrébb van, mint a másik sorbeli. A

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

mátrix inverzióinak száma például 4, ugyanis az első-második, első-negyedik, második-negyedik, harmadik-negyedik sorpárok inverzióban vannak.

**6.9. TÉTEL (PERMUTÁLÓ MÁTRIX DETERMINÁNSA).** *A permutáló mátrix aszerint  $+1$  vagy  $-1$ , hogy inverzióban álló sorpárjainak száma páros vagy páratlan.*

**BIZONYÍTÁS.** Elég megmutatni, hogy egy sorcsere mindig megváltoztatja az inverziók számának paritását, vagyis azok száma párosból páratlanra, páratlanból párosra változik. Így ha egy permutáló mátrix

inverzióinak száma páros, akkor csak páros sok sorcserével vihető az identikus mátrixba. Hasonlóan, ha az inverziók száma páratlan, akkor csak páratlan sokkal.

Ha a két megcserélendő sor szomszédos, akkor a sorcsere megváltoztatja e két sor viszonyát: ha inverzióban álltak, akkor ezután nem fognak, és fordítva. Az előttük és mögöttük álló sorokhoz való viszonyuk nem változott. Eszerint az inverziók száma eggyel nőtt vagy eggyel csökkent, azaz paritása megváltozott.

Ezután cseréljük fel az  $i$ -edik és  $j$ -edik sorokat (legyen  $i < j$ ). Az inverziók számának nyomon követése érdekében ezt szomszédos sorok cseréjével valószínűsítjük meg. Cseréljük ki az  $i$ -ediket az  $(i+1)$ -edikkel, majd azt az  $(i+2)$ -edikkel, ..., míg az eredetileg  $i$ -edik sor a  $j$ -edik helyére nem kerül. Ehhez  $j-i$  sorcserére van szükség. Ezután az eredetileg  $j$ -edik sort  $j-i-1$  sorcserével az  $i$ -edik helyre visszük. Ez összesen  $2(j-i)-1$ , azaz páratlan sok sorcsere, ami a paritást valóban ellenkezőjére változtatja.  $\square$

**Mátrixműveletek és determináns** Kérdés, hogy milyen kapcsolat van a mátrixműveletek és a determináns között. Fontos megjegyezni, hogy a determinánsfüggvénynek *nincs* a mátrixösszeadásra és a skalárral való szorzásra nézve művelettartó tulajdonsága, azaz általában  $\det(\mathbf{A} + \mathbf{B}) \neq \det(\mathbf{A}) + \det(\mathbf{B})$ , és  $\det(c\mathbf{A}) \neq c \det(\mathbf{A})$ .

A skalárral való szorzás esetén mondható valami: mivel egy mátrix  $c$ -szeresének determinánsa minden sorából kiemelhető  $c$ , ez annyi kiemelést jelent, ahány sora van a mátrixnak. Így tetszőleges  $n \times n$ -es  $\mathbf{A}$  mátrixra és tetszőleges  $c$  skalárra  $\det(c\mathbf{A}) = c^n \det(\mathbf{A})$ . Ez az  $\mathbb{R}^2$ - vagy  $\mathbb{R}^3$ -beli geometriai interpretációból is világos: egy paralelogramma előjeles területe 4-szeresére, egy paralelepipedon előjeles térfogata 8-szorosára nő, ha minden élét 2-szeresére növeljük.

A determináns művelettartó a négyzetes mátrixok szorzására nézve. Ezt mondja ki a következő állítás.

**6.10. ÁLLÍTÁS (DETERMINÁNSOK SZORZÁSSZABÁLYA).** *Ha  $\mathbf{A}$  és  $\mathbf{B}$  azonos méretű négyzetes mátrixok, akkor  $\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$ .*

**BIZONYÍTÁS.** A 6.8. tétel következtében egy elemi mátrixszal való balról szorzás egy mátrixon olyan sorműveletet hajt végre, mely determinánsát épp annyszorozásra változtatja, amennyi az elemi mátrix determinánsa. Így egy  $\mathbf{E}$  elemi mátrix és egy tetszőleges négyzetes  $\mathbf{B}$  mátrix szorzatának determinánsa megegyezik determinánsaik szorzatával, azaz

$$\det(\mathbf{EB}) = \det(\mathbf{E}) \det(\mathbf{B}).$$

Tudjuk, hogy ha  $\mathbf{A}$  szinguláris, akkor  $\mathbf{AB}$  is, azaz ha  $\det(\mathbf{A}) = 0$ , akkor  $\det(\mathbf{AB})$  is 0, tehát  $\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$ . Ha  $\mathbf{A}$  nem szinguláris, akkor felbontható elemi mátrixok szorzatára:  $\mathbf{A} = \mathbf{E}_1 \mathbf{E}_2 \dots \mathbf{E}_k$ ,

így  $\mathbf{AB} = \mathbf{E}_1 \mathbf{E}_2 \dots \mathbf{E}_k \mathbf{B}$ . A  $\det(\mathbf{EB}) = \det(\mathbf{E}) \det(\mathbf{B})$  összefüggést az  $\mathbf{E}_1 \mathbf{E}_2 \dots \mathbf{E}_k$ -ra és  $\mathbf{E}_1 \mathbf{E}_2 \dots \mathbf{E}_k \mathbf{B}$ -re is használva kapjuk, hogy

$$\begin{aligned} \det(\mathbf{A}) \det(\mathbf{B}) &= \det(\mathbf{E}_1 \mathbf{E}_2 \dots \mathbf{E}_k) \det(\mathbf{B}) = \det(\mathbf{E}_1) \det(\mathbf{E}_2) \dots \det(\mathbf{E}_k) \det(\mathbf{B}) \\ &= \det(\mathbf{E}_1) \det(\mathbf{E}_2) \det(\mathbf{E}_3 \dots \mathbf{E}_k) \det(\mathbf{B}) = \dots = \\ &= \det(\mathbf{E}_1) \det(\mathbf{E}_2) \dots \det(\mathbf{E}_k) \det(\mathbf{B}), \text{ másrészt} \\ \det(\mathbf{AB}) &= \det(\mathbf{E}_1 \mathbf{E}_2 \dots \mathbf{E}_k \mathbf{B}) = \det(\mathbf{E}_1) \det(\mathbf{E}_2 \dots \mathbf{E}_k \mathbf{B}) \\ &= \det(\mathbf{E}_1) \det(\mathbf{E}_2) \det(\mathbf{E}_3 \dots \mathbf{E}_k \mathbf{B}) = \dots = \\ &= \det(\mathbf{E}_1) \det(\mathbf{E}_2) \dots \det(\mathbf{E}_k) \det(\mathbf{B}), \end{aligned}$$

ami bizonyítja az állítást. Egy másik, nagyon szép bizonyítás található a 6.12. feladatban.  $\square$

A determinánsok szorzásszabályának egy fontos alkalmazása a determináns értékének kiszámítása PLU-felbontással (ld. 6.5. kód).

**6.11. PÉLDA (DETERMINÁNS KISZÁMOLÁSA PLU-FELBONTÁSBÓL).** *Hogyan határozzuk meg egy  $\mathbf{A}$  mátrix determinánsát, ha ismerjük PLU-felbontását? Konkrétan mennyi a következő mátrix determinánsa?*

$$\begin{bmatrix} 0 & 1 & 2 \\ 3 & 5 & 6 \\ 4 & 7 & 9 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3/4 & -1/4 & 1 \end{bmatrix} \begin{bmatrix} 4 & 7 & 9 \\ 0 & 1 & 2 \\ 0 & 0 & -1/4 \end{bmatrix}$$

**MEGOLDÁS.** Egy PLU-felbontásban szereplő mindegyik mátrix determinánsa könnyen meghatározható.  $\mathbf{P}$  két sorcserével egységmátrixszá válik, tehát  $\det \mathbf{P} = 1$ .  $\mathbf{L}$  és  $\mathbf{U}$  háromszögmátrixok, amelyek determinánsa a főátlóbeli elemek szorzata, ami  $\mathbf{L}$  esetén mindig 1. A megadott konkrét esetben tehát  $\det \mathbf{A} = 4 \cdot 1 \cdot (-1/4) = -1$ .  $\square$

Mátrix determinánsa és transzponáltjának determinánsa megegyezik. Ez lehetővé teszi, hogy a determináns kiszámításához nem csak az elemi sor-, de az elemi oszlopműveleteket is használjuk, hisz egy mátrixon végzett oszlopművelet a transzponált sorművelete.

**6.12. ÁLLÍTÁS (TRANSPONÁLT DETERMINÁNSA).** *Mátrix determinánsa megegyezik transzponáltjának determinánsával, azaz bármely négyzetes  $\mathbf{A}$  mátrixra  $\det(\mathbf{A}) = \det(\mathbf{A}^T)$ .*

**BIZONYÍTÁS.** Az  $\mathbf{A}$  mátrix redukált lépcsős alakra hozásának mátrixszorzatos alakja legyen  $\mathbf{A} = \mathbf{E}_1 \mathbf{E}_2 \dots \mathbf{E}_k \mathbf{R}$ , ahol  $\mathbf{E}_i$  elemi mátrix,  $\mathbf{R}$  az  $\mathbf{A}$  redukált lépcsős alakja. A transzponált determinánsa

$$|\mathbf{A}^T| = |\mathbf{R}^T \mathbf{E}_k^T \dots \mathbf{E}_2^T \mathbf{E}_1^T| = |\mathbf{R}^T| |\mathbf{E}_k^T| \dots |\mathbf{E}_2^T| |\mathbf{E}_1^T|.$$

Könnyen ellenőrizhető, hogy minden elemi mátrix determinánsa megegyezik transzponáltjának determinánsával (ellenőrizzük!). Mivel  $\mathbf{R}$

```
sage: M = matrix(3, range(9))
sage: M[2,2]=9
sage: N=M.change_ring(RDF)
sage: N
[0.0 1.0 2.0]
[3.0 4.0 5.0]
[6.0 7.0 9.0]
sage: N.det()
-3.0
sage: P,L,U = N.LU()
sage: P
[0.0 0.0 1.0]
[1.0 0.0 0.0]
[0.0 1.0 0.0]
sage: U
[ 6.0  7.0  9.0]
[ 0.0  1.0  2.0]
[ 0.0  0.0 -0.5]
sage: P.det()
1.0
sage: U.det()
-3.0
```

6.5. kód: Determináns kiszámítása a PLU-felbontásból. A felbontás az egészek gyűrűjében nem működik, ezért gyűrűt váltunk és dupla pontosságú lebegőpontos számokkal számolunk (RDF).



redukált lépcsős alak, ezért  $\mathbf{R} = \mathbf{I}$ , vagy  $\mathbf{R}$ -nek van egy zérus sora. Ha  $\mathbf{R} = \mathbf{I}$ , akkor  $|\mathbf{R}^T| = |\mathbf{R}| = |\mathbf{I}| = 1$ , ha pedig  $\mathbf{R}$ -nek van zérus sora, akkor  $\mathbf{R}^T$ -nak zérus oszlopa, és egy ilyen mátrix nem alakítható elemi sorműveletekkel egységmátrixszá, tehát determinánsa 0. Azaz  $|\mathbf{R}| = |\mathbf{R}^T|$  ekkor is fennáll. Ekkor pedig

$$\begin{aligned} |\mathbf{A}^T| &= |\mathbf{R}^T| |\mathbf{E}_k^T| \dots |\mathbf{E}_2^T| |\mathbf{E}_1^T| = |\mathbf{R}| |\mathbf{E}_k| \dots |\mathbf{E}_2| |\mathbf{E}_1| \\ &= |\mathbf{E}_1| |\mathbf{E}_2| \dots |\mathbf{E}_k| |\mathbf{R}| = |\mathbf{E}_1 \mathbf{E}_2 \dots \mathbf{E}_k \mathbf{R}| = |\mathbf{A}|. \end{aligned}$$

Tehát  $|\mathbf{A}^T| = |\mathbf{A}|$ . □

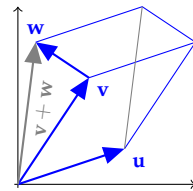
**6.13. PÉLDA (DETERMINÁNS KISZÁMÍTÁSA ELEMI OSZLOPMŰVELETEKKEL).** Az alábbi determinánst elemi sor- és oszlopműveletek alkalmazásával 2 lépésben is kiszámíthatjuk:

$$\begin{vmatrix} 1 & 0 & 0 & 1 & 0 \\ 2 & 2 & 1 & 3 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 2 & 0 \\ 1 & 1 & 1 & 2 & 1 \end{vmatrix} \xrightarrow{S_2 - S_5} \begin{vmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 2 & 0 \\ 1 & 1 & 1 & 2 & 1 \end{vmatrix} \xrightarrow{O_4 - O_1} \begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{vmatrix} = 1$$

**Determinánsok soronkénti additivitása** A determinánsok egy fontos tulajdonságát a paralelogramma előjeles területével szemléltetünk. Tekintsük a síkban az  $\mathbf{u}$ ,  $\mathbf{v}$  és  $\mathbf{w}$  vektorokat, valamint az  $\mathbf{u}$  és  $\mathbf{v}$  által, valamint az  $\mathbf{u}$  és  $\mathbf{w}$  által kifeszített paralelogrammákat, ahogy azt a 6.6. ábra mutatja. Igazolható, de az ábráról is leolvasható, hogy (előjeles) területük összege megegyezik az  $\mathbf{u}$  és a  $\mathbf{v} + \mathbf{w}$  vektorok által kifeszített paralelogramma (előjeles) területével. Az előjeles területet  $f$ -fel jelölve igaz tehát az  $f(\mathbf{u}, \mathbf{v}) + f(\mathbf{u}, \mathbf{w}) = f(\mathbf{u}, \mathbf{v} + \mathbf{w})$  összefüggés. Hasonlóképp igaz, hogy  $f(\mathbf{u}, \mathbf{v}) + f(\mathbf{w}, \mathbf{v}) = f(\mathbf{u} + \mathbf{w}, \mathbf{v})$  bármely három  $\mathbf{u}$ ,  $\mathbf{v}$ ,  $\mathbf{w}$  vektorra. Szavakban kifejezve  $f$  additív mindkét változójában.

**6.14. TÉTEL (SORONKÉNTI ADDITIVITÁS).** Legyen  $\mathbf{A}$ ,  $\mathbf{B}$  és  $\mathbf{C}$  három olyan mátrix, melyek  $i$ -edik sorukat kivéve megegyeznek egymással. A három mátrix  $i$ -edik sorvektora legyen rendre  $\mathbf{a}_i$ ,  $\mathbf{b}_i$  és  $\mathbf{a}_i + \mathbf{b}_i$ . Ekkor  $|\mathbf{A}| + |\mathbf{B}| = |\mathbf{C}|$ , azaz

$$\begin{vmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_i \\ \vdots \\ \mathbf{a}_n \end{vmatrix} + \begin{vmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{b}_i \\ \vdots \\ \mathbf{a}_n \end{vmatrix} = \begin{vmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_i + \mathbf{b}_i \\ \vdots \\ \mathbf{a}_n \end{vmatrix}$$



6.6. ábra: Az  $f(\mathbf{u}, \mathbf{v}) + f(\mathbf{u}, \mathbf{w}) = f(\mathbf{u}, \mathbf{v} + \mathbf{w})$  összefüggés szemléltetése.

**BIZONYÍTÁS.** Ha az  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{i-1}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n$  vektorok lineárisan összefüggők, akkor  $|\mathbf{A}| = |\mathbf{B}| = |\mathbf{C}| = 0$ , így a tétel állítása fönnáll. Ugyanez igaz akkor is, ha  $\mathbf{a}_i$  és  $\mathbf{b}_i$  is előáll a fenti vektorok lineáris kombinációjaként, mert akkor összegük is előáll, és így ismét mindhárom determináns 0. Feltesszük tehát, hogy  $\mathbf{a}_i$  és  $\mathbf{b}_i$  legalább egyike független a többi sorvektortól. Az általánosság megszorítása nélkül feltehető, hogy  $\mathbf{a}_i$  független a többi sorvektortól, vagyis  $\mathbf{A}$  sorvektorai függetlenek, így bázist alkotnak. Ekkor  $\mathbf{b}_i$  előáll lineáris kombinációjuként:

$$\mathbf{b}_i = b_1 \mathbf{a}_1 + b_2 \mathbf{a}_2 + \dots + b_i \mathbf{a}_i + \dots + b_n \mathbf{a}_n.$$

Vonjuk ki a  $|\mathbf{B}|$  determináns  $i$ -edik sorából a  $b_k \mathbf{a}_k$  sorokat, ahol  $k = 1, 2, \dots, i-1, i+1, \dots, n$ . E műveletek közben a determináns értéke nem változik, és az  $i$ -edik sorban csak a  $b_i \mathbf{a}_i$  vektor marad. Ezután emeljük ki az  $i$ -edik sorból a  $b_i$  konstans, kapjuk, hogy  $|\mathbf{B}| = b_i |\mathbf{A}|$ . A  $\mathbf{C}$  mátrixszal is megismételjük e műveleteket, csak ott a végén az  $i$ -edik sorban az  $(1 + b_i) \mathbf{a}_i$  vektor marad, így kapjuk, hogy  $|\mathbf{C}| = (1 + b_i) |\mathbf{A}|$ . Innen pedig

$$|\mathbf{A}| + |\mathbf{B}| = |\mathbf{A}| + b_i |\mathbf{A}| = (1 + b_i) |\mathbf{A}| = |\mathbf{C}|,$$

és ezzel kész a bizonyítás. □

► Az előbbi tételt használva a determinánsok második soraira, a determinánsok kiszámítása nélkül is látjuk, hogy

$$\begin{vmatrix} 2 & 2 & 2 \\ 1 & 2 & 3 \\ 9 & 8 & 6 \end{vmatrix} + \begin{vmatrix} 2 & 2 & 2 \\ 2 & 1 & 1 \\ 9 & 8 & 6 \end{vmatrix} = \begin{vmatrix} 2 & 2 & 2 \\ 3 & 3 & 4 \\ 9 & 8 & 6 \end{vmatrix}.$$

► A tétel indukciójak kettőnél több sorra is igazolható, így például a következő egyenlőség is igaz:

$$\begin{vmatrix} 2 & 2 & 2 \\ 1 & 2 & 3 \\ 9 & 8 & 6 \end{vmatrix} + \begin{vmatrix} 2 & 2 & 2 \\ 2 & 1 & 1 \\ 9 & 8 & 6 \end{vmatrix} + \begin{vmatrix} 2 & 2 & 2 \\ 1 & 1 & 2 \\ 9 & 8 & 6 \end{vmatrix} = \begin{vmatrix} 2 & 2 & 2 \\ 4 & 4 & 6 \\ 9 & 8 & 6 \end{vmatrix}.$$

► A tételbeli képletet fordított irányban is használni fogjuk, nevezetesen egy determináns fölbontható több determináns összegére. Például:

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 9 & 0 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 4 & 5 & 6 \\ 7 & 9 & 0 \end{vmatrix} + \begin{vmatrix} 0 & 2 & 0 \\ 4 & 5 & 6 \\ 7 & 9 & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 & 3 \\ 4 & 5 & 6 \\ 7 & 9 & 0 \end{vmatrix}.$$

mivel  $(1, 2, 3) = (1, 0, 0) + (0, 2, 0) + (0, 0, 3)$ .

► Mivel a transzponálás nem változtat a determináns értékén, e tétel sorvektorok helyett oszlopvektorokra is kimondható.

► E tétel a determináns definíciójának  $D1$ . feltételével együtt azt mondja, hogy a determináns olyan függvénye bármelyik sorának (a többi sor rögzítése mellett), mely megőrzi a lineáris kombinációt. Ezen azt értjük, hogy ha egy determináns  $i$ -edik sorvektora egyenlő a  $c\mathbf{x} + d\mathbf{y}$  vektorral, akkor a determináns felbontható a következő képlet szerint:

$$\begin{vmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{i-1} \\ c\mathbf{x} + d\mathbf{y} \\ \mathbf{a}_{i+1} \\ \vdots \\ \mathbf{a}_n \end{vmatrix} = c \begin{vmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{i-1} \\ \mathbf{x} \\ \mathbf{a}_{i+1} \\ \vdots \\ \mathbf{a}_n \end{vmatrix} + d \begin{vmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{i-1} \\ \mathbf{y} \\ \mathbf{a}_{i+1} \\ \vdots \\ \mathbf{a}_n \end{vmatrix}.$$

E tulajdonsággal rendelkező függvényeket *lineárisaknak* fogjuk nevezni, azokat a többváltozós függvényeket pedig, amelyek minden változójukban lineárisak, *multilineárisaknak*. Tehát a determináns, mint  $n$ -változós függvényt multilineáris, ugyanis bármely  $i$ -re ( $i = 1, 2, \dots, n$ ):

$$\begin{aligned} & \det(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, c\mathbf{x} + d\mathbf{y}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n) \\ &= c \det(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{x}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n) + d \det(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{y}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n). \end{aligned}$$

## Feladatok

6.1• Melyek igazak az alábbi állítások közül? (Az  $\mathbf{A}$  itt mindig négyzetes mátrixot jelöl.)

1. Ha egy determináns értéke 0, akkor van két azonos sora.
2. Ha egy determináns értéke nem 0, akkor oszlopvektorai lineárisan függetlenek.
3. Ha az  $\mathbf{Ax} = \mathbf{0}$  egyenletrendszernek van nemtriviális megoldása, akkor  $|\mathbf{A}| \neq 0$ .
4.  $|\mathbf{A}| \neq 0$  pontosan akkor igaz, ha az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszer nem oldható meg.
5.  $|\mathbf{A}| = 0$  pontosan akkor igaz, ha az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszer egyértelműen megoldható.

6.2♥ Számítsuk ki az alábbi determinánsok értékét fejben!

$$\begin{array}{lll} a) \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} & b) \begin{vmatrix} 0 & 0 & 0 \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{vmatrix} & c) \begin{vmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{vmatrix} \\ d) \begin{vmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{vmatrix} & e) \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix} & f) \begin{vmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 3 & 3 & 3 \end{vmatrix} \\ g) \begin{vmatrix} 1 & -2 & 1 & 2 \\ 4 & 3 & 4 & -1 \\ 5 & 1 & 5 & 4 \\ 6 & 5 & 6 & 0 \end{vmatrix} & & \end{array}$$

6.3♥ Mutassuk meg – lineáris összefüggőséget keresve a sorok közt –, hogy az alábbi determinánsok értéke 0.

$$\begin{array}{lll} a) \begin{vmatrix} 1 & 1 & 1 \\ 2 & 3 & 5 \\ 3 & 4 & 6 \end{vmatrix} & b) \begin{vmatrix} 1 & -2 & 3 \\ -2 & 4 & -6 \\ 3 & 6 & 9 \end{vmatrix} & c) \begin{vmatrix} 2 & 1 & 0 \\ 3 & 2 & 1 \\ 5 & 3 & 1 \end{vmatrix} \\ d) \begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{vmatrix} & e) \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} & f) \begin{vmatrix} 1 & 1 & -2 \\ 1 & -2 & 1 \\ -2 & 1 & 1 \end{vmatrix} \\ g) \begin{vmatrix} \sin \alpha & \cos \alpha & \sin(\alpha + \delta) \\ \sin \beta & \cos \beta & \sin(\beta + \delta) \\ \sin \gamma & \cos \gamma & \sin(\gamma + \delta) \end{vmatrix} & h) \begin{vmatrix} \ln 10 & \ln 4 & \ln 40 \\ \ln 5 & \ln 4 & \ln 20 \\ \ln 2 & 0 & \ln 2 \end{vmatrix} \end{array}$$

6.4♥ Fölhasználva, hogy

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = -2.$$

számítsuk ki az alábbi determinánsok értékét:

$$\begin{array}{ll} a) \begin{vmatrix} a & b & c \\ g & h & i \\ d & e & f \end{vmatrix} & b) \begin{vmatrix} a & b & c \\ 2d & 2e & 2f \\ g & h & i \end{vmatrix} \\ c) \begin{vmatrix} a & b & c \\ a+d & b+e & c+f \\ g & h & i \end{vmatrix} & d) \begin{vmatrix} a & d & g \\ b & e & h \\ c & f & i \end{vmatrix} \end{array}$$

$$\begin{array}{ll} e) \begin{vmatrix} a & 3b & c \\ d & 3e & f \\ g & 3h & i \end{vmatrix} & f) \begin{vmatrix} a & b & c+a \\ d & e & f+d \\ g & h & i+g \end{vmatrix} \\ g) \begin{vmatrix} 2a & 3b & c+a \\ 2d & 3e & f+d \\ 2g & 3h & i+g \end{vmatrix} & h) \begin{vmatrix} 2a & 2b & 2c \\ 3d & 3e & 3f \\ g+4a & h+4b & i+4c \end{vmatrix} \end{array}$$

6.5♥ Legyen  $\mathbf{A}$  és  $\mathbf{B}$  két  $3 \times 3$ -as mátrix, és legyen  $\det(\mathbf{A}) = 5$ ,  $\det(\mathbf{B}) = 4$ . Számítsuk ki a következő determinánsok értékét!

$$\begin{array}{lll} a) \det(\mathbf{A}^2) & b) \det(2\mathbf{A}) & c) \det((2\mathbf{A})^2) \\ d) \det(\mathbf{A}^{-1}) & e) \det(5\mathbf{A}^{-1}) & f) \det((5\mathbf{A})^{-1}) \\ g) \det(\mathbf{AB}^{-1}) & h) \det(\mathbf{A}^T \mathbf{B}) & i) |\mathbf{A}^{-1}| |\mathbf{B}^{-1} \mathbf{A}| |\mathbf{B}| \end{array}$$

6.6♥ Csak sorcserék segítségével hozzuk egyszerűbb alakra (például háromszög alakra) az alábbi determinánsokat, és így számítsuk ki értéküket:

$$\begin{array}{ll} a) \begin{vmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 1 & 0 & 0 \end{vmatrix} & b) \begin{vmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 4 & 0 \\ 5 & 0 & 0 & 0 & 0 \end{vmatrix} \\ c) \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}, \begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix}, \begin{vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{vmatrix} \\ d) \begin{vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 5 \\ 0 & 3 & 6 & 8 \\ 4 & 7 & 9 & 2 \end{vmatrix} & e) \begin{vmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 0 \\ 3 & 3 & 0 & 0 \\ 4 & 0 & 0 & 0 \end{vmatrix} \\ f) \begin{vmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 1 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{vmatrix} & g) \begin{vmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 1 & \dots & 1 & 1 \\ 1 & 1 & \dots & 1 & 1 \end{vmatrix} \end{array}$$

6.7♥ Számítsuk ki elemi sorműveletekkel az alábbi determinánsokat!

$$\begin{array}{ll} a) \begin{vmatrix} 1 & 2 & 3 \\ 1 & 3 & 5 \\ 1 & 3 & 6 \end{vmatrix} & b) \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 5 & 7 \\ 1 & 4 & 7 & 10 \end{vmatrix} \\ c) \begin{vmatrix} 3 & 8 & 6 & 3 \\ 1 & 2 & 0 & 1 \\ 1 & 1 & -1 & 2 \\ 2 & 5 & 1 & 5 \end{vmatrix} & d) \begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 0 & 0 & 0 & 4 \\ 3 & 0 & 1 & 0 & 3 \\ 4 & 0 & 0 & 0 & 2 \\ 5 & 4 & 3 & 2 & 1 \end{vmatrix} \end{array}$$

6.8. MELLÉKÁTLÓBAN EGYESEK Hány sor áll inverzióban abban a mátrixban, melynek mellékátlójában egyese, egyebütt nullák állnak, és mennyi ennek determinánsa?

6.9\* Számítsuk ki elemi sorműveletekkel az alábbi  $n$ -edrendű determinánsokat!

$$a) \begin{vmatrix} 1+x_1y_1 & 1+x_1y_2 & \dots & 1+x_1y_n \\ 1+x_2y_1 & 1+x_2y_2 & \dots & 1+x_2y_n \\ \vdots & \vdots & \ddots & \vdots \\ 1+x_ny_1 & 1+x_ny_2 & \dots & 1+x_ny_n \end{vmatrix}$$

$$b) \begin{vmatrix} 1 & a & a^2 & \dots & a^{n-1} \\ a^{n-1} & 1 & a & \dots & a^{n-2} \\ a^{n-2} & a^{n-1} & 1 & \dots & a^{n-3} \\ \vdots & \vdots & & \ddots & \vdots \\ a & a^2 & a^3 & \dots & 1 \end{vmatrix}$$

$$c) \begin{vmatrix} a & b & b & \dots & b \\ b & a & b & \dots & b \\ b & b & a & \dots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \dots & a \end{vmatrix} \quad d) \begin{vmatrix} a & b & b & \dots & b \\ c & a & b & \dots & b \\ c & c & a & \dots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c & c & c & \dots & a \end{vmatrix}$$

6.10. Számítsuk ki a Petersen-gráf szomszédsági mátrixának (ld. ?? feladat) determinánsát!

6.11. Írjuk fel a determináns definícióját oly módon, hogy  $\det$  egy  $n$ -változós,  $n$ -dimenziós vektorokon értelmezett skalár értékű függvény legyen.

6.12. Adjunk új bizonyítást a determinánsok szorzásszabályára azt igazolva, hogy az  $\mathbf{A} \mapsto \det(\mathbf{AB}) / \det(\mathbf{B})$  leképezés eleget tesz a determináns definíciójában kirótt feltételeknek.

6.13. Bizonyítsuk be az LU-felbontás fölhasználásával a transzponált determinánsára vonatkozó 6.12. állítást.

6.14. Fejezzük ki az elemi mátrixokra használt jelöléseket használva ( $\mathbf{E}_{S_i+cS_j}$ ,  $\mathbf{E}_{S_i \leftrightarrow S_j}$ ,  $\mathbf{E}_{cS_i}$ ) azok determinánsát!

6.15. Számítsuk ki a 6.7. példa Pascal-háromszöget tartalmazó

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{vmatrix}$$

determinánsának értékét úgy, hogy az első oszlop főátló alatti elemeinek kinullázásához először vonjuk ki az utolsó előtti sort az utolsóból, majd a második sort a harmadikból, végül az első a másodikból, és kövessük e módszert a többi főátló alatti elemre is.

6.16. Számítsuk ki az

$$\left| \binom{i+j-2}{j-1} \right|_{n \times n} = \begin{vmatrix} \binom{0}{0} & \binom{1}{1} & \dots & \binom{n-1}{n-1} \\ \binom{1}{0} & \binom{2}{1} & \dots & \binom{n}{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ \binom{n-1}{0} & \binom{n}{1} & \dots & \binom{2n-2}{n-1} \end{vmatrix}$$

determináns értékét! (Útmutatás: az utolsó sorral kezdve mindegyik sorból vonjuk ki az előzőt, majd mindegyik oszlopból az előzőt!)

6.17. Mutassuk meg, hogy egy legalább 3-adrendű determináns értéke 0, ha elemei sorfolytonosan olvasva számtani sorozatot adnak. Például

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 0.$$

6.18. Mutassuk meg, hogy egy legalább 3-adrendű determináns értéke 0, ha minden sora számtani sorozat, például

$$\begin{vmatrix} 1 & 2 & 3 \\ 1 & 4 & 7 \\ 1 & 3 & 5 \end{vmatrix} = 0.$$

6.19. Mutassuk meg, hogy ha egy determináns elemei sorfolytonosan olvasva mértani sorozatot alkotnak, akkor értéke 0. Például

$$\begin{vmatrix} \frac{1}{8} & \frac{1}{4} & \frac{1}{2} \\ 1 & 2 & 4 \\ 8 & 16 & 32 \end{vmatrix} = 0.$$

6.20. Mutassuk meg, hogy tetszőleges  $a, b, c$  és  $d$  valósokra

$$\begin{vmatrix} a^2 & (a+1)^2 & (a+2)^2 & (a+3)^2 \\ b^2 & (b+1)^2 & (b+2)^2 & (b+3)^2 \\ c^2 & (c+1)^2 & (c+2)^2 & (c+3)^2 \\ d^2 & (d+1)^2 & (d+2)^2 & (d+3)^2 \end{vmatrix} = 0.$$

6.21. Mutassuk meg, hogy ha  $\mathbf{C}$  invertálható, akkor  $\det(\mathbf{CAC}^{-1}) = \det(\mathbf{A})$  tetszőleges azonos méretű  $\mathbf{A}$  mátrixra fennáll.

6.22. VEKTOROK DETERMINÁNSA MÁSIK BÁZISBAN Igazoljuk, hogy ha  $\mathbf{A}_{C \leftarrow B}$  az áttérés mátrixa, akkor a  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  vektorok  $B$ - és  $C$ -beli koordinátas alakjaiból képzett  $\mathbf{V}_B$  és  $\mathbf{V}_C$  mátrixok determinánsára  $|\mathbf{V}_C| = |\mathbf{A}_{C \leftarrow B}| |\mathbf{V}_B|$ .

6.23. Igazoljuk, hogy páratlan rendű ferdén szimmetrikus mátrix determinánsa 0.

6.24. MÁTRIX NÉGYZETÉNEK DETERMINÁNSA Igazoljuk, hogy bármely négyzetes  $\mathbf{A}$  mátrixra  $|\mathbf{A}^2| = |\mathbf{A}\mathbf{A}^T|$ .

6.25. A determináns négyzetének kiszámításával (6.24. feladat) és a determinánsok szorzástételének alkalmazásával

számítsuk ki az alábbi determinánsok értékét:

$$\begin{vmatrix} a & b \\ -b & a \end{vmatrix}, \quad \begin{vmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{vmatrix},$$

$$\begin{vmatrix} a & b & c & d & e & f & g & h \\ b & -a & -d & c & f & -e & h & -g \\ c & d & -a & -b & g & -h & -e & f \\ d & -c & b & -a & h & g & -f & -e \\ e & -f & -g & -h & -a & b & c & d \\ f & e & h & -g & -b & -a & d & -c \\ g & -h & e & f & -c & -d & -a & b \\ h & g & -f & e & -d & c & -b & -a \end{vmatrix}.$$

6.26. Mutassuk meg, hogy az  $(x_1^2 + x_2^2)(y_1^2 + y_2^2)$  szorzat előállítható két szám négyzetének összegeként, azaz

$$(x_1^2 + x_2^2)(y_1^2 + y_2^2) = (z_1^2 + z_2^2),$$

ahol  $z_1$  és  $z_2$  mindegyike külön az  $x_i$  és külön az  $y_i$  változóknak is lineáris kifejezése ( $i = 1, 2$ ). (Hasonló össze-

függések bizonyíthatóak négy illetve nyolc négyzetszám összegéről is. Például a négy szám négyzetösszegére vonatkozó képlet

$$(x_1^2 + x_2^2 + x_3^2 + x_4^2)(y_1^2 + y_2^2 + y_3^2 + y_4^2) = (z_1^2 + z_2^2 + z_3^2 + z_4^2),$$

ahol  $z_i$  az  $x_i$  és az  $y_i$  ( $i = 1, 2, 3, 4$ ) változóknak lineáris. A megoldáshoz használjuk fel az előző feladat állítását.)

6.27. **TÉGLALAP KÉPÉNEK TERÜLETE** Legyen egy téglalap négy csúcsa  $(p, q)$ ,  $(p + x, q)$ ,  $(p, q + y)$ ,  $(p + x, q + y)$ , ahol  $x, y > 0$ . Tehát a téglalap oldalhossza  $x$  és  $y$ , területe  $xy$ . Mekkora lesz a területe annak a síkidomnak, mely e téglalapból keletkezik az  $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$  mátrixú lineáris transzformáció hatására.

6.28\* **ORIENTÁCIÓ** Azt mondjuk, hogy  $\mathbb{R}^n$  két bázisa azonos orientációjú, ha az egyiket a másikba vivő lineáris transzformáció determinánsa pozitív. Mutassuk meg, hogy  $\mathbb{R}^n$  bázisain az „azonos orientációjú” reláció ekvivalencia reláció, amely így az összes bázist két osztályba sorolja.

## A determináns mint elemeinek függvénye

A determinánst eddig sorvektorainak függvényeként kezeltük, a következőkben elemeinek függvényeként fogjuk. Ehhez két olyan módszerrel fogunk megismerkedni, melyekben a determináns kiszámítását egyszerűbb determinánsok kiszámítására vezetjük vissza.

Eddig nagyvonalúan bántunk a determináns elemeinek mibenlétével. Annyit feltételeztünk róluk kimondatlanul, hogy azonos algebrai struktúrából valók, és az összeadás, kivonás, szorzás és osztás elvégezhető köztük. Az elemi sorműveletek elvégzéséhez épp e négy műveletre volt szükség. E szakaszban ki fog derülni, hogy a determináns kiszámolható osztás nélkül is. Ennek következménye például, hogy egészelemű mátrix determinánsa egész szám, akkor is, ha számolás közben racionálisokba botlunk.<sup>1</sup>

**Kígyók determinánsa** A  $2 \times 2$ -es determináns kiszámítására ismerjük azt a formulát, amely a determináns értékét a determináns elemeinek függvényében írja fel:  $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$ . Itt tehát csak az összeadásra, kivonásra és a szorzásra van szükség. Hasonló formulát keresünk az  $n$ -edrendű determinánsokra. Ehhez az ún. *kígyókat* – más nevén *transzverzálisokat* – használjuk. A kígyók a diagonális mátrixok sorainak permutációjával származtatott mátrixok, azaz minden  $\mathbf{K}$  kígyó felírható  $\mathbf{K} = \mathbf{P} \text{diag}(a_1, a_2, \dots, a_n)$  alakban, ahol  $\mathbf{P}$  egy permutáló mátrix. Ezt a kígyóhoz tartozó permutáló mátrixnak fogjuk nevezni. Mivel  $\mathbf{P}$  determinánsa 1 vagy  $-1$ , ezért  $|\mathbf{K}| = a_1 a_2 \dots a_n$  vagy  $|\mathbf{K}| = -a_1 a_2 \dots a_n$ .

A determinánsok soronkénti linearitását használva érdekes felbontását kapjuk a determinánsnak. Tekintsük példaként az

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$$

determinánst. Első sorvektorának  $(a, b, c) = (a, 0, 0) + (0, b, 0) + (0, 0, c)$  felbontását fölhasználva bontsuk fel a determinánst három determináns összegére:

$$\begin{vmatrix} a+0+0 & 0+b+0 & 0+0+c \\ d & e & f \\ g & h & i \end{vmatrix} = \begin{vmatrix} a & 0 & 0 \\ d & e & f \\ g & h & i \end{vmatrix} + \begin{vmatrix} 0 & b & 0 \\ d & e & f \\ g & h & i \end{vmatrix} + \begin{vmatrix} 0 & 0 & c \\ d & e & f \\ g & h & i \end{vmatrix}$$

Ezután folytassuk e felbontást a második sorvektorral, így már az eredeti determinánst 9 determináns összegére bontottuk. Végül tegyük ugyanezt az utolsó sorral is. Az így kapott 27 determinánst nem írjuk föl, de szemléltetésül egy sematikus ábrán megmutatjuk a felbontás

<sup>1</sup> Általánosan fogalmazva: nem csak testek pl. a valós  $\mathbb{R}$ , a racionális  $\mathbb{Q}$ , a komplex  $\mathbb{C}$  számtestek vagy a véges  $\mathbb{F}_q$  testek elemeiből álló determinánst számolhatunk ki az adott struktúrán belül, hanem pl. az egészek  $\mathbb{Z}$  gyűrűjének vagy a polinomok gyűrűjének elemeiből képzett determinánsokat is. További részletekért lásd a függelék 13 szakaszát.

lépéseit (6.7 ábra). Tömör négyzet jelöli azokat a helyeket, ahol megtartjuk a determináns eredeti elemét, üres kör azokat, ahová zérust írunk. A 27 determináns mindegyikének minden sorában egy elem az eredeti determinánsból való, a többi zérus. Közöttük azonban csak 6 kígyó van. A többinek van zérus oszlopa, így azok értéke 0, vagyis az eredeti determinánst 6 kígyó összegére bontottuk (a 0 értékű determinánsokat szürke színnel jeleztük).

Hasonló módon bármely  $n$ -edrendű determináns fölbomlik  $n^n$  olyan determináns összegére, melynek minden sorában egyetlen elem az eredeti determinánsból való, a többi 0, de ezek közül csak azok lesznek kígyók determinánsai, melyek minden oszlopában is van egy elem az eredetiből. (Ezeket nevezzük a mátrixból/determinánsból kiválasztható kígyóknak.) Ezek száma  $n!$ , mert az első sorból  $n$ -féleképp választhatunk egy elemet, a második sorból minden esetben már csak  $n - 1$ -féleképp, ..., és ez összesen  $n(n - 1) \dots 3 \cdot 2 \cdot 1 = n!$  eset. Igaz tehát a következő állítás:

**6.15. ÁLLÍTÁS (FELBONTÁS KÍGYÓK DETERMINÁNSAINAK ÖSSZEGÉRE).** Minden  $n$ -edrendű determináns fölbomlik az összes belőle kiválasztható kígyó determinánsának összegére. Jelölje  $d_{j_1 j_2 \dots j_n}$  annak a permutáló mátrixnak a determinánsát, mely az  $a_{1j_1}, a_{2j_2}, \dots, a_{nj_n}$  elemekből álló kígyóhoz tartozik. Ekkor

$$\det([a_{ij}]) = \sum d_{j_1 j_2 \dots j_n} a_{1j_1} a_{2j_2} \dots a_{nj_n},$$

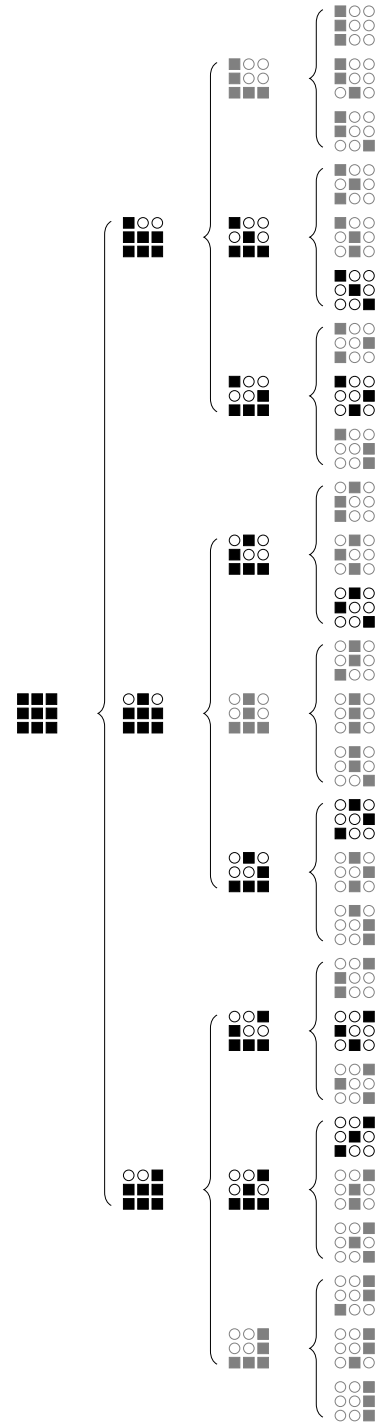
ahol az összegzés az  $\{1, 2, \dots, n\}$  halmaz összes lehetséges  $\{j_1, j_2, \dots, j_n\}$  permutációján végigfut.

Az  $n!$  az  $n$  növekedtével rendkívül gyorsan nő (pl.  $10! = 3628800$ ,  $20! = 2432902008176640000$ ), determináns ilyen módon való számítása viszonylag kis rend esetén már számítógéppel sem lehetséges emberi idő alatt. E felbontást a determinánsok tulajdonságainak vizsgálatában használjuk. Számításhoz csak az  $n = 2$  és  $n = 3$  esetekben használjuk, igaz, azokra gyakran.  $n = 2$  esetén az előző állítás szerint

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix} = ad - bc,$$

mivel a második determináns egyetlen sorcserével hozható diagonális alakra.  $n = 3$  esetén – felhasználva a 6.7 ábrát is – kapjuk, hogy

$$\begin{aligned} \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} &= \begin{vmatrix} a & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & i \end{vmatrix} + \begin{vmatrix} a & 0 & 0 \\ 0 & 0 & f \\ 0 & h & 0 \end{vmatrix} + \begin{vmatrix} 0 & b & 0 \\ d & 0 & 0 \\ 0 & 0 & i \end{vmatrix} + \begin{vmatrix} 0 & b & 0 \\ 0 & 0 & f \\ g & 0 & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 & c \\ d & 0 & 0 \\ 0 & h & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 & c \\ 0 & e & 0 \\ g & 0 & 0 \end{vmatrix} \\ &= aei - afh - bdi + bfg + cdh - ceg \\ &= aei + bfg + cdh - afh - bdi - ceg \end{aligned}$$



6.7. ábra: Egy  $3 \times 3$ -as determináns felbontása  $3^3$  determináns összegére, melyek közül  $3! = 6$  darabot kivéve mindegyikben van egy zérusoszlop – ezek sematikus ábráját szürke szín jelöli.



E két formula könnyen megjegyezhető egy egyszerű szabállyal, amelyet az  $n = 2$  és  $n = 3$  esetben *Sarrus-szabálynak* is neveznek: a főátló irányú szorzatok összegéből vonjuk ki a mellékátló irányú szorzatokat. (Hogy mit értsünk főátló és mellékátló irányú szorzaton, a 6.8. és a 6.9. ábrákról megérthető.) Fontos, hogy hasonló szabály  $n > 3$  esetén már *nem érvényes* (ld. a 6.33. feladatot)!

A determináns 6.15. tételbeli felbontása a determináns értékét a determináns elemeinek függvényeként állítja elő. Ennek sok szép és fontos következménye van.

**6.16. KÖVETKEZMÉNY (DETERMINÁNSFÜGGVÉNY LÉTEZÉSE).** *A determinánsfüggvény létezik, és egyértelmű.*

**BIZONYÍTÁS.** Mivel a definícióból következőleg minden determináns egyenlő a belőle kiválasztható kigyók determinánsainak összegével, és minden kigyó determinánsa létezik és egyértelmű, ezért minden négyzetes mátrix determinánsa létezik és egyértelmű.  $\square$

Íme további két fontos következménye a kigyókra bontásnak:

► Egy algebrai következmény: a determináns kiszámolásához elég csak az összeadás és szorzás művelete, az osztásra, melyet az elemi sorműveletek során használhatunk, nincs szükség. Eszerint egész számokból álló determináns értéke egész szám.

► Egy függvényanalízis körébe tartozó következmény: a determináns értéke folytonos függvénye elemeinek. Eszerint bármely kis pozitív  $\varepsilon$ -hoz van olyan  $\delta > 0$  szám, hogy ha a determináns bármely eleme legfeljebb  $\delta$  értékkel megváltozik, akkor a determináns értéke legfeljebb  $\varepsilon$ -nyit változik. Sőt, mivel a determináns kifejtésében csak az összeadás és a szorzás művelete szerepel, a determináns differenciálható függvénye elemeinek.

**Előjeles alldetermináns** Ha az első sor elemeit kiemeljük a  $3 \times 3$ -as determinánst kifejtő képletből, érdekes sejtést fogalmazhatunk meg:

$$\begin{aligned} \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} &= aei + bfg + cdh - afh - bdi - ceg \\ &= a(ei - fh) - b(fg - di) + c(dh - eg) \\ &= a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}. \end{aligned}$$

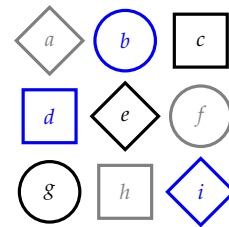
Mielőtt ezt megtennénk, némi előkészítés következik.

**6.17. DEFINÍCIÓ (ELŐJELES ALDETERMINÁNS).** *Az  $n$ -edrendű  $|\mathbf{A}|$  determináns  $i$ -edik sorának és  $j$ -edik oszlopának elhagyásával kapott  $(n - 1)$ -*

(a)  $ad - bc$

(b)  $aei + bfg + cdh - afh - bdi - ceg$

6.8. ábra: Az (a) másod- és a (b) harmadrendű determináns kiszámítása: a főátló irányú szorzatok összegéből vonjuk ki a mellékátló irányú szorzatokat. Harmadrendű esetben kezdetben könnyíthetünk magunknak a determináns első két oszlopának a determináns utáni megismétlésével.



6.9. ábra: A harmadrendű determináns kiszámítására egy – IQ-tesztek típuskérdésére emlékeztető – másik módszer: az egyforma alakúak szorzatának összegéből ki kell vonni az egyforma színűek szorzatait.

*edrendű determináns  $(-1)^{i+j}$ -szeresét az  $|\mathbf{A}|$  determináns  $a_{ij}$  eleméhez tartozó előjeles aldeterminánsának nevezzük.*

Az előjeles aldeterminánshoz kiszámítandó előjel a mátrixon sakk-táblaszerűen változik, azaz a bal felső sarokban  $+$ , és két egymás melletti vagy alatti mezőben ellenkező előjelű. Ezt nevezik *sakktáblaszabálynak*.

**6.18. PÉLDA (ELŐJELES ALDETERMINÁNS).** Számítsuk ki az

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 9 & 1 \\ 2 & 2 & 2 & 2 \\ 0 & 1 & 2 & 0 \end{vmatrix}$$

*determináns második sor harmadik eleméhez tartozó előjeles aldeterminánsát!*

**MEGOLDÁS.** A determináns második sorát és harmadik oszlopát kiemeltük

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 9 & 1 \\ 2 & 2 & 2 & 2 \\ 0 & 1 & 3 & 0 \end{vmatrix}$$

Az ezek elhagyása után megmaradó aldetermináns és  $-1$  megfelelő hatványának szorzata, vagyis a kért előjeles aldetermináns

$$(-1)^{2+3} \begin{vmatrix} 1 & 2 & 4 \\ 2 & 2 & 2 \\ 0 & 1 & 0 \end{vmatrix} = -1 \cdot 6 = -6.$$

Tehát a determináns második sor harmadik eleméhez tartozó előjeles aldeterminánsa  $-6$ .  $\square$

**6.19. ÁLLÍTÁS (DETERMINÁNS RENDJÉNEK CSÖKKENTÉSE).** *Tegyük fel, hogy az  $n$ -edrendű  $|\mathbf{A}|$  determináns  $a_{ij}$  elemének sorában vagy oszlopában minden további elem 0. Jelölje  $A_{ij}$  az  $a_{ij}$  elemhez tartozó előjeles aldeterminánst. Ekkor*

$$|\mathbf{A}| = a_{ij} A_{ij}.$$

**BIZONYÍTÁS.** Legyen az  $|\mathbf{A}|$  determináns  $i$ -edik sorában az  $a_{ij}$ -n kívül minden elem 0 (hasonlóan tárgyalható, ha a  $j$ -edik oszlopban vannak nullák). Cseréljük ki a  $j$ -edik oszlopot a  $(j-1)$ -edikkel, majd ezt a  $(j-2)$ -edikkel. ..., addig, míg az  $\mathbf{A}_*$  oszlop az első oszlopba nem kerül. Ez  $j-1$  oszlopcserét jelent, azaz a determináns értéke  $(-1)^{j-1}$ -szeresére változik. Ezután hasonlóképp vigyük az  $i$ -edik sort szomszédos sorok cseréjével az első sorba. Ehhez  $i-1$  csere szükséges,

|   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|
| + | - | + | - | + | - | + | - |
| - | + | - | + | - | + | - | + |
| + | - | + | - | + | - | + | - |
| - | + | - | + | - | + | - | + |
| + | - | + | - | + | - | + | - |
| - | + | - | + | - | + | - | + |
| + | - | + | - | + | - | + | - |
| - | + | - | + | - | + | - | + |

6.10. ábra: Sakktáblaszabály: a  $(-1)^{i+j}$  előjele a bal felső sarokban, vagyis az első sor első oszlopában  $+$ , él mentén szomszédos mezőkben pedig ellentétes.

miközben a determináns értéke  $(-1)^{i-1}$ -szeresére változik.

$$\begin{aligned}
 & \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & a_{ij} & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{nn} \end{vmatrix} = (-1)^{j-1} \begin{vmatrix} a_{1j} & a_{11} & a_{12} & \dots & a_{1n} \\ a_{2j} & a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{ij} & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ a_{nj} & a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \\
 & = (-1)^{i-1}(-1)^{j-1} \begin{vmatrix} a_{ij} & 0 & 0 & \dots & 0 \\ a_{1j} & a_{11} & a_{12} & \dots & a_{1n} \\ a_{2j} & a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{nj} & a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \\
 & \stackrel{*}{=} (-1)^{i+j} a_{ij} \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ a_{1j} & a_{11} & a_{12} & \dots & a_{1n} \\ a_{2j} & a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{nj} & a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \\
 & \stackrel{**}{=} (-1)^{i+j} a_{ij} \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \\
 & = a_{ij} A_{ij}.
 \end{aligned}$$

Az \*-os egyenlőségnél kihasználtuk, hogy  $i + j - 2$  és  $i + j$  paritása azonos, tehát  $-1$  kitevőjeként is azonos eredményt adnak, továbbá kiemeltük  $a_{ij}$ -t az első sorból. A \*\*-os egyenlőség előtt álló determináns kiszámításához csak a másodiktól lefelé lévő sorokat kell használni, a végeredményt az első oszlop elemei nem befolyásolják, így az első sor és első oszlop elhagyásával kapott determináns értéke ugyanaz. Végül az így kapott determináns az előjellel együtt épp  $A_{ij}$ , és ezzel bizonyítottuk az állítást.  $\square$

**6.20. PÉLDA (DETERMINÁNS RENDJÉNEK CSÖKKENTÉSE).** A determináns rendjének csökkentésével számítsuk ki az alábbi determináns értékét!

$$\begin{vmatrix} 1 & 2 & 0 & 3 & 4 \\ 1 & 2 & 0 & 8 & 4 \\ 6 & 0 & 0 & 7 & 0 \\ 8 & 9 & 8 & 7 & 6 \\ 5 & 4 & 0 & 3 & 2 \end{vmatrix}.$$

**MEGOLDÁS.** Minden lépésben – esetleg egy apró átalakítás után – találunk egy sort vagy oszlopot, melyben csak egy nemnulla szám áll, így a determináns könnyen számolható:

$$\begin{vmatrix} 1 & 2 & 0 & 3 & 4 \\ 1 & 2 & 0 & 8 & 4 \\ 6 & 0 & 0 & 7 & 0 \\ 8 & 9 & \boxed{8} & 7 & 6 \\ 5 & 4 & 0 & 3 & 2 \end{vmatrix} = (-1)^{4+3} \cdot 8 \begin{vmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 8 & 4 \\ 6 & 0 & 7 & 0 \\ 5 & 4 & 3 & 2 \end{vmatrix}$$

$$\begin{aligned}
 (S_2 - S_1) &= (-8) \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & \boxed{5} & 0 \\ 6 & 0 & 7 & 0 \\ 5 & 4 & 3 & 2 \end{vmatrix} \\
 &= (-8) \cdot (-1)^{2+3} \cdot 5 \begin{vmatrix} 1 & 2 & 4 \\ \boxed{6} & 0 & 0 \\ 5 & 4 & 2 \end{vmatrix} \\
 &= (-8) \cdot (-5) \cdot (-1)^{2+1} \cdot 6 \begin{vmatrix} 2 & 4 \\ 4 & 2 \end{vmatrix} \\
 &= (-8) \cdot (-5) \cdot (-6) \cdot (-12) \\
 &= 2880.
 \end{aligned}$$

□

**Determináns kifejtése** Ritkán adódik, hogy a determináns rendje az előző (6.19.) állítás segítségével csökkenthető, viszont fölhasználásával a determinánsok egy gyönyörű kifejtési tételét kapjuk.

**6.21. TÉTEL (DETERMINÁNSOK KIFEJTÉSI TÉTELE).** Egy determináns értéke megkapható úgy, hogy egy tetszőleges sorának vagy oszlopának minden elemét beszorozzuk a hozzá tartozó előjeles aldeterminánssal, és e szorzatokat összeadjuk. Képletben, az  $n$ -edrendű  $|\mathbf{A}|$  determináns értéke  $i$ -edik sora szerint kifejtve

$$|\mathbf{A}| = \sum_{k=1}^n a_{ik} A_{ik},$$

és  $j$ -edik oszlopa szerint kifejtve

$$|\mathbf{A}| = \sum_{k=1}^n a_{kj} A_{kj}.$$

E kifejtési tételt egyes könyvek *Laplace-féle kifejtési tételnek* nevezik, míg más könyvek csak ennek egy – a feladatok közt megtalálható – általánosítását hívják így, sok könyv pedig e tételbeli összefüggéssel definiálja a determinánst.

**BIZONYÍTÁS.** Hasonlóan a korábbiakban látottakhoz, az  $i$ -edik sorvektor felbontásával a determinánst  $n$  olyan determináns összegére bontjuk, amelyek  $i$ -edik sorában csak egy elem származik az eredeti determinánsból, a többi 0. Az egyszerűség kedvéért e felbontást csak  $n = 3$  és  $i = 2$  esetére írjuk fel, de tetszőleges  $n$ -re ugyanígy megy. Ezután

a 6.19. állítást alkalmazzuk mindegyik új determinánusra:

$$\begin{aligned}
 |\mathbf{A}| &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & 0 & 0 \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\
 &= a_{21}A_{21} + a_{22}A_{22} + a_{23}A_{23} \\
 &= \sum_{k=1}^3 a_{2k}A_{2k}.
 \end{aligned}$$

A bizonyítás ugyanígy megy az oszlopokra is, amit példaként az  $n = 3, j = 3$  esettel szemléltetünk:

$$\begin{aligned}
 |\mathbf{A}| &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & 0 \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & 0 \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\
 &= a_{13}A_{13} + a_{23}A_{23} + a_{33}A_{33} \\
 &= \sum_{k=1}^3 a_{k3}A_{k3}. \quad \square
 \end{aligned}$$

**6.22. PÉLDA (KIFEJTÉSI TÉTEL).** Számítsuk ki az alábbi determináns értékét a kifejtési tételt használva!

$$\begin{vmatrix} 3 & 2 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 2 \end{vmatrix}.$$

**MEGOLDÁS.** Érdemes e determinánst a harmadik oszlopa szerint kifejtetni, mert ott két 0 is van, így a velük megszorozott aldeterminánsokat le sem kell írni.

$$\begin{vmatrix} 3 & 2 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 2 \end{vmatrix} = 1 \cdot \begin{vmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{vmatrix} - 1 \cdot \begin{vmatrix} 3 & 2 & 2 \\ 2 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = 1 - 0 = 1. \quad \square$$

**Vandermonde-determináns** Bemutatunk egy fontos determinánst. Számítalan alkalmazása van, melyek egyike a polinominterpoláció.

**6.23. PÉLDA (INTERPOLÁCIÓ MÁSODFOKÚ POLINOMOKRA).** Legyen  $x, y$  és  $z$  három különböző valós,  $a, b$  és  $c$  három tetszőleges valós. Mutassuk meg, hogy egyetlen olyan legfőbb másodfokú  $f$  polinom létezik, melyre  $f(x) = a, f(y) = b$  és  $f(z) = c$ .

**MEGOLDÁS.** Legyen  $f : x \mapsto p + qx + rx^2$ , ahol  $p, q$  és  $r$  a polinom ismeretlen együtthatói. Az  $f(x) = a$ ,  $f(y) = b$  és  $f(z) = c$  egyenlőségek a következő egyenletrendszerre vezetnek:

$$\begin{bmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{bmatrix} \begin{bmatrix} p \\ q \\ r \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Ez az egyenletrendszer a 6.5. tétel szerint pontosan akkor oldható meg egyértelműen, ha az együtthatómátrix determinánsa nem 0. Oszlop-műveletekkel kezdjük az átalakítást:

$$\begin{aligned} & \begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix} \stackrel{O_3 - xO_2}{=} \begin{vmatrix} 1 & x & 0 \\ 1 & y & y^2 - xy \\ 1 & z & z^2 - xz \end{vmatrix} \stackrel{O_2 - xO_1}{=} \begin{vmatrix} 1 & 0 & 0 \\ 1 & y - x & y^2 - xy \\ 1 & z - x & z^2 - xz \end{vmatrix} \\ &= \begin{vmatrix} y - x & y^2 - xy \\ z - x & z^2 - xz \end{vmatrix} = (y - x) \begin{vmatrix} 1 & y \\ z - x & z^2 - xz \end{vmatrix} = (y - x)(z - x) \begin{vmatrix} 1 & y \\ 1 & z \end{vmatrix} \\ &= (y - x)(z - x)(z - y) \end{aligned}$$

Mivel  $x, y$  és  $z$  három különböző valós, ezért a determináns értéke nem 0, tehát az egyenletrendszer egyértelműen megoldható, vagyis egyetlen olyan polinom létezik, mely a feltételeket teljesíti.  $\square$

E probléma, és a benne szereplő determináns általánosítása a következő definícióhoz vezet:

**6.24. DEFINÍCIÓ (VANDERMONDE-DETERMINÁNS).** Az  $x_1, x_2, \dots, x_n$  számokhoz tartozó Vandermonde-determináns  $a$

$$V_n(x_1, x_2, \dots, x_n) = \begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix} \quad (6.1)$$

determinánst vagy ennek

$$\begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{vmatrix}$$

transzponáltját értjük. A hozzá tartozó mátrixot Vandermonde-mátrixnak nevezzük.

Mivel egy determináns értéke megegyezik transzponáltjának értékével, ezért a definícióbeli két determináns értéke is azonos, így mindig melyik alakot használjuk.

6.25. TÉTEL (VANDERMONDE-DETERMINÁNS ÉRTÉKE). Az  $x_1, x_2, \dots, x_n$  ( $n > 1$ ) számokhoz tartozó Vandermonde-determináns értéke megegyezik az olyan  $(x_j - x_i)$  alakú különbségek szorzatával, ahol  $i < j$ , azaz

$$V_n(x_1, x_2, \dots, x_n) = \prod_{i < j} (x_j - x_i).$$

**BIZONYÍTÁS.** A determináns utolsó oszlopával kezdve minden oszlop-ból vonjuk ki az előző oszlop  $x_1$ -szeresét.

$$\begin{aligned} V_n(x_1, x_2, \dots, x_n) &= \begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{vmatrix} \\ &= \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & x_2 - x_1 & x_2^2 - x_1x_2 & \dots & x_2^{n-1} - x_1x_2^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n - x_1 & x_n^2 - x_1x_n & \dots & x_n^{n-1} - x_1x_n^{n-2} \end{vmatrix} \end{aligned}$$

ami az első sora szerinti kifejtés, majd minden sorból kiemelve az első oszlopbeli elemet, a következő alakra vezet:

$$\begin{aligned} &= \begin{vmatrix} x_2 - x_1 & x_2^2 - x_1x_2 & \dots & x_2^{n-1} - x_1x_2^{n-2} \\ \vdots & \vdots & & \vdots \\ x_n - x_1 & x_n^2 - x_1x_n & \dots & x_n^{n-1} - x_1x_n^{n-2} \end{vmatrix} \\ &= (x_2 - x_1)(x_3 - x_1) \dots (x_n - x_1) \begin{vmatrix} 1 & x_2 & x_2^2 & \dots & x_2^{n-2} \\ 1 & x_3 & x_3^2 & \dots & x_3^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-2} \end{vmatrix} \\ &= (x_2 - x_1)(x_3 - x_1) \dots (x_n - x_1) V_{n-1}(x_2, \dots, x_n) \\ &= \prod_{j>1} (x_j - x_1) \cdot V_{n-1}(x_2, \dots, x_n). \end{aligned}$$

Eredményül egy rekurzív képletet kaptunk, melyet önmagába helyettesítve, és a  $V_2(x_{n-1}, x_n) = x_n - x_{n-1}$  képletet is fölhasználva a tételbeli összefüggésre jutunk.  $\square$

**Cramer-szabály és a mátrix inverze** Eddig akár az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszer megoldására, akár az  $\mathbf{A}$  mátrix inverzének kiszámítására olyan módszert használtunk, mely az elemi sorműveletek használatával csak egy algoritmust ad a számításokra, de nem adja meg a kapcsolatot (képletet) az adatok és a kiszámítandók közt. E paragrafusban ezt pótoljuk!

Jelölje  $\mathbf{A}_{i,\mathbf{b}}$  azt a mátrixot, melyet akkor kapunk, ha az  $\mathbf{A}$  mátrix  $i$ -edik oszlopának helyére a  $\mathbf{b}$  vektort írjuk. Kifejtve

$$\mathbf{A}_{i,\mathbf{b}} = [\mathbf{a}_{*1} \ \dots \ \mathbf{a}_{*,i-1} \ \mathbf{b} \ \mathbf{a}_{*,i+1} \ \dots \ \mathbf{a}_{*n}].$$

E jelöléssel  $\mathbf{I}_{i,\mathbf{x}}$  mátrixon az  $[\mathbf{e}_{*1} \ \dots \ \mathbf{e}_{*,i-1} \ \mathbf{x} \ \mathbf{e}_{*,i+1} \ \dots \ \mathbf{e}_{*n}]$  mátrixot értjük.

**6.26. TÉTEL (CRAMER-SZABÁLY).** Legyen  $\mathbf{A}$  egy  $n \times n$ -es mátrix. Az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszer pontosan akkor oldható meg egyértelműen, ha  $\det \mathbf{A} \neq 0$ . Ekkor a megoldás:

$$x_i = \frac{\det \mathbf{A}_{i,\mathbf{b}}}{\det \mathbf{A}}, \quad (i = 1, 2, \dots, n)$$

**BIZONYÍTÁS.** Az állítás első felét már bizonyítottuk a 6.5. tételben. Eből felhasználjuk, hogy mivel az egyenletrendszer megoldható,  $\det \mathbf{A} \neq 0$ . Kihasználva, hogy  $\mathbf{Ax} = \mathbf{b}$ , továbbá hogy  $\mathbf{Ae}_i = \mathbf{a}_{*i}$ , kapjuk, hogy

$$\begin{aligned} \mathbf{AI}_{i,\mathbf{x}} &= \mathbf{A}[\mathbf{e}_{*1} \ \dots \ \mathbf{e}_{*,i-1} \ \mathbf{x} \ \mathbf{e}_{*,i+1} \ \dots \ \mathbf{e}_{*n}] \\ &= [\mathbf{Ae}_{*1} \ \dots \ \mathbf{Ae}_{*,i-1} \ \mathbf{Ax} \ \mathbf{Ae}_{*,i+1} \ \dots \ \mathbf{Ae}_{*n}] \\ &= [\mathbf{a}_{*1} \ \dots \ \mathbf{a}_{*,i-1} \ \mathbf{b} \ \mathbf{a}_{*,i+1} \ \dots \ \mathbf{a}_{*n}] \\ &= \mathbf{A}_{i,\mathbf{b}} \end{aligned}$$

Mivel az  $\mathbf{I}_{i,\mathbf{x}}$  mátrix  $i$ -edik sorának és oszlopának elhagyása után egy identikus mátrix marad, ezért az  $i$ -edik sora szerint kifejtve

$$\det \mathbf{I}_{i,\mathbf{x}} = \begin{vmatrix} 1 & 0 & \dots & x_1 & \dots & 0 \\ 0 & 1 & \dots & x_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \dots & x_i & \dots & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & x_n & \dots & 1 \end{vmatrix} = (-1)^{i+i} x_i = x_i.$$

Így a determinánsok szorzási szabályát is használva  $\det(\mathbf{AI}_{i,\mathbf{x}}) = \det \mathbf{A}_{i,\mathbf{b}}$ , amiből  $x_i \det \mathbf{A} = \det \mathbf{A}_{i,\mathbf{b}}$ , azaz  $x_i = \det \mathbf{A}_{i,\mathbf{b}} / \det \mathbf{A}$ .  $\square$

**6.27. PÉLDA (CRAMER-SZABÁLY).** Oldjuk meg az

$$2x + 5y = 4$$

$$5x + 3y = 6$$

egyenletrendszert a Cramer-szabállyal!

**MEGOLDÁS.** A kiszámolandó determinánsok a  $\mathbf{b} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$  jelöléssel:

$$\mathbf{A} = \begin{vmatrix} 2 & 5 \\ 5 & 3 \end{vmatrix} = -19, \quad \mathbf{A}_{1,\mathbf{b}} = \begin{vmatrix} 4 & 5 \\ 6 & 3 \end{vmatrix} = -18, \quad \mathbf{A}_{2,\mathbf{b}} = \begin{vmatrix} 2 & 4 \\ 5 & 6 \end{vmatrix} = -8.$$

Gabriel Cramer (1704–1752) genfi születésű svájci matematikus, akinek az algebrai görbékről szóló „Introduction à l’analyse des lignes courbes algébriques” című, 1750-ben publikált munkájában szerepelt a ma Cramer-szabály néven ismert tétel. A szabályt korábban már mások is ismerték.



Innen  $x = \frac{-18}{-19} = \frac{18}{19}$ ,  $y = \frac{-8}{-19} = \frac{8}{19}$ .  $\square$

Ha egyenletrendszert meg tudunk oldani, akkor szimultán egyenletrendszert is, és így pl. az  $\mathbf{AX} = \mathbf{I}$  megoldásával a mátrix inverzét is ki tudjuk számítani. Az  $x_{ij}$  elem kiszámításához az  $\mathbf{Ax}_{*j} = \mathbf{e}_j$  egyenletrendszert kell megoldani. A megoldás  $i$ -edik koordinátája az  $x_{ij}$  elem. A Cramer-szabály szerint

$$x_{ij} = \frac{\det \mathbf{A}_{i,\mathbf{e}_j}}{\det \mathbf{A}}$$

Mivel az  $\mathbf{A}_{i,\mathbf{e}_j}$  mátrix  $i$ -edik oszlopában csak egy elem nem 0, a kifejtési tétel szerint

$$\det \mathbf{A}_{i,\mathbf{e}_j} = \begin{vmatrix} a_{11} & a_{12} & \dots & 0 & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & 0 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \dots & 1 & \dots & a_{jn} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & 0 & \dots & a_{nn} \end{vmatrix} = A_{ji},$$

vagyis a determináns megegyezik az  $\mathbf{A}$  egy előjeles aldeterminánsával, tehát

$$x_{ij} = \frac{\det \mathbf{A}_{i,\mathbf{e}_j}}{\det \mathbf{A}} = \frac{\det \mathbf{A}_{ji}}{\det \mathbf{A}}.$$

Mint látjuk, az  $\mathbf{X} = \mathbf{A}^{-1}$  előállításához az  $\mathbf{A}$  előjeles aldeterminánsai mátrixának transzponáltjára van szükség. E mátrixot az  $\mathbf{A}$  klasszikus adjungáltjának nevezzük és  $\text{adj}(\mathbf{A})$ -val jelöljük. A klasszikus jelzőre azért van szükség, mert az adjungált szót komplex elemű mátrix konjugált transzponáltjára is használjuk, és ez félreértésekhez vezethet. Képletben tehát

$$\text{adj } \mathbf{A} = [A_{ij}]^T = [A_{ji}]. \quad (6.2)$$

Így a következő tételt kapjuk:

**6.28. TÉTEL (MÁTRIX INVERZÉNEK ELEMEI).** *Tegyünk fel, hogy  $\mathbf{A}$  egy invertálható mátrix. Ekkor inverzének  $ij$  indexű eleme az  $a_{ji}$  elemhez tartozó előjeles aldetermináns és az  $\mathbf{A}$  mátrix determinánsának hányadosa, azaz*

$$[\mathbf{A}^{-1}]_{ij} = \frac{A_{ji}}{\det \mathbf{A}}.$$

Így az inverz mátrix az

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} [\mathbf{A}_{ij}]^T = \frac{1}{\det \mathbf{A}} \text{adj } \mathbf{A}. \quad (6.3)$$

alakba írható.

► Könnyen ellenőrizhető, hogy az  $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  mátrix klasszikus adjungáltja

$$\begin{bmatrix} d & -c \\ -b & a \end{bmatrix}^T = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix},$$

így inverze

$$\mathbf{A}^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{\det \mathbf{A}} \operatorname{adj} \mathbf{A} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

► A mátrix inverzének e kifejezése azt mutatja, hogy az inverz mátrix minden eleme folytonos függvénye a mátrix minden elemének minden olyan helyen, ahol a determináns nem 0, azaz minden olyan helyen, ahol az inverz egyáltalán létezik.

► Az előző megjegyzésből az is következik, hogy egy  $n$ -ismeretlenes  $n$  egyenletből álló egyenletrendszer megoldásvektorának minden koordinátája folytonos függvénye az egyenletrendszer együtthatóinak és a jobb oldalán álló vektor koordinátáinak, hisz a megoldás az inverzzel való szorzással megkapható.

► Egészelemű mátrix inverze pontosan akkor egészelemű, ha determinánsa 1 vagy  $-1$ . Ez abból adódik, hogy  $\det(\mathbf{A}) \det(\mathbf{A}^{-1}) = \det \mathbf{I} = 1$ , tehát ha  $|\det \mathbf{A}| \neq 1$ , akkor  $\det(\mathbf{A}^{-1})$  nem egész szám, tehát  $\mathbf{A}^{-1}$  nem lehet egészelemű, ha pedig  $|\det \mathbf{A}| = 1$ , akkor a (6.3) képlet szerint  $\mathbf{A}^{-1}$  minden eleme egész szám.

► A tételbeli képlet könnyen kiterjeszthető szinguláris mátrixokra is, vagyis amikor a determináns 0, ugyanis

$$\mathbf{A} \operatorname{adj} \mathbf{A} = \det(\mathbf{A}) \mathbf{I} \quad (6.4)$$

minden négyzetes mátrixra fennáll (ld. a 6.44. feladatot).

6.29. PÉLDA (MÁTRIX INVERZE). Számítsuk ki a szemléltetés céljából csupa különböző elemet tartalmazó

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 5 & 6 \\ 4 & 7 & 9 \end{bmatrix}$$

mátrix inverzét!

**MEGOLDÁS.** Az  $\operatorname{adj} \mathbf{A}$  determinánst olyan alakba írjuk föl, ahonnan látszik minden elem kiszámításának módja. Szürke színnel szedjük az

elhagyandó elemeket:

$$\begin{aligned}
 \operatorname{adj} \mathbf{A} &= \begin{bmatrix} + \begin{vmatrix} 0 & 1 & 2 \\ 3 & 5 & 6 \\ 4 & 7 & 9 \end{vmatrix} & - \begin{vmatrix} 0 & 1 & 2 \\ 3 & 5 & 6 \\ 4 & 7 & 9 \end{vmatrix} & + \begin{vmatrix} 0 & 1 & 2 \\ 3 & 5 & 6 \\ 4 & 7 & 9 \end{vmatrix} \\ - \begin{vmatrix} 0 & 1 & 2 \\ 3 & 5 & 6 \\ 4 & 7 & 9 \end{vmatrix} & + \begin{vmatrix} 0 & 1 & 2 \\ 3 & 5 & 6 \\ 4 & 7 & 9 \end{vmatrix} & - \begin{vmatrix} 0 & 1 & 2 \\ 3 & 5 & 6 \\ 4 & 7 & 9 \end{vmatrix} \\ + \begin{vmatrix} 0 & 1 & 2 \\ 3 & 5 & 6 \\ 4 & 7 & 9 \end{vmatrix} & - \begin{vmatrix} 0 & 1 & 2 \\ 3 & 5 & 6 \\ 4 & 7 & 9 \end{vmatrix} & + \begin{vmatrix} 0 & 1 & 2 \\ 3 & 5 & 6 \\ 4 & 7 & 9 \end{vmatrix} \end{bmatrix}^T \\
 &= \begin{bmatrix} + \begin{vmatrix} 5 & 6 \\ 7 & 9 \end{vmatrix} & - \begin{vmatrix} 3 & 6 \\ 4 & 9 \end{vmatrix} & + \begin{vmatrix} 3 & 5 \\ 4 & 7 \end{vmatrix} \\ - \begin{vmatrix} 1 & 2 \\ 7 & 9 \end{vmatrix} & + \begin{vmatrix} 0 & 2 \\ 4 & 9 \end{vmatrix} & - \begin{vmatrix} 0 & 1 \\ 4 & 7 \end{vmatrix} \\ + \begin{vmatrix} 1 & 2 \\ 5 & 6 \end{vmatrix} & - \begin{vmatrix} 0 & 2 \\ 3 & 6 \end{vmatrix} & + \begin{vmatrix} 0 & 1 \\ 3 & 5 \end{vmatrix} \end{bmatrix}^T \\
 &= \begin{bmatrix} 3 & -3 & 1 \\ 5 & -8 & 4 \\ -4 & 6 & -3 \end{bmatrix}^T
 \end{aligned}$$

Mivel  $\det \mathbf{A} = -1$ , ezért

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \operatorname{adj} \mathbf{A} = - \begin{bmatrix} 3 & -3 & 1 \\ 5 & -8 & 4 \\ -4 & 6 & -3 \end{bmatrix}^T = \begin{bmatrix} -3 & -5 & 4 \\ 3 & 8 & -6 \\ -1 & -4 & 3 \end{bmatrix} \quad \square$$

Már ezekből az egyszerű példákból is látszik, hogy mátrix invertálása e módszerrel igen műveletigényes. Valóban, gyakorlati számításokhoz nem használjuk, elméleti okfejtésekben vesszük nagy hasznát.

**Blokkmátrixok determinánsa\*** Az  $\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}$  mátrix általában még négyzetes részmátrixok esetén sem számítható az  $\mathbf{AD} - \mathbf{BC}$  képlettel (ld. a ?? feladatban)! Először egy speciális, de fontos esettel kezdjük.

6.30. TÉTEL (DETERMINÁNSOK SZORZATA BLOKKMÁTRIXBAN). *Legyenek  $\mathbf{A}$  és  $\mathbf{D}$  négyzetes mátrixok. Ekkor*

$$\begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{D} \end{vmatrix} = \begin{vmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{C} & \mathbf{D} \end{vmatrix} = |\mathbf{A}| |\mathbf{D}|.$$

**BIZONYÍTÁS.** Megmutatjuk, hogy minden olyan kígyó, melynek nincs eleme a  $\mathbf{O}$ -mátrixból, egy  $\mathbf{A}$ -beli és egy  $\mathbf{D}$ -beli kígyó szorzata. Ehhez

elég megmutatni, hogy ha egy kígyónak van eleme a  $\mathbf{B}$ , illetve a  $\mathbf{C}$  mátrixból, akkor az  $\mathbf{O}$ -ból is. Valóban, ha pl.  $\mathbf{B}$  egy eleme benne van egy kígyóban, akkor oszlopában nincs elem  $\mathbf{D}$ -ben, így  $\mathbf{D}$ -ben marad egy sor is üresen, amelyet csak egy  $\mathbf{O}$ -beli elem foghat le. Ellenőrizni kell még, hogy az  $\mathbf{A}$ - és  $\mathbf{D}$ -beli kígyók előjeleinek szorzata megegyezik-e az egyesítésükkel kapott kígyó előjelével. Ez nyilván igaz, hisz egy  $\mathbf{A}$ -t és egy  $\mathbf{D}$ -t metsző sor nem lehet inverzióban, így az egyesített kígyó inverzióinak száma megegyezik a két kígyó inverzióinak összegével, az előjelet pedig a  $-1$ -nek az inverziók számára emelt hatványa adja.  $\square$

6.31. TÉTEL ( $2 \times 2$ -ES BLOKKMÁTRIX DETERMINÁNSA). *Legyen*

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix},$$

ahol  $\mathbf{A}$  és  $\mathbf{D}$  négyzetes mátrixok.

1. Ha  $|\mathbf{A}| \neq 0$ , akkor  $|\mathbf{M}| = |\mathbf{A}| |\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}|$ .
2. Ha  $|\mathbf{D}| \neq 0$ , akkor  $|\mathbf{M}| = |\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}| |\mathbf{D}|$ .

**BIZONYÍTÁS.** Ha  $\mathbf{A}$  invertálható, akkor  $\mathbf{M}$  alábbi alsó és felső blokkháromszögmátrix szorzatára való bontása segít:

$$\begin{aligned} \mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} &= \begin{bmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{C} & \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{O} & \mathbf{I} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{C}\mathbf{A}^{-1} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{O} & \mathbf{I} \end{bmatrix} \end{aligned}$$

Az utóbbi három mátrix közül a szélsők determinánsa 1, a középsőé pedig a bizonyítandó kifejezés. Az

$$\mathbf{M} = \begin{bmatrix} \mathbf{I} & \mathbf{B}\mathbf{D}^{-1} \\ \mathbf{O} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C} & \mathbf{O} \\ \mathbf{O} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{D}^{-1}\mathbf{C} & \mathbf{I} \end{bmatrix}$$

felbontás bizonyítja a második összefüggést.  $\square$

## Feladatok

6.29. Melyek igazak az alábbi állítások közül? (Az  $A$  itt mindig négyzetes mátrixot jelöl.)

1. A determináns folytonos függvénye minden elemének.
2. A determináns differenciálható függvénye minden elemének.
3. Ha egy determináns minden eleme racionális szám, akkor értéke is racionális.
4. Ha egy determináns minden sorában és minden oszlopában pontosan egy elem nem 0, akkor a determináns értéke nem 0.
5. Ha egy mátrix két kígyó összege, akkor determinánsa is két kígyó determinánsának összege.
6. Ha  $i + j$  páratlan szám, akkor az előjeles  $A_{ij}$  aldetermináns negatív.
7. Ha egy determináns minden eleme pozitív, akkor értéke nem lehet negatív.
8. Mátrix inverze folytonos függvénye minden elemének.

### Felbontás kígyók determinánsainak összegére

6.30. Válasszuk ki az alábbi determinánsokból az összes nemnulla determinánsú kígyót, és ezek segítségével számítsuk ki a determináns értékét!

$$\begin{aligned} a) & \begin{vmatrix} 0 & 1 & 0 \\ 2 & 3 & 4 \\ 5 & 0 & 6 \end{vmatrix} \\ b) & \begin{vmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 2 & 0 \\ 0 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{vmatrix} \\ c) & \begin{vmatrix} 1 & 0 & 2 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 & 0 \\ 2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 2 \end{vmatrix} \end{aligned}$$

6.31. A determináns értékének kiszámítása nélkül mutassuk meg, hogy osztható 30-cal:

$$\begin{vmatrix} 24 & 40 & 68 \\ 27 & 15 & 31 \\ 51 & 55 & 53 \end{vmatrix}$$

6.32. Az alábbi – lottótípekből álló – determináns elemeinek csak a paritását vizsgálva minden számolás nélkül igazoljuk, hogy

$$\begin{vmatrix} 12 & 25 & 28 & 44 & 56 \\ 21 & 34 & 54 & 68 & 80 \\ 10 & 40 & 52 & 69 & 72 \\ 24 & 36 & 53 & 56 & 84 \\ 18 & 24 & 28 & 58 & 87 \end{vmatrix} \neq 0.$$

6.33. A 4-edrendű determinánsok  $4! = 24$  kígyó determinánsának összegére bonthatók. Soroljuk fel közülük azt a 12 darabot, melyet elemei szorzata után  $-1$ -gyel kell szorozni! (A Sarrus-szabály 4-edrendű determinánsra csak 8 kígyóból állna, ezért nem használható!)

### Kifejtési tétel

6.34. Tudjuk, hogy 504, 747 és 855 egyaránt oszthatók 9-cel. Ezt fölhasználva, a determináns értékének kiszámítása nélkül mutassuk meg, hogy az alábbi determináns osztható 9-cel:

$$\begin{vmatrix} 5 & 0 & 4 \\ 7 & 4 & 7 \\ 8 & 5 & 5 \end{vmatrix}.$$

6.35. Konstruáljunk olyan nemnulla értékű determinánst, melynek van olyan eleme, amelyet tetszőlegesen változtatva a determináns értéke nem változik.

### Blokkdeterminánsok

6.36. Számítsuk ki az alábbi determinánsok értékét kihasználva blokkstruktúrájukat!

$$\begin{aligned} a) & \begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & 3 & 3 & 3 \end{vmatrix} \\ b) & \begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 3 & 3 & 0 & 0 \\ 0 & 2 & 2 & 2 & 0 \\ 0 & 1 & 1 & 1 & 1 \end{vmatrix} \end{aligned}$$

### Speciális mátrixok determinánsa

6.37. Számítsuk ki az alábbi determinánsok értékét!

$$\begin{aligned} a) & \begin{vmatrix} 1 & 1 & 1 & 1 \\ 2 & -1 & -2 & 1 \\ 4 & 1 & 4 & 1 \\ 8 & -1 & -8 & 1 \end{vmatrix} \\ b) & \begin{vmatrix} 1 & -3 & 9 & -27 & 81 \\ 1 & 2 & 4 & 8 & 16 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & -2 & 4 & -8 & 16 \\ 1 & -1 & 1 & -1 & 1 \end{vmatrix} \\ c) & \begin{vmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d+e & d^2+e^2 & d^3+e^3 \end{vmatrix} \end{aligned}$$

6.38. Bizonyítsuk be, hogy

$$\begin{aligned} D &= \begin{vmatrix} p^2 & p & 1 & qrs \\ q^2 & q & 1 & prs \\ r^2 & r & 1 & pqs \\ s^2 & s & 1 & pqr \end{vmatrix} \\ &= (p-q)(p-r)(p-s)(q-r)(q-s)(r-s). \end{aligned}$$

6.39. Igazoljuk, hogy az  $a_1 = 1$ ,  $a_2 = 2$ ,  $a_n = a_{n-1} + a_{n-2}$  képletekkel definiált Fibonacci-sorozat  $n$ -edik eleme egyen-

lő az alábbi  $n \times n$ -es tridiagonális determinánssal:

$$a_n = \begin{vmatrix} 1 & -1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 1 & -1 & 0 & \dots & 0 & 0 \\ 0 & 1 & 1 & -1 & \dots & 0 & 0 \\ 0 & 0 & 1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & -1 \\ 0 & 0 & 0 & 0 & \dots & 1 & 1 \end{vmatrix}$$

6.40. Legyen

$$P_n = \begin{vmatrix} a_n & -1 & 0 & 0 & \dots & 0 & 0 \\ 1 & a_{n-1} & -1 & 0 & \dots & 0 & 0 \\ 0 & 1 & a_{n-2} & -1 & \dots & 0 & 0 \\ 0 & 0 & 1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & a_2 & -1 \\ 0 & 0 & 0 & 0 & \dots & 1 & a_1 \end{vmatrix}$$

Mutassuk meg, hogy

$$\frac{P_k}{P_{k-1}} = a_k + \frac{1}{a_{k-1} + \frac{1}{a_{k-2} + \frac{1}{\ddots + \frac{1}{a_2 + \frac{1}{a_1}}}}}$$

### Vegyes feladatok

6.41. Elérhető-e egyetlen elem megváltoztatásával, hogy egy tetszőleges  $n \times n$ -es nem szinguláris mátrix determinánsa 0-vá váljon?

6.42. **FERDE KIFEJTÉS** Vegyük egy determináns egy sorának elemeit, és szorozzuk meg mindegyiket egy másik sor azonos oszlopbeli eleméhez tartozó előjeles aldeterminánsával, majd képezzük ezek összegét. Ez mindig 0. Hasonló állítás igaz a determináns minden oszlopára is. Tehát az  $i$ -edik és  $u$ -adik sorra ( $i \neq u$ ) és a  $j$ -edik és  $v$ -edik oszlopára ( $j \neq v$ ):

$$\sum_{k=1}^n a_{ik} A_{uk} = 0, \quad \sum_{k=1}^n a_{kj} A_{kv} = 0.$$

6.43. Foglalkozunk egyetlen állításba a kifejtési és a ferde kifejtési tételekkel!

6.44. **MÁTRIX INVERZE A KIFEJTÉSI TÉTELEKKEL** A kifejtési és a ferde kifejtési (ld. az előző és a 6.42. feladatokat) segítségével adjunk új bizonyítást a **mátrix inverzére vonatkozó (6.3) formulára!**

6.45. Legyen  $x_0, x_1, x_2, \dots, x_n$   $n+1$  darab különböző valós,  $y_0, y_1, \dots, y_n$  ugyanannyi tetszőleges valós. Mutassuk meg, hogy egyetlen olyan legfőbb  $n$ -edfokú  $p$  polinom van, melyre  $p(x_i) = y_i$  minden  $i = 0, \dots, n$  esetén.

### Cramer-szabály és mátrix inverze

6.46♥ Oldjuk meg Cramer-szabállyal az alábbi egyenlet-rendszereket!

$$\begin{array}{ll} a) \begin{cases} x + y = 1 \\ x - 2y = 4 \end{cases} & b) \begin{cases} 2x - y - z = 4 \\ 3x + 4y - 2z = 11 \\ 3x - 2y + 4z = 11 \end{cases} \\ c) \begin{cases} x + 2y + 4z = 31 \\ 5x + y + 2z = 29 \\ 3x - y + z = 10 \end{cases} & d) \begin{cases} x + y = 1 \\ x + 2y + z = 2 \\ y + 2z + w = 3 \\ z + 2w = 4 \end{cases} \end{array}$$

6.47♥ Határozzuk meg a megadott mátrixok inverzének megadott indexű elemét!

$$a) \begin{bmatrix} 1 & 4 & 7 \\ 2 & 4 & 6 \\ 3 & 2 & 3 \end{bmatrix}, a_{23} = ? \quad b) \begin{bmatrix} 1 & 2 & 5 & 7 \\ 0 & 1 & 3 & 6 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix} a_{24} = ?$$

6.48♥ Határozzuk meg a megadott mátrixok inverzét a klasszikus adjungált kiszámolásával:

$$\begin{array}{ll} a) \begin{bmatrix} 3 & 1 & 4 \\ -7 & 2 & 7 \\ 2 & 1 & 4 \end{bmatrix} & b) \begin{bmatrix} 1 & 2 & 3 \\ 2 & 0 & 2 \\ 3 & 2 & 1 \end{bmatrix} \\ c) \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix} & d) \begin{bmatrix} 0 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{bmatrix} \\ e) \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} (abc \neq 0) & f) \begin{bmatrix} 1+i & i \\ i & i \end{bmatrix} \\ g) \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 3 & 0 \\ 4 & 0 & 0 & 0 \end{bmatrix} & h) \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

6.49. Igazoljuk, hogy tetszőleges négyzetes mátrixra  $\mathbf{A} \operatorname{adj}(\mathbf{A}) = \det(\mathbf{A}) \mathbf{I}$ .

### Véges testek fölötti mátrixok determinánsa\*

6.50. A determináns kiszámításának megismert technikái véges testek fölött is működnek. Számítsuk ki az alábbi – a megadott test fölött értelmezett – mátrixok determinánsát!

$$a) \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \mathbb{F}_2, \mathbb{F}_3, \mathbb{F}_5$$

$$b) \begin{bmatrix} 3 & 2 & 3 \\ 5 & 7 & 6 \\ 2 & 7 & 2 \end{bmatrix}, \mathbb{F}_{11}$$

**6.51. VÉLETLEN BITMÁTRIX DETERMINÁNSA** Számítsuk ki  $\mathbb{F}_2$  fölötti véletlen mátrixok determinánsát! Egy  $\mathbb{F}_2^{5 \times 5}$ -beli mátrix determinánsa mekkora valószínűséggel 0? Kísérletezzünk számítógéppel, majd válaszoljuk meg a kérdést pontosan.

*Projekt: a vektori szorzás általánosítása*

**6.52.** Bizonyított tény, hogy nem lehet olyan bináris vektorműveletet definiálni az  $n$ -dimenziós tér vektorain ( $n > 3$ ), mely eredményül ugyanannak a térnek egy vektorát adja és rendelkezik a vektori szorzás műveleti tulajdonságaival. E feladatsorban egy másik irányú általánosítást dolgozunk fel, mely nem a bináris műveleti tulajdonságokat, hanem az eredménynek a vektorokra való merőlegességét tarja meg.

a) Fogalmazzuk meg, hogy mit kapunk eredményül, ha a vektori szorzásra vonatkozó formális

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ \mathbf{i} & \mathbf{j} & \mathbf{k} \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & \mathbf{i} \\ a_2 & b_2 & \mathbf{j} \\ a_3 & b_3 & \mathbf{k} \end{vmatrix}$$

összefüggést  $2 \times 2$ -es vagy  $4 \times 4$ -es formális determinánssokra írjuk föl, vagyis mit ad eredményül az

$$\begin{vmatrix} a_1 & a_2 \\ \mathbf{i} & \mathbf{j} \end{vmatrix} \quad \text{és az} \quad \begin{vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 \end{vmatrix}$$

kifejezés?

b) Igazoljuk, hogy az  $n$ -dimenziós

$$\mathbf{a}_1 = (a_{11}, a_{12}, \dots, a_{1n}),$$

$$\mathbf{a}_2 = (a_{21}, a_{22}, \dots, a_{2n}),$$

$$\vdots$$

$$\mathbf{a}_{n-1} = (a_{n-1,1}, a_{n-1,2}, \dots, a_{n-1,n})$$

vektorok által kifeszített  $n - 1$ -dimenziós paralelepipedon térfogata megegyezik az

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n-1,1} & a_{n-1,2} & \dots & a_{n-1,n} \\ \mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_n \end{vmatrix}$$

vektor abszolút értékével.

c) Ha a fentiek alapján általánosított képlettel  $n - 1$  darab  $n$ -dimenziós vektorhoz egy  $n$ -ediket rendelünk, akkor mit mondhatunk az így kapott  $n$  vektor körüljárásáról?

**6.53.♥** Határozzuk meg azt a vektort, mely merőleges az  $(1, 1, 1, 1)$ ,  $(1, 2, 2, 2)$ ,  $(1, 2, 3, 3)$  vektorokra, hossza megegyezik a három vektor által kifeszített paralelepipedon térfogatával, és e három vektor mellé negyedeknek véve velük jobbrendszert alkot.

## Megoldások

6.1. 1. Hamis. 2. Igaz. 3. Hamis. 4. Hamis. Az  $|\mathbf{A}| \neq 0$  azzal ekvivalens, hogy az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszer nem oldható meg egyértelműen, vagyis vagy nem oldható meg, vagy több megoldása is van. 5. Hamis.

6.2.

- a)  $-2$ .
- b)  $0$ , mert van két azonos sora.
- c)  $0$ , mert a második sor az első konstansszorosa.
- d)  $1$ , mert háromszögmátrix determinánsa a főátlóbeli elemek szorzata.
- e)  $6$ , mert háromszögmátrix determinánsa a főátlóbeli elemek szorzata.
- f)  $0$ , mert van két azonos oszlopa.

6.3.

- a) a második sor az első  $-1$ -szerese.
- b) a harmadik sor egyenlő az első kettő összegével.
- c) a harmadik sor egyenlő az első kettő összegével.
- d) a második sor az első és a harmadik számtani közepe (másként: a harmadik sorból kivonva a másodikat, majd a másodikból az első, mindkétszer az  $(1, 1, 1)$  vektort kapjuk, azaz így van két azonos sor).
- e) a második sor az első és a harmadik számtani közepe.
- f) a három sorvektor összege a zérusvektor.
- g)  $\sin(\xi + \delta) = \sin \xi \cos \delta + \cos \xi \sin \delta$ , így a harmadik oszlop az első és a második oszlop lineáris kombinációja, vagyis az oszlopvektorok lineárisan összefüggőek, tehát a determináns értéke  $0$ .
- h) Az első és második oszlop összege a harmadik oszlop (ill. az első és a második sor különbsége a harmadik sor), tehát az oszlopvektorok (ill. sorvektorok) lineárisan összefüggőek.

6.5. a)  $125$ , b)  $40$ , c)  $1600$ , d)  $1/5$ , e)  $25$ , f)  $1/625$ , g)  $5/4$ , h)  $20$ , i)  $1$ .

6.6.

$$a) \begin{vmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 1 & 0 & 0 \end{vmatrix} = - \begin{vmatrix} 0 & 2 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{vmatrix} = 6$$

b) Az 1. és 2., azután az 1. és 3., végül az 1. és 5. sorokat

felcserélve:

$$\begin{vmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 4 & 0 \\ 5 & 0 & 0 & 0 & 0 \end{vmatrix} = - \begin{vmatrix} 0 & 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 4 & 0 \\ 5 & 0 & 0 & 0 & 0 \end{vmatrix} =$$

$$\begin{vmatrix} 0 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 5 & 0 & 0 & 0 & 0 \end{vmatrix} = - \begin{vmatrix} 5 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{vmatrix} = -120.$$

- c)  $-1, -1, 1$ .
- d)  $24$ .
- e)  $24$ .
- f) Az első sort cseréljük fel az utolsóval, a másodikat az utolsó elöttivel, ..., így  $\lfloor \frac{n}{2} \rfloor$  sorcserét hajtottunk végre, tehát a determináns értéke  $(-1)^{\lfloor n/2 \rfloor}$  (itt  $\lfloor \cdot \rfloor$  az egészrész-függvényt jelöli). Ennek értéke  $1$ , ha  $n = 4k$  vagy  $n = 4k + 1$ , és  $-1$ , ha  $n = 4k + 2$  vagy  $n = 4k + 3$  valamilyen  $k$  természetes számra. Más alakban kapjuk meg az eredményt, ha csak szomszédos sorokat cserélünk: először az első sort visszük (szomszédos sorok cseréjével) az utolsóba, majd az eredeti determináns második sorát az utolsó elöttibe, ..., azaz az alábbi sorpárok cseréjét hajtjuk végre:  
 $(1, 2), (2, 3), (3, 4), \dots, (n-1, n),$   
 $(1, 2), (2, 3), \dots, (n-2, n-1),$   
 $\dots$   
 $(1, 2), (2, 3),$   
 $(1, 2).$   
Ez összesen  $(n-1) + (n-2) + \dots + 2 + 1 = \frac{n(n-1)}{2}$  sorcseré. Minden sorcserével  $(-1)$ -szeresére változik a determináns értéke, így a végeredmény  $(-1)^{n(n-1)/2}$ . Természetesen e hatvány értéke is akkor  $1$ , ha  $n = 4k$  vagy  $4k + 1$ , és akkor  $-1$ , ha  $n = 4k + 2$  vagy  $4k + 3$ . (Ugyanilyen gondolatmenettel kimutatható, hogy ha egy determináns mellékátlója felett csupa  $0$  áll, akkor a determináns értéke a mellékátlóbeli elemek szorzatának  $(-1)^{\lfloor n/2 \rfloor}$ -szerese vagy más alakban  $(-1)^{n(n-1)/2}$ -szerese.)
- g) ld. az előző pontot.

6.7.

- a) Az első sort kivonjuk a másodikból és a harmadikból, majd a másodikat a harmadikból:

$$\begin{vmatrix} 1 & 2 & 3 \\ 1 & 3 & 5 \\ 1 & 3 & 6 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 1 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{vmatrix} = 1.$$



- b) Az első sort kivonva a többi sorból, majd a második sor kétszeresét kivonva a harmadikból, kapjuk, hogy

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 2 & 4 & 6 \\ 0 & 3 & 6 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 6 & 9 \end{vmatrix} = 0.$$

$$\begin{aligned} c) \quad & \begin{vmatrix} 3 & 8 & 6 & 3 \\ 1 & 2 & 0 & 1 \\ 1 & 1 & -1 & 2 \\ 2 & 5 & 1 & 5 \end{vmatrix} = - \begin{vmatrix} 1 & 2 & 0 & 1 \\ 3 & 8 & 6 & 3 \\ 1 & 1 & -1 & 2 \\ 2 & 5 & 1 & 5 \end{vmatrix} = \\ & - \begin{vmatrix} 1 & 2 & 0 & 1 \\ 0 & 2 & 6 & 0 \\ 0 & -1 & -1 & 1 \\ 0 & 1 & 1 & 3 \end{vmatrix} = -2 \begin{vmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 3 & 0 \\ 0 & -1 & -1 & 1 \\ 0 & 1 & 1 & 3 \end{vmatrix} = \\ & -2 \begin{vmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & -2 & 3 \end{vmatrix} = -2 \begin{vmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 4 \end{vmatrix} = -16. \end{aligned}$$

tezzük a megoldás lépéseit:

1. lépés: Cseréljük ki az első és második sort, hogy az első sor első eleme 1 legyen, s így ne kelljen törtekkel számolni. A determináns értéke  $(-1)$ -szeresére változik.
2. lépés: Az első sor  $(-3)$ -,  $(-1)$ - ill.  $(-2)$ -szeresét adjuk a második, harmadik ill. negyedik sorhoz.
3. lépés: Hogy a második sor második eleme is 1 legyen, emeljük ki 2-t a második sorból.
4. lépés: A második sort ill.  $(-1)$ -szeresét adjuk a harmadik ill. negyedik sorhoz.
5. lépés: Adjuk a harmadik sort a negyedikhez. A determináns értéke  $-16$ .

d) 144.

6.8. E mátrixban bármely két sor inverzióban áll egymással, így ha a sorok száma  $n$ , a sorpároké  $n(n-1)/2$ . Eszerint e mátrix determinánsa  $(-1)^{n(n-1)/2}$ . (Az egység mátrix  $\lfloor \frac{n}{2} \rfloor$  sorcserével is megkapható e mátrixból, így determinánsát  $(-1)^{\lfloor \frac{n}{2} \rfloor}$  alakban is ki lehet fejezni, ld. még a 6.6. feladatban).

6.9.

- a)  $n = 1$  esetén  $1 + x_1 y_1$ ,  $n = 2$  esetén  $x_1 y_1 + x_2 y_2 - x_1 y_2 - x_2 y_1$  a determináns értéke. Ha  $n \geq 3$ , akkor a determináns értéke 0. Ezt úgy bizonyítjuk, hogy először a determinánst két determináns összegére bontjuk, majd mindkettőről belátjuk, hogy értékük 0. Az első determináns csupa 1-esből álló első sorát kivonjuk az összes többi sorból, az így kapott determináns értéke pedig valóban 0, hisz ha  $x_2 = 0$ , akkor a második sor csupa 0-ból áll, ha pedig  $x_2 \neq 0$ , akkor a második sorának  $x_3/x_2$ -szerese egyenlő a harmadik sorral. A második

determináns értéke is 0, hiszen ha  $x_1 = 0$ , akkor az első sor csupa 0-ból áll, ha pedig  $x_1 \neq 0$ , akkor az első sor  $x_i/x_1$ -szeresét kivonva az  $i$ -edik sorból egy olyan determinánst kapunk, amelyben a második sortól kezdve minden sor 1-esekből áll, tehát a determinánsnak van két azonos sora.

$$\begin{aligned} & \begin{vmatrix} 1 & 1 & \dots & 1 \\ 1+x_2 y_1 & 1+x_2 y_2 & \dots & 1+x_2 y_n \\ \vdots & \vdots & \ddots & \vdots \\ 1+x_n y_1 & 1+x_n y_2 & \dots & 1+x_n y_n \end{vmatrix} \\ & + \begin{vmatrix} x_1 y_1 & x_1 y_2 & \dots & x_1 y_n \\ 1+x_2 y_1 & 1+x_2 y_2 & \dots & 1+x_2 y_n \\ \vdots & \vdots & \ddots & \vdots \\ 1+x_n y_1 & 1+x_n y_2 & \dots & 1+x_n y_n \end{vmatrix} \\ & = \begin{vmatrix} 1 & 1 & \dots & 1 \\ x_2 y_1 & x_2 y_2 & \dots & x_2 y_n \\ \vdots & \vdots & \ddots & \vdots \\ x_n y_1 & x_n y_2 & \dots & x_n y_n \end{vmatrix} \\ & + \begin{vmatrix} x_1 y_1 & x_1 y_2 & \dots & x_1 y_n \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{vmatrix} \\ & = 0. \end{aligned}$$

- b)  $(1-a^n)^{n-1}$ . Vonjuk ki az első sor  $a^{n-1}$ -szeresét a második sorból,  $a^{n-2}$ -szeresét a harmadik sorból, ...,  $a$ -szorosát az utolsó sorból: így a főátló alatt csak nullák lesznek, a determináns értéke.
- c)  $(a+(n-1)b)(a-b)^{n-1}$ . Első megoldás: adjunk minden sort az elsőhöz, emeljük ki a közös  $a+(n-1)b$  értéket, majd e sor  $b$ -szeresét vonjuk ki minden sorból. Másik megoldás: az utolsó sorral kezdve mindegyik sorból vonjuk ki a fölötte lévőt, majd jobbról kezdve mindegyik oszlopot adjuk a megelőzőhöz.

6.10. Az eredmény 48. Megoldás Sage-ben:

```
g = graphs.PetersenGraph()
G = matrix(g)
G.det()
```

6.13. Az  $A$  előáll PLU alakban, ahol  $P$  permutáló mátrix,  $L$  alsó,  $U$  felső háromszögmátrix. Az  $L$  és az  $U$  háromszögmátrixok, így determinánsuk megegyezik transzponáltjuk determinánsával, hisz a főátlóbeli elemek helyben maradnak a transzponálás során. A  $P$  permutáló mátrix determinánsa 1 vagy  $-1$ , transzponáltja pedig megegyezik inverzével, így  $\det(I) = \det(PP^T) = \det(P) \det(P^T) = 1$ , azaz  $P$  és  $P^{-1}$  egyszerre 1 vagy  $-1$ , tehát megegyeznek. Végül  $\det(A) = \det(PLU) = \det(P) \det(L) \det(U)$ , és  $\det(A^T) =$

$\det((\mathbf{PLU})^T) = \det(\mathbf{U}^T \mathbf{L}^T \mathbf{P}^T) = \det(\mathbf{U}) \det(\mathbf{L}) \det(\mathbf{P})$  összevetése bizonyítja az állítást.

6.14.  $\det(\mathbf{E}_{S_i+cS_j}) = 1$ ,  $\det(\mathbf{E}_{S_i \leftrightarrow S_j}) = -1$ ,  $\det(\mathbf{E}_{cS_i}) = c$ .

6.15.

$$\begin{aligned} & \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{vmatrix} \xrightarrow{S_4-S_3} \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 0 & 1 & 4 & 10 \end{vmatrix} \\ & \xrightarrow{S_3-S_2} \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 0 & 1 & 3 & 6 \\ 0 & 1 & 4 & 10 \end{vmatrix} \xrightarrow{S_2-S_1} \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 3 & 6 \\ 0 & 1 & 4 & 10 \end{vmatrix} \\ & \xrightarrow{\substack{S_4-S_3 \\ S_3-S_2}} \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 1 & 4 \end{vmatrix} \xrightarrow{S_4-S_3} \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 1. \end{aligned}$$

Ld. még a 6.16. feladatot!

6.16. Felhasználva, hogy  $\binom{n}{k} - \binom{n-1}{k} = \binom{n-1}{k-1}$ , elvégezve az ajánlott sor-, majd oszlopműveleteket, majd azt megismételve az egyre kisebb bal alsó részdeterminánssal kapjuk, hogy

$$\begin{aligned} D &= \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & \binom{1}{0} & \binom{2}{0} & \dots & \binom{n-1}{0} \\ 1 & \binom{2}{1} & \binom{3}{1} & \dots & \binom{n}{1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \binom{n-1}{n-2} & \binom{n}{n-2} & \dots & \binom{2n-3}{n-2} \end{vmatrix} \\ &= \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & \binom{0}{0} & \binom{1}{0} & \dots & \binom{n-2}{0} \\ 0 & \binom{1}{1} & \binom{2}{1} & \dots & \binom{n-1}{1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \binom{n-2}{n-2} & \binom{n-1}{n-2} & \dots & \binom{2n-4}{n-2} \end{vmatrix} \\ &= \dots = \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \binom{0}{0} \end{vmatrix} = 1. \end{aligned}$$

6.17. Az első sort kivonva a másodikból és a harmadikból két konstans sort kapunk, melyek egymás konstansszorosai, tehát a determináns értéke 0.

6.18. Vonjuk ki az első oszlopot a másodikból és a harmadikból. Az így kapott harmadik oszlop kétszerese a másodiknak, tehát a determináns értéke 0.

6.20. Első megoldás: vonjuk ki az első oszlopot a többiből, ezzel eltüntetve azokból a négyzetes tagot, majd vonjuk

a második oszlop megfelelő skalárszorását a harmadik és negyedik oszlopból, hogy elimináljuk azok lineáris tagját, végül a harmadik oszlop konstansszorosát vonjuk ki a negyedikből, hogy ott csak 0-k maradjanak.

Második megoldás: Elég megmutatnunk, hogy a determináns oszlopai lineárisan összefüggőek. Az  $a^2x + (a+1)^2y + (a+2)^2z + (a+3)^2w = 0$  egyenlet a homogén

$$x + y + z + w = 0$$

$$2y + 4z + 6w = 0$$

$$y + 4z + 9w = 0$$

egyenletrendszerre vezet, aminek biztosan van nemtriviális megoldása, hisz 4 ismeretlenhez csak 3 egyenlet van adva. (Megoldani már szükségtelen, elég a megoldás létezését igazolni, de például az  $(x, y, z, w) = (1, -3, 3, -1)$  egy megoldás).

6.22. Mivel a koordináták báziscserében való változásáról szóló 4.25. állításban láttuk, hogy a koordinátas alakokat a  $[\mathbf{v}_i]_C = \mathbf{A}_{C \leftarrow B}[\mathbf{v}_i]_B$  képlet kapcsolja össze, ezért a  $\mathbf{v}$  vektorok koordinátas alakjaiból, mint oszlopvektorokból képzett mátrixokra  $\mathbf{V}_C = \mathbf{A}_{C \leftarrow B} \mathbf{V}_B$ , így determinánsaikra  $|\mathbf{V}_C| = |\mathbf{A}_{C \leftarrow B}| |\mathbf{V}_B|$ .

6.23. Egyrészt  $\det(\mathbf{A}) = \det(\mathbf{A}^T)$ , másrészt mivel  $\mathbf{A}^T = -\mathbf{A}$ , ezért  $\det(\mathbf{A}^T) = (-1)^n \det(\mathbf{A})$ , azaz  $\det(\mathbf{A}) = -\det(\mathbf{A})$ , amiből  $\det(\mathbf{A}) = 0$ .

6.24.  $|\mathbf{A}^2| = |\mathbf{A}|^2 = |\mathbf{A}| |\mathbf{A}| = |\mathbf{A}| |\mathbf{A}^T| = |\mathbf{A} \mathbf{A}^T|$ .

6.25. Mindhárom determinánst a következőképpen számíthatjuk ki. Legyen  $\mathbf{A}$  a determinánshoz tartozó mátrix. Tekintsük az  $|\mathbf{A} \mathbf{A}^T|$  determinánst. Ezt könnyű kiszámítani (hisz a főátlón kívül csak nullák állnak), s ennek négyzetgyöke lesz a determináns értéke. Ezek alapján a három determináns értéke:  $a^2 + b^2$ ,  $(a^2 + b^2 + c^2 + d^2)^2$ ,  $(a^2 + b^2 + c^2 + d^2 + e^2 + f^2 + g^2 + h^2)^4$ .

6.26. A determinánsok szorzási szabályát is felhasználva:

$$\begin{aligned} (x_1^2 + x_2^2)(y_1^2 + y_2^2) &= \begin{vmatrix} x_1 & x_2 \\ -x_2 & x_1 \end{vmatrix} \begin{vmatrix} y_1 & y_2 \\ -y_2 & y_1 \end{vmatrix} \\ &= \begin{vmatrix} x_1 y_1 - x_2 y_2 & x_1 y_2 + x_2 y_1 \\ -x_2 y_1 - x_1 y_2 & -x_2 y_2 + x_1 y_1 \end{vmatrix} \\ &= (x_1 y_1 - x_2 y_2)^2 + (x_1 y_2 + x_2 y_1)^2. \end{aligned}$$

A négy illetve a nyolc négyzet összegére vonatkozó analóg összefüggések hasonlóan bizonyíthatóak. (Hurwitz bebizonyította, hogy ha  $n$  négyzetszám összegére igaz a feladatbelivel analóg összefüggés, akkor  $n = 1, 2, 4$  vagy 8.)

6.27. Minthogy lineáris transzformáció alteret alterbe, eltolt alteret eltolt alterbe visz, e téglalap képe egy (esetleg elfajuló) téglalap lesz. Ezért elég kiszámolni csak a téglalap

4 csúcsának képét. Ez kiszámolható egyetlen mátrixszorzással:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} p & p+x & p & p+x \\ q & q & q+y & q+y \end{bmatrix} = \begin{bmatrix} ap+bq & ap+ax+bq & ap+bq+by & ap+ax+bq+by \\ cp+dq & cp+cx+dq & cp+dq+dy & cp+cx+dq+dy \end{bmatrix}$$

Innen leolvasható, hogy a téglalap képeként kapott parallelogramma oldalvektorai  $(ax, cx)$  és  $(by, dy)$ , és így területe

$$|(ax)(dy) - (cx)(by)| = |ad - bc|xy.$$

Eszerint tehát a téglalap képének területe független a téglalap helyzetétől, és mindig a téglalap területének  $|ad - bc|$ -szerese.

6.29. 1. Igaz. 2. Igaz. 3. Igaz. 4. Igaz. 5. Hamis. Mátrixok összegének determinánsa általában nem egyenlő determinánsaik összegével (ld. a 6.30. feladatot). 6. Hamis. Egy aldetermináns értéke bármilyen előjelű lehet, az előjeles aldeterminánst belőle úgy kapjuk, hogy páratlan  $i + j$  esetén megszorozzuk  $-1$ -gyel. 7. Hamis. 8. Hamis. Csak azokon a helyeken folytonos függvénye a mátrix elemeinek, ahol a determinánsa nem 0.

6.30.

$$a) \begin{vmatrix} 0 & 1 & 0 \\ 2 & 3 & 4 \\ 5 & 0 & 6 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 4 \\ 5 & 0 & 0 \end{vmatrix} + \begin{vmatrix} 0 & 1 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 6 \end{vmatrix} = 8$$

$$b) \begin{vmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 2 & 0 \\ 0 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} 0 & 0 & 0 & 2 \\ 0 & 0 & 2 & 0 \\ 0 & 2 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{vmatrix} +$$

$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 \end{vmatrix} = 1 + 16 - 4 - 4 = 9$$

$$c) \begin{vmatrix} 1 & 0 & 2 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 & 0 \\ 2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 & 2 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{vmatrix}$$

6.31. Az első sor minden eleme páros, az első oszlop minden eleme osztható 3-mal, a második oszlop minden eleme osztható 5-tel, tehát minden kígyó osztható  $2 \cdot 3 \cdot 5 = 30$ -cal, így az összegük is.

6.32. Csak egyetlen kígyó áll csupa páratlan számból, így a kígyók összegére bontásnál csak annak determinánsa páratlan, a többi páros, összegük tehát páratlan, vagyis nem lehet 0.

6.33. Megadjuk, hogy melyik sorban hányadik elem lesz a kígyóba választva. A 12 kígyó: 1243, 1324, 1432, 2134, 2341,

2413, 3142, 3214, 3421, 4123, 4231, 4312. Ezek alapján a 12 determináns – a kígyó elemeit négyzettel jelölve:

$$\begin{vmatrix} \blacksquare & 0 & 0 & 0 \\ 0 & \blacksquare & 0 & 0 \\ 0 & 0 & \blacksquare & 0 \\ 0 & 0 & 0 & \blacksquare \end{vmatrix} \begin{vmatrix} \blacksquare & 0 & 0 & 0 \\ 0 & 0 & \blacksquare & 0 \\ 0 & \blacksquare & 0 & 0 \\ 0 & 0 & 0 & \blacksquare \end{vmatrix} \begin{vmatrix} \blacksquare & 0 & 0 & 0 \\ 0 & 0 & 0 & \blacksquare \\ 0 & 0 & \blacksquare & 0 \\ 0 & \blacksquare & 0 & 0 \end{vmatrix}$$

$$\begin{vmatrix} 0 & \blacksquare & 0 & 0 \\ \blacksquare & 0 & 0 & 0 \\ 0 & 0 & \blacksquare & 0 \\ 0 & 0 & 0 & \blacksquare \end{vmatrix} \begin{vmatrix} 0 & \blacksquare & 0 & 0 \\ 0 & 0 & \blacksquare & 0 \\ 0 & 0 & 0 & \blacksquare \\ \blacksquare & 0 & 0 & 0 \end{vmatrix} \begin{vmatrix} 0 & \blacksquare & 0 & 0 \\ 0 & 0 & 0 & \blacksquare \\ \blacksquare & 0 & 0 & 0 \\ 0 & 0 & \blacksquare & 0 \end{vmatrix}$$

$$\begin{vmatrix} 0 & 0 & \blacksquare & 0 \\ \blacksquare & 0 & 0 & 0 \\ 0 & 0 & 0 & \blacksquare \\ 0 & \blacksquare & 0 & 0 \end{vmatrix} \begin{vmatrix} 0 & 0 & \blacksquare & 0 \\ 0 & \blacksquare & 0 & 0 \\ \blacksquare & 0 & 0 & 0 \\ 0 & 0 & 0 & \blacksquare \end{vmatrix} \begin{vmatrix} 0 & 0 & \blacksquare & 0 \\ 0 & 0 & 0 & \blacksquare \\ \blacksquare & 0 & 0 & 0 \\ 0 & \blacksquare & 0 & 0 \end{vmatrix}$$

$$\begin{vmatrix} 0 & 0 & 0 & \blacksquare \\ \blacksquare & 0 & 0 & 0 \\ 0 & \blacksquare & 0 & 0 \\ 0 & 0 & \blacksquare & 0 \end{vmatrix} \begin{vmatrix} 0 & 0 & 0 & \blacksquare \\ 0 & \blacksquare & 0 & 0 \\ 0 & 0 & \blacksquare & 0 \\ \blacksquare & 0 & 0 & 0 \end{vmatrix} \begin{vmatrix} 0 & 0 & 0 & \blacksquare \\ 0 & 0 & 0 & 0 \\ \blacksquare & 0 & 0 & 0 \\ 0 & \blacksquare & 0 & 0 \end{vmatrix}$$

6.34. Adjuk a harmadik oszlophoz az első 100-szorosát és a második 10-szeresét. Így az utolsó sorban a megadott, 9-cel osztható számok szerepelnek. Ha e sor szerint fejtjük ki a determinánst, akkor minden összeadandó osztható lesz 9-tel, tehát a determináns is.

6.35. Egy olyan determinánst kell konstruálni, melynek van egy nulla értékű aldeterminánsa. például a

$$\begin{vmatrix} 1 & 2 & 1 \\ 1 & 9 & 2 \\ 1 & 1 & 1 \end{vmatrix}$$

determináns értéke 1, de a 9-hez tartozó aldetermináns értéke 0, így a második sor vagy oszlop szerinti kifejtésben e szám 0-val szorozódik, vagyis nem befolyásolja a determináns értékét.

6.36.

a)  $-6 \cdot 6 = -36$ , mert a blokkmátrixok determinánsára vonatkozó tétel szerint a bal felső  $2 \times 2$ -es és a jobb alsó  $3 \times 3$ -as determinánsok szorzata adja az eredményt.

b) 24, mert a bal felső  $1 \times 1$ -es és a jobb alsó  $4 \times 4$ -es determinánsok értéke 1, illetve 24, és ezek szorzata 24. Másik megoldáshoz jutunk, ha a determinánst az első oszlopa, az egyetlen kiszámítandó aldeterminánst az első sora... szerint fejtjük ki.

6.37.

a) A determináns a 2,  $-1$ ,  $-2$ , 1 számokból képezett Vandermonde-determináns, így értéke:  $(-1 - 2)(-2 - 2)(1 - 2)(-2 - (-1))(1 - (-1))(1 - (-2)) = 72$ .

b) Vandermonde-determináns; értéke  $-2880$ .

c) A determináns két Vandermonde-determináns összegére bomlik:

$$\begin{vmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{vmatrix} + \begin{vmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & e & e^2 & e^3 \end{vmatrix} \\ = (b-a)(c-a)(c-b) \\ \times [(d-a)(d-b)(d-c) + (e-a)(e-b)(e-c)].$$

6.38. Ha  $pqrs \neq 0$ , akkor szorozzuk be az első sort  $p$ -vel, a másodikat  $q$ -val, a harmadikat  $r$ -rel, a negyediket  $s$ -sel, majd a negyedik oszlopból emeljük ki  $pqrs$ -t; így egy Vandermonde-determinánst kapunk:

$$D = \frac{pqrs}{pqrs} \begin{vmatrix} p^3 & p^2 & p & 1 \\ q^3 & q^2 & q & 1 \\ r^3 & r^2 & r & 1 \\ s^3 & s^2 & s & 1 \end{vmatrix} \\ = (q-p)(r-p)(s-p)(r-q)(s-q)(s-r).$$

Ha  $pqrs = 0$ , például  $s = 0$ , akkor az eredeti determináns negyedik oszlopa szerinti kifejtéssel kapjuk, hogy

$$D = pqr \begin{vmatrix} p^2 & p & 1 \\ q^2 & q & 1 \\ r^2 & r & 1 \end{vmatrix}.$$

Ezekből rövid átalakítás után látható, hogy az összefüggés ebben az esetben is fennáll.

6.39.  $a_1 = \det[1] = 1$ ,  $a_2 = \begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} = 2$ , az  $(n \times n)$ -es determinánst első sora szerint kifejtve kapjuk, hogy  $a_n = a_{n-1} + a_{n-2}$ .

6.41. Igen. Tekintsük a determináns első sor szerinti kifejtését! Ha mindegyik elemhez tartozó előjeles alldetermináns 0 lenne, akkor a mátrix szinguláris lenne, így valamelyik elemhez tartozó alldetermináns nem 0. Legyen pl.  $A_{1j} \neq 0$ . Ekkor a kifejtés összes többi tagját összevonva kapjuk, hogy  $\det \mathbf{A} = a_{1j}A_{1j} + c$ . Mivel  $A_{1j} \neq 0$ , ezért az  $a_{1j}A_{1j} + c = 0$  egyenlet megoldható  $a_{ij}$ -re, tehát ennek az elemnek a megváltoztatása 0-vá teszi a determinánst.

6.42. Ha az  $i$ -edik sor elemeit az  $u$ -adik sorhoz tartozó előjeles alldeterminánsokkal szorozzuk, akkor az  $u$ -adik sor elemeit nem használjuk, tehát szabadon megváltoztathatjuk. Másoljuk az  $i$ -edik sort az  $u$ -adik helyére, tehát minden  $k$ -ra  $a_{ik} = a_{uk}$ . Ekkor egyrészt  $\sum_{k=1}^n a_{ik}A_{uk} = \sum_{k=1}^n a_{uk}A_{uk}$ , azaz e determináns  $u$ -adik sor szerinti kifejtését kaptuk, másrészt e determinánsnak van két azonos sora, tehát determinánsa 0. Az oszlopokra vonatkozó állítás egy transzponálással visszavezethető erre.

6.43. A két tétel képletei közös képletbe foglalhatók. Sorokra:

$$\sum_{k=1}^n a_{ik}A_{uk} = \begin{cases} \det \mathbf{A}, & \text{ha } i = u, \\ 0, & \text{ha } i \neq u, \end{cases} \quad (6.5)$$

oszlopokra:

$$\sum_{k=1}^n a_{kj}A_{kv} = \begin{cases} \det \mathbf{A}, & \text{ha } j = v, \\ 0, & \text{ha } j \neq v. \end{cases} \quad (6.6)$$

6.44. A két kifejtési tételből adódik, hogy

$$[a_{ij}][A_{ij}]^T = \det(\mathbf{A})\mathbf{I},$$

ugyanis  $[a_{ij}]$   $i$ -edik sorának és  $[A_{ij}]^T$   $u$ -adik oszlopának, azaz  $[A_{ij}]$   $u$ -adik sorának skaláris szorzata a (6.5) képlet szerint  $\det(\mathbf{A})$ , ha  $i = u$ , azaz a szorzat főátlójában, egyebütt pedig 0. Ebből pedig mindkét képlet adódik.

6.48.

a) Az előjeles alldeterminánsok mátrixának transzponáltja:

$$\begin{bmatrix} \begin{vmatrix} 2 & 7 \\ 1 & 4 \end{vmatrix} & -\begin{vmatrix} -7 & 7 \\ 2 & 4 \end{vmatrix} & \begin{vmatrix} -7 & 2 \\ 2 & 1 \end{vmatrix} \\ -\begin{vmatrix} 1 & 4 \\ 1 & 4 \end{vmatrix} & \begin{vmatrix} 3 & 4 \\ 2 & 4 \end{vmatrix} & -\begin{vmatrix} 3 & 1 \\ 2 & 1 \end{vmatrix} \\ \begin{vmatrix} 1 & 4 \\ 2 & 7 \end{vmatrix} & -\begin{vmatrix} 3 & 4 \\ -7 & 7 \end{vmatrix} & \begin{vmatrix} 3 & 1 \\ -7 & 2 \end{vmatrix} \end{bmatrix}^T = \begin{bmatrix} 1 & 0 & -1 \\ 42 & 4 & -49 \\ -11 & -1 & 13 \end{bmatrix}$$

Mivel a mátrix determinánsa 1, ezért inverze megegyezik az előjeles alldeterminánsok előbb kiszámolt mátrixával.

b) Az előjeles alldeterminánsok mátrixának transzponáltja:

$$\begin{bmatrix} \begin{vmatrix} 0 & 2 \\ 2 & 1 \end{vmatrix} & -\begin{vmatrix} 2 & 2 \\ 3 & 1 \end{vmatrix} & \begin{vmatrix} 2 & 0 \\ 3 & 2 \end{vmatrix} \\ -\begin{vmatrix} 2 & 3 \\ 2 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 3 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 3 & 2 \end{vmatrix} \\ \begin{vmatrix} 2 & 3 \\ 0 & 2 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ 2 & 2 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 2 & 0 \end{vmatrix} \end{bmatrix}^T = \begin{bmatrix} -4 & 4 & 4 \\ 4 & -8 & 4 \\ 4 & 4 & -4 \end{bmatrix}.$$

Mivel a mátrix determinánsa 16, ezért az inverz mátrix

$$\frac{1}{4} \begin{bmatrix} -1 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -1 \end{bmatrix}.$$

c) Mivel  $\det(\mathbf{A}) = 1$ , ezért  $\mathbf{A}^{-1}$  megegyezik az előjeles alldeterminánsok mátrixának transzponáltjával. Ennek mind a 16 elemét nem kell kiszámolni, mert felső háromszögmátrix inverze felső háromszögmátrix. Hasonlóan könnyen látható, hogy a főátlóbeli elemekhez tartozó előjeles alldeterminánsok értéke 1. Tehát csak a főátló alatti elemek előjeles alldeterminánsait kell kiszámolni. Példaként egyet mutatunk:

$$A_{32} = (-1)^{3+2} \begin{vmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & 1 \end{vmatrix} = -2.$$

Hasonlóan kiszámolva a többi is kapjuk, hogy

$$\mathbf{A}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ -1 & 3 & -3 & 1 \end{bmatrix}^T = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Mi lehet e feladat általánosítása, és mi a válasz?

- d) A mátrixból csak egy nemnulla kigyó választható ki, így determinánsa könnyen számolható:  $\det \mathbf{B} = 16$ . Az inverz kiszámításához nem kell sok aldeterminánst számolni, mert nagy részük láthatóan 0 értékű. Vegyük figyelembe a számolásnál azt is, hogy  $\mathbf{B}$  szimmetrikus, így egyrészt a szimmetrikusan elhelyezkedő elemek közül csak az egyiket kell kiszámolni, másrészt a szimmetria miatt a végén szükségtelen a transzponálás.

$$\mathbf{A}^{-1} = \frac{1}{16} \begin{bmatrix} 0 & 8 & 0 & -8 \\ 8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8 \\ -8 & 0 & 8 & 0 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2} & 0 & -\frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 & \frac{1}{2} & 0 \end{bmatrix}.$$

- e) Az inverz

$$\frac{1}{abc} \begin{bmatrix} bc & 0 & 0 \\ 0 & ac & 0 \\ 0 & 0 & ab \end{bmatrix}^T = \begin{bmatrix} \frac{1}{a} & 0 & 0 \\ 0 & \frac{1}{b} & 0 \\ 0 & 0 & \frac{1}{c} \end{bmatrix},$$

ha  $abc \neq 0$ . Az  $abc = 0$  esetben a mátrix nem invertálható.

f)  $\begin{bmatrix} 1 & -1 \\ -1 & 1-i \end{bmatrix}.$

g) Az inverz  $\begin{bmatrix} 0 & 0 & 0 & 1/4 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1/3 & 0 \\ 0 & 1/2 & 0 & 0 \end{bmatrix}$

h)  $\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

6.50. a) A három eredmény: 1, 2, 4. Mivel mindhárom test esetén ugyanazokat a számolásokat kell elvégezni, csak más modulus szerinti maradék lesz az eredmény, legegyszerűbb, ha az egészek fölött számolunk, és annak maradékait tekintjük. Valóban, az egészek fölött  $-1$  a determináns, és  $-1 \bmod 2 = 1$ ,  $-1 \bmod 3 = 2$ ,  $-1 \bmod 5 = 4$ .

- b) 5. Legegyszerűbb, ha az első sor 2-szeresét hozzáadjuk a második, és 3-szorosát a harmadik sorhoz.

6.51. Sage-kód egy  $\mathbb{F}_2$  fölötti véletlen mátrix kiírására:

```
sage: random_matrix(GF(2), 5)
```

```
[1 0 0 1 1]
```

```
[1 1 1 0 1]
```

```
[1 1 1 0 0]
```

```
[1 0 0 0 0]
```

```
[0 0 1 0 0]
```

```
sage: _.det()
```

1

Hány olyan  $\mathbb{F}_2^{5 \times 5}$ -beli mátrix van, amelynek determinánsa nem 0? Az első sora bármelyik vektor lehet, kivéve a 0-vektort, így  $2^5 - 1$  lehetőség van. A második sor nem lehet ez a vektor és a 0-vektor, ez  $2^5 - 2$  lehetőség. A harmadik vektor nem lehet az előző két vektor által kifeszített altér, melynek a 0-vektorral együtt  $2^2 = 4$  eleme van, e vektor kiválasztására tehát  $2^5 - 2^2$  lehetőség adódik. Hasonlóan folytatva kapjuk, hogy az összes független vektorötösök – azaz a nemnulla értékű determinánsok – száma  $(2^5 - 2^0)(2^5 - 2^1)(2^5 - 2^2)(2^5 - 2^3)(2^5 - 2^4)$ . Ha ezt elosztjuk az összes  $\mathbb{F}_2^{5 \times 5}$ -beli mátrixok számával, 0.2980-t kapunk, így a determináns 0.7020 valószínűséggel lesz 0.

6.52.

- a) Egy vektort, mely merőleges a megadott vektorokra, és azzal/azokkal jobbrendszer alkot.

- b) Segítség: használjuk fel, hogy egy  $n - 1$ -dimenziós  $P$  paralelepipedon térfogata megegyezik annak az  $n$ -dimenziós  $Q$  paralelepipedonnak a térfogatával, melyet a  $P$ -ből úgy kapunk, hogy a  $P$ -t kifeszítő vektorokhoz  $n$ -edik vektorként egy egységvektort adunk, mely merőleges a többire.

- c) Pozitív. Ehhez épp az kellett, hogy a bázisvektorokat ne az első, hanem az utolsó sorba vagy oszlopba írjuk.

6.53. Az előző feladat szerint a kért vektort a következőképp kaphatjuk meg:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 \end{bmatrix} = \mathbf{e}_4 - \mathbf{e}_3$$

azaz a negyedik vektor  $(0, 0, -1, 1)$ .



## 7

# Mátrixleképezések és geometriájuk

E fejezetet a lineáris leképezések mátrixműveletekből való származtatását, majd általános fogalmának főként szemléletes, geometriai indítatású megalapozását tárgyalja. Merőlegesség, távolság, vetítés, forgatás, és ezek általánosításai lesznek az alapfogalmak.

### Mátrixleképezés, lineáris leképezés

---

Minden  $\mathbf{A}$  mátrixhoz tartozik egy  $\mathbf{x} \mapsto \mathbf{Ax}$  leképezés. E leképezések épp egybeesnek a lineáris kombinációt megtartó leképezésekkel, melyeket lineáris leképezéseknek nevezünk. A lineáris leképezés nem csak a lineáris algebrának, de az egész matematikának egyik legfontosabb fogalma.

---

**A mátrixleképezés fogalma** Mátrixhoz tartozó leképezésen, vagy egyszerűen mátrixleképezésen az  $\mathbf{x} \mapsto \mathbf{Ax}$  leképezést értjük, ahol  $\mathbf{A}$  egy mátrix. Egy  $m \times n$ -es  $\mathbf{A} \in \mathbb{R}^{m \times n}$  mátrixhoz így egy  $\mathbb{R}^n \rightarrow \mathbb{R}^m$  leképezés tartozik, ugyanis ha  $\mathbf{x} \in \mathbb{R}^n$  és  $\mathbf{y} = \mathbf{Ax}$ , akkor  $\mathbf{y} \in \mathbb{R}^m$ .

A mátrixok jelölésére félkövér betűket használunk, a leképezésekre dőlt (kurzív) betűket. A továbbiakban azt a konvenciót követjük, hogy egy mátrixhoz tartozó mátrixleképezést ugyanannak a betűnek a dőlt változatával jelöljük, például az  $\mathbf{A}$  mátrixhoz tartozó mátrixleképezést  $A$  jelöli, azaz

$$A : \mathbf{x} \mapsto A(\mathbf{x}) = \mathbf{Ax}.$$

Az  $A(\mathbf{x})$  mellett az  $\mathbf{Ax}$  jelölés is használatos.

Az  $A$  leképezés értékkészletét  $\text{Im}(A)$  jelöli, mely az  $\mathbb{R}^m$  altere (gondoljuk meg, miért?). Ezt szokás *képtérnek* is nevezni, minthogy ez az  $\mathbb{R}^n$  tér képe. Ez megegyezik az  $\mathbf{A}$  mátrix oszlopterével, azaz  $\mathcal{O}(\mathbf{A})$ -val. Azoknak a vektoroknak az alterét, melyet  $A$  a nullvektorba visz, az  $A$  leképezés *magterének* nevezzük. Magtérre a *kernel* szó is használatos.  $\text{Ker}(A)$ -val jelöljük. Ez megegyezik a hozzá tartozó  $\mathbf{A}$  mátrix

nullterével. Tehát

$$\text{Im}(A) = \mathcal{O}(\mathbf{A}), \quad \text{Ker}(A) = \mathcal{N}(\mathbf{A}).$$

Az  $\text{Im}$  rövidítés a kép jelentésű *image*, a  $\text{Ker}$  a mag jelentésű *kernel* szóból származik.

**7.1. PÉLDA (VEKTORI SZORZÁSSAL DEFINIÁLT MÁTRIXLEKÉPEZÉS).** Legyen  $\mathbf{a} = (a_1, a_2, a_3)$  egy adott  $\mathbb{R}^3$ -beli vektor. Legyen  $A$  az a transzformáció, mely a tér tetszőleges  $\mathbf{x}$  vektorához az  $\mathbf{a} \times \mathbf{x}$  vektort rendeli. Tehát

$$A : \mathbb{R}^3 \rightarrow \mathbb{R}^3 : \mathbf{x} \mapsto \mathbf{a} \times \mathbf{x}.$$

Mutassuk meg, hogy az  $A$  függvény egy mátrixleképezés, azaz létezik egy olyan  $\mathbf{A}$  mátrix, hogy  $A(\mathbf{x}) = \mathbf{A}\mathbf{x}$ .

**MEGOLDÁS.** Az  $\mathbf{a} \times \mathbf{x}$  vektori szorzat koordinátás alakban:

$$\mathbf{y} = \mathbf{a} \times \mathbf{x} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} a_2x_3 - a_3x_2 \\ a_3x_1 - a_1x_3 \\ a_1x_2 - a_2x_1 \end{bmatrix}.$$

Az eredményből azonnal látszik, hogy e transzformáció mátrixleképezés, hisz  $\mathbf{y}$  minden koordinátája  $\mathbf{x}$  koordinátáinak lineáris kifejezése. A szorzatot  $\mathbf{x}$  koordinátái szerint rendezzük, ahonnan azonnal leolvasható a transzformáció mátrixa, amit a továbbiakban  $[\mathbf{a}]_{\times}$  jelöl. Segítségével fölírható a transzformáció mátrixszorzatos alakja:

$$\mathbf{a} \times \mathbf{x} = \begin{bmatrix} -a_3x_2 + a_2x_3 \\ a_3x_1 - a_1x_3 \\ -a_2x_1 + a_1x_2 \end{bmatrix} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Tehát

$$[\mathbf{a}]_{\times} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}. \quad (7.1)$$

E feladat eredménye különösen fontos a 3-dimenziós tér transzformációinak vizsgálatánál, így pl. az anyagtranszformációk fizikai/mérnöki vizsgálatában.  $\square$

**Műveletek mátrixleképezések között** A következőkben megvizsgáljuk, hogy mi a kapcsolat a mátrixműveletek, és a mátrixokhoz tartozó mátrixleképezések közti műveletek között.

**7.2. TÉTEL (MÁTRIXLEKÉPEZÉSEK ALAPMŰVELETEI).** Legyen  $\mathbf{A}$ ,  $\mathbf{B}$  és  $\mathbf{C}$  három  $m \times n$ -es mátrix, legyen  $A$ ,  $B$  és  $C$  a hozzájuk tartozó három mátrixleképezés és legyen  $c$  egy skálár. Ekkor

- $\mathbf{A} + \mathbf{B} = \mathbf{C}$  pontosan akkor igaz, ha  $A + B = C$ , és
- $c\mathbf{A} = \mathbf{C}$  pontosan akkor igaz, ha  $cA = C$ .



Ha  $\mathbf{X}$ ,  $\mathbf{Y}$  és  $\mathbf{Z}$  típusa rendre  $m \times k$ ,  $k \times n$ , illetve  $m \times n$ , és  $X$ ,  $Y$  és  $Z$  a hozzájuk tartozó három mátrixleképezés, akkor

c)  $\mathbf{XY} = \mathbf{Z}$  pontosan akkor igaz, ha  $X \circ Y = Z$ , azaz mátrixok szorzásának a függvények kompozíciója felel meg.

Az  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  függvénynek a  $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$  függvény inverze, ha minden  $\mathbf{x} \in \mathbb{R}^n$  helyen  $(f(g(\mathbf{x}))) = \mathbf{x}$  és  $(g(f(\mathbf{x}))) = \mathbf{x}$ , azaz ha kompozícióik az  $f \circ g$  és a  $g \circ f$  függvények megegyeznek az identikus leképezéssel.

**7.3. TÉTEL (INVERZ MÁTRIXLEKÉPEZÉSEK).** Legyenek  $A$  és  $B$  az  $n \times n$ -es  $\mathbf{A}$  és  $\mathbf{B}$  mátrixokhoz tartozó mátrixleképezések. Ekkor az  $\mathbf{A}$  mátrix inverze pontosan akkor a  $\mathbf{B}$  mátrix, ha az  $A$  leképezés inverze a  $B$  leképezés.

A fenti két tétel bizonyítását az Olvasóra hagyjuk (ld. 7.5. és 7.6. feladatok)!

**Mátrixleképezések tulajdonságai** A mátrixleképezések megőrzik a lineáris kombinációt, a nullvektort nullvektorba, alteret altérbe visznek.

**7.4. TÉTEL (MÁTRIXLEKÉPEZÉSEK ALAPTULAJDONSÁGAI).** Legyen  $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$  egy tetszőleges mátrixleképezés,  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ ,  $c, d \in \mathbb{R}$ .

a)  $A(c\mathbf{x} + d\mathbf{y}) = cA(\mathbf{x}) + dA(\mathbf{y})$ , azaz  $A$  megőrzi a lineáris kombinációt.

b) Az  $A$  homogén és additív leképezés, azaz

$$\begin{aligned} A(c\mathbf{x}) &= cA(\mathbf{x}), & (a \text{ leképezés homogén}), \text{ és} \\ A(\mathbf{x} + \mathbf{y}) &= A(\mathbf{x}) + A(\mathbf{y}), & (a \text{ leképezés additív}). \end{aligned}$$

c)  $A\mathbf{0} = \mathbf{0}$ .

d) Tetszőleges altér képe altér.

e) Tetszőleges affin altér képe affin altér.

**BIZONYÍTÁS.** a) Bármely  $\mathbf{x}$  és  $\mathbf{y}$  vektorra és  $c, d \in \mathbb{R}$  valósra

$$A(c\mathbf{x} + d\mathbf{y}) = \mathbf{A}(c\mathbf{x} + d\mathbf{y}) = c\mathbf{Ax} + d\mathbf{Ay} = cA(\mathbf{x}) + dA(\mathbf{y}).$$

b) Az előző egyenlőség  $d = 0$  esetén a homogenitást,  $c = d = 1$  esetén az additivitást bizonyítja. c) igaz, mert bármely  $\mathbf{x}$  vektorra  $A\mathbf{0} = A(0\mathbf{x}) = 0A(\mathbf{x}) = \mathbf{0}$ . d) abból következik, hogy ha  $\mathbf{b}_1, \dots, \mathbf{b}_k$  egy  $\mathcal{U}$  altér bázisa, akkor ezek összes lineáris kombinációjának, vagyis az altér vektorainak képe

$$A(c_1\mathbf{b}_1 + \dots + c_k\mathbf{b}_k) = c_1A(\mathbf{b}_1) + \dots + c_kA(\mathbf{b}_k).$$

Világos, hogy e vektorok kiadják az  $A\mathbf{b}_1, \dots, A\mathbf{b}_k$  vektorok által kifesztett altér minden vektorát, tehát  $A(\mathcal{U})$  altér. Hasonlóan e)-ben, ha

$\mathbf{u} \in \mathbb{R}^n$  egy tetszőleges vektor és  $\mathcal{U}$  a fenti altér, akkor

$$\begin{aligned} A(\mathbf{u} + \mathcal{U}) &= A(u + c_1 \mathbf{b}_1 + \dots + c_k \mathbf{b}_k) \\ &= A(\mathbf{u}) + c_1 A(\mathbf{b}_1) + \dots + c_k A(\mathbf{b}_k) \\ &= A(\mathbf{u}) + A(\mathcal{U}), \end{aligned}$$

ami egy altér eltoltja, azaz affin altér.  $\square$

**Lineáris leképezés** A mátrixleképezések alaptulajdonságai a lineáris leképezés fogalmához vezetnek.

**7.5. DEFINÍCIÓ (LINEÁRIS LEKÉPEZÉS).** Legyen  $H_1$  és  $H_2$  mindegyike olyan halmaz, melynek elemein értelmezve van egy asszociatív összeadás és egy „skalárral való szorzás” művelet. Azt mondjuk, hogy egy  $A : H_1 \rightarrow H_2$  leképezés lineáris, ha homogén és additív, azaz ha tetszőleges  $\mathbf{x}, \mathbf{y} \in H_1$  elemre és  $c$  skalárra

$$\begin{aligned} A(c\mathbf{x}) &= cA(\mathbf{x}) & (A \text{ homogén,}) \\ A(\mathbf{x} + \mathbf{y}) &= A\mathbf{x} + A\mathbf{y} & (A \text{ additív.}) \end{aligned}$$

$H_1 = H_2$  esetén a lineáris leképezéseket lineáris transzformációknak is nevezzük.

Azt a későbbiekben fogjuk részletezni, hogy milyen algebrai tulajdonságokat érdemes a skalárokról, valamint  $H_1$  és  $H_2$  elemeiről föltenni. Most csak néhány példát mutatunk lineáris leképezésekre.

**7.6. PÉLDA (A DERIVÁLÁS ÉS AZ INTEGRÁLÁS LINEÁRIS LEKÉPEZÉS).** Legyen  $H_1$  az egyváltozós valós, és minden valós helyen differenciálható függvények halmaza,  $H_2$  pedig az egyváltozós valós függvények halmaza. Véljük, hogy  $H_1$  és  $H_2$  is olyan halmaz, melynek elemei közt értelmezve van az összeadás és a skalárral való szorzás művelete. A deriválás, azaz a  $D : H_1 \rightarrow H_2 : f \mapsto D(f) = f'$  leképezés lineáris. Fogalmazzunk meg hasonló állítást az integrálra is.

**MEGOLDÁS.** Tetszőleges  $c \in \mathbb{R}$  skalárra és  $f, g \in H_1$  függvényre

$$\begin{aligned} D(cf) &= (cf)' = cf' = cD(f), \text{ és} \\ D(f + g) &= (f + g)' = f' + g' = D(f) + D(g). \end{aligned}$$

Hasonló összefüggések állnak fenn az integrálokra is, például legyen  $H_1$  a  $[0, 1]$  intervallumon Riemann-integrálható függvények halmaza, és legyen  $H_2 = \mathbb{R}$ . Ekkor az  $f \mapsto \int_0^1 f$  leképezés lineáris, ugyanis tetszőleges  $c \in \mathbb{R}$  skalárra és tetszőleges  $f, g \in H_1$  függvényre

$$\int_0^1 cf = c \int_0^1 f, \text{ és } \int_0^1 (f + g) = \int_0^1 f + \int_0^1 g. \quad \square$$

**7.7. ÁLLÍTÁS (SÍKBELI FORGATÁS, TÜKRÖZÉS, VETÍTÉS).** A síkbeli vektorok egy rögzített  $O$  pont körüli forgatása, egy egyenesre való tükrözése és merőleges vetítése lineáris leképezés.

**BIZONYÍTÁS.** A precíz bizonyításokat mellőzzük, csak a tényt szemléltetjük. A síkbeli vektorok pont körüli forgatása lineáris leképezés, ugyanis könnyen látható, hogy egy vektor  $c$ -szeresének ( $c \in \mathbb{R}$ ) elforgatottja megegyezik a vektor elforgatottjának  $c$ -szeresével, valamint hogy két vektor összegének elforgatottja megegyezik a vektorok elforgatottjainak összegével (ld. 7.1. ábra).

Hasonlóan egyszerűen látszik, hogy egy egyenesre való tükrözés egy vektor  $c$ -szeresét a vektor tükröképének  $c$ -szeresébe viszi, és két vektor összegét a két vektor tükröképének összegébe (ld. 7.2 ábra).

Végül ugyanígy megmutatható, hogy egy egyenesre való merőleges vetítés egy vektor  $c$ -szeresét vetületének  $c$ -szeresébe viszi, és két vektor összegét a két vektor vetületének összegébe (ld. 7.3 ábra).  $\square$

**$\mathbb{R}^n$ -ből  $\mathbb{R}^m$ -be képző lineáris leképezések** E fejezet további részében csak lineáris  $\mathbb{R}^n \rightarrow \mathbb{R}^m$  leképezésekkel foglalkozunk, ahol a skalárok a valós számok. Megmutatjuk, hogy ezek mátrixleképezések.

**7.8. TÉTEL (LINEÁRIS LEKÉPEZÉS EKVIVALENS DEFINÍCIÓI).** Egy tetszőleges  $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$  leképezésre az alábbi állítások ekvivalensek:

1. A lineáris, azaz homogén és additív.
2. Tetszőleges  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  és  $c, d \in \mathbb{R}$  esetén

$$A(c\mathbf{x} + d\mathbf{y}) = cA(\mathbf{x}) + dA(\mathbf{y})$$

3. Tetszőleges  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  és  $c \in \mathbb{R}$  esetén

$$A(c\mathbf{x} + \mathbf{y}) = cA(\mathbf{x}) + A(\mathbf{y})$$

4. „Megőrzi” a lineáris kombinációt, azaz tetszőleges  $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$  vektorokra és  $c_1, c_2, \dots, c_k \in \mathbb{R}$  skalárakra

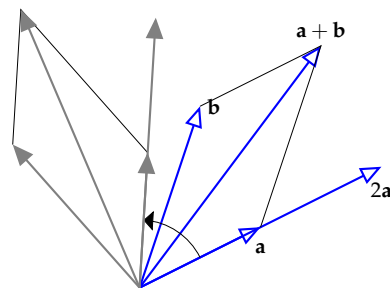
$$A(c_1\mathbf{x}_1 + \dots + c_k\mathbf{x}_k) = c_1A\mathbf{x}_1 + \dots + c_kA\mathbf{x}_k.$$

A bizonyítást az Olvasóra hagyjuk (ld. 7.7. feladat).

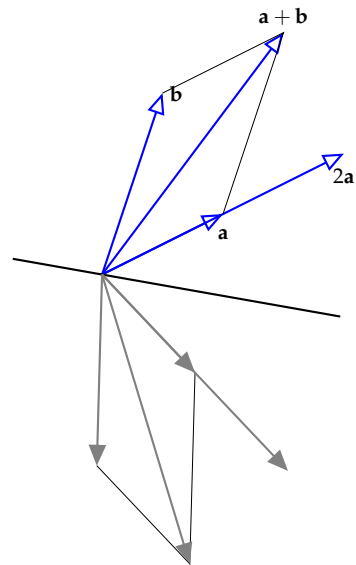
**7.9. TÉTEL (AZ  $\mathbb{R}^n \rightarrow \mathbb{R}^m$  LINEÁRIS LEKÉPEZÉSEK MÁTRIXLEKÉPEZÉSEK).** Legyen  $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$  egy tetszőleges függvény. Az  $A$  pontosan akkor lineáris, ha létezik egy olyan  $m \times n$ -es  $\mathbf{A}$  mátrix, hogy az  $A$  függvény megegyezik az  $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$  leképezéssel. Ekkor

$$\mathbf{A} = [\mathbf{A}\mathbf{e}_1 | \mathbf{A}\mathbf{e}_2 | \dots | \mathbf{A}\mathbf{e}_n],$$

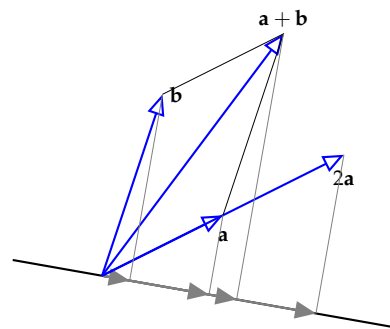
ahol  $\mathbf{e}_i$  az  $i$ -edik standard egységvektor ( $i = 1, 2, \dots, n$ ).



7.1. ábra: A pont körüli elforgatás lineáris leképezés



7.2. ábra: Az egyenesre való tükrözés lineáris leképezés



7.3. ábra: Az egyenesre való merőleges vetítés lineáris leképezés

**BIZONYÍTÁS.** Minden mátrixleképezés lineáris, ez bizonyítja az állítás egyik felét. A állítás másik felének bizonyításához tekintsük  $\mathbb{R}^n$  standard bázisát és az  $A\mathbf{e}_i$  vektorokból képzett

$$\mathbf{A} = [A\mathbf{e}_1 | A\mathbf{e}_2 | \dots | A\mathbf{e}_n] \quad (7.2)$$

mátrixot, valamint legyen  $\mathbf{x} \in \mathbb{R}^n$  egy tetszőleges vektor. Ha  $A$  lineáris leképezés, azaz megőrzi a lineáris kombinációt, akkor

$$\begin{aligned} A\mathbf{x} &= A(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \dots + x_n\mathbf{e}_n) \\ &= x_1A\mathbf{e}_1 + x_2A\mathbf{e}_2 + \dots + x_nA\mathbf{e}_n \\ &= \begin{bmatrix} A\mathbf{e}_1 & A\mathbf{e}_2 & \dots & A\mathbf{e}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= \mathbf{A}\mathbf{x} \end{aligned}$$

Tehát valóban létezik olyan  $\mathbf{A}$  mátrix, hogy  $A\mathbf{x} = \mathbf{A}\mathbf{x}$ . Ráadásul ilyen mátrix csak ez az egy van, mert bármely  $\mathbf{e}_i$  bázisvektorra és bármely  $\mathbf{A}$  mátrixra  $A\mathbf{e}_i = \mathbf{A}_{*i}$ , tehát az  $\mathbf{A}_{*i}$  oszlopvektor csak  $A\mathbf{e}_i$  lehet.  $\square$

**7.10. PÉLDA.** Mutassuk meg, hogy az

$$\begin{aligned} A : \mathbb{R}^2 &\rightarrow \mathbb{R}^3; (x, y) \mapsto (x - y, 2x + y, -x + 1) \text{ és} \\ L : \mathbb{R}^2 &\rightarrow \mathbb{R}^3; (x, y) \mapsto (x - y, 2x + y, -x) \end{aligned}$$

leképezések közül az  $A$  nem lineáris leképezés, de az  $L$  igen. Utóbbinak írjuk föl a mátrixát!

**MEGOLDÁS.** Az  $A$  leképezés nem lineáris, mert a 7.4. tétel következtében  $A\mathbf{0} = \mathbf{0}$  kellene, de  $A : (0, 0) \mapsto (0, 0, 1)$ .

Alakítsuk át a függvényértékül kapott vektort:

$$L \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x - y \\ 2x + y \\ -x \end{bmatrix} = x \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + y \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 2 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

ami igazolja, hogy  $L$  mátrixleképezés, és mátrixa

$$\mathbf{L} = \begin{bmatrix} 1 & -1 \\ 2 & 1 \\ -1 & 0 \end{bmatrix}.$$

A mátrixleképezések pedig lineáris leképezések, tehát  $L$  is.  $\square$

► Mint azt a 7.6. példa mutatja, lineáris leképezésekről olyan esetben is beszélhetünk, amikor a leképezésnek nincs mátrixa, azaz a lineáris leképezés általánosabb fogalom.

► Különbség van a lineáris leképezés és a mátrixleképezés közt  $\mathbb{R}^n \rightarrow \mathbb{R}^m$  függvények esetén is. A lineáris leképezés független a bázistól, az csak maga a függvény, mely megadja, hogy melyik vektornak melyik vektor a képe. A mátrixleképezés mindig valamely bázisra vonatkozik. Egy lineáris leképezéshez minden bázisban tartozik egy mátrixleképezés, melynek mátrixa függ a bázistól.

► Keressük meg a lineáris  $\mathbb{R} \rightarrow \mathbb{R}$  transzformációkat. Itt  $\mathbb{R}$  elemei az 1-dimenziós vektorok (azonosíthatók a számokkal). E térben az  $e = 1$  vektor (szám) a bázis. Az előző tétel szerint egy lineáris  $L : \mathbb{R} \rightarrow \mathbb{R}$  transzformáció mátrixa  $[Le] = [L(1)]$ , ami egy szám, jelölje ezt  $c := L(1)$ . Így  $L(x) = L(1x) = L(x1) = xc = cx$ , azaz a lineáris  $\mathbb{R} \rightarrow \mathbb{R}$  transzformációk azonosak az  $x \mapsto cx$  függvényekkel, ahol  $c$  egy tetszőleges konstans. Az ilyen leképezések grafikonja egy origón átmenő (függőlegestől különböző) egyenes. (Az  $\mathbb{R} \rightarrow \mathbb{R}$  lineáris leképezések tehát nem azonosak a lineáris  $\mathbb{R} \rightarrow \mathbb{R}$  függvényekkel, melyek általános alakja  $f(x) = cx + d$ , ahol  $c, d \in \mathbb{R}$ .)

**A mátrixleképezés hatásának szemléltetése** Egy mátrixszal való szorzás hatásának megértését még egy adott konkrét alkalmazásban is segítheti, ha vizuálisan is megjeleníthető képünk van róla.

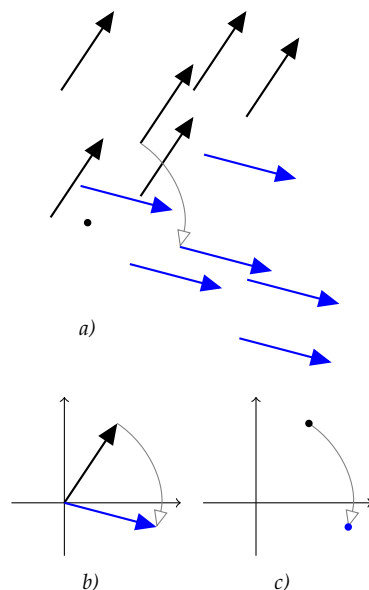
Egy vektor és egy mátrixleképezés általi képe – akár szabad vektorokkal, akár helyvektorokkal, akár a helyvektorok végpontjával – egyszerűen ábrázolható. A 7.4. ábra például a forgatómátrix hatását szemlélteti e három módon. (Szabad vektorok esetén a szabad vektorok mellett a forgatás középpontja is szabadon megválasztható, helyvektorok esetén a forgatás középpontja az origó!)

$\mathbb{R}^2 \rightarrow \mathbb{R}^2$  leképezések esetén a legegyszerűbb szemléltetéshez elég csak az egységnégyzet képét megrajzolni, amint azt a 7.5. ábra mutatja. A kép mindig egy parallelogramma (esetleg elfajuló), melynek körüljárását is jelölni kell valahogy. Ezt az ábrán az oldalak különböző színezése megteszi. A parallelogramma területe és körüljárása a mátrix determinánsból olvasható ki. Az egységnégyzetrács képe parallelogrammarács. Ennek segítségével egy tetszőleges vektor képének megszerkesztése egyszerű, hisz a lineáris leképezés megtartja a lineáris kombinációt (ld. 7.5. ábra).

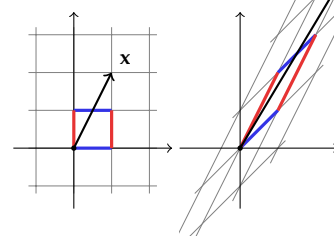
**7.11. PÉLDA (MÁTRIXLEKÉPEZÉS ÁBRÁZOLÁSA AZ EGYSÉGNÉGYZETRÁCS KÉPÉVEL).** Ábrázoljuk az egységnégyzet és az egységnégyzetrács, valamint az  $(1, 2)$  vektor képét az

$$\mathbf{A} = \begin{bmatrix} 5 & 3 \\ 4 & 4 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 3 & 5 \\ 4 & 3 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} -5 & 3 \\ -4 & 5 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} -3 & 5 \\ -4 & 4 \end{bmatrix}$$

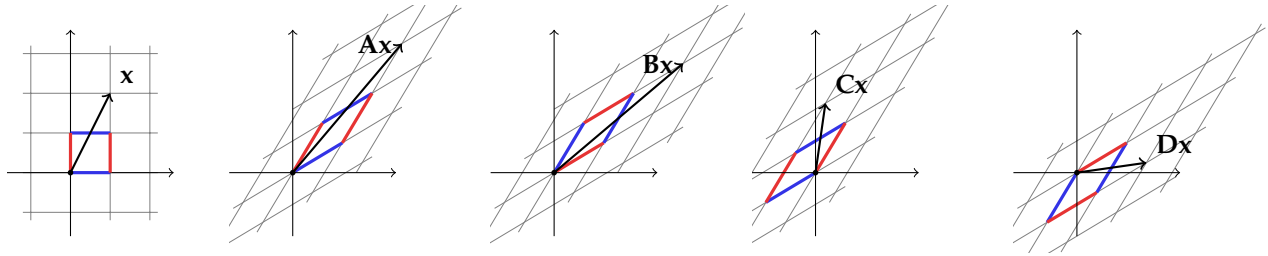
mátrixokkal megadott leképezések esetén.



7.4. ábra: Vektorok elforgatásának szemléltetése a vektor különböző ábrázolásai szerint: a) szabad vektorok, b) helyvektorok, c) pontok.

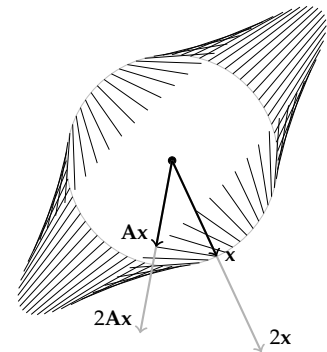


7.5. ábra: Az  $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$  mátrix hatása az egységnégyzetrácson, és az  $\mathbf{x} = (1, 2)$  vektoron. Az  $\mathbf{Ax}$  vektor végpontja a parallelogrammarácson 1 lépés az  $x$ -tengely képének irányába, és 2 lépés az  $y$ -tengely képének irányába.



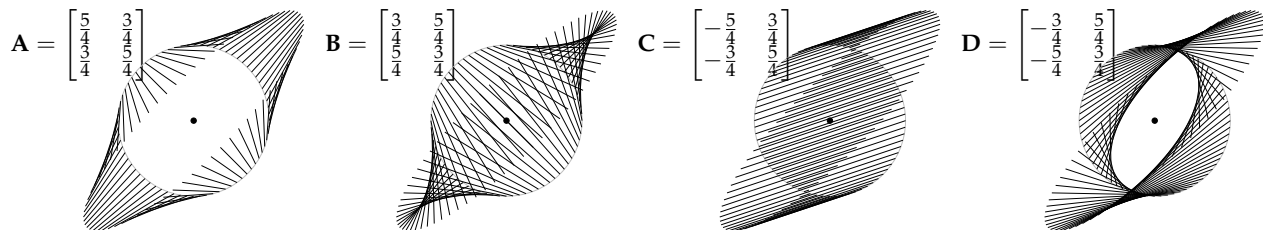
7.6. ábra: Az egységnyezet és az egységnyezetrács képe a megadott négy mátrix esetén, és az  $(1, 2)$  vektor képe.

Egy másik ábrázolási lehetőséget kapunk az egységvektorok képének megszerkesztésével. Mivel a mátrixleképezés homogén, azaz egy vektor  $c$ -szereséhez a vektor képének  $c$ -szeresét rendeli, ezért elég minden irányból egyetlen vektor – például az egységvektor – képének megszerkesztése. Hogy a kép áttekinthető legyen, helyvektorok helyett csak az őket reprezentáló pontokat tekintjük, és az egységkör összes pontja helyett csak néhányat (pl. 50–100-at). Mivel a kör lineáris leképezés általi képe mindig egy ellipszis (esetleg elfajuló), ezért a leképezés szemléltetéséhez elég összekötni az egységkört kiválasztott pontot a képével ahhoz, hogy nagyjából a sík bármelyik vektorának „lássuk”, hogy mi a képe. Az így kialakuló ábra sokat elmond a leképezésről. Ezt mutatja a 7.7 ábra, ahol a  $-65^\circ$ -os irányhoz tartozó  $x$  egységvektort és  $Ax$  képét külön berajzoltuk, és azt is megmutattuk, hogy pl. hogyan kapható meg  $2x$  képe.



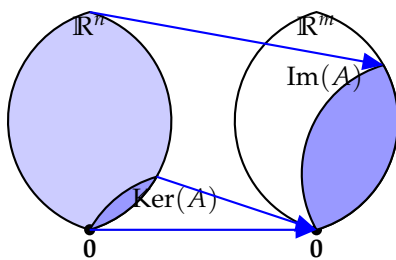
7.7. ábra: Az  $A = \begin{bmatrix} 5/4 & 3/4 \\ 3/4 & 5/4 \end{bmatrix}$  mátrix hatását szemlélteti oly módon, hogy az egységkör néhány pontját összeköti a képükkel. Az ábrán egy egységnyi  $x$  vektort és  $Ax$  képét, valamint a  $2x$  vektort és képét a  $A(2x) = 2Ax$  vektort kiemeltük.

A 7.8 ábra az előző példabeli mátrixleképezések egységkör-ábráját mutatja.



7.8. ábra: Négy leképezés egységkör-ábrája.

Egy általános  $\mathbb{R}^n \rightarrow \mathbb{R}^m$  leképezés megjelenítéséhez a levéldiagrammot hívjuk segítségül. Itt elsőként a képteret és a magteret tudjuk szemléltetni, amint azt a 7.9. ábra mutatja.



7.9. ábra: Egy  $A : \mathbb{R}^n \rightarrow \mathbb{R}^m, x \mapsto Ax$  mátrixleképezés levéldiagrammja. Az ábrán három altér színezéssel ki van emelve – az értelmezési tartomány  $(\mathbb{R}^n)$ , az értékkészlet  $(\text{Im}(A) = \mathcal{O}(A))$  és a magtér  $(\text{Ker}(A) = \mathcal{N}(A))$ .

*Lineáris leképezés mátrixa különböző bázisokban* Tegyük fel, hogy az  $L$  lineáris leképezés mátrixa az  $\mathcal{A}$  bázisban  $L_{\mathcal{A}}$ , a  $\mathcal{B}$  bázisban  $L_{\mathcal{B}}$ , és az  $\mathcal{A}$  bázisról a  $\mathcal{B}$ -re való áttérés mátrixa  $C_{\mathcal{B} \leftarrow \mathcal{A}}$ . Kérdés, hogy mi a kapcsolat e három mátrix között.

A válasz egyszerűen megadható, ha megvizsgáljuk egy tetszőleges  $x$  vektornak és  $Lx$  képének koordinátás alakját. Ezeket jelölje  $[x]_{\mathcal{A}}$ ,  $[x]_{\mathcal{B}}$ ,  $[Lx]_{\mathcal{A}}$ ,  $[Lx]_{\mathcal{B}}$ . Az áttérés mátrixa köztük a következő kapcsolatokat létesíti:

$$[x]_{\mathcal{B}} = C_{\mathcal{B} \leftarrow \mathcal{A}}[x]_{\mathcal{A}}, \quad [Lx]_{\mathcal{B}} = C_{\mathcal{B} \leftarrow \mathcal{A}}[Lx]_{\mathcal{A}},$$

az  $L_{\mathcal{A}}$  és  $L_{\mathcal{B}}$  mátrixok pedig a következőket:

$$L_{\mathcal{A}}[x]_{\mathcal{A}} = [Lx]_{\mathcal{A}}, \quad L_{\mathcal{B}}[x]_{\mathcal{B}} = [Lx]_{\mathcal{B}}.$$

Ezeket összevetve kapjuk, hogy

$$L_{\mathcal{B}}[x]_{\mathcal{B}} = [Lx]_{\mathcal{B}} = C_{\mathcal{B} \leftarrow \mathcal{A}}[Lx]_{\mathcal{A}} = C_{\mathcal{B} \leftarrow \mathcal{A}}L_{\mathcal{A}}[x]_{\mathcal{A}},$$

azaz

$$L_{\mathcal{B}}C_{\mathcal{B} \leftarrow \mathcal{A}}[x]_{\mathcal{A}} = C_{\mathcal{B} \leftarrow \mathcal{A}}L_{\mathcal{A}}[x]_{\mathcal{A}}$$

vagyis csak mátrixokra fölírva:

$$L_{\mathcal{B}}C_{\mathcal{B} \leftarrow \mathcal{A}} = C_{\mathcal{B} \leftarrow \mathcal{A}}L_{\mathcal{A}} \text{ vagy átrendezve } L_{\mathcal{A}} = C_{\mathcal{B} \leftarrow \mathcal{A}}^{-1}L_{\mathcal{B}}C_{\mathcal{B} \leftarrow \mathcal{A}}.$$

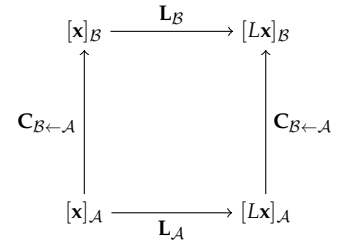
A bizonyítás – és általában e mátrixok közti összefüggés – egy diagrammon is szemléltethető. A diagramm csúcaiban az  $x$  és az  $Lx$  vektorok szerepelnek. A függőleges nyilak az  $\mathcal{A}$  bázisról a  $\mathcal{B}$ -re való áttérés irányát mutatják. Ebbe az irányba lépve az eredmény a  $C_{\mathcal{B} \leftarrow \mathcal{A}}$  mátrixszal való szorzással kapható meg. A vízszintes nyilak az  $L$  leképezés hatását mutatják. Ezirányba lépve az  $L_{\mathcal{A}}$ , illetve  $L_{\mathcal{B}}$  mátrixszal való szorzás adja meg az eredményt. A bal alsó sarokban lévő  $[x]_{\mathcal{A}}$  vektorból a jobb felsőben lévő  $[Lx]_{\mathcal{B}}$  vektorba kétféleképp juthatunk: vagy először hat az  $L$  leképezés, aztán áttérünk a  $\mathcal{B}$  bázisra, vagy előbb áttérünk a  $\mathcal{B}$  bázisra, és azután hat  $L$ . Tehát

$$\begin{aligned} [x]_{\mathcal{A}} &\longrightarrow L_{\mathcal{A}}[x]_{\mathcal{A}} \longrightarrow C_{\mathcal{B} \leftarrow \mathcal{A}}L_{\mathcal{A}}[x]_{\mathcal{A}}, \\ [x]_{\mathcal{A}} &\longrightarrow C_{\mathcal{B} \leftarrow \mathcal{A}}[x]_{\mathcal{A}} \longrightarrow L_{\mathcal{B}}C_{\mathcal{B} \leftarrow \mathcal{A}}[x]_{\mathcal{A}}. \end{aligned}$$

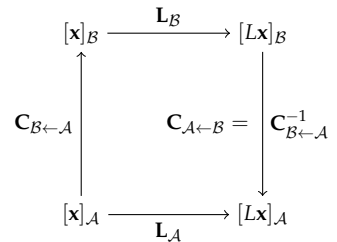
A két végeredménynek meg kell egyeznie. Így ugyanazt kaptuk, amit behelyettesítésekkel. Összefoglalva tehát bizonyítottuk az alábbi tételt:

**7.12. TÉTEL (LINEÁRIS LEKÉPEZÉS MÁTRIXAI KÖZTI KAPCSOLAT).** Legyen az  $L$  lineáris leképezés mátrixa az  $\mathcal{A}$  bázisban  $L_{\mathcal{A}}$ , a  $\mathcal{B}$  bázisban  $L_{\mathcal{B}}$ , és az  $\mathcal{A}$  bázisról a  $\mathcal{B}$ -re való áttérés mátrixa  $C_{\mathcal{B} \leftarrow \mathcal{A}}$ . Ekkor

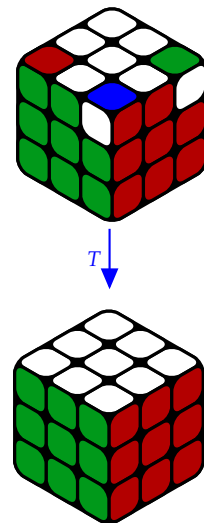
$$L_{\mathcal{A}} = C_{\mathcal{A} \leftarrow \mathcal{B}}L_{\mathcal{B}}C_{\mathcal{B} \leftarrow \mathcal{A}}, \text{ azaz } L_{\mathcal{A}} = C_{\mathcal{B} \leftarrow \mathcal{A}}^{-1}L_{\mathcal{B}}C_{\mathcal{B} \leftarrow \mathcal{A}}.$$



7.10. ábra: A vízszintes nyilak az  $L$  leképezés hatását mutatják. E hatás elérhető az  $L_{\mathcal{A}}$ , illetve az  $L_{\mathcal{B}}$  mátrixszal való szorzással. A függőleges nyilak a báziscsere irányát mutatják. A  $C_{\mathcal{B} \leftarrow \mathcal{A}}$  mátrixszal való szorzással megvalósítható. Az ábráról leolvasható a  $L_{\mathcal{B}}C_{\mathcal{B} \leftarrow \mathcal{A}} = C_{\mathcal{B} \leftarrow \mathcal{A}}L_{\mathcal{A}}$  összefüggés.



7.11. ábra: Ezt az ábrát az előzőből egyetlen nyíl és felírata megváltoztatásával kaptuk. Erről közvetlenül leolvasható az  $L_{\mathcal{A}} = C_{\mathcal{A} \leftarrow \mathcal{B}}L_{\mathcal{B}}C_{\mathcal{B} \leftarrow \mathcal{A}}$ , illetve az  $L_{\mathcal{A}} = C_{\mathcal{B} \leftarrow \mathcal{A}}^{-1}L_{\mathcal{B}}C_{\mathcal{B} \leftarrow \mathcal{A}}$  összefüggés. Ehhez az  $[x]_{\mathcal{A}}$ -ból az  $[Lx]_{\mathcal{A}}$ -ba vezető két utat kell bejárunk, és közben a megfelelő mátrixokat összeszorozunk.



7.12. ábra: A  $T$  transzformáció a bűvös kocka egy lapon lévő 3 csúcsát ciklikusan fölcseréli egymással (a jobb felsőt sarokban lévő a bal sarokba, azt a középsőbe viszi), az összes többi helyben hagyja. Mit tegyünk ha 3 nem egy lapon lévő csúcsot kell ciklikusan fölcserélni?

**7.13. PÉLDA (LINEÁRIS LEKÉPEZÉS MÁTRIXA MÁSIK BÁZISBAN).** Az  $L$  lineáris leképezés mátrixa  $\mathbf{L} = \begin{bmatrix} -1 & 6 \\ -2 & 6 \end{bmatrix}$ . Írjuk fel mátrixát az  $\mathcal{B} = \{(-2, -1), (3, 2)\}$  bázisban!

**MEGOLDÁS.** A megadott  $\mathcal{B}$  bázisról a standard  $\mathcal{E}$  bázisra való áttérés mátrixa a bázisvektorokból, mint oszlopvektorokból áll, azaz

$$\mathbf{C}_{\mathcal{E} \leftarrow \mathcal{B}} = \begin{bmatrix} -2 & 3 \\ -1 & 2 \end{bmatrix}.$$

Ekkor az  $\mathbf{L} = \mathbf{L}_{\mathcal{E}}$  jelölést használva

$$\mathbf{L}_{\mathcal{B}} = \mathbf{C}_{\mathcal{E} \leftarrow \mathcal{B}}^{-1} \mathbf{L}_{\mathcal{E}} \mathbf{C}_{\mathcal{E} \leftarrow \mathcal{B}} = \begin{bmatrix} -2 & 3 \\ -1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} -1 & 6 \\ -2 & 6 \end{bmatrix} \begin{bmatrix} -2 & 3 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}. \quad \square$$

**Hasonlóság** Tudjuk, hogy ha egy  $L$  lineáris transzformációnak egy  $\mathcal{A}$  bázisban  $\mathbf{A}$  a mátrixa, egy  $\mathcal{B}$  bázisban  $\mathbf{B}$ , akkor a két mátrix kapcsolata

$$\mathbf{B} = \mathbf{C}_{\mathcal{B} \leftarrow \mathcal{A}} \mathbf{A} \mathbf{C}_{\mathcal{A} \leftarrow \mathcal{B}},$$

ahol  $\mathbf{C}_{\mathcal{B} \leftarrow \mathcal{A}} = \mathbf{C}_{\mathcal{A} \leftarrow \mathcal{B}}^{-1}$  az áttérés mátrixa. E tény motiválja a következő definíciót:

**7.14. DEFINÍCIÓ (HASONLÓSÁG).** Azt mondjuk, hogy az  $n \times n$ -es  $\mathbf{A}$  mátrix hasonló a  $\mathbf{B}$  mátrixhoz, ha létezik olyan invertálható  $\mathbf{C}$  mátrix, hogy

$$\mathbf{B} = \mathbf{C}^{-1} \mathbf{A} \mathbf{C}. \quad (7.3)$$

Jelölés:  $\mathbf{A} \sim \mathbf{B}$ .

► Ha  $\mathbf{A}$  hasonló  $\mathbf{B}$ -hez, akkor  $\mathbf{B}$  is hasonló  $\mathbf{A}$ -hoz. Legyen ugyanis  $\hat{\mathbf{C}} = \mathbf{C}^{-1}$ . Akkor

$$\mathbf{A} = \mathbf{C} \mathbf{B} \mathbf{C}^{-1} = (\mathbf{C}^{-1})^{-1} \mathbf{B} \mathbf{C}^{-1} = \hat{\mathbf{C}}^{-1} \mathbf{B} \hat{\mathbf{C}}.$$

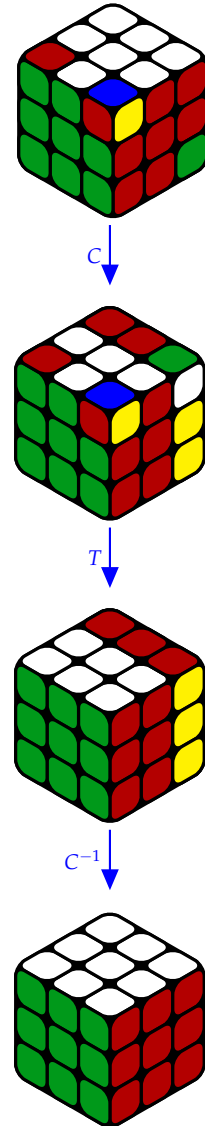
Így tehát mondható az, hogy  $\mathbf{A}$  és  $\mathbf{B}$  hasonlóak, mivel a hasonlóság szimmetrikus reláció.

► Például  $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \sim \begin{bmatrix} -6 & 4 \\ -9 & 6 \end{bmatrix}$ , ugyanis

$$\begin{bmatrix} -6 & 4 \\ -9 & 6 \end{bmatrix} = \begin{bmatrix} -2 & -1 \\ -3 & -2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 3 & -2 \end{bmatrix}, \quad \left( \mathbf{C} = \begin{bmatrix} -2 & 1 \\ 3 & -2 \end{bmatrix} \right).$$

► A  $\mathbf{B} = \mathbf{C}^{-1} \mathbf{A} \mathbf{C}$  összefüggés ekvivalens az  $\mathbf{C} \mathbf{B} = \mathbf{A} \mathbf{C}$  összefüggéssel, amit még egyszerűbb lehet ellenőrizni. Példánk esetében

$$\begin{bmatrix} -2 & 1 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} -6 & 4 \\ -9 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 3 & -2 \end{bmatrix} \quad \left( = \begin{bmatrix} 3 & -2 \\ 0 & 0 \end{bmatrix} \right).$$



7.13. ábra: A felső kockán látható három csúcskockát kell kicserélni. Először a három csúcsot egy síkba mozgatjuk ( $\mathbf{C}$  transzformáció), majd az előző ábrán látható  $\mathbf{T}$  transzformációval kicseréljük őket, végül  $\mathbf{C}$  inverzének alkalmazása után minden kocka a helyére kerül. Így a transzformációk egymás utáni elvégzését szorzásnak tekintve a megoldást a  $\mathbf{C} \mathbf{T} \mathbf{C}^{-1}$  transzformáció adja.



► A  $\mathbf{C}^{-1}\mathbf{A}\mathbf{C}$  alakú kifejezést az  $\mathbf{A}$  mátrix  $\mathbf{C}$ -vel való *konjugáltjának* nevezik. A konjugált más algebrai struktúrákban is fontos szerepet kap. Példaként véges halmazok permutációinak struktúráját említjük. Ekkor a permutációk közti művelet a permutációk egymás után való elvégzése, aminek eredményeként megintcsak a halmaz egy permutációját kapjuk. Konkrétan a Rubik kockán mutatjuk meg a konjugált szerepét a 7.12 és a 7.13 ábrák segítségével.

**7.15. TÉTEL (HASONLÓ MÁTRIXOK HATÁSA).** *Két mátrix pontosan akkor hasonló, ha van két olyan bázis, melyekben e két mátrix ugyanannak a lineáris leképezésnek a mátrixa.*

**BIZONYÍTÁS.** Ha  $\mathbf{A}$  és  $\mathbf{B}$  hasonlók, azaz  $\mathbf{B} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$ , akkor  $\mathbf{C}$ -t, mint a  $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n\}$  bázisról az  $\mathcal{E}$  standard bázisra való áttérés mátrixát tekintve azt kapjuk, hogy

$$\mathbf{B} = \mathbf{C}_{\mathcal{E} \leftarrow \mathcal{C}}^{-1} \mathbf{A} \mathbf{C}_{\mathcal{E} \leftarrow \mathcal{C}}.$$

Eszerint ha  $L$  az  $\mathbf{A}$  mátrixhoz tartozó mátrixleképezés, azaz  $\mathbf{A}$  az  $L$  mátrixa a standard bázisban, akkor  $\mathbf{B}$  az  $L$  mátrixa a  $\mathcal{C}$  bázisban. A fordított állítást a bevezetőben igazoltuk.  $\square$

**7.16. TÉTEL (HASONLÓSÁGRA INVARIÁNS TULAJDONSÁGOK).** *Ha  $\mathbf{A}$  és  $\mathbf{B}$  hasonló mátrixok, azaz  $\mathbf{A} \sim \mathbf{B}$ , akkor*

- $r(\mathbf{A}) = r(\mathbf{B})$ ,
- $\dim(\mathcal{N}(\mathbf{A})) = \dim(\mathcal{N}(\mathbf{B}))$ ,
- $\det(\mathbf{A}) = \det(\mathbf{B})$ ,
- $\text{trace}(\mathbf{A}) = \text{trace}(\mathbf{B})$ .

**BIZONYÍTÁS.** A bizonyítás során föltesszük, hogy valamely invertálható  $\mathbf{C}$  mátrixszal  $\mathbf{A} = \mathbf{C}^{-1}\mathbf{B}\mathbf{C}$ .

- Mivel mátrix rangja megegyezik az oszloptér dimenziójával, az oszloptér pedig megegyezik a mátrixleképezés képterével, ami a két mátrixra azonos, ezért a rangok is megegyeznek.
- Hasonlóan bizonyítható az is, hogy a két mátrix nullterének dimenziója megegyezik, hisz a nulltér megegyezik a lineáris leképezés magterével, ami pedig közös.
- $\det(\mathbf{A}) = \det(\mathbf{C}^{-1}\mathbf{B}\mathbf{C}) = \det(\mathbf{C}^{-1}) \det(\mathbf{B}) \det(\mathbf{C}) = \det(\mathbf{B})$ , mivel  $\det(\mathbf{C}) \det(\mathbf{C}^{-1}) = 1$ .
- Bizonyítása (ld. ?? feladat) egyszerűbb lesz a sajátértékek használatával.  $\square$

► Egy fontos következménye e tételnek, hogy a rang fogalma természetes módon átvihető véges dimenziós terek közötti lineáris leképezésekre. A lineáris  $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$  leképezés *rangján* képterének dimenzióját értjük, azaz  $r(L) = \dim(\text{Im}(L))$ .

A latin eredetű *invariáns* szó jelentése: *átalakulás közben változatlanul maradó*. Matematikában valamilyen művelet, átalakítás, leképezés során változatlanul maradó kifejezés, mennyiség, érték. Esetünkben a bázis megváltoztatása után is változatlanul maradó mennyiségeket jelenti.

## Feladatok

### Mátrixleképezések

Döntsük el, hogy az alábbi leképezések mátrixleképezések-e! Amennyik igen, annak írjuk fel a mátrixát! Legyen  $\mathbf{a} = (a_1, a_2, a_3)$  egy tetszőleges vektor.

7.1.  $A : \mathbf{x} \mapsto \mathbf{a} \cdot \mathbf{x},$

7.2.  $A : \mathbf{x} \mapsto \mathbf{a} + \mathbf{x},$

7.3.  $A : \mathbf{x} \mapsto \frac{1}{2}(\mathbf{a} \times \mathbf{x}),$

7.4.  $A : \mathbf{x} \mapsto \mathbf{a}(\mathbf{a} \cdot \mathbf{x}).$

7.5. MÁTRIXLEKÉPEZÉSEK KÖZTI MŰVELETEK Bizonyítsuk be a 7.2. tétel állításait!

7.6. INVERZ MÁTRIXLEKÉPEZÉSEK Bizonyítsuk be a 7.3. tételt!

### Lineáris leképezések

7.7. LINEÁRIS LEKÉPEZÉS EKVIVALENS DEFINÍCIÓI Igazoljuk a 7.8. tétel állításainak ekvivalenciáját!

Döntsük el, hogy az alábbi leképezések lineáris leképezések-e!

7.8.  $A : (x, y) \mapsto (x + 2y, x - y).$

7.9. Legyen  $\mathcal{P}_3$  a legfőbb 3-adfokú polinomok halmaza, és legyen  $D : \mathcal{P}_3 \rightarrow \mathcal{P}_3 : p(x) \mapsto p'(x).$

7.10. Legyen  $\mathcal{D}_{[0,1]}$  a  $[0, 1]$  intervallumon differenciálható függvények halmaza, és  $\mathcal{F}_{[0,1]}$  a  $[0, 1]$  intervallumon értelmezett függvények halmaza. Legyen továbbá  $A : \mathcal{D}_{[0,1]} \rightarrow \mathcal{F}_{[0,1]}; f(x) \mapsto xf'(x).$

### Lineáris leképezés mátrixa

#### Hasonló mátrixok

7.11. Mutassuk meg, hogy a nyom invariáns a mátrixok hasonlóságára nézve!

### Alkalmazás: differenciálhatóság\*

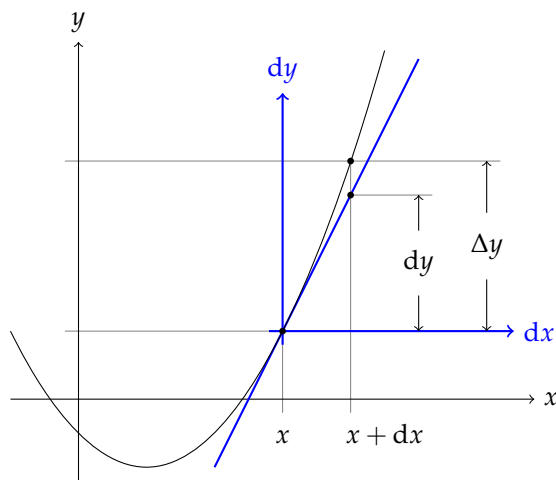
A lineáris leképezés fogalma az alkalmazott matematika sok területén bukkan föl, aminek az az egyik oka, hogy tetszőleges vektor-vektor függvény differenciálhatósága azt jelenti, hogy létezik a függvény megváltozását „jól közelítő” lineáris leképezés.

**Vektor-vektor függvények differenciálhatósága** Az  $\mathbb{R}^n$ -ből  $\mathbb{R}^m$ -be képző lineáris leképezések egy igen fontos alkalmazása a vektor-vektor függvények differenciálhatóságának fogalma.

A differenciálhatóság szokásos definíciója a következő: azt mondjuk, hogy az  $f: \mathbb{R} \rightarrow \mathbb{R}$  függvény differenciálható az  $x$  helyen, ha létezik és véges a

$$D = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

határérték. A  $D$  számnak fontos jelentése van: az  $f$  függvény  $x$  körüli megváltozása jól közelíthető a  $dx \mapsto D dx$  függvény 0 körüli megváltozásával. Szemléltetve ez azt jelenti, hogy ha az  $f$  grafikonján az  $(x, f(x))$  pontra helyezünk egy  $dx$  és  $dy$  változójú koordinátarendszert, akkor a  $dx \mapsto dy = D dx$  grafikonja az  $f$  függvény grafikonjának érintője (ld. a 7.14 ábrát). Eszerint, kicsit leegyszerűsítve a megfogalmazást, a differenciálhatóság azt jelenti, hogy a függvény „jól közelíthető” egy  $\mathbb{R} \rightarrow \mathbb{R}$  lineáris leképezéssel, hisz a  $dx \mapsto D dx$  leképezés ilyen.



7.14. ábra: A  $dx$  és  $dy$  koordinátarendszereket és a  $dy = D dx$  függvény grafikonját színezéssel kiemeltük. Az ábra egyúttal a  $\Delta y \approx dy$  kapcsolatot is szemlélteti.

A „jól közelítés” szemléletesen azt jelenti, hogy az  $f$  grafikonjára „zoomolva”, azaz azt folyamatosan nagyítva, a grafikon kiegyenesedni

látszik. Ez az az egyenes, melyet a grafikon érintőjének nevezünk, és amelynek  $dy = D dx$  az egyenlete az új koordinátarendszerben.

Ez a definíció ekvivalens módon átfogalmazható: azt mondjuk, hogy az  $f : \mathbb{R} \rightarrow \mathbb{R}$  függvény *differentiálható* az  $x$  helyen, ha van olyan  $D$  szám, hogy

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x) - Dh}{h} = 0.$$

Ez utóbbi alak azzal az előnnyel is jár, hogy könnyen általánosítható. Az általánosítás legfőbb nehézsége az, hogy a vektorral való osztás nem definiálható megfelelően, ezért e formulán még egy apró, de még mindig ekvivalens változtatást teszünk: nem  $h$ -val, hanem annak abszolút értékével osztunk:

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x) - Dh}{|h|} = 0.$$

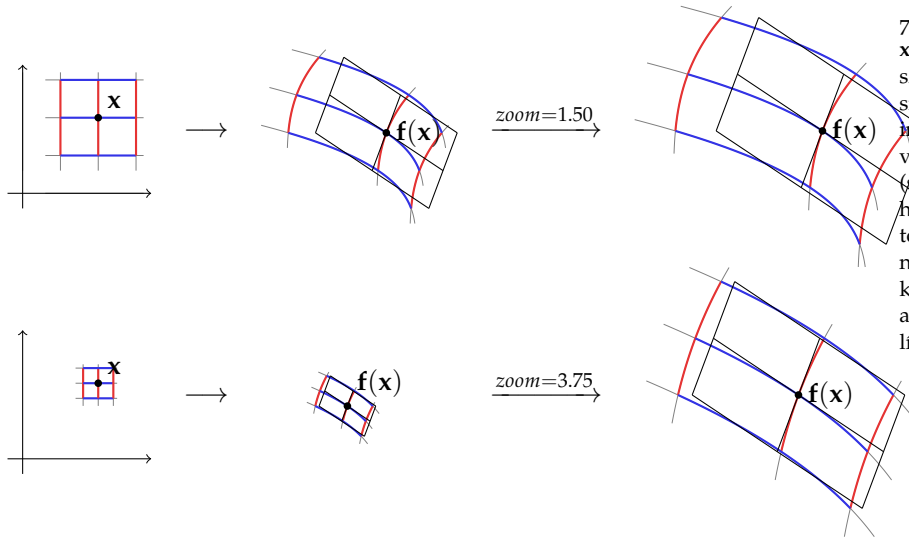
Mindezek a következő definícióhoz vezetnek:

**7.17. DEFINÍCIÓ (DIFFERENCIÁLHATÓSÁG).** Azt mondjuk, hogy az  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  függvény *differentiálható* az  $\mathbf{x}$  helyen, ha létezik olyan  $D_{f,\mathbf{x}} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  lineáris leképezés, melyre

$$\lim_{\mathbf{h} \rightarrow 0} \frac{f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) - D_{f,\mathbf{x}}\mathbf{h}}{|\mathbf{h}|} = 0.$$

A  $D_{f,\mathbf{x}}$  leképezést az  $f$  függvény  $\mathbf{x}$  ponthoz tartozó deriváltleképezésének nevezzük.

- A  $D_{f,\mathbf{x}}$  jelölés arra utal, hogy a deriváltleképezés az  $f$  függvénytől és az  $\mathbf{x}$  helytől is függ, maga viszont mint leképezés egy  $\mathbf{h}$  vektorhoz a  $D_{f,\mathbf{x}}\mathbf{h}$  vektort rendeli.
- Elterjedtebb a  $D_{\mathbf{x}}(f)$  jelölés, itt didaktikai okból választottunk olyat, mely jobban világossá teszi, hogy ez egy lineáris leképezés, mely majd hat valamely  $\mathbf{h}$  vektoron, és annak képe  $D_{\mathbf{x}}(f)\mathbf{h}$  vagy  $D_{\mathbf{x}}(f)(\mathbf{h})$  – az általunk használt jelölésben  $D_{f,\mathbf{x}}\mathbf{h}$ .
- Egy  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$  függvényen könnyen szemléltethető a derivált jelentése. Tekintsük az értelmezési tartomány egy négyzetrácsát, annak középpontja legyen  $\mathbf{x}$ . Tekintsük e rács képét az  $f$  függvény által, és a  $D_{f,\mathbf{x}}$  deriváltleképezés hatását e rácson, ha az origót  $\mathbf{x}$ -be tesszük. A rács méretét folyamatosan csökkentve, a képeket pedig arányosan fölnagyítva azt látjuk, hogy a két kép egyre jobban „összesimul” (ld. 7.15 ábra). Ez emlékeztet arra – bár nem tökéletesen analóg vele –, ahogy az egyváltozós függvény grafikonjának egy pontjára „zoomolva” a grafikon az érintőhöz közelít, rásimul.



7.15. ábra: Egy  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$  függvény egy  $\mathbf{x}$  pontban való differenciálhatóságának szemléltetésére tekintjük az értelmezési tartomány egyre sűrűbb négyzetrácsának az  $\mathbf{x}$  pontot körülvevő négyzeteit, valamint ezek  $\mathbf{f}$  függvény általi képét (színes rács), és a  $\mathbf{D}_{\mathbf{f},\mathbf{x}}$  deriváltleképezés hatását e rácson, ha az értelmezési tartományának origóját  $\mathbf{x}$ -be, értékkészletének origóját  $\mathbf{f}(\mathbf{x})$ -be tesszük. Az egyre kisebb képeket fölnagyítva látható, hogy a függvény általi kép egyre jobban közelít a deriváltleképezés általi képhez.

**Jacobi-mátrix** A deriváltleképezés mátrixa könnyen megkapható a koordinátafüggvények parciális deriváltjai segítségével.

**7.18. TÉTEL (JACOBI-MÁTRIX).** Ha az  $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m; (x_1, x_2, \dots, x_n) \mapsto (f_1, f_2, \dots, f_m)$  függvény differenciálható az  $\mathbf{x}$  helyen, akkor a lineáris  $\mathbf{D}_{\mathbf{f},\mathbf{x}}$  deriváltleképezés mátrixa a következő, ún. Jacobi-mátrix:

$$\mathbf{D}_{\mathbf{f},\mathbf{x}} = \frac{\partial(f_1, f_2, \dots, f_m)}{\partial(x_1, x_2, \dots, x_n)}(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}) & \frac{\partial f_2}{\partial x_2}(\mathbf{x}) & \dots & \frac{\partial f_2}{\partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}) & \frac{\partial f_m}{\partial x_2}(\mathbf{x}) & \dots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}) \end{bmatrix}$$

**BIZONYÍTÁS.** Ha  $\mathbf{f}$  differenciálható, akkor a definícióbeli határérték akkor is fönnáll, ha  $\mathbf{h}$  speciális módon tart a nullvektorhoz, például ha  $\mathbf{h} = t\mathbf{e}_j$ , és  $t \rightarrow 0$ . Ekkor

$$\lim_{t \rightarrow 0} \frac{\mathbf{f}(\mathbf{x} + t\mathbf{e}_j) - \mathbf{f}(\mathbf{x}) - \mathbf{D}_{\mathbf{f},\mathbf{x}}(t\mathbf{e}_j)}{|t|} = \mathbf{0}.$$

Az  $\mathbf{f}$  függvény  $i$ -edik koordinátafüggvénye  $f_i$ , a  $\mathbf{D}_{\mathbf{f},\mathbf{x}}(t\mathbf{e}_j)$  vektor  $i$ -edik koordinátája  $\mathbf{e}_i^T \mathbf{D}_{\mathbf{f},\mathbf{x}}(t\mathbf{e}_j)$ . Ennek alapján

$$\lim_{t \rightarrow 0} \frac{f_i(\mathbf{x} + t\mathbf{e}_j) - f_i(\mathbf{x}) - \mathbf{e}_i^T \mathbf{D}_{\mathbf{f},\mathbf{x}}(t\mathbf{e}_j)}{|t|} = 0.$$

Ez a határérték viszont már egy egyváltozós függvény deriváltja, ami nem más, mint az  $f_i$  függvény  $j$ -edik parciális deriváltja, ugyanis átrendezve az egyenlőséget és  $t$  előjelével is osztva kapjuk, hogy

$$\lim_{t \rightarrow 0} \frac{f_i(\mathbf{x} + t\mathbf{e}_j) - f_i(\mathbf{x})}{t} = \mathbf{e}_i^T \mathbf{D}_{\mathbf{f},\mathbf{x}} \mathbf{e}_j, \text{ azaz } \mathbf{e}_i^T \mathbf{D}_{\mathbf{f},\mathbf{x}} \mathbf{e}_j = \frac{\partial f_i}{\partial x_j}(\mathbf{x}).$$

Ez bizonyítja állításunkat.  $\square$

- A gyakorlatban az  $\mathbb{R}^n \rightarrow \mathbb{R}$  függvények, vagyis az  $n$ -változós skálárértékű függvények esetén az egyetlen sorból álló Jacobi-mátrix helyett annak vektoralakját használják, melyet *gradiensvektornak* neveznek, és  $\nabla f$ -fel jelölnek.
- Hasonlóképp, mivel az  $\mathbb{R} \rightarrow \mathbb{R}^n$  függvények Jacobi-mátrixa egyetlen oszlopból áll, gyakran használják annak vektoralakját. Ha például egy  $\mathbf{r} : \mathbb{R} \rightarrow \mathbb{R}^3; t \mapsto \mathbf{r}(t)$  függvény a térben mozgó tárgy mozgását az idő függvényében írja le, e vektor épp a mozgás sebességvektora.

**7.19. PÉLDA (JACOBI-MÁTRIX KISZÁMÍTÁSA).** *Határozzuk meg az alábbi függvények egy általános ponthoz és a megadott ponthoz tartozó Jacobi-mátrixát!*

1.  $f(x, y) = x^2y - xy^3 + 1, (x, y) = (0, 1)$ .
2.  $\mathbf{f}(x, y) = (-x^3/2 + y^3/8, x + y), (x, y) = (1, 1)$ .
3.  $\mathbf{r}(t) = (t^3, t^2, t), t = 2$ .
4.  $\mathbf{f}(x_1, x_2, x_3) = (2x_1 + 3x_2, x_1 - x_2 - x_3), (x_1, x_2, x_3) = (1, 2, 0)$ .

**MEGOLDÁS.** a)  $f(x, y) = x^2y - xy^3$ , parciális deriváltjai  $\frac{\partial}{\partial x}f(x, y) = 2xy - y^3$ ,  $\frac{\partial}{\partial y}f(x, y) = x^2 - 3xy^2$ . A deriváltleképezés mátrixa, azaz a Jacobi-mátrix itt

$$\begin{bmatrix} 2xy - y^3 & x^2 - 3xy^2 \end{bmatrix}$$

E mátrix vektor alakja, azaz a gradiensvektor

$$\nabla f(x, y) = (2xy - y^3, x^2 - 3xy^2).$$

Ennek értéke a  $(0, 1)$  helyen  $\nabla f(0, 1) = (-1, 0)$ , illetve a Jacobi-mátrix e helyen  $\begin{bmatrix} -1 & 0 \end{bmatrix}$ .

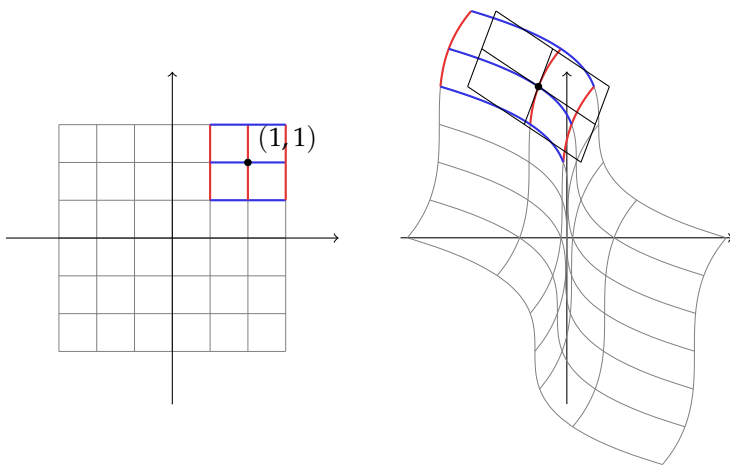
b) Az  $\mathbf{f}(x, y) = (-x^3/2 + y^3/8, x + y)$  függvény Jacobi-mátrixa és annak értéke a megadott  $(x, y) = (1, 1)$  pontban

$$\begin{bmatrix} -\frac{3}{2}x^2 & \frac{3}{8}y^2 \\ 1 & 1 \end{bmatrix}, \text{ illetve } \begin{bmatrix} -\frac{3}{2} & \frac{3}{8} \\ 1 & 1 \end{bmatrix}.$$

Például az első sor első eleme  $\frac{\partial}{\partial x}(-x^3/2 + y^3/8) = -\frac{3}{2}x^2$ . Az  $\mathbf{f}$  függvény deriváltleképezésének, vagyis Jacobi-mátrixának hatását szemlélteti a 7.16 és a 7.15 ábra.

c) Az  $\mathbf{r}(t) = (t^3, t^2, t)$  függvény Jacobi-mátrixa

$$\begin{bmatrix} 3t^2 \\ 2t \\ 1 \end{bmatrix}, \text{ ami a } t = 2 \text{ helyen } \begin{bmatrix} 12 \\ 4 \\ 1 \end{bmatrix}.$$



7.16. ábra: A bal ábra az  $f(x,y) = (-x^3/2 + y^3/8, x+y)$  függvény értelmezési tartományán megadott rácsot, és annak egy kis  $2 \times 2$ -es részét mutatja, melynek középpontja az  $(1,1)$  pont. Az alsó ábra egyrészt halványan jelöli e rács és színesen a kiemelt rács képét, valamint az  $(1,1)$  ponthoz tartozó deriváltleképezés hatását e kiemelt rácson.

A térben mozgó pont (test) mozgásának leírására is  $\mathbb{R} \rightarrow \mathbb{R}^3$  függvényt használunk. Ha e függvény egy ilyen mozgást ír le, akkor sebességvektora egy tetszőleges pontban

$$\dot{\mathbf{r}}(t) = (3t^2, 2t, 1),$$

a  $t = 2$  paraméterhez tartozó pontban  $\dot{\mathbf{r}}(2) = (12, 4, 1)$ .

d) Az utolsó példa fontos állítást szemléltet, nevezetesen azt, hogy egy lineáris leképezés deriváltja minden  $\mathbf{x}$  helyen megegyezik magával a leképezéssel, azaz a deriváltja önmaga. Világos, hogy a megadott leképezés egy lineáris leképezés, melynek mátrixszorzatos alakja:

$$\mathbf{f}(x_1, x_2, x_3) = \begin{bmatrix} 2 & 3 & 0 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Ennek Jacobi-mátrixa valóban bármely  $(x_1, x_2, x_3)$  helyen

$$\begin{bmatrix} 2 & 3 & 0 \\ 1 & -1 & -1 \end{bmatrix},$$

ugyanis az  $i$ -edik koordinátafüggvény  $j$ -edik parciális deriváltja épp az együtthatómátrix  $i$ -edik sor-,  $j$ -edik oszlopbeli eleme, azaz egy konstans. Így minden helyen e mátrix lesz a Jacobi-mátrix, speciálisan az  $(x_1, x_2, x_3) = (1, 2, 0)$  helyen is.  $\square$

**7.20. PÉLDA (FÜGGVÉNYÉRTÉK BECSLÉSE JACOBI-MÁTRIXSZAL).** Ismerjük egy differenciálható függvény értelmezési tartományának egy pontjához tartozó Jacobi-mátrixát és a függvényértéket ugyan ebben a pontban. Becsüljük meg a függvény értékét egy e ponthoz közeli helyen az alábbi adatok ismeretében!

$$1. \quad f(0,1) = 1, \mathbf{D}_{f,(0,1)} = \begin{bmatrix} -1 & 0 \end{bmatrix}, (x,y) = (-0.05, 1.1),$$

$$2. \mathbf{f}(1,1) = \left(-\frac{3}{8}, 2\right), \mathbf{D}_{\mathbf{f},(1,1)} = \begin{bmatrix} -3/2 & 3/8 \\ 1 & 1 \end{bmatrix}, (x,y) = (0.8, 1.1).$$

Mennyire lennének jók e becslések, ha a függvények az előző feladatbeli a) és b) függvényei lennének?

**MEGOLDÁS.** A függvény megváltozásának becsléséhez az  $\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x})$  értéket kell megbecsülni. A differenciálhatóság definíciója szerint erre a  $\mathbf{D}_{\mathbf{f},\mathbf{x}}\mathbf{h}$  mennyiség alkalmas, ha a függvény differenciálható az  $\mathbf{x}$  pontban. Eszerint tehát

$$\mathbf{f}(\mathbf{x} + \mathbf{h}) \approx \mathbf{f}(\mathbf{x}) + \mathbf{D}_{\mathbf{f},\mathbf{x}}\mathbf{h}.$$

E képletet felhasználva az alábbi megoldásokra jutunk:

a) E feladatban  $\mathbf{h} = (-0.05, 0.1)$ , így a függvény megváltozása a

$$\mathbf{D}_{f,(0,1)}\mathbf{h} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -0.05 \\ 0.1 \end{bmatrix} = 0.05$$

értékkel becsülhető, tehát a függvény értéke

$$f(\mathbf{x} + \mathbf{h}) = f(-0.05, 1.1) \approx f(0, 1) + \mathbf{D}_{f,(0,1)} \begin{bmatrix} -0.05 \\ 0.1 \end{bmatrix} = 1.05,$$

azaz  $f(-0.05, 1.1) \approx 1.05$ . Ha  $f$  az előző a) feladatbeli függvény, azaz  $f(x, y) = x^2y - xy^3 + 1$ , akkor a pontos érték  $f(-0.05, 0.1) = 1.0693$ .

b) Itt  $\mathbf{h} = (-0.2, 0.1)$ , így a függvény megváltozása a

$$\mathbf{D}_{\mathbf{f},(1,1)}\mathbf{h} = \begin{bmatrix} -\frac{3}{2} & \frac{3}{8} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -0.2 \\ 0.1 \end{bmatrix} = \begin{bmatrix} \frac{3}{2} \cdot \frac{2}{10} + \frac{3}{8} \cdot \frac{1}{10} \\ -\frac{2}{10} + \frac{1}{10} \end{bmatrix} = \begin{bmatrix} 0.3375 \\ -0.1 \end{bmatrix}$$

értékkel becsülhető, tehát a függvény értéke  $\mathbf{f}(0.8, 1.1) \approx \mathbf{f}(1, 1) + (0.3375, -0.1) = (-0.0375, 1.9)$ . Ha  $\mathbf{f}$  az előző b) feladatbeli függvény, azaz  $\mathbf{f}(x, y) = (-x^3/2 + y^3/8, x + y)$ , akkor a pontos érték  $\mathbf{f}(0.8, 1.1) = (-0.089625, 1.9)$ .  $\square$

**Jacobi-determináns és az integrál transzformációja** A 2- és 3-dimenziós tér leírására leggyakrabban használt koordinátarendszerek közötti váltás a többváltozós integrálok kiszámításában fontos szerepet kap. Az a kérdés, hogy az integrálközelítő összegben szereplő „téglányoknak” mennyi a mértékük. E szakasz kalkulus-előismereteket igényel.

Felidézzük a síkbeli polárkoordináta-rendszernek, a térbeli henger- és gömbi koordinátarendszereknek a derékszögű koordinátarendszerrel való kapcsolatát:

| (a) Polár              | (b) Henger             | (c) Gömbi                              |
|------------------------|------------------------|----------------------------------------|
| $x = r \cos \vartheta$ | $x = r \cos \vartheta$ | $x = \rho \sin \varphi \cos \vartheta$ |
| $y = r \sin \vartheta$ | $y = r \sin \vartheta$ | $y = \rho \sin \varphi \sin \vartheta$ |
|                        | $z = m$                | $z = \rho \cos \varphi$                |



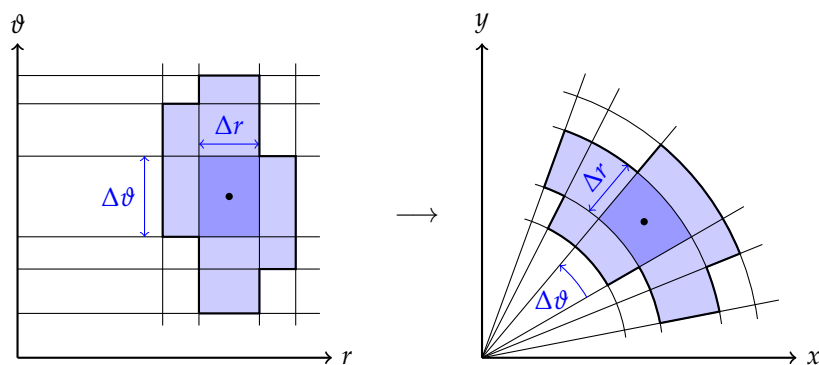
A felsorolt változók jelentése:  $r$  az  $xy$ -síkban az origótól való távolság,  $\rho$  a térben az origótól való távolság,  $\vartheta$  az  $x$ -tengely pozitív felével bezárt szög az  $xy$ -síkban,  $\varphi$  a  $z$ -tengely pozitív felével bezárt szög.

Jacobi-determinánsnak nevezzük egy  $\mathbb{R}^n \rightarrow \mathbb{R}^n$  függvény deriváltleképezésének determinánsát.

A síkbeli polárkoordináta-rendszerről a derékszögűre való áttérés egy  $\mathbb{R}^2 \rightarrow \mathbb{R}^2; (r, \vartheta) \mapsto (x, y)$  függvény, melyet a fönti (a)-beli képletek definiálnak. Ennek deriváltleképezése, pontosabban a leképezés  $\mathbf{D}$  mátrixa (szokás Jacobi-mátrixnak is hívni), és annak determinánsa, a Jacobi-determináns:

$$\mathbf{D} = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \vartheta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \vartheta} \end{bmatrix} = \begin{bmatrix} \cos \vartheta & -r \sin \vartheta \\ \sin \vartheta & r \cos \vartheta \end{bmatrix} \quad |\mathbf{D}| = \begin{vmatrix} \cos \vartheta & -r \sin \vartheta \\ \sin \vartheta & r \cos \vartheta \end{vmatrix} = r.$$

Az, hogy a Jacobi-determináns értéke  $r$ , azt jelenti, hogy egy „kicsiny”  $\Delta r \times \Delta \vartheta$  méretű téglány – melynek területe  $\Delta r \Delta \vartheta$  – a transzformáció után, azaz a polárkoordináta-rendszerben „nagyjából”  $r$ -szerese lesz az eredetinek, azaz  $r \Delta r \Delta \vartheta$ , ahol  $r$  a téglány egy pontjának origótól való távolsága. Ezt a leképezést a 7.17 ábrával szemléltetjük.



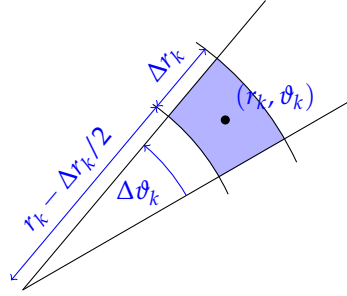
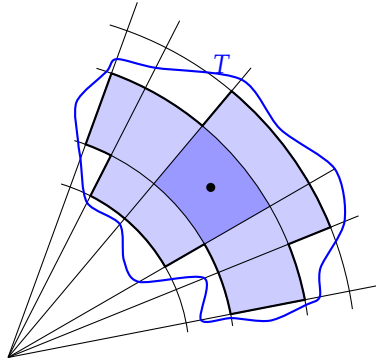
7.17. ábra: A síkbeli polárkoordináta-rendszerre való áttérést megadó leképezés szemléltetése egy téglányokból álló tartomány képének ábrázolásával.

Az  $r$ -szereződés geometriailag is könnyen igazolható, ahogy azt a 7.18 ábra mutatja. Kiszámoljuk egy polár-rendszerbeli téglány területét. Ez két körcikk területének különbsége. A nagyobbik sugara  $r_k + \Delta r_k/2$ , a határoló ív hossza  $(r_k + \Delta r_k/2)\Delta \vartheta_k$ , így területe  $\frac{1}{2}(r_k + \Delta r_k/2)^2 \Delta \vartheta_k$ . Hasonlóan kiszámolva a kisebbik körcikk területét, majd kivonva a nagyobbikéból kapjuk, hogy a téglány  $\Delta A_k$  területe

$$\Delta A_k = \frac{1}{2} \left( r_k + \frac{\Delta r_k}{2} \right)^2 \Delta \vartheta_k - \frac{1}{2} \left( r_k - \frac{\Delta r_k}{2} \right)^2 \Delta \vartheta_k = r_k \Delta r_k \Delta \vartheta_k.$$

Eszerint egy  $T$  tartományon értelmezett  $f(r, \vartheta)$  függvény integrálközelítő összege és annak határértéke, amint a legnagyobb átmérőjű téglány átmérője tart 0-hoz (ld. 7.19 ábra):

$$\sum_k f(r_k, \vartheta_k) \Delta A_k = \sum_k f(r_k, \vartheta_k) r_k \Delta r_k \Delta \vartheta_k \rightarrow \int_T f(r, \vartheta) r \, dr \, d\vartheta.$$

7.18. ábra: A síkbeli polárkoordináta-rendszer téglányának területe  $r_k \Delta r_k \Delta \theta_k$ .7.19. ábra: Egy  $T$  tartományba eső téglányok, és a  $k$ -adik téglány kiemelve.

A két térbeli koordinátarendszerre való áttérés hasonló módon való megértését és a leképezések elképzelését már az Olvasóra hagyjuk, de a leképezések deriváltjának determinánsát még fölírjuk. A hengerkoordináták esetén az  $(r, \vartheta, m) \mapsto (x, y, z)$  leképezésre ez

$$|\mathbf{D}| = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \vartheta} & \frac{\partial x}{\partial m} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \vartheta} & \frac{\partial y}{\partial m} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \vartheta} & \frac{\partial z}{\partial m} \end{vmatrix} = \begin{vmatrix} \cos \vartheta & -r \sin \vartheta & 0 \\ \sin \vartheta & r \cos \vartheta & 0 \\ 0 & 0 & 1 \end{vmatrix} = r.$$

A gömbi koordinátarendszer esetén a leképezés  $(\rho, \varphi, \vartheta) \mapsto (x, y, z)$ , amelynek Jacobi-determinánsa:

$$\begin{vmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \varphi} & \frac{\partial x}{\partial \vartheta} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \varphi} & \frac{\partial y}{\partial \vartheta} \\ \frac{\partial z}{\partial \rho} & \frac{\partial z}{\partial \varphi} & \frac{\partial z}{\partial \vartheta} \end{vmatrix} = \begin{vmatrix} \sin \varphi \cos \vartheta & \rho \cos \varphi \cos \vartheta & -\rho \sin \varphi \sin \vartheta \\ \sin \varphi \sin \vartheta & \rho \cos \varphi \sin \vartheta & \rho \sin \varphi \cos \vartheta \\ \cos \varphi & -\rho \sin \varphi & 0 \end{vmatrix} = \rho^2 \sin \varphi.$$

Így tehát az integrál kiszámításának képletei e három koordinátarendszerre:

Polár: 
$$\iint_T f(r, \vartheta) \, dA = \iint_T f(r, \vartheta) r \, dr \, d\vartheta$$

Henger: 
$$\iiint_T f(r, \vartheta, m) \, dV = \iiint_T f(r, \vartheta, m) r \, dm \, dr \, d\vartheta$$

Gömbi: 
$$\iiint_T f(\rho, \varphi, \vartheta) \, dV = \iiint_T f(\rho, \varphi, \vartheta) \rho^2 \sin \varphi \, d\rho \, d\varphi \, d\vartheta.$$

**Függvények kompozíciójának deriváltja** E paragrafusnak nem célja a függvényanalízis területére tartozó témák feldolgozása, de a többváltozós függvények kompozíciójának deriváltleképezése az egyváltozós függvények láncszabályához hasonló módon számolható, és erre érdemes egy pillantást vetnünk, mert a megoldást a deriváltleképezések kompozíciója, azaz a Jacobi-mátrixok szorzata adja.

Bizonyítás nélkül közöljük a következő tételt.

**7.21. TÉTEL (LÁNCSZABÁLY).** Legyen  $f : \mathbb{R}^k \rightarrow \mathbb{R}^m$ ,  $g : \mathbb{R}^n \rightarrow \mathbb{R}^k$  két függvény. Ha  $g$  differenciálható az  $x$  helyen, és  $f$  a  $g(x)$  helyen, akkor  $f \circ g$  differenciálható az  $x$  helyen, és deriváltleképezése, illetve annak mátrixa:

$$D_{f \circ g, x} = D_{f, g(x)} \circ D_{g, x}, \quad \text{illetve} \quad D_{f \circ g, x} = D_{f, g(x)} D_{g, x}.$$

**7.22. PÉLDA (LÁNCSZABÁLY).** Írjuk fel a láncszabály általános képleteit a megadott függvénytípusokra, az összetett függvény deriváltját pedig a láncszabállyal és behelyettesítéssel is számítsuk ki!

1.  $f : (x, y) \mapsto x^2 - y$ ,  $g : u \mapsto (u^2 + u, u - 1)$ ,  $u = 1$ .
2.  $f : \mathbb{R} \rightarrow \mathbb{R}^2; x \mapsto (x^2, x - 1)$ ,  $g : \mathbb{R}^2 \rightarrow \mathbb{R}; (u, v) \mapsto x = u^2 v$ ,  $(u, v) = (1, 2)$ .
3.  $f(x, y) = (xy^2 - 1, x - y)$ ,  $g(u, v) = (u + 1, u - v)$ ,  $(u, v) = (0, 1)$ .

**MEGOLDÁS.** Az  $a)$  esetben az  $f$ -hez, illetve  $g$ -hez tartozó láncszabály általános alakja

$$\frac{df}{du} = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{dg_1}{du} \\ \frac{dg_2}{du} \end{bmatrix} = \frac{\partial f}{\partial x} \frac{dg_1}{du} + \frac{\partial f}{\partial y} \frac{dg_2}{du},$$

a függvények parciális deriváltjait kiszámolva és a helyet megadva

$$\frac{df}{du}(1) = \begin{bmatrix} 2x & -1 \end{bmatrix}_{g(1)=(2,0)} \begin{bmatrix} 2u+1 \\ 1 \end{bmatrix}_{u=1},$$

végül a behelyettesítést is elvégezve:

$$\begin{bmatrix} 4 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 11.$$

Ugyanezt az eredményt kapjuk, ha a deriválás előtt elvégezzük a helyettesítést:  $(f \circ g)(u) = (u^2 + u)^2 - (u - 1) = u^4 + 2u^3 + u^2 - u + 1$ , ennek  $u$  szerinti deriváltja  $4u^3 + 6u^2 + 2u - 1$ , és ennek értéke az  $u = 1$  helyen 11.

A  $b)$  esetben  $\mathbf{f} : \mathbb{R} \rightarrow \mathbb{R}^2$ ,  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ , így  $\mathbf{f} \circ g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , és

$$\begin{bmatrix} \frac{\partial f_1}{\partial u} & \frac{\partial f_1}{\partial v} \\ \frac{\partial f_2}{\partial u} & \frac{\partial f_2}{\partial v} \end{bmatrix} = \begin{bmatrix} \frac{df_1}{dx} \\ \frac{df_2}{dx} \end{bmatrix} \begin{bmatrix} \frac{\partial g}{\partial u} & \frac{\partial g}{\partial v} \end{bmatrix} = \begin{bmatrix} \frac{df_1}{dx} \frac{\partial g}{\partial u} & \frac{df_1}{dx} \frac{\partial g}{\partial v} \\ \frac{df_2}{dx} \frac{\partial g}{\partial u} & \frac{df_2}{dx} \frac{\partial g}{\partial v} \end{bmatrix}$$

A megadott függvényekre és a helyettesítendő értékeket is megadva:

$$\begin{bmatrix} 2x \\ 1 \end{bmatrix}_{x=g(1,2)=2} \begin{bmatrix} 2uv & u^2 \end{bmatrix}_{u=1, v=2} = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \begin{bmatrix} 4 & 1 \end{bmatrix} = \begin{bmatrix} 16 & 4 \\ 4 & 1 \end{bmatrix}.$$

Behelyettesítés után a függvény  $(u, v) \mapsto (u^4v^2, u^2v - 1)$ , aminek deriváltja az  $(u, v) = (1, 2)$  helyen

$$\begin{bmatrix} 4u^3v^2 & 2u^4v \\ 2uv & u^2 \end{bmatrix}_{(1,2)} = \begin{bmatrix} 16 & 4 \\ 4 & 1 \end{bmatrix},$$

ami természetesen megegyezik az előző eredménnyel.

Végül a  $c)$  esetben az általános alak

$$\begin{bmatrix} \frac{\partial f_1}{\partial u} & \frac{\partial f_1}{\partial v} \\ \frac{\partial f_2}{\partial u} & \frac{\partial f_2}{\partial v} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{\partial g_1}{\partial u} & \frac{\partial g_1}{\partial v} \\ \frac{\partial g_2}{\partial u} & \frac{\partial g_2}{\partial v} \end{bmatrix}.$$

A parciális deriváltakat kiszámolva és a helyettesítési értékeket is megadva kapjuk, hogy

$$\begin{aligned} \begin{bmatrix} \frac{\partial f_1}{\partial u} & \frac{\partial f_1}{\partial v} \\ \frac{\partial f_2}{\partial u} & \frac{\partial f_2}{\partial v} \end{bmatrix}_{(0,1)} &= \begin{bmatrix} y^2 & 2xy \\ 1 & -1 \end{bmatrix}_{(1,-1)} \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}_{(0,1)} \\ &= \begin{bmatrix} 1 & -2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

Itt fölhasználtuk, hogy  $\mathbf{g}(0,1) = (1, -1)$ . Ha a deriválás előtt elvégezzük a függvények kompozícióját, akkor ugyanerre az eredményre jutunk, ugyanis

$$(\mathbf{f}(\mathbf{g}(u, v))) = ((u+1)(u-v)^2 - 1, v+1),$$

aminek a deriváltmátrixa

$$\begin{bmatrix} (u-v)^2 + 2(u+1)(u-v) & -2(u+1)(u-v) \\ 0 & 1 \end{bmatrix}_{(0,1)} = \begin{bmatrix} -1 & 2 \\ 0 & 1 \end{bmatrix}. \quad \square$$

## 2- és 3-dimenziós geometriai transzformációk mátrixa

A 7.9. tétel bizonyításában megkonstruáltuk egy lineáris  $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$  leképezéshez azt a mátrixot, mely ezt a leképezést generálja, nevezetesen bizonyítottuk, hogy minden  $\mathbf{x}$  vektorra

$$A\mathbf{x} = [A\mathbf{e}_1 | A\mathbf{e}_2 | \dots | A\mathbf{e}_n] \mathbf{x}.$$

Először vizsgáljunk meg néhány geometriailag leírható  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$  és  $\mathbb{R}^3 \rightarrow \mathbb{R}^3$  leképezést.

**Forgatás** Vizsgáljuk meg a síbeli pont körüli és a térbeli egyenes körüli forgatások mátrixát!

**7.23. ÁLLÍTÁS (A FORGATÁS MÁTRIXA).** *A sík vektorait egy pont körül  $\alpha$  szöggel elforgató leképezés mátrixa*

$$\begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}.$$

**BIZONYÍTÁS.** A 7.7. állítás szerint a forgatás lineáris leképezés, így van mátrixa, melynek alakja  $[A\mathbf{i} \ A\mathbf{j}]$ , ahol  $\mathbf{i}$  és  $\mathbf{j}$  jelöli az  $\mathbb{R}^2$  standard bázisának elemeit. E vektorokat szemlélteti a 7.20 ábra.

Az  $A\mathbf{i}$  vektor megegyezik  $\mathbf{i}$  elforgatottjával, amelynek ismerjük koordinátáit:  $A\mathbf{i} = \begin{bmatrix} \cos \alpha \\ \sin \alpha \end{bmatrix}$ . A  $\mathbf{j}$  vektor  $\alpha$  szöggel való elforgatottja megegyezik az  $A\mathbf{i}$  vektor  $\pi/2$  szöggel való elforgatottjával, azaz  $A\mathbf{j} = \begin{bmatrix} -\sin \alpha \\ \cos \alpha \end{bmatrix}$ . Így az  $A$ -hoz tartozó mátrix

$$A = [A\mathbf{i} \ A\mathbf{j}] = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}$$

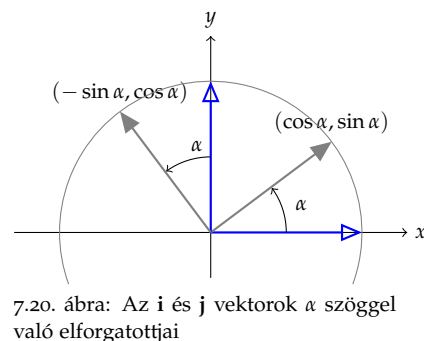
Tehát egy  $\mathbf{x}$  vektor  $\alpha$  szöggel való elforgatottja  $A\mathbf{x} = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \mathbf{x}$ .  $\square$

**7.24. PÉLDA (FORGATÁS EGY TETSZŐLEGES PONT KÖRÜL).** *Határozzuk meg a koordinátáit a  $(4,3)$  pont  $(2,1)$  körül  $\pi/3$  radiánnal való elforgatásával kapott pontnak!*

**MEGOLDÁS.** A forgatás középpontját toljuk az origóba, így a  $(4,3)$  pont a  $(4,3) - (2,1) = (2,2)$  pontba kerül. E pontot, illetve az oda mutató helyvektort forgassuk el  $\pi/3$  radiánnal, azaz  $60^\circ$ -kal. Ez a forgatás mátrixával való beszorzással megkapható:

$$\begin{bmatrix} \cos \frac{\pi}{3} & -\sin \frac{\pi}{3} \\ \sin \frac{\pi}{3} & \cos \frac{\pi}{3} \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 - \sqrt{3} \\ 1 + \sqrt{3} \end{bmatrix}$$

E pontot a  $(2,1)$  vektorral eltoljuk, hogy ne az origó, hanem a  $(2,1)$



7.20. ábra: Az  $\mathbf{i}$  és  $\mathbf{j}$  vektorok  $\alpha$  szöggel való elforgatottjai

pont körüli elforgatottat kapjuk meg:

$$\begin{bmatrix} 1 - \sqrt{3} \\ 1 + \sqrt{3} \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 - \sqrt{3} \\ 2 + \sqrt{3} \end{bmatrix} \quad \square$$

**7.25. PÉLDA (KOORDINÁTATENGELY KÖRÜLI FORGATÁS A TÉRBEN).** Írjuk fel a koordinátatengelyek körüli  $\alpha$  szöggel való forgatás mátrixát.

**MEGOLDÁS.** Tekintsük először a  $z$ -tengely körüli forgatást. Ekkor az  $\mathbf{i}$  és  $\mathbf{j}$  vektorok úgy transzformálódnak, mint a sík elforgatásánál, míg a  $\mathbf{k}$  vektor helyben marad, tehát a bázisvektorok így transzformálódnak:

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} \cos \alpha \\ \sin \alpha \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} -\sin \alpha \\ \cos \alpha \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Így a  $z$ -tengely körüli forgatás mátrixa:

$$\begin{bmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Hasonlóképp kapjuk az  $x$ - és az  $y$ -tengely körüli forgatás mátrixát is:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix}, \quad \begin{bmatrix} \cos \alpha & 0 & \sin \alpha \\ 0 & 1 & 0 \\ -\sin \alpha & 0 & \cos \alpha \end{bmatrix}.$$

Ez utóbbi mátrix előjelhibásnak tűnhet, de nem az, ha itt is a forgatás tengelyiránya felől nézve pozitív a forgásirány, azaz  $\mathbf{k}$ -t forgatjuk  $\mathbf{i}$ -be, és nem fordítva.  $\square$

**7.26. TÉTEL (TENGEY KÖRÜLI FORGATÁS – RODRIGUES-FORMULA).** Ha  $\mathbf{e} \in \mathbb{R}^3$  egységvektor, akkor az  $\mathbf{e}$  körüli  $\alpha$  szögű forgatás tetszőleges  $\mathbf{x}$  vektort az

$$\mathbf{x} \cos \alpha + (\mathbf{e} \times \mathbf{x}) \sin \alpha + \mathbf{e}(\mathbf{e} \cdot \mathbf{x})(1 - \cos \alpha) \quad (7.4)$$

vektorba visz. E leképezés mátrixa

$$\begin{aligned} \mathbf{R} &= \mathbf{I} + \sin \alpha [\mathbf{e}]_{\times} + (1 - \cos \alpha) [\mathbf{e}]_{\times}^2 \\ &= \mathbf{I} + \sin \alpha [\mathbf{e}]_{\times} + (1 - \cos \alpha) (\mathbf{e} \mathbf{e}^T - \mathbf{I}) \end{aligned} \quad (7.5)$$

**BIZONYÍTÁS.** Ha  $\mathbf{x}$  párhuzamos  $\mathbf{e}$ -vel, akkor elforgatottja önmaga, és valóban, ekkor  $(\mathbf{e} \times \mathbf{x}) = \mathbf{0}$  és  $\mathbf{e}(\mathbf{e} \cdot \mathbf{x}) = \mathbf{x}$ , így a (7.4) képlet  $\mathbf{x}$ -et ad eredményül.

A továbbiakban legyen tehát  $\mathbf{x}$  az  $\mathbf{e}$ -vel nem párhuzamos vektor. Jelölje  $\mathbf{x}_e$ -nek az  $\mathbf{e}$ -re eső merőleges vetületét  $\mathbf{x}_e$ , azaz legyen

$$\mathbf{x}_e = (\mathbf{e} \cdot \mathbf{x})\mathbf{e}.$$

Jelölje továbbá az  $\mathbf{x}$  vektornak az  $\mathbf{e}$ -re merőleges síkra eső merőleges vetületét  $\mathbf{x}_1$ , azaz

$$\mathbf{x}_1 = \mathbf{x} - (\mathbf{e} \cdot \mathbf{x})\mathbf{e}.$$

Végül legyen  $\mathbf{x}_2 = \mathbf{e} \times \mathbf{x}$ . Világos, hogy  $\mathbf{x}_1 \perp \mathbf{x}_2$ . E két vektor hossza:

$$|\mathbf{x}_1| = |\mathbf{x}| \sin \gamma,$$

$$|\mathbf{x}_2| = |\mathbf{e}||\mathbf{x}| \sin \gamma = |\mathbf{x}| \sin \gamma,$$

tehát  $|\mathbf{x}_1| = |\mathbf{x}_2|$ . Ha  $R$  jelöli a forgató leképezést, akkor

$$\begin{aligned} R\mathbf{x}_1 &= \mathbf{x}_1 \cos \alpha + \mathbf{x}_2 \sin \alpha \\ &= (\mathbf{x} - (\mathbf{e} \cdot \mathbf{x})\mathbf{e}) \cos \alpha + (\mathbf{e} \times \mathbf{x}) \sin \alpha. \end{aligned}$$

Mivel  $R\mathbf{x}_e = \mathbf{x}_e$ , és  $\mathbf{x} = \mathbf{x}_e + \mathbf{x}_1$ , ezért

$$\begin{aligned} R\mathbf{x} &= R\mathbf{x}_e + R\mathbf{x}_1 \\ &= (\mathbf{e} \cdot \mathbf{x})\mathbf{e} + (\mathbf{x} - (\mathbf{e} \cdot \mathbf{x})\mathbf{e}) \cos \alpha + (\mathbf{e} \times \mathbf{x}) \sin \alpha \\ &= \mathbf{x} \cos \alpha + (\mathbf{e} \times \mathbf{x}) \sin \alpha + \mathbf{e}(\mathbf{e} \cdot \mathbf{x})(1 - \cos \alpha). \end{aligned}$$

Ezzel igazoltuk a (7.4) formulát. A leképezés könnyen átírható mátrix-szorzat alakba:

$$\cos \alpha \mathbf{I} \mathbf{x} + [\mathbf{e}]_{\times} \mathbf{x} \sin \alpha + (1 - \cos \alpha)(\mathbf{e}\mathbf{e}^T)\mathbf{x},$$

így a forgatás  $\mathbf{R}$  mátrixa

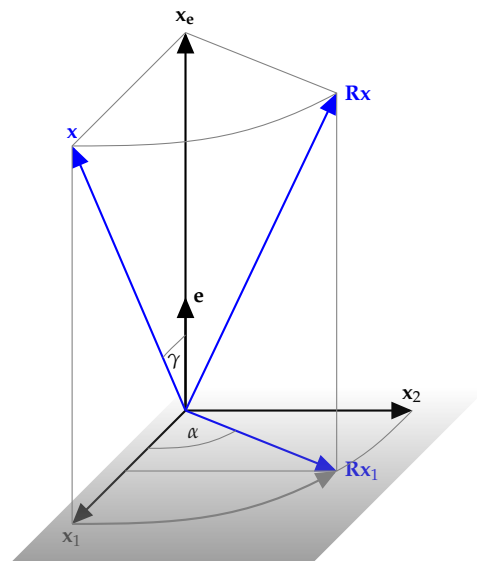
$$\mathbf{R} = \cos \alpha \mathbf{I} + \sin \alpha [\mathbf{e}]_{\times} + (1 - \cos \alpha)(\mathbf{e}\mathbf{e}^T).$$

Egyszerű számolással igazolható, hogy  $\mathbf{e}\mathbf{e}^T - \mathbf{I} = [\mathbf{e}]_{\times}^2$  (ld. ?? feladat), amiből azonnal adódnak a (7.5) képletei.  $\square$

**7.27. PÉLDA (FORGATÁS MÁTRIXA).** Írjuk fel annak a leképezésnek a mátrixát, mely az  $(2, 0, 1)$  vektor egyenese körül  $\alpha$  szöggel forgat, ahol  $\cos \alpha = \frac{2}{3}$ . Határozzuk meg a  $(3, 2, -1)$  vektor elforgatottját! Más eredményt kapnánk-e, ha a  $(-2, 0, -1)$  vektor egyenese körül kéne forgatnunk  $\alpha$  szöggel?

**MEGOLDÁS.** Az  $\mathbf{e} = \frac{1}{\sqrt{5}}(2, 0, 1)$  egységvektorral

$$[\mathbf{e}]_{\times} = \frac{1}{\sqrt{5}} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & -2 \\ 0 & 2 & 0 \end{bmatrix}, \quad [\mathbf{e}]_{\times}^2 = \frac{1}{5} \begin{bmatrix} -1 & 0 & 2 \\ 0 & -5 & 0 \\ 2 & 0 & -4 \end{bmatrix}$$



Így rövid számolás után a forgatás mátrixa

$$\begin{aligned} \mathbf{R} &= \mathbf{I} + \sin \alpha [\mathbf{e}]_{\times} + (1 - \cos \alpha) [\mathbf{e}]_{\times}^2 \\ &= \begin{bmatrix} 14/15 & -1/3 & 2/15 \\ 1/3 & 2/3 & -2/3 \\ 2/15 & 2/3 & 11/15 \end{bmatrix} = \frac{1}{15} \begin{bmatrix} 14 & -5 & 2 \\ 5 & 10 & -10 \\ 2 & 10 & 11 \end{bmatrix} \end{aligned}$$

és  $\mathbf{R} \cdot (3, 2, -1) = (2, 3, 1)$ .

A  $(-2, 0, -1)$  vektor körüli forgatással más eredményt kapnánk, hisz a forgatás iránya a vektor irányától is függ, és mivel az az ellenkezőjére változott, így a forgásirány is ellenkező irányú lesz.  $\square$

**7.28. PÉLDA (A FORGATÁS MÁTRIXÁNAK INVERZE).** *Határozzuk meg a síkot  $\alpha$  szöggel elforgató mátrix inverzét!*

**MEGOLDÁS.** Először megállapítjuk, hogy a forgatás mátrixa invertálható, ugyanis determinánsa nem 0, hiszen  $\begin{vmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{vmatrix} = \cos^2 \alpha + \sin^2 \alpha = 1$ . Az egyik lehetséges megoldás, hogy egyszerűen a  $2 \times 2$ -es mátrixok 5.13. tételben leírt invertálási technikáját alkalmazva határozzuk meg az inverzet, mely szerint

$$\begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}^{-1} = \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix}.$$

Egy másik megoldás arra épül, hogy a ?? állítás szerint két mátrix pontosan akkor inverze egymásnak, ha a hozzájuk tartozó lineáris leképezések is inverzei egymásnak. Az  $\alpha$  szöggel való elforgatásnak, mint leképezésnek az inverze nyilvánvalóan a  $-\alpha$  szöggel való elforgatás, tehát mátrixaik is egymás inverzei. Eszerint

$$\begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}^{-1} = \begin{bmatrix} \cos(-\alpha) & -\sin(-\alpha) \\ \sin(-\alpha) & \cos(-\alpha) \end{bmatrix} = \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix}. \quad \square$$

**Merőleges vetítés** A merőlegesség mind az elméleti matematika, mind az alkalmazások fontos fogalma.

**7.29. ÁLLÍTÁS (EGYENESRE VALÓ MERŐLEGES VETÍTÉS MÁTRIXA).** *A sík vagy a tér vektorait egy  $\mathbf{b}$  irányvektorú egyenesre merőlegesen vetítő leképezés mátrixa*

$$\mathbf{P} = \frac{1}{\mathbf{b}^T \mathbf{b}} \mathbf{b} \mathbf{b}^T. \quad (7.6)$$

*Speciálisan  $\mathbf{e}$  mátrix alakja*

$$\mathbf{P} = \mathbf{e} \mathbf{e}^T, \quad (7.7)$$

*ha az egyenes irányvektora az  $\mathbf{e}$  egységvektor.*



**BIZONYÍTÁS.** A 7.7. állítás szerint a merőleges vetítés lineáris leképezés, van tehát mátrixa. Az 1.24. tétel szerint ha  $\mathbf{x}$  egy tetszőleges vektor és  $\mathbf{e}$  egy egységvektor, akkor  $\mathbf{x}$ -nek a  $\mathbf{e}$  egyenesére eső merőleges vetülete

$$\text{proj}_{\mathbf{e}} \mathbf{x} = (\mathbf{x} \cdot \mathbf{e}) \cdot \mathbf{e}.$$

Ennek mátrixszorzással való átírása:

$$(\mathbf{x} \cdot \mathbf{e})\mathbf{e} = \mathbf{e}(\mathbf{e} \cdot \mathbf{x}) = \mathbf{e}(\mathbf{e}^T \mathbf{x}) = (\mathbf{e}\mathbf{e}^T)\mathbf{x},$$

tehát

$$\text{proj}_{\mathbf{e}} \mathbf{x} = (\mathbf{e}\mathbf{e}^T)\mathbf{x}.$$

Ebből kiolvasható, hogy az  $\mathbf{e}$  egységvektor-irányú egyenesre való merőleges vetítés mátrixa

$$\mathbf{P} = \mathbf{e}\mathbf{e}^T.$$

Ha  $\mathbf{b}$  egy tetszőleges zérustól különböző vektor, akkor az  $\mathbf{e} = \mathbf{b}/|\mathbf{b}|$  jelölés mellett  $\mathbf{P} = \mathbf{e}\mathbf{e}^T = \mathbf{b}\mathbf{b}^T/|\mathbf{b}|^2$ , ami  $|\mathbf{b}|^2 = \mathbf{b}^T\mathbf{b}$  behelyettesítésével bizonyítja a tételt.  $\square$

► A tételből következik, hogy a sík vektorait az  $x$ -tengellyel  $\alpha$  szöget bezáró egyenesre merőlegesen vetítő lineáris leképezés mátrixa

$$\mathbf{P} = \begin{bmatrix} \cos \alpha \\ \sin \alpha \end{bmatrix} \begin{bmatrix} \cos \alpha & \sin \alpha \end{bmatrix} = \begin{bmatrix} \cos^2 \alpha & \sin \alpha \cos \alpha \\ \sin \alpha \cos \alpha & \sin^2 \alpha \end{bmatrix}, \quad (7.8)$$

mivel ekkor  $\mathbf{e} = (\cos \alpha, \sin \alpha)$ .

**7.30. ÁLLÍTÁS (SÍKRA VALÓ MERŐLEGES VETÍTÉS MÁTRIXA).** A tér vektorait az  $\mathbf{n}$  normálvektorú síkra merőlegesen vetítő leképezés mátrixa

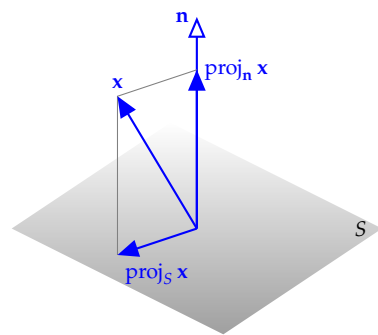
$$\mathbf{P} = \mathbf{I} - \mathbf{n}\mathbf{n}^T.$$

**BIZONYÍTÁS.** Egy tetszőleges  $\mathbf{x}$  vektornak a normálvektor egyenesére eső merőleges vetülete a 7.29. állítás szerint  $\text{proj}_{\mathbf{n}} \mathbf{x} = (\mathbf{n}\mathbf{n}^T)\mathbf{x}$ . Az  $\mathbf{n}$  normálvektorú  $S$  síkra eső merőleges vetületre  $\text{proj}_S \mathbf{x} = \mathbf{x} - \text{proj}_{\mathbf{n}} \mathbf{x} = \mathbf{x} - (\mathbf{n}\mathbf{n}^T)\mathbf{x}$  (lásd a 7.21. ábrát). Ebből következik, hogy a síkra való merőleges vetítés mátrixa  $\mathbf{I} - \mathbf{n}\mathbf{n}^T$ .

**7.31. PÉLDA (SÍKRA ESŐ MERŐLEGES VETÜLET KISZÁMÍTÁSA).** Határozzuk meg a  $(-2, 1, 3)$  vektornak a  $2x + y - 2z = 0$  egyenletű síkra eső merőleges vetületét! (ld. később a 7.40. példát)

**MEGOLDÁS.** A sík egy normálvektora  $(2, 1, -2)$ , így az egységnyi hosszú normálvektor  $\mathbf{n} = (2/3, 1/3, -2/3)$ . A  $\mathbf{P}$  vetítő mátrix

$$\mathbf{P} = \mathbf{I}_3 - \mathbf{n}\mathbf{n}^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{1}{9} \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} \begin{bmatrix} 2 & 1 & -2 \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{bmatrix}.$$



7.21. ábra: Vektor vetülete egy síkra

Így a  $(-2, 1, 3)$  vektor merőleges vetülete

$$\frac{1}{9} \begin{bmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}.$$

□

**Tükrözés** Vizsgáljuk meg síkban az egyenesre való és térben a síkra való tükrözés mátrixát!

**7.32. ÁLLÍTÁS (SÍKBELI TÜKRÖZÉS MÁTRIXA).** A sík vektorait az  $x$ -tengellyel  $\alpha/2$  szöget bezáró egyenesre tükröző lineáris leképezés mátrixa

$$\begin{bmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{bmatrix}.$$

**BIZONYÍTÁS.** a 7.7. állítás szerint a tükrözés lineáris leképezés. A vektorok tükörképe csak a tükrözés tengelyének állásától függ, ami most  $\alpha/2$ . Helyvektorokban gondolkodva a tükrözés tengelyének át kell mennie az origón.

A mellékelt ábráról leolvasható, hogy  $\mathbf{i}$  tükörképe  $A\mathbf{i} = \begin{bmatrix} \cos \alpha \\ \sin \alpha \end{bmatrix}$ , míg a  $\mathbf{j}$  vektoré  $A\mathbf{j} = \begin{bmatrix} \sin \alpha \\ -\cos \alpha \end{bmatrix}$ . Így a sík vektorait az első tengellyel  $\alpha/2$  szöget bezáró egyenesre tükröző leképezés mátrixa  $\begin{bmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{bmatrix}$ . □

A térben egy síkra való tükrözés feladata a síkra való vetítéshez hasonlóan adódik:

**7.33. ÁLLÍTÁS (SÍKRA VALÓ TÜKRÖZÉS MÁTRIXA).** Igazoljuk, hogy a tér vektorait az  $\mathbf{n}$  normálvektorú síkra tükröző leképezés mátrixa

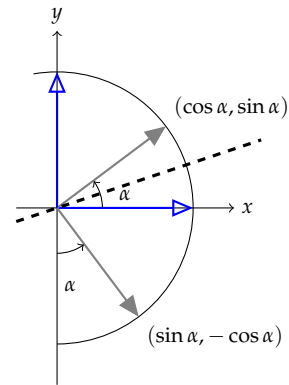
$$\mathbf{P} = \mathbf{I} - 2\mathbf{nn}^T.$$

**BIZONYÍTÁS.** A 7.30. állításhoz hasonlóan minden leolvasható a mellékelt 7.23. ábráról: ha  $\mathbf{x}$ -ből kivonjuk a  $\text{proj}_{\mathbf{n}} \mathbf{x}$  vektort, akkor a síkra eső vetületet kapjuk, így ha a kétszeresét vonjuk ki, a tükörképhez jutunk. E leképezés mátrixa az  $\mathbf{x} - 2(\mathbf{nn}^T)\mathbf{x} = (\mathbf{I} - 2\mathbf{nn}^T)\mathbf{x}$  összefüggésből adódik.

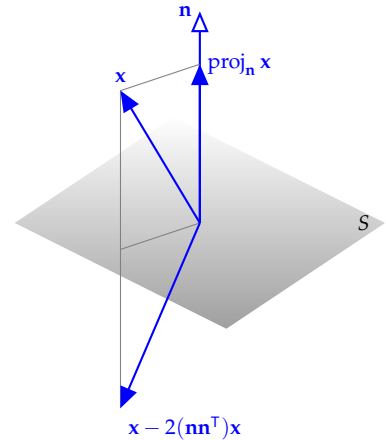
**Vetítés** Tárgyaltuk a merőleges vetítést. Vetíteni azonban másként is lehet.

**7.34. PÉLDA (VETÍTÉS SÍKRA).** Határozzuk meg annak a lineáris leképezésnek a mátrixát, mely a tér összes pontját az  $(1, -2, 1)$  vektorral párhuzamos irányban az  $x + y + 2z = 0$  egyenletű síkra vetíti.

**MEGOLDÁS.** Elemi geometriai eszközökkel könnyen látható, hogy e leképezés valóban lineáris. Világos, hogy a képtér az  $x + y + 2z = 0$



7.22. ábra: Az  $\mathbf{i}$  és  $\mathbf{j}$  vektorok egy egyenesre való tükröképe



7.23. ábra: Vektor tükörképe egy síkra

egyenletű sík összes vektora lesz. E teret megkapjuk a sík egyenletéből, ha azt mint egyenletrendszert megoldjuk. A megoldás  $(-s - 2t, s, t)$ , azaz e tér egy bázisa a  $(-1, 1, 0)$ , és a  $(-2, 0, 1)$  vektorokból áll. Könnyen látható az is, hogy a nulltérbe pontosan azok a vektorok tartoznak, amelyek párhuzamosak a vetítő vektorral, azaz az  $(1, -2, 1)$  vektorral. A vetítés  $\mathbf{P}$  mátrixa tehát eleget kell hogy tegyen az alábbi feltételeknek:

$$\mathbf{P} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{P} \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{P} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

E három feltétel egyetlen mátrixszorzásba foglalható:

$$\mathbf{P} \begin{bmatrix} -1 & -2 & 1 \\ 1 & 0 & -2 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & -2 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \text{ ahonnan } \mathbf{P} = \begin{bmatrix} 0 & -1 & -2 \\ 2 & 3 & 4 \\ -1 & -1 & -1 \end{bmatrix} \quad \square$$

► Az előző feladatban kapott  $\mathbf{P}$  mátrix eleget tesz a  $\mathbf{P}^2 = \mathbf{P}$  összefüggésnek. Ez szemléletesen világos is, hisz ha  $P$  jelöli a lineáris transzformációt, akkor arra is igaz, hogy  $P^2 = P$ . Ez abból következik, hogy a  $P$  vetítés az  $x + y + 2z = 0$  egyenletű sík minden vektorát helyben hagyja, másrészt bármely  $\mathbf{x}$  vektor esetén  $P\mathbf{x}$  ebben a síkban van, így a második vetítés már minden vektort helyben hagy.

**Eltolás** Az eltolás nem lineáris leképezés, hisz minden vektorhoz egy konstans vektort ad, tehát a nullvektort nem a nullvektorba képi. Egy szellemes ötlettel mégis megvalósítható lineáris leképezéssel.

El szeretnénk tolni a síkot egy  $(a, b)$  vektorral. Az ötlet az, hogy beágyazzuk a síkot a térbe, és ott keresünk egy olyan térbeli lineáris leképezést, amely ezt a síkot eltolja (hogy másutt meg mit csinál, nem is érdekes). Legyen tehát a vizsgált sík a  $z = 1$  egyenletű sík, és keressük azt a lineáris  $T$  leképezést, melyre

$$T \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x + a \\ y + b \\ 1 \end{bmatrix}.$$

Ez ugyan még mindig nem tűnik lineárisnak, de mivel  $z = 1$ , ezért a

$$T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x + az \\ y + bz \\ z \end{bmatrix}$$

leképezés már minden tekintetben megfelel. E leképezés mátrixa

$$\mathbf{T} = T \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \end{bmatrix} = \begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix}.$$

Hasonló ötlettel a tér eltolása is megvalósítható. A tér tetszőleges  $(x, y, z) \mapsto (x + a, y + b, z + c)$  eltolása megvalósítható a következő mátrixleképezéssel:

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & a \\ 0 & 1 & 0 & b \\ 0 & 0 & 1 & c \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{T} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & a \\ 0 & 1 & 0 & b \\ 0 & 0 & 1 & c \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x + a \\ y + b \\ z + c \\ 1 \end{bmatrix}.$$

## Feladatok

### Bizonyítások

**7.12.** Mutassuk meg, hogy tetszőleges  $H \subseteq \mathbb{R}^n$  vektorhalmazra a következő két állítás ekvivalens:

1. bármely véges sok  $H$ -beli vektor bármely lineáris kombinációja  $H$ -ban van;
2. bármely  $H$ -beli vektor tetszőleges skalárszorosa, és bármely két  $H$ -beli vektor összege  $H$ -ban van.

**7.13. LINEÁRIS FÜGGETLENSÉG EGY SZÜKSÉGES ÉS ELÉGSÉGES FELTÉTELE** A  $V$  vektorrendszer pontosan akkor lineárisan független, ha  $\text{span}(V)$  bármely vektora csak egyféleképp áll elő  $V$  lineáris kombinációjaként.

**7.14. SORTÉR ÉS NULLTÉR** A 2.37. példában megoldottuk a

$$x_1 + 2x_2 + x_3 + 2x_4 + x_5 = 0$$

$$x_1 + 2x_2 + 3x_3 + 3x_4 + x_5 = 0$$

$$3x_1 + 6x_2 + 7x_3 + 8x_4 + 3x_5 = 0$$

egyenletrendszert. A megoldása

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2s - \frac{3}{2}t - u \\ s \\ -\frac{1}{2}t \\ t \\ u \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} s + \begin{bmatrix} -\frac{3}{2} \\ 0 \\ -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} t + \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} u.$$

Ezt felhasználva fejben számolva adjunk meg egy olyan vektorrendszert, amely kifeszíti az

$$-2x_1 + x_2 = 0$$

$$-3x_1 - x_3 + 2x_4 = 0$$

$$-x_1 + x_5 = 0$$

homogén lineáris egyenletrendszer megoldásai alterét.

**7.15.** Általánosítsuk az előző feladat eredményét tetszőleges homogén lineáris egyenletrendszerre!

## Merőleges vetítés és a legjobb közelítés

A legjobb közelítés, a legkisebb négyzetek elve, vagy a lineáris regresszió az alkalmazásokban igen gyakran előforduló fontos fogalmak. Lényegük az  $\mathbb{R}^n$  egy alterére való merőleges vetítésének fogalmával jól megvilágítható.

**Alterek összege és direkt összege** A koordinátázás általánosításaként tekinthetünk arra a gondolatra, hogy a tér minden vektorát különböző alterekbe eső vektorok összegeként állítsuk elő, és egyértelműen. Ez vezet az alterek direkt összegének fogalmához.

Ha  $\mathcal{U}$  és  $\mathcal{V}$  ugyanannak a vektortérnek az alterei, akkor az egyesítésük által generált alteret  $\mathcal{U} + \mathcal{V}$ -vel jelöljük, és a két altér összegének nevezzük.

**7.35. ÁLLÍTÁS (ALTEREK ÖSSZEGE).** Ha  $\mathcal{U}$  és  $\mathcal{V}$  a  $\mathcal{W}$  altér két altere, akkor az egyesítésük által generált  $\mathcal{U} + \mathcal{V}$  altér pontosan azokból a vektorokból áll, melyek egy  $\mathcal{U}$ - és egy  $\mathcal{V}$ -beli vektor összegeként előállnak.

**BIZONYÍTÁS.** Ha  $\mathbf{x}$  egy  $\mathcal{U} + \mathcal{V}$ -beli vektor, akkor előáll néhány  $\mathcal{U}$ - és  $\mathcal{V}$ -beli vektor lineáris kombinációjaként. De a lineáris kombináció  $\mathcal{U}$ -beli vektorokat tartalmazó része egy  $\mathcal{U}$ -beli  $\mathbf{u}$  vektort ad, míg a többi egy  $\mathcal{V}$ -beli  $\mathbf{v}$  vektort, így  $\mathbf{x} = \mathbf{u} + \mathbf{v}$ . Fordítva, minden  $\mathbf{u} + \mathbf{v}$  alakú vektor  $\mathcal{U}$ - és  $\mathcal{V}$ -beli vektorok lineáris kombinációja, így benne van  $\mathcal{U} + \mathcal{V}$ -ben.  $\square$

Szemléltetésül: ha például  $\mathcal{W} = \mathbb{R}^3$ , és  $\mathcal{U}$  és  $\mathcal{V}$  egy-egy egymástól különböző 1-dimenziós altere, akkor az egyesítésük által generált 2-dimenziós altérbe pontosan azok a vektorok tartoznak, melyek egy  $\mathcal{U}$ -beli  $\mathbf{u}$  és egy  $\mathcal{V}$ -beli  $\mathbf{v}$  vektor összegei (ld. 7.24. ábra).

Legyen  $\mathcal{V}$  és  $\mathcal{W}$  az  $\mathcal{U}$  vektortér két tetszőleges altere. Azt mondjuk, hogy  $\mathcal{W}$  a  $\mathcal{V}$  kiegészítő altere, vagy komplementer altere vagy hogy  $\mathcal{V}$  és  $\mathcal{W}$  egymás kiegészítő (komplementer) alterei, ha

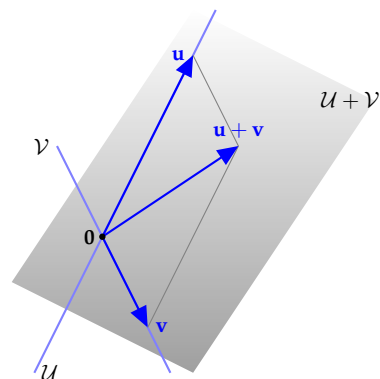
$$\mathcal{V} \cap \mathcal{W} = \{\mathbf{0}\}, \quad \mathcal{V} + \mathcal{W} = \mathcal{U},$$

azaz a két altérnek a zérusvektoron kívül nincs közös eleme, és  $\mathcal{U}$  minden vektora előáll  $\mathcal{V}$ - és  $\mathcal{W}$ -beli elemek összegeként!

E fogalom a sík koordinátázására emlékeztető dolog: a síkban az origón átmenő két koordinátatengely vektorai a két alteret adják, melyekben csak a zérusvektor közös, és a sík minden vektora (egyértelműen) előáll az egyikből és a másikkól vett vektor összegeként.

**7.36. TÉTEL (KIEGÉSZÍTŐ ALTEREK TULAJDONSÁGAI).** Legyen  $\mathcal{V}$  és  $\mathcal{W}$  az  $\mathcal{U}$  vektortér két altere és legyen  $\mathcal{V}$  egy bázisa  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r\}$ ,  $\mathcal{W}$  egy bázisa  $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k\}$ . Az alábbi állítások ekvivalensek:

a)  $\mathcal{V} \cap \mathcal{W} = \{\mathbf{0}\}$  és  $\mathcal{V} + \mathcal{W} = \mathcal{U}$ , azaz  $\mathcal{V}$  és  $\mathcal{W}$  kiegészítő alterek,



7.24. ábra:  $\mathcal{U} + \mathcal{V}$  bármely vektora előáll  $\mathbf{u} + \mathbf{v}$  alakban

- b)  $\mathcal{U}$  minden vektora egyértelműen előáll egy  $\mathcal{V}$ - és egy  $\mathcal{W}$ -beli vektor összegeként,  
 c)  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r\} \cup \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k\}$  az  $\mathcal{U}$  vektortér egy bázisa,  
 d)  $\mathcal{V} \cap \mathcal{W} = \{\mathbf{0}\}$  és  $\dim \mathcal{V} + \dim \mathcal{W} = \dim \mathcal{U}$ .

**BIZONYÍTÁS.**  $a) \Rightarrow b)$  : Meg kell mutatnunk, hogy minden vektor egyértelműen áll elő egy  $\mathcal{V}$ - és egy  $\mathcal{W}$ -beli vektor összegeként. Legyen  $\mathbf{u} \in \mathcal{U}$  olyan vektor, hogy  $\mathbf{u} = \mathbf{v}_1 + \mathbf{w}_1 = \mathbf{v}_2 + \mathbf{w}_2$ , ahol  $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{V}$  és  $\mathbf{w}_1, \mathbf{w}_2 \in \mathcal{W}$ . Átrendezés után  $\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{w}_2 - \mathbf{w}_1$ . Ennek bal oldalán  $\mathcal{V}$ -beli, jobb oldalán  $\mathcal{W}$ -beli vektor áll, amik csak a nullvektor esetén lehetnek azonosak, mivel  $\mathcal{V} \cap \mathcal{W} = \{\mathbf{0}\}$ . Így  $\mathbf{v}_1 = \mathbf{v}_2$  és  $\mathbf{w}_1 = \mathbf{w}_2$ .

$b) \Rightarrow c)$  : Mivel bármely  $\mathbf{u} \in \mathcal{U}$  vektor előáll  $\mathbf{v} + \mathbf{w}$  alakban, ahol  $\mathbf{v} \in \mathcal{V}$  és  $\mathbf{w} \in \mathcal{W}$ , e két vektor pedig előáll a bázisvektorok lineáris kombinációjaként, ezért a két bázis egyesítésével kapott vektorrendszer kifeszíti  $\mathcal{U}$ -t. Másrészt megmutatjuk, hogy a vektorrendszer független vektorokból áll. Tegyük fel, hogy

$$c_1 \mathbf{v}_1 + \dots + c_r \mathbf{v}_r + d_1 \mathbf{w}_1 + \dots + d_k \mathbf{w}_k = \mathbf{0}.$$

Mivel a  $\mathbf{0}$  vektor egyértelműen áll elő  $\mathbf{v} + \mathbf{w}$  alakban, és egy előállítása a  $\mathbf{0} + \mathbf{0}$ , ezért

$$c_1 \mathbf{v}_1 + \dots + c_r \mathbf{v}_r = \mathbf{0}, \text{ és } d_1 \mathbf{w}_1 + \dots + d_k \mathbf{w}_k = \mathbf{0}.$$

Innen pedig a bázisvektorok lineáris függetlenségéből következik, hogy minden együttható 0. Tehát a két altér bázisának egyesítése lineárisan független vektorokból áll, így  $\mathcal{U}$  egy bázisát adja.

$c) \Rightarrow d)$  :  $\mathcal{U}$  egy bázisa  $r + k$  elemű, így  $\dim \mathcal{U} = r + k = \dim \mathcal{V} + \dim \mathcal{W}$ .  $\mathcal{V} \cap \mathcal{W} = \{\mathbf{0}\}$  az előzőekben látottakhoz hasonlóan igazolható.

$d) \Rightarrow a)$  : Csak azt kell megmutatni, hogy ha  $\mathcal{V} \cap \mathcal{W} = \{\mathbf{0}\}$  és  $\dim \mathcal{V} + \dim \mathcal{W} = \dim \mathcal{U}$ , akkor  $\mathcal{U} = \mathcal{V} + \mathcal{W}$ . Ehhez legyen  $\mathcal{V}$  egy bázisa  $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ ,  $\mathcal{W}$  egy bázisa  $\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ . Ha egyesítésük bázis  $\mathcal{U}$ -ban, akkor kész vagyunk, hisz minden vektor e bázis vektorainak lineáris kombinációja, mely felbomlik  $\mathcal{V}$ -beli és  $\mathcal{W}$ -beli részre. Ezért tegyük fel, hogy e vektorok lineárisan összefüggők, azaz a nullvektor megkapható valamely nemtriviális lineáris kombinációjukkal:

$$c_1 \mathbf{v}_1 + \dots + c_r \mathbf{v}_r + d_1 \mathbf{w}_1 + \dots + d_k \mathbf{w}_k = \mathbf{0}.$$

Átrendezve

$$c_1 \mathbf{v}_1 + \dots + c_r \mathbf{v}_r = -d_1 \mathbf{w}_1 - \dots - d_k \mathbf{w}_k,$$

ami ellentmond a  $\mathcal{V} \cap \mathcal{W} = \{\mathbf{0}\}$  feltételnek, mivel indirekt feltevésünk szerint nem minden együttható 0.  $\square$

**7.37. DEFINÍCIÓ (DIREKT ÖSSZEG).** Ha a  $\mathcal{V}$  és  $\mathcal{W}$  alterek  $\mathcal{U}$  kiegészítő alterei, akkor azt mondjuk, hogy  $\mathcal{U}$  a  $\mathcal{V}$  és  $\mathcal{W}$  alterek direkt összege, amit az alterek egyszerű összegétől megkülönböztetendő  $\mathcal{V} \oplus \mathcal{W}$  jelöl.

Már láttunk példát kiegészítő alterekre, hisz a sortér és a nulltér dimenziójának összege  $n$ , és a két altérnek a nullvektoron kívül nincs közös eleme, így  $\mathcal{S}(\mathbf{A})$  és  $\mathcal{N}(\mathbf{A})$  kiegészítő alterek, azaz  $\mathbb{R}^n = \mathcal{S}(\mathbf{A}) \oplus \mathcal{N}(\mathbf{A})$  bármely valós  $m \times n$ -es  $\mathbf{A}$  mátrixra.

Egy  $\mathcal{W}$  altér esetén  $\mathcal{W}^\perp$  jelölte a  $\mathcal{W}$ -re merőleges vektorok alterét. Ezt merőleges kiegészítő altérnek neveztük, de azt, hogy ez valóban kiegészítő altér-e, még nem mutattuk meg.

**7.38. TÉTEL (A MERŐLEGES KIEGÉSZÍTŐ ALTÉR TULAJDONSÁGAI).** Legyen  $\mathcal{W}$  az  $n$ -dimenziós valós vagy komplex  $\mathcal{U}$  vektortér egy altere. Ekkor

- a)  $\mathcal{W} \cap \mathcal{W}^\perp = \{\mathbf{0}\}$ ,
- b)  $\mathcal{W} + \mathcal{W}^\perp = \mathcal{U}$ ,
- c)  $\mathcal{U}$  minden vektora egyértelműen előáll egy  $\mathcal{W}$ - és egy  $\mathcal{W}^\perp$ -beli vektor összegeként,
- d)  $(\mathcal{W}^\perp)^\perp = \mathcal{W}$ .

**BIZONYÍTÁS.** a) igaz, hisz ha  $\mathbf{x} \in \mathcal{W} \cap \mathcal{W}^\perp$ , akkor  $\mathbf{x} \cdot \mathbf{x} = 0$ , ami csak a  $\mathbf{0}$  vektorra áll fenn.

b) abból adódik, hogy az [előző](#), a [kiegészítő alterekről](#) szóló tétel szerint, ha két altér dimenzióinak összege  $n$ , és a két altér metszete csak a zérusvektorból áll, akkor a két altér összege  $\mathbb{R}^n$ . Esetünkben a két altér  $\mathcal{W}$  és  $\mathcal{W}^\perp$ . Ha  $\mathcal{W}$  egy bázisának vektoraiból, mint sorvektorokból mátrixot képzünk, annak sortere épp  $\mathcal{W}$ , nulltere  $\mathcal{W}^\perp$  lesz, és a sortér és nulltér dimenzióinak összege valóban  $n$  a [dimenziótétel](#) szerint.

Ugyancsak az [előző](#) tétel és az a) és b) állítások következménye, hogy a „merőleges kiegészítő alterek” valóban kiegészítő alterek, ami bizonyítja c)-t.

d) bizonyításához megmutatjuk, hogy  $\mathcal{W} \subseteq (\mathcal{W}^\perp)^\perp$  és  $\mathcal{W} \supseteq (\mathcal{W}^\perp)^\perp$ , ami bizonyítja, hogy  $\mathcal{W} = (\mathcal{W}^\perp)^\perp$ .

Legyen  $\mathbf{w}$  a  $\mathcal{W}$  tér egy tetszőleges vektora. Mivel  $\mathcal{W}^\perp$  épp azokból a vektorokból áll, melyek merőlegesek  $\mathcal{W}$  minden vektorára, ezért  $\mathbf{w}$  merőleges  $\mathcal{W}^\perp$  minden vektorára. Ez viszont épp azt jelenti, hogy  $\mathbf{w}$  benne van a  $(\mathcal{W}^\perp)^\perp$  altérben, tehát  $\mathcal{W} \subseteq (\mathcal{W}^\perp)^\perp$ .

A fordított tartalmazás bizonyításához legyen  $\mathbf{w} \in (\mathcal{W}^\perp)^\perp$ . A b) pont szerint e vektor előáll  $\mathbf{w} = \mathbf{v} + \mathbf{v}^\perp$  alakban, ahol  $\mathbf{v} \in \mathcal{W}$  és  $\mathbf{v}^\perp \in \mathcal{W}^\perp$ . Elég lenne megmutatnunk, hogy  $\mathbf{v}^\perp = \mathbf{0}$ . A  $(\mathcal{W}^\perp)^\perp$  és  $\mathcal{W}^\perp$  merőlegessége miatt  $\mathbf{w} \cdot \mathbf{v}^\perp = 0$ , így

$$0 = \mathbf{w} \cdot \mathbf{v}^\perp = \mathbf{w} \cdot (\mathbf{v} + \mathbf{v}^\perp) = \mathbf{v} \cdot \mathbf{v}^\perp + \mathbf{v}^\perp \cdot \mathbf{v}^\perp = \mathbf{v}^\perp \cdot \mathbf{v}^\perp,$$

hisz  $\mathbf{v} \cdot \mathbf{v}^\perp = 0$ . A  $\mathbf{v}^\perp \cdot \mathbf{v}^\perp = 0$  egyenlőség viszont csak  $\mathbf{v}^\perp = \mathbf{0}$  esetén



áll fönn. Így tehát  $\mathbf{w} = \mathbf{v}$ , azaz  $\mathbf{w} \in \mathcal{W}$ , ami bizonyítja az állítást.  $\square$

**Merőleges vetítés  $\mathbb{R}^n$  egy altérére** Azt mondjuk, hogy az  $\mathcal{V}$  vektortér egy  $\mathbf{v}$  vektorának a  $\mathcal{W} \leq \mathcal{V}$  altérre eső merőleges vetülete a  $\mathbf{w}$  vektor, ha  $\mathbf{w} \in \mathcal{W}$ , és  $\mathbf{v} - \mathbf{w}$  merőleges a  $\mathcal{W}$  altérre, azaz  $\mathbf{v} - \mathbf{w} \in \mathcal{W}^\perp$ . A  $\mathbf{v} - \mathbf{w}$  vektort a  $\mathbf{v}$  vektor  $\mathcal{W}$  altérre merőleges összetevőjének nevezzük.

Kérdés, hogy létezik-e minden vektornak egy altérre eső merőleges vetülete, és hogy egyértelmű-e. A 7.38. tétel c) pontja szerint ha  $\mathcal{W} \leq \mathcal{V}$ , akkor minden  $\mathbf{v} \in \mathcal{V}$  vektor egyértelműen felbomlik egy  $\mathcal{W}$ -beli  $\mathbf{w}$  és egy  $\mathcal{W}^\perp$ -beli  $\mathbf{w}^\perp$  vektor összegére. Ez azt jelenti, hogy a  $\mathbf{w}$  vektor épp a  $\mathbf{v}$  vektor  $\mathcal{W}$  altérre eső merőleges vetülete. Ezt az egyértelműen létező vektort – összhangban korábbi jelölésünkkel – jelölje  $\text{proj}_{\mathcal{W}} \mathbf{v}$ .

Egy mátrix teljes oszlopangú, ha oszlopai lineárisan függetlenek, azaz rangja megegyezik oszlopainak számával, azaz ha oszlopai az oszlopterének bázisát alkotják. Hasonlóan definiálható a teljes sorrangú mátrix fogalma.

**7.39. TÉTEL (ALTÉRRE VALÓ VETÍTÉS MÁTRIXA).** Ha  $\mathcal{W}$  az  $\mathbb{R}^n$  egy altéré, és az  $\mathbf{A}$  mátrix oszlopvektorai a  $\mathcal{W}$  egy bázisát alkotják (tehát  $\mathbf{A}$  teljes oszlopangú), akkor a  $\mathcal{W}$  altérre való merőleges vetítés, azaz a  $\text{proj}_{\mathcal{W}}$  leképezés mátrixa

$$\mathbf{A}(\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top.$$

**BIZONYÍTÁS.** Legyen a  $\mathbf{v} \in \mathbb{R}^n$  vektor  $\mathcal{W}$ -re eső merőleges vetülete  $\mathbf{w}$ . Mivel  $\mathbf{A}$  definíciója szerint  $\mathbf{A}$  oszloptere  $\mathcal{W}$ , ezért létezik olyan  $\mathbf{x}$  vektor, hogy  $\mathbf{Ax} = \mathbf{w}$ . Másrészt  $\mathcal{W} = \mathcal{O}(\mathbf{A})$  miatt  $\mathcal{W}^\perp = \mathcal{N}(\mathbf{A}^\top)$ , így  $\mathbf{v} - \mathbf{w}$  benne van  $\mathbf{A}^\top$  nullterében, mivel a merőleges vetület definíciója szerint  $\mathbf{v} - \mathbf{w}$  merőleges  $\mathcal{W}$ -re. Eszerint  $\mathbf{A}^\top(\mathbf{v} - \mathbf{w}) = \mathbf{0}$ , azaz  $\mathbf{A}^\top(\mathbf{v} - \mathbf{Ax}) = \mathbf{0}$ . Átrendezve kapjuk, hogy

$$\mathbf{A}^\top \mathbf{Ax} = \mathbf{A}^\top \mathbf{v}.$$

Az  $\mathbf{A}$  mátrix teljes oszlopangú, így a ?? tétel szerint  $\mathbf{A}^\top \mathbf{A}$  invertálható, azaz  $\mathbf{x} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{v}$ , amiből kapjuk, hogy  $\text{proj}_{\mathcal{W}} \mathbf{v} = \mathbf{w} = \mathbf{Ax} = \mathbf{A}(\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{v}$ , ami bizonyítja az állítást.  $\square$

► A tételbeli képlet könnyen megjegyezhető, hisz összhangban van az egyenesre való merőleges vetítés (7.6) képletével. Ha ugyanis az  $\mathbf{A}$  mátrix egyetlen oszlopból áll,  $(\mathbf{A}^\top \mathbf{A})^{-1}$  egyetlen szám, ami kiemelhető, azaz az  $\mathbf{A} = \mathbf{b}$  jelöléssel  $\mathbf{b}(\mathbf{b}^\top \mathbf{b})^{-1} \mathbf{b}^\top = \frac{1}{\mathbf{b}^\top \mathbf{b}} \mathbf{b} \mathbf{b}^\top$ .

**7.40. PÉLDA (MERŐLEGES VETÜLET KISZÁMÍTÁSA).** Határozzuk meg a  $(-2, 1, 3)$  vektornak az  $(1, 0, 1)$  és a  $(-1, 2, 0)$  vektorok által kifeszített síkra eső merőleges vetületét! (ld. még a 7.31. példát)

**MEGOLDÁS.** Az altér bázisvektoraiból képzett mátrix

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 0 & 2 \\ 1 & 0 \end{bmatrix}, \text{ amiből } \mathbf{A}(\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top = \frac{1}{9} \begin{bmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{bmatrix}.$$

Így a  $(-2, 1, 3)$  vektor merőleges vetülete

$$\frac{1}{9} \begin{bmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}.$$

Ez a feladat megegyezik a 7.31. példabelivel, mivel ennek a síknak is  $2x + y - 2z = 0$  az egyenlete, ugyanis  $(1, 0, 1) \times (-1, 2, 0) = (-2, -1, 2)$ .  $\square$

**Melyik mátrix merőleges vetítés mátrixa?** Olyan – könnyen ellenőrizhető – feltételeket keresünk egy lineáris leképezés mátrixára, melyek segítségével azonnal megállapítható, hogy a mátrixleképezés merőleges vetítés-e.

**7.41. TÉTEL (MERŐLEGES VETÍTÉS MÁTRIXAI).** Egy  $\mathbf{P}$  mátrix pontosan akkor merőleges vetítés mátrixa, ha  $\mathbf{P} = \mathbf{P}^\top = \mathbf{P}^2$ .

**BIZONYÍTÁS.** A  $\mathbf{P} = \mathbf{P}^2$  feltétel szükségessége szemléletesen világos, hisz minden  $P$  lineáris leképezés, mely az egész  $\mathbb{R}^n$  teret egy altérre – nevezetesen  $\text{Im } P$ -re – vetíti, az altér vektorait helyben hagyja. Tehát  $P^2\mathbf{x} = P\mathbf{x}$  minden  $\mathbf{x}$ -re fennáll, így ennek az összefüggésnek  $P$  minden mátrixára is igaznak kell lennie.

$(\implies)$  Tegyük fel, hogy  $\mathbf{P}$  egy  $P$  merőleges vetítés mátrixa  $\mathbb{R}^n$  standard bázisában. Tekintsük  $\text{Im}(P) = \mathcal{O}(\mathbf{P})$  egy tetszőleges bázisát, és legyen  $\mathbf{A}$  az a mátrix, melynek e bázis elemei az oszlopai. A 7.39. tétel szerint ekkor  $\mathbf{P} = \mathbf{A}(\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top$ . Erre viszont könnyen ellenőrizhető, a tételbeli feltétel.

$$\begin{aligned} \mathbf{P}^2 &= \left( \mathbf{A}(\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \right)^2 = \mathbf{A}(\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{A}(\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \\ &= \mathbf{A}(\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top = \mathbf{P}, \end{aligned}$$

másrészt

$$\begin{aligned} \mathbf{P}^\top &= \left( \mathbf{A}(\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \right)^\top = \mathbf{A} \left( (\mathbf{A}^\top \mathbf{A})^{-1} \right)^\top \mathbf{A}^\top \\ &= \mathbf{A}(\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top = \mathbf{P}. \end{aligned}$$

$(\impliedby)$  Tegyük fel, hogy  $\mathbf{P} = \mathbf{P}^\top = \mathbf{P}^2$ . Megmutatjuk, hogy  $\mathbf{P}$  az  $\mathcal{O}(\mathbf{P})$ -re való merőleges vetítés mátrixa. Ehhez elég megmutatnunk, hogy az  $\mathbf{x} - \mathbf{P}\mathbf{x}$  vektor merőleges  $\mathcal{O}(\mathbf{P})$ -re bármely  $\mathbf{x}$  vektor esetén. A  $\mathbf{P}^2 = \mathbf{P}$  feltétel miatt  $\mathbf{P}(\mathbf{x} - \mathbf{P}\mathbf{x}) = \mathbf{P}\mathbf{x} - \mathbf{P}^2\mathbf{x} = \mathbf{0}$ , tehát  $\mathbf{x} - \mathbf{P}\mathbf{x} \in \mathcal{N}(\mathbf{P})$ , de  $\mathbf{P} = \mathbf{P}^\top$ , így  $\mathbf{x} - \mathbf{P}\mathbf{x} \in \mathcal{N}(\mathbf{P}^\top)$ . Ez épp azt jelenti, hogy  $\mathbf{x} - \mathbf{P}\mathbf{x}$  merőleges  $\mathcal{O}(\mathbf{P})$ -re, és ezt akartuk belátni.  $\square$

- A  $\mathbf{P} = \mathbf{P}^T$  összefüggés azt jelenti, hogy  $\mathbf{P}$  szimmetrikus. A  $\mathbf{P}^2 = \mathbf{P}$  tulajdonságnak eleget tevő mátrixokat *idempotensnek* nevezzük. A tétel tehát úgy is fogalmazható, hogy egy mátrix pontosan akkor egy merőleges vetítés mátrixa, ha idempotens és szimmetrikus.
- Később látni fogjuk, hogy a – később definiálandó – nem feltétlenül merőleges vetítés mátrixai egybe esnek az idempotens mátrixokkal, tehát a vetítő lineáris leképezések az idempotens lineáris leképezésekkel.
- Azt, hogy egy vetítés hány dimenziós térre vetít, annak rangja mondja meg, hisz az megegyezik a képtér dimenziójával.

7.42. PÉLDA. Igazoljuk, hogy az

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad \frac{1}{4} \begin{bmatrix} 3 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ -1 & -1 & -1 & 3 \end{bmatrix}$$

mátrixok merőleges vetítés mátrixai! Hány dimenziós térre vetítenek?

**MEGOLDÁS.** Könnyen ellenőrizhető, hogy mindegyik mátrix szimmetrikus és idempotens, azaz kielégíti a  $\mathbf{P}^T = \mathbf{P}$  és a  $\mathbf{P}^2 = \mathbf{P}$  egyenlőségeket. Az első két mátrixról átalakítás nélkül is leolvasható, hogy rangjuk 2. A harmadik mátrix rangja 3, ugyanis egyrészt legalább 3, hisz ha kivonjuk az utolsó sort az első háromból, egy  $3 \times 3$ -as egységmátrixot kapunk benne, másrészt nem lehet 4, mert a négy sorvektor összege a zérusvektor, azaz lineárisan összefüggők.  $\square$

**Altértől való távolság** Adva van az  $\mathbb{R}^n$  tér egy  $\mathbf{x}$  vektora és egy  $\mathcal{W}$  altere.  $\mathbf{x}$ -nek a  $\mathcal{W}$  altértől való távolságán a  $\mathcal{W}$  altér  $\mathbf{x}$ -hez legközelebbi  $\mathbf{w}$  vektorának tőle való távolságát értjük. Kérdés azonban, hogy létezik-e ilyen vektor egyáltalán! Meg fogjuk mutatni, hogy ilyen  $\mathbf{w}$  vektor létezik és egyértelmű. E vektort az  $\mathbf{x}$  vektor  $\mathcal{W}$ -beli legjobb közelítésének nevezzük.

7.43. TÉTEL (LEGJOBB KÖZELÍTÉS TÉTELE). Adva van az  $\mathbb{R}^n$  tér egy  $\mathbf{x}$  vektora és egy  $\mathcal{W}$  altere. Az  $\mathbf{x}$  vektornak egyetlen  $\mathcal{W}$ -beli legjobb  $\hat{\mathbf{x}}$  közelítése van, nevezetesen  $\hat{\mathbf{x}} = \text{proj}_{\mathcal{W}} \mathbf{x}$ .

**BIZONYÍTÁS.** Legyen  $\mathbf{w}$  a  $\mathcal{W}$  egy tetszőleges vektora. Ekkor

$$\mathbf{x} - \mathbf{w} = (\mathbf{x} - \text{proj}_{\mathcal{W}} \mathbf{x}) + (\text{proj}_{\mathcal{W}} \mathbf{x} - \mathbf{w}).$$

A merőleges vetítés definíciója miatt az egyenlőség jobb oldalán álló első kifejezés  $\mathcal{W}^\perp$  eleme, míg a második  $\mathcal{W}$  eleme. Tehát az  $\mathbf{x} - \text{proj}_{\mathcal{W}} \mathbf{x}$  és a  $\text{proj}_{\mathcal{W}} \mathbf{x} - \mathbf{w}$  vektorok merőlegesek egymásra, így alkalmazható rájuk Pithagorász tétele:

$$|\mathbf{x} - \mathbf{w}|^2 = |\mathbf{x} - \text{proj}_{\mathcal{W}} \mathbf{x}|^2 + |\text{proj}_{\mathcal{W}} \mathbf{x} - \mathbf{w}|^2.$$

Ebből világos, hogy

$$|\mathbf{x} - \mathbf{w}|^2 \geq |\mathbf{x} - \text{proj}_{\mathcal{W}} \mathbf{x}|^2,$$

és egyenlőség csak akkor állhat fenn, ha  $\mathbf{w} = \hat{\mathbf{x}} = \text{proj}_{\mathcal{W}} \mathbf{x}$ , ami egyúttal a legjobb közelítés egyértelműségét is bizonyítja.  $\square$

► E tétel egyik következménye, hogy  $\mathbb{R}^n$  minden vektora felbontható egy  $\mathcal{W}$ -beli és egy rá merőleges vektor összegére, ugyanis

$$\mathbf{x} = \text{proj}_{\mathcal{W}} \mathbf{x} + \mathbf{w}^\perp, \text{ ahol } \mathbf{w}^\perp = \mathbf{x} - \text{proj}_{\mathcal{W}} \mathbf{x}.$$

Ennél azonban több is igaz, nevezetesen az, hogy e felbontás egyértelmű.

**7.44. TÉTEL (VEKTOR FELBONTÁSA ÖSSZETEVŐKRE).** *Adva van az  $\mathbb{R}^n$  tér egy  $\mathbf{x}$  vektora és egy  $\mathcal{W}$  altere. Az  $\mathbf{x}$  vektor egyértelműen felbomlik egy  $\mathcal{W}$ -beli  $\mathbf{w}$  és egy  $\mathcal{W}$ -re merőleges  $\mathbf{w}^\perp$  vektor összegére, nevezetesen  $\mathbf{w} = \text{proj}_{\mathcal{W}} \mathbf{x}$  és  $\mathbf{w}^\perp = \mathbf{x} - \mathbf{w}$ .*

**BIZONYÍTÁS.** Tegyük fel, hogy létezik  $\mathbf{x}$ -nek egy másik ilyen tulajdonságú felbontása is, tehát  $\mathbf{x} = \mathbf{w} + \mathbf{w}^\perp$  és  $\mathbf{x} = \mathbf{v} + \mathbf{v}^\perp$ . A második egyenletet az elsőből kivonva, majd átrendezve kapjuk, hogy

$$\mathbf{v} - \mathbf{w} = \mathbf{w}^\perp - \mathbf{v}^\perp.$$

A bal oldal eleme  $\mathcal{W}$ -nek a jobb oldali vektor viszont merőleges rá, hisz mindkét vektor eleme a  $\mathcal{W}^\perp$  altérnek. Ez viszont csak akkor állhat fenn, ha mindkét oldal egyenlő a zérusvektorral, tehát  $\mathbf{v} = \mathbf{w}$ .  $\square$

**7.45. PÉLDA.** *Tekintsük az  $\mathbb{R}^4$  tér  $(1, -1, 1, 0)$  és  $(0, 1, -1, 0)$  vektorai által kifeszített  $\mathcal{W}$  alterét és legyen  $\mathbf{x} = (8, 4, 2, 1)$ . Bontsuk fel az  $\mathbf{x}$  vektort  $\mathcal{W}$ -be eső és  $\mathcal{W}$ -re merőleges vektorok összegére.*

**MEGOLDÁS.** A  $\mathcal{W}$ -re való merőleges vetítés mátrixa  $\mathbf{P} = \mathbf{W}(\mathbf{W}^\top \mathbf{W})^{-1} \mathbf{W}^\top$ , ahol  $\mathbf{W}$  két oszlopa a megadott két bázisvektor, tehát

$$\mathbf{W} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 1 & -1 \\ 0 & 0 \end{bmatrix}, \text{ amiből } \mathbf{P}\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & -1/2 & 0 \\ 0 & -1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 8 \\ 4 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 1 \\ -1 \\ 0 \end{bmatrix}.$$

Így  $\text{proj}_{\mathcal{W}} \mathbf{x} = \mathbf{P}\mathbf{x} = (8, 1, -1, 0)$  és  $\mathbf{x} - \text{proj}_{\mathcal{W}} \mathbf{x} = (0, 3, 3, 1)$ . Egyszerű számítással ellenőrizhető, hogy a  $(8, 1, -1, 0) \in \mathcal{W}$  és hogy  $(0, 3, 3, 1) \perp \mathcal{W}$ , azaz merőleges a  $\mathcal{W}$ -t kifeszítő bázisvektorok mindegyikére.  $\square$

**Egyenletrendszer optimális megoldása** Az altérre való merőleges vetítés és a legjobb közelítés fogalmával olyan eszközhöz jutottunk, amellyel a lineáris egyenletrendszerek elmélete e szinten teljessé tehető. A gyakorlatban rendkívül gyakran előfordul, hogy az ismeretlen mennyiségek meghatározására méréseket végzünk, de az elkerülhetetlen mérési hibák ellenmondó egyenletrendszerre vezetnek. Hogyan határozható meg ekkor a valóságban bizonyosan létező megoldás, egy ellentmondásos, tehát nem megoldható egyenletrendszerből?

Tudjuk, hogy az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszer pontosan akkor oldható meg, ha  $\mathbf{b}$  benne van az oszloptérben, azaz  $\mathcal{O}(\mathbf{A})$ -ban. Természetes ötlet, hogy  $\mathbf{b}$  helyett az azt legjobban közelítő oszloptérbeli  $\hat{\mathbf{b}} = \text{proj}_{\mathcal{O}(\mathbf{A})} \mathbf{b}$  vektossal oldjuk meg az egyenletrendszert. Ez már biztosan megoldható lesz, és olyan megoldásokat szolgáltat, melyekre  $\mathbf{Ax}$  ugyan nem lesz egyenlő  $\mathbf{b}$ -vel, de attól a lehető legkisebb távolságra van. Az ilyen megoldásokat az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszer *optimális megoldásainak* vagy a *legkisebb négyzetek elve szerinti megoldásainak* nevezzük. Világos, hogy ha egy egyenletrendszer konzisztens, akkor optimális megoldásai megegyeznek a megoldásaival. E definícióból azt is látjuk, mit tegyünk, ha egy egyenletrendszer ellentmondásos (azaz inkonzisztens): határozzuk meg a  $\hat{\mathbf{b}} = \text{proj}_{\mathcal{O}(\mathbf{A})} \mathbf{b}$  vektort, és az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszer helyett oldjuk meg az  $\mathbf{Ax} = \hat{\mathbf{b}}$  egyenletrendszert. Ez kiindulásul jó, de adódik egy egyszerűbb módszer is.

Az természetes ötlet, hogy  $\mathbf{b}$  helyett  $\hat{\mathbf{b}}$ -vel oldjuk meg az egyenletrendszert, de vajon nincs-e jobb ötlet, végülis miért épp e merőleges vetület adja számunkra a „legjobb megoldást” és mit is jelent itt a „legjobb”. Az igazi választ a valószínűségszámítási előismeretet igénylő Gauss–Markov-tétel adja meg, melyet a feladatok közt ismertetünk.

**7.46. TÉTEL (EGYENLETRENDSZER OPTIMÁLIS MEGOLDÁSA).** Az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszer optimális megoldásai megegyeznek az

$$\mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b} \quad (7.9)$$

egyenletrendszer megoldásaival. Ezek közül egyetlen egy esik az  $\mathbf{A}$  mátrix sorterébe, a legkisebb abszolút értékű.

► A (7.9) egyenletrendszert az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszerhez tartozó *normálegyenlet-rendszernek* nevezzük. (A *normálegyenlet* kifejezés is helyes, ha a (7.9) kifejezésre, mint mátrixegyenletre gondolunk.)

**BIZONYÍTÁS.** Az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszer optimális megoldásai megegyeznek az  $\mathbf{Ax} = \text{proj}_{\mathcal{O}(\mathbf{A})} \mathbf{b}$  egyenletrendszer megoldásaival. Ezeket fogjuk tehát keresni.

Először megmutatjuk, hogy ha  $\hat{\mathbf{x}}$  egy optimális megoldás, akkor  $\hat{\mathbf{x}}$  kielégíti a (7.9) egyenletet. Mivel  $\mathbf{b} - \text{proj}_{\mathcal{O}(\mathbf{A})} \mathbf{b}$  a vetítés definíciója miatt merőleges  $\mathcal{O}(\mathbf{A})$ -ra, ezért  $\mathbf{A}^T$  nullterében van, tehát

$$\mathbf{A}^T (\mathbf{b} - \text{proj}_{\mathcal{O}(\mathbf{A})} \mathbf{b}) = \mathbf{0}.$$

Másrészt felhasználva, hogy  $\mathbf{Ax} = \text{proj}_{\mathcal{O}(\mathbf{A})} \mathbf{b}$ , kapjuk, hogy

$$\mathbf{A}^T (\mathbf{b} - \mathbf{Ax}) = \mathbf{0},$$

azaz átrendezés után

$$\mathbf{A}^T \mathbf{A} \hat{\mathbf{x}} = \mathbf{A}^T \mathbf{b}.$$

Ezután megmutatjuk, hogy a (7.9) egyenletet kielégítő minden  $\hat{\mathbf{x}}$  vektor optimális megoldás. Ha (7.9) teljesül, akkor

$$\mathbf{A}^T (\mathbf{b} - \mathbf{A} \hat{\mathbf{x}}) = \mathbf{0},$$

tehát  $\mathbf{b} - \mathbf{A} \hat{\mathbf{x}}$  benne van  $\mathbf{A}^T$  nullterében, így merőleges  $\mathbf{A}$  oszlopterére. Ezért a

$$\mathbf{b} = \mathbf{A} \hat{\mathbf{x}} + (\mathbf{b} - \mathbf{A} \hat{\mathbf{x}})$$

felbontás két merőleges kiegészítő altérbe eső vektor, hisz  $\mathbf{A} \hat{\mathbf{x}}$  az  $\mathbf{A}$  oszlopterébe esik. Így a [merőleges vetület definíciója](#) szerint

$$\mathbf{A} \hat{\mathbf{x}} = \text{proj}_{\mathcal{O}(\mathbf{A})} \mathbf{b},$$

azaz  $\hat{\mathbf{x}}$  optimális megoldás.

Végül meg kell mutatnunk, hogy a megoldások közt egyetlen van, mely  $\mathbf{A}$  sorterébe esik. Ez azonnal következik abból, hogy a normál-egyenlet megoldásai közt egyetlen egy van, mely  $\mathbf{A}^T \mathbf{A}$  sorterébe esik, az  $\mathbf{A}^T \mathbf{A}$  és  $\mathbf{A}$  sorterei pedig megegyeznek.  $\square$

**Lineáris és polinomiális regresszió** Az egyenletrendszerek optimális megoldásainak egyik fontos alkalmazása a lineáris regresszió. Tegyük fel, hogy két változó mennyiség, az  $x$  és az  $y$  között fennáll az  $y = a + bx$  kapcsolat. Méréseket végzünk, melyek eredménye az  $(x_i, y_i)$  ( $i = 1, 2, \dots, n$ ) párok sorozata. Keressük az  $a$  és  $b$  értékét, mely kielégíti az  $y_i = a + bx_i$  ( $i = 1, 2, \dots, n$ ) egyenletek mindegyikét! Ez egy kétismeretlenes lineáris egyenletrendszer, melynek mátrixalakja:

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

A hozzá tartozó normálegyenlet-rendszer

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},$$

amely a mátrixműveletek elvégzése után a következő alakra vezet:

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$

Ennek  $\hat{a}$  és  $\hat{b}$  megoldása adja az eredeti egyenletrendszer optimális megoldását! Az ilyen módon kapott  $y = \hat{a} + \hat{b}x$  egyenest *regressziós egyenesnek* nevezzük, mely a megadott adatokra a legkisebb négyzetek elve szerinti legjobban illeszkedő egyenes.

Összefoglalva:

7.47. ÁLLÍTÁS (LINEÁRIS REGRESSZIÓ). Az  $(x_i, y_i)$  ( $i = 1, 2, \dots, n$ ) párokhoz tartozó,  $y = \hat{a} + \hat{b}x$  egyenletű regressziós egyenes paraméterei kielégítik az

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$

egyenletet. Ez egyértelműen megoldható, ha van legalább két különböző  $x_i$  érték.

**BIZONYÍTÁS.** Az összefüggést már fent igazoltuk, csak a egyértelmű megoldhatóság igazolása maradt hátra. A számtani és négyzetes közép közti összefüggés szerint bármely  $x_i$  ( $i = 1, 2, \dots, n$ ) valósokra

$$\frac{x_1 + \dots + x_n}{n} \leq \sqrt{\frac{x_1^2 + \dots + x_n^2}{n}},$$

és egyenlőség csak akkor állhat fenn, ha  $x_1 = \dots = x_n$ . Mivel az együtthatómátrix determinánsa  $n \sum x_i^2 - (\sum x_i)^2$ , ezért a számtani és négyzetes közép közti összefüggés miatt ez csak akkor lehet 0, ha az  $x_i$  értékek mind azonosak.  $\square$

A lineáris regresszió gyakran egyéb függvénykapcsolat esetén is alkalmazható:

7.48. ÁLLÍTÁS (LINEARIZÁLHATÓ REGRESSZIÓS MODELLEK). Ha az  $x$  és  $y$  mennyiségek között az alábbi táblázat szerinti függvénykapcsolatok valamelyike áll, akkor a táblázatban megadott helyettesítéssel a kapcsolat  $Y = a + bX$  alakúvá, azaz lineárisrá válik, így lineáris regresszió végezhető.

| Modell          | Függvénykapcsolat | Helyettesítés                               |
|-----------------|-------------------|---------------------------------------------|
| hatványfüggvény | $y = cx^b$        | $X = \ln x \quad Y = \ln y \quad a = \ln c$ |
| exponenciális   | $y = ce^{bx}$     | $X = x \quad Y = \ln y \quad a = \ln c$     |
| logaritmikus    | $y = a + b \ln x$ | $X = \ln x \quad Y = y$                     |

**BIZONYÍTÁS.** Az  $y = cx^b$  egyenlőség mindkét oldalának logaritmusát véve az  $\ln y = \ln c + b \ln x$  egyenlőséget kapjuk, ami a megadott helyettesítésekkel az  $Y = a + bX$  kifejezést adja. Ugyanígy, az  $y = ce^{bx}$  egyenlet logaritmusát véve az  $\ln y = \ln c + bx$  egyenletet kapjuk. A szükséges helyettesítés a harmadik esetben még nyilvánvalóbb.  $\square$

A regresszió hasonló módon más függvényekkel is végezhető, ezek közül a polinomiálist emeljük ki:

**7.49. PÉLDA.** Keressünk az  $a_0 + a_1x + \dots + a_kx^k$  polinom együtthatóira optimális becslést a legkisebb négyzetek módszerével, ha az  $(x_i, y_i)$  ( $i = 1, 2, \dots, n$ ) párok sorozatát ismerjük.

**MEGOLDÁS.** Keresendő az  $n$  egyenletből álló  $k+1$ -ismeretlenes

$$\begin{aligned} a_0 + a_1x_1 + \dots + a_kx_1^k &= y_1 \\ a_0 + a_1x_2 + \dots + a_kx_2^k &= y_2 \\ \vdots & \quad \quad \quad \vdots \\ a_0 + a_1x_n + \dots + a_kx_n^k &= y_n \end{aligned}$$

egyenletrendszer megoldása az  $a_0, a_1, \dots, a_k$  ismeretlenekre. Mátrixalakja

$$\begin{bmatrix} 1 & x_1 & \dots & x_1^k \\ 1 & x_2 & \dots & x_2^k \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \dots & x_n^k \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Ha az együtthatómátrixot  $\mathbf{X}$  jelöli, az ismeretlenek vektorát  $\mathbf{a}$ , az  $y_i$  értékek vektorát  $\mathbf{y}$ , akkor az egyenletrendszer az  $\mathbf{Xa} = \mathbf{y}$  alakba írható. Ez biztosan megoldható, mégpedig egyértelműen, ha az  $x_i$  értékek különbözőek, és  $n = k+1$ , ekkor ugyanis az együtthatómátrix négyzetes, determinánsa Vandermonde-determináns, melynek értéke nem 0. Egyéb esetekben a normálegyenlet-rendszert kell felírni, melynek mátrixszorzatos alakja

$$\mathbf{X}^T \mathbf{Xa} = \mathbf{X}^T \mathbf{y}.$$

Ez egyértelműen megoldható, ha  $\mathbf{X}$  teljes oszloprangú, mert akkor  $\mathbf{X}^T \mathbf{X}$  invertálható, így a megoldás

$$\mathbf{a} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}.$$

Ez pontosan akkor áll fenn, ha van legalább  $k+1$  különböző  $x_i$  érték, ekkor ugyanis  $\mathbf{X}$ -ben van egy  $(k+1) \times (k+1)$ -es nemnulla determinánsú részmátrix, nevezetesen a különböző  $x_i$  értékekhez tartozó sorokból álló Vandermonde-mátrix.  $\square$

**Vetítés** Ha  $\mathcal{V}$  és  $\mathcal{W}$  kiegészítő alterek, akkor természetes módon értelmezhető a  $\mathcal{V}$  altérre való és a  $\mathcal{W}$  altérrel párhuzamos vetítés fogalma. E transzformációk halmaza egybeesik a  $P^2 = P$  feltételt kielégítő lineáris transzformációk halmazával.

**7.50. DEFINÍCIÓ (VETÍTÉS ALTÉRRE).** Tudjuk, hogy ha  $\mathbb{R}^n = \mathcal{V} \oplus \mathcal{W}$ , azaz  $\mathcal{V}$  és  $\mathcal{W}$  kiegészítő alterek, akkor a tér bármely  $\mathbf{u}$  vektora egyértelműen előáll  $\mathbf{u} = \mathbf{v} + \mathbf{w}$  alakban, ahol  $\mathbf{v} \in \mathcal{V}$ ,  $\mathbf{w} \in \mathcal{W}$ . Azt mondjuk, hogy a  $\mathbf{v}$  vektor az  $\mathbf{u}$  vektornak a  $\mathcal{V}$  altérre  $\mathcal{W}$  mentén való vetülete, vagy  $\mathcal{W}$ -vel párhuzamosan vett,  $\mathcal{V}$ -re való vetülete.



- Természetesen ugyanígy a  $\mathbf{w}$  vektor az  $\mathcal{U}$ -val párhuzamosan vett,  $\mathcal{W}$ -re való vetülete. Könnyen látható, hogy a  $P : \mathbf{u} \mapsto \mathbf{v}$  leképezés lineáris transzformáció (ellenőrizzük!). E lineáris transzformációt *vetítésnek* vagy *projekciónak* nevezzük.
- Világos, hogy minden  $P$  vetítés az  $\text{Im } P$ -re  $\text{Ker } P$  mentén való vetítés.

Határozzuk meg e transzformáció mátrixát! A 7.36. tétel utáni megjegyzés szerint  $\mathcal{V}$  és  $\mathcal{W}$  dimenzióinak összege  $n$ , és ha  $\mathcal{V}$  egy bázisa  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r\}$ ,  $\mathcal{W}$  egy bázisa  $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{n-r}\}$ , akkor a két bázis diszjunkt (metszetük üres) és egyesítésük az egész tér egy bázisa. E vektorokból képezzük az alábbi mátrixot:

$$\mathbf{U} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_r | \mathbf{w}_1 \ \mathbf{w}_2 \ \dots \ \mathbf{w}_{n-r}] = [\mathbf{V} | \mathbf{W}].$$

Mivel  $P\mathbf{v}_i = \mathbf{v}_i$  ( $i = 1, 2, \dots, r$ ) és  $P\mathbf{w}_j = \mathbf{0}$  ( $j = 1, 2, \dots, n-r$ ), ezért a  $P$  leképezés  $\mathbf{P}$  mátrixára

$$\mathbf{P}\mathbf{U} = \mathbf{P}[\mathbf{V} | \mathbf{W}] = [\mathbf{P}\mathbf{V} | \mathbf{P}\mathbf{W}] = [\mathbf{V} | \mathbf{O}].$$

Mivel pedig  $\mathbf{U}$  invertálható, ezért a vetítés mátrixa

$$\mathbf{P} = [\mathbf{V} | \mathbf{O}]\mathbf{U}^{-1} = [\mathbf{V} | \mathbf{O}][\mathbf{V} | \mathbf{W}]^{-1}.$$

**7.51. TÉTEL (A PROJEKCIÓ TULAJDONSÁGAI).** Legyen  $P : \mathbb{R}^n \rightarrow \mathbb{R}^n$  egy projekció. Ekkor

- $\mathbb{R}^n$ -nek van olyan bázisa, melyben a mátrixa  $\mathbf{P} = \text{diag}(1, 1, \dots, 1, 0, \dots, 0)$ .
- $I - P$  is projekció, mégpedig  $\text{Ker}(I - P) = \text{Im } P$ ,  $\text{Im}(I - P) = \text{Ker } P$ ,
- $r(P) = \text{trace}(\mathbf{P})$ .

**BIZONYÍTÁS.** a) A fenti jelölésekkel a  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r, \mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{n-r}\}$  bázisban nyilván  $\text{diag}(1, 1, \dots, 1, 0, \dots, 0)$  a mátrix. Ez a

$$\begin{aligned} [\mathbf{V} | \mathbf{O}][\mathbf{V} | \mathbf{W}]^{-1} &= [\mathbf{V} | \mathbf{W}] \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix} [\mathbf{V} | \mathbf{W}]^{-1} \\ &= [\mathbf{V} | \mathbf{W}] \text{diag}(1, 1, \dots, 1, 0, \dots, 0) [\mathbf{V} | \mathbf{W}]^{-1} \end{aligned}$$

átalakításból is látható.

- Ha  $\mathbf{u} = \mathbf{v} + \mathbf{w}$ , és  $P : \mathbf{u} \mapsto \mathbf{v}$  a  $\mathcal{W}$  mentén  $\mathcal{V}$ -re való vetítés, akkor  $I - P : \mathbf{u} \mapsto \mathbf{u} - \mathbf{v} = \mathbf{w}$  a  $\mathcal{V}$  mentén  $\mathcal{W}$ -re való vetítés. Világos, hogy  $\text{Im } P = \mathcal{V}$ ,  $\text{Ker } P = \mathcal{W}$ , és így  $\text{Im}(I - P) = \mathcal{W}$ ,  $\text{Ker}(I - P) = \mathcal{V}$ .  $\square$

**7.52. PÉLDA (PROJEKCIÓ MÁTRIXA).** Határozzuk meg az  $\mathbb{R}^3$  tér  $\mathcal{W} = \text{span}((1, -2, 1))$  altérrel párhuzamos,  $\mathcal{V} = \text{span}((0, 2, -1), (2, 0, -1))$  altérre való vetítésének és a  $\mathcal{V}$ -vel párhuzamos  $\mathcal{W}$ -re való vetítés mátrixát! (ld. még a 7.34. példát!)

**MEGOLDÁS.** Miután

$$\mathbf{V} = \begin{bmatrix} 0 & 2 \\ 2 & 0 \\ -1 & -1 \end{bmatrix}, \quad \mathbf{W} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix},$$

a transzformáció mátrixa egyszerű behelyettesítés után:

$$\begin{aligned} \mathbf{P} &= [\mathbf{V}|\mathbf{O}][\mathbf{V}|\mathbf{W}]^{-1} = \begin{bmatrix} 0 & 2 & 0 \\ 2 & 0 & 0 \\ -1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 2 & 1 \\ 2 & 0 & -2 \\ -1 & -1 & 1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 0 & -1 & -2 \\ 2 & 3 & 4 \\ -1 & -1 & -1 \end{bmatrix}. \end{aligned}$$

A 7.34. példában ugyanezt a kérdést tettük fel, másként megfogalmazva. Ráadásul a megoldás is lényegében ugyanaz, csak a  $\mathcal{V}$  altérnek most más bázist kerestünk. A másik vetítés mátrixa

$$\mathbf{I} - \mathbf{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & -1 & -2 \\ 2 & 3 & 4 \\ -1 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 2 \\ -2 & -2 & -4 \\ 1 & 1 & 2 \end{bmatrix}. \quad \square$$

**7.53. TÉTEL (A VETÍTÉS EKVIVALENS DEFINÍCIÓJA).** A  $P : \mathbb{R}^n \rightarrow \mathbb{R}^n$  lineáris transzformáció pontosan akkor vetítés, ha  $P^2 = P$ , azaz ha  $P$  idempotens.

**BIZONYÍTÁS.** A vetítés definíciója szerinti  $P\mathbf{u} = \mathbf{v}$ , ahol  $\mathbf{u} = \mathbf{v} + \mathbf{w}$  és  $\mathbf{v} \in \mathcal{V}$ ,  $\mathbf{w} \in \mathcal{W}$ . Mivel  $\mathbf{v} = \mathbf{v} + \mathbf{0}$  a  $\mathbf{v}$  felbontása, ezért  $P\mathbf{v} = \mathbf{v}$ , tehát  $P^2\mathbf{u} = P\mathbf{v} = \mathbf{v}$ , azaz minden  $\mathbf{u}$  vektorra  $P^2\mathbf{u} = P\mathbf{u}$ , tehát  $P^2 = P$ .

Minden  $\mathbf{u}$  vektor felbontható a következőképp:  $\mathbf{u} = P\mathbf{u} + (\mathbf{u} - P\mathbf{u})$ . Itt  $P\mathbf{u} \in \text{Im } P$ ,  $\mathbf{u} - P\mathbf{u} \in \text{Ker } P$ , ugyanis  $P(\mathbf{u} - P\mathbf{u}) = P\mathbf{u} - P^2\mathbf{u} = \mathbf{0}$ . Tehát  $\mathbb{R}^n = \text{Im } P + \text{Ker } P$ .  $\text{Im } P \cap \text{Ker } P = \{\mathbf{0}\}$ , ugyanis ha  $\mathbf{u} \in \text{Im } P \cap \text{Ker } P$ , akkor  $\mathbf{u} \in \text{Im } P$  miatt van olyan  $\mathbf{x}$ , hogy  $P\mathbf{x} = \mathbf{u}$ , így  $P^2\mathbf{x} = P\mathbf{u} = \mathbf{0}$ , másrészt  $P^2 = P$  miatt  $P^2\mathbf{x} = P\mathbf{x} = \mathbf{u}$ . Tehát  $\mathbf{u} = \mathbf{0}$ , vagyis  $\text{Im } P \cap \text{Ker } P = \{\mathbf{0}\}$ . Eszerint  $\mathbb{R}^n = \text{Im } P \oplus \text{Ker } P$ . Legyen tehát  $\mathcal{V} = \text{Im } P$ ,  $\mathcal{W} = \text{Ker } P$ . Minden  $\mathbf{u}$  vektor egyértelműen felírható  $\mathbf{u} = \mathbf{v} + \mathbf{w}$  alakban, ahol  $\mathbf{v} \in \mathcal{V}$ , és  $\mathbf{w} \in \mathcal{W}$ . Mivel  $\mathcal{V} = \text{Im } P$ , van olyan  $\mathbf{x}$ , hogy  $\mathbf{v} = P\mathbf{x}$ . Ezért  $\mathbf{u} = P\mathbf{x} + \mathbf{w}$ , így  $P\mathbf{u} = P^2\mathbf{x} + P\mathbf{w} = P\mathbf{x} + \mathbf{0} = \mathbf{v}$ . Tehát  $P$  vetítés.  $\square$

A 7.41. tételben láttuk, hogy egy  $\mathbf{P}$  mátrix pontosan akkor mátrixa egy merőleges vetítésnek, ha  $\mathbf{P}^2 = \mathbf{P}$  és  $\mathbf{P}^\top = \mathbf{P}$ , azaz ha  $\mathbf{P}$  idempotens és szimmetrikus. Így nyilvánvaló az alábbi állítás:

**7.54. KÖVETKEZMÉNY (MIKOR MERŐLEGES EGY VETÍTÉS?).** Legyen  $\mathbf{P}$  egy vetítés mátrixa.  $\mathbf{P}$  pontosan akkor egy merőleges vetítés mátrixa, ha  $\mathbf{P}$  szimmetrikus.

### Feladatok

7.16. Igazoljuk a síkbeli egyenesre merőlegesen vetítő mátrix (7.8) képletét az  $\mathbf{i}$  és  $\mathbf{j}$  vektorok vetületének meghatározásával! Kétféleképp is számolhatunk, *a)* legyen az egyenes hajlásszöge az  $x$ -tengellyel  $\alpha$ , *b)* legyen az egye-

nes irányvektora  $(b_1, b_2)$ . Vessük össze a két eredményt, és igazoljuk azonosságukat.

7.17. **MÍKOR MERŐLEGES EGY VETÍTÉS?** Legyen  $P : \mathbb{R}^n \rightarrow \mathbb{R}^n$  egy vetítés. Mutassuk meg, hogy  $P$  pontosan akkor merőleges vetítés, ha minden  $\mathbf{v} \in \mathbb{R}^n$  vektorra  $|P\mathbf{v}| \leq |\mathbf{v}|$ .

## Pszéudoinverz\*

*A mátrix inverzének olyan – pszeudoinverznek nevezett – általánosítását keressük, mely képes lesz bármely  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszerből a minimális abszolút értékű optimális  $\hat{\mathbf{x}}$  megoldást a mátrix inverzéhez hasonló módon megadni. Tehát ha  $\mathbf{A}^+$  jelöli a pszeudoinvertet, akkor  $\mathbf{A}^+\mathbf{b} = \hat{\mathbf{x}}$ .*

**A pszeudoinverz fogalma** Tudjuk, hogy az  $A : \mathbf{x} \mapsto \mathbf{Ax}$  mátrixleképezés a sorteret kölcsönösen egyértelmű módon viszi az oszloptérbe. Ezért ha  $\mathbf{Ax} = \mathbf{b}$  egy konzisztens egyenletrendszer, és a sortérbe eső megoldás  $\hat{\mathbf{x}}$ , akkor fenn kell álljon az  $\mathbf{A}^+\mathbf{b} = \hat{\mathbf{x}}$  egyenlőség. Ha viszont  $\mathbf{Ax} = \mathbf{b}$  inkonzisztens, akkor a  $\hat{\mathbf{b}} = \text{proj}_{\mathcal{O}(\mathbf{A})} \mathbf{b}$  jelöléssel  $\mathbf{Ax} = \hat{\mathbf{b}}$  konzisztens, így fenn kell álljon az  $\mathbf{A}^+\mathbf{b} = \mathbf{A}^+\hat{\mathbf{b}} = \hat{\mathbf{x}}$  egyenlőség. Ez azt jelenti, hogy  $\mathbf{A}^+(\mathbf{b} - \hat{\mathbf{b}}) = \mathbf{0}$ , vagyis a  $\mathcal{N}(\mathbf{A}^\top)$  minden  $\mathbf{z}$  vektorára  $\mathbf{A}^+\mathbf{z} = \mathbf{0}$ . Mindez a következő definícióhoz vezet.

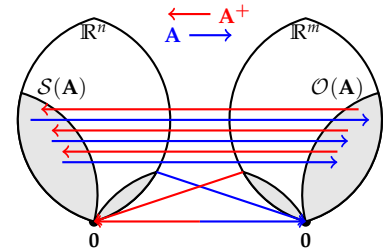
**7.55. DEFINÍCIÓ (A MOORE–PENROSE-FÉLE PSZÉUDOINVERZ).** Legyen  $\mathbf{A}$  egy  $m \times n$ -es valós mátrix. Pszeudoinverzén vagy Moore–Penrose-féle pszeudoinverzén azt az  $\mathbf{A}^+$ -szal jelölt mátrixot értjük, amellyel

- a sortér minden  $\mathbf{x}$  vektorára  $\mathbf{A}^+(\mathbf{Ax}) = \mathbf{x}$ , továbbá
- az oszloptérre merőleges minden  $\mathbf{z}$  vektorra  $\mathbf{A}^+\mathbf{z} = \mathbf{0}$ .

- Azonnal látható, hogy  $m \times n$ -es mátrix pszeudoinverze  $n \times m$ -es.
- Mivel  $\mathbf{Ax} \in \mathcal{O}(\mathbf{A})$ , és a definícióban szereplő  $\mathbf{z}$  vektor eleme  $\mathcal{O}(\mathbf{A})^\perp$ -nek, ezért az  $\mathbf{A}^+$ -hoz tartozó leképezés hatását ismerjük az  $\mathcal{O}(\mathbf{A})$  altérén és merőleges kiegészítő alterén. E leképezés a megadott altereken lineáris, hisz az egyiken egy lineáris leképezés inverze, a másikon a zérusleképezés. Ebből következik, hogy a definícióban megadott leképezés a lineáritás megtartásával egyértelműen kiterjeszthető az egész térre, hisz a tér minden vektora egyértelműen áll elő egy  $\mathcal{O}(\mathbf{A})$ -beli és egy rá merőleges vektor összegeként. Ebből tehát következik, hogy a definícióbeli leképezés létezik, egyértelmű és lineáris, így van mátrixa.
- A definícióból azonnal adódik az is, hogy  $\mathcal{N}(\mathbf{A}^+) = \mathcal{N}(\mathbf{A}^\top)$ , és  $\mathcal{S}(\mathbf{A}^+) = \mathcal{N}(\mathbf{A}^+)^\perp$ , így  $\mathcal{S}(\mathbf{A}^+) = \mathcal{S}(\mathbf{A}^\top) = \mathcal{O}(\mathbf{A})$ .

**7.56. PÉLDA (NÉHÁNY PSZÉUDOINVERZ).** A definíció alapján igazoljuk az alábbi összefüggéseket!

- $\mathbf{A}^+ = \mathbf{A}^{-1}$ , ha  $\mathbf{A}$  invertálható,
- $\mathbf{O}_{m \times n}^+ = \mathbf{O}_{n \times m}$ ,
- $[a]^+ = [1/a]$ , ha  $a \neq 0$ , és  $[0]^+ = [0]$ ,
- $(\mathbf{A}^+)^+ = \mathbf{A}$ ,



7.25. ábra: Az  $\mathbf{A}$  mátrix sortere és oszloptere közti kölcsönösen egyértelmű leképezés a pszeudoinverz fogalmának alapja.

A mátrixinverz fogalmának több más gyengítése is létezik, melyeket itt nem tárgyalunk. A továbbiakban pszeudoinverzén az itt definiálandó Moore–Penrose-féle pszeudoinvertet fogjuk érteni.

e) ha  $a_{ii} \neq 0$  ( $i = 1, 2, \dots, r$ ), akkor

$$\left[ \begin{array}{cccc|c} a_{11} & 0 & \dots & 0 & 0 \\ 0 & a_{22} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & \dots & a_{rr} & 0 \\ \hline & & & & 0 \end{array} \right]_{m \times n}^+ = \left[ \begin{array}{cccc|c} \frac{1}{a_{11}} & 0 & \dots & 0 & 0 \\ 0 & \frac{1}{a_{22}} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & \dots & \frac{1}{a_{rr}} & 0 \\ \hline & & & & 0 \end{array} \right]_{n \times m}$$

**MEGOLDÁS.** a) Ha  $\mathbf{A}$  egy  $n \times n$ -es méretű invertálható mátrix, akkor sortere és oszloptere is a teljes  $n$ -dimenziós tér, és tetszőleges  $\mathbf{x}$  vektorra  $\mathbf{A}^+ \mathbf{A} \mathbf{x} = \mathbf{x}$ , tehát  $\mathbf{A}^+ = \mathbf{A}^{-1}$ .

b) Zérusmátrix oszloptere a zérusvektorból áll, így pszeudo inverze annak merőleges kiegészítő alterét, vagyis az egész teret a nullvektorba viszi, tehát  $\mathbf{O}_{m \times n}^+ = \mathbf{O}_{n \times m}$ .

c) Az előző két eredményből azonnal következik.

d) Ha  $\mathbf{x} \in \mathcal{S}(\mathbf{A})$  és  $\mathbf{y} = \mathbf{A} \mathbf{x}$ , akkor  $\mathbf{A}^+ \mathbf{y} = \mathbf{x}$ , és mivel  $\mathbf{y} \in \mathcal{O}(\mathbf{A}) = \mathcal{S}(\mathbf{A}^+)$ , ezért  $(\mathbf{A}^+)^+ \mathbf{x} = \mathbf{y}$ . Ha pedig  $\mathbf{z} \perp \mathcal{O}(\mathbf{A}^+) = \mathcal{S}(\mathbf{A})$ , akkor  $(\mathbf{A}^+)^+ \mathbf{z} = \mathbf{0}$ , azaz az  $\mathbf{A}$  és az  $(\mathbf{A}^+)^+$  mátrixokhoz tartozó leképezések megegyeznek az  $\mathcal{S}(\mathbf{A})$  és az  $\mathcal{N}(\mathbf{A})$  altereken, így az általuk kifesztett téren, azaz  $\mathbb{R}^n$ -en is. Tehát  $\mathbf{A} = (\mathbf{A}^+)^+$ .

e) Jelölje  $\mathbb{R}^n$  standard bázisának elemeit  $\mathbf{e}_i$  ( $i = 1, 2, \dots, n$ ),  $\mathbb{R}^m$  standard bázisának elemeit  $\mathbf{f}_i$  ( $i = 1, 2, \dots, m$ ). Ha  $a_{ii} \neq 0$ , akkor  $\mathbf{A} \mathbf{e}_i = a_{ii} \mathbf{f}_i$ , és  $\mathbf{e}_i$  a sortérben,  $\mathbf{f}_i$  az oszloptérben van. Ha  $a_{ii} = 0$ , akkor  $\mathbf{A} \mathbf{e}_i = \mathbf{0}$ . Így  $\mathbf{A}^+ \mathbf{f}_i = 1/a_{ii} \mathbf{e}_i$  ha  $a_{ii} \neq 0$  és  $\mathbf{A}^+ \mathbf{f}_i = \mathbf{0}$  egyébként.  $\mathbf{A}^+$  mátrixának ezek az oszlopvektorai, így a főátlójában  $1/a_{ii}$  áll, ha  $a_{ii} \neq 0$ , egyebütt 0. Gondoljuk meg, miért elég megadni  $\mathbf{A}^+$  hatását csak a báziselemeken?  $\square$

Először konstruktív módon megmutatjuk, hogy minden mátrixnak létezik pszeudo inverze, majd bizonyítjuk annak egyértelműségét is.

**7.57. TÉTEL (A PSZEUDOINVERZ MÁTRIXA).** Ha a valós  $\mathbf{A}$  mátrix teljes oszloprangú, akkor

$$\mathbf{A}^+ = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T, \quad (7.10)$$

ha teljes sorrangú, akkor

$$\mathbf{A}^+ = \mathbf{A}^T (\mathbf{A} \mathbf{A}^T)^{-1}. \quad (7.11)$$

Legyen  $\mathbf{A} = \mathbf{B} \mathbf{C}$ , ahol  $\mathbf{B}$  egy teljes oszloprangú,  $\mathbf{C}$  egy teljes sorrangú mátrix (ilyen felbontás mindig létezik, pl. ilyen a *bázisfelbontás*). Ekkor

$$\mathbf{A}^+ = \mathbf{C}^+ \mathbf{B}^+ = \mathbf{C}^T (\mathbf{C} \mathbf{C}^T)^{-1} (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \quad (7.12)$$

$$= \mathbf{C}^T (\mathbf{B}^T \mathbf{A} \mathbf{C}^T)^{-1} \mathbf{B}^T. \quad (7.13)$$

E példa a) pontja mutatja, hogy a pszeudo inverz név nem igazán jó, hisz itt nem *álinverzről*, nem *hamis inverzről* van szó, hanem az inverz általánosításáról, tehát az *általánosított inverz* helyesebb kifejezés. Szokás ezt is használni, de a Moore–Penrose pszeudo inverz kifejezés sokkal elterjedtebb (angol nyelvű művekben is leginkább a pseudoinverse szót használják).

**BIZONYÍTÁS.** Ha  $\mathbf{A}$  teljes oszloprangú, akkor értelmezési tartományának minden vektora a sortérben van, így minden  $\mathbf{x}$  vektorra az  $\mathbf{Ax}$  vektorból vissza kell kapnunk  $\mathbf{x}$ -et egy megfelelő mátrixszal való szorzással. Mivel teljes oszloprangú  $\mathbf{A}$  mátrixra  $\mathbf{A}^T\mathbf{A}$  invertálható, ezért a (7.10)-beli mátrix megfelelő, hisz

$$(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{Ax} = \mathbf{x}.$$

Meg kell még mutatnunk, hogy ha  $\mathbf{z}$  merőleges az oszloptérre, azaz ha  $\mathbf{z} \in \mathcal{N}(\mathbf{A}^T)$ , vagyis ha  $\mathbf{A}^T\mathbf{z} = \mathbf{0}$ , akkor  $\mathbf{A}^+\mathbf{z} = \mathbf{0}$ , de ez valóban igaz, hisz ekkor  $(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{z} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{0} = \mathbf{0}$ .

Ha  $\mathbf{A}$  teljes sorrangú, akkor az oszloptér megegyezik az egész térrel, így az oszloptér bármely  $\mathbf{y}$  vektora esetén az  $\mathbf{Ax} = \mathbf{y}$  egyenletrendszer konzisztens. Ha  $\hat{\mathbf{x}}$  jelöli az egyetlen sortérbe eső megoldást, akkor minden más  $\mathbf{x}$  megoldás esetén  $\text{proj}_{\mathcal{S}(\mathbf{A})}\mathbf{x} = \hat{\mathbf{x}}$ . Így  $\mathbf{A}^+$ -ra fenn kell állnia az  $\mathbf{A}^+\mathbf{y} = \hat{\mathbf{x}}$  összefüggésnek. Ez pedig fönnáll, hisz

$$\text{proj}_{\mathcal{S}(\mathbf{A})}\mathbf{x} = \mathbf{A}^T(\mathbf{AA}^T)^{-1}\mathbf{Ax} = (\mathbf{A}^T(\mathbf{AA}^T)^{-1})(\mathbf{Ax}) = \mathbf{A}^+\mathbf{y}.$$

Végül az általános esetben legyen  $\mathbf{A} = \mathbf{BC}$ , továbbá  $\mathbf{y} = \mathbf{Ax}$ ,  $\mathbf{w} = \mathbf{Cx}$  és  $\mathbf{y} = \mathbf{Bw}$ . Mivel  $\mathbf{B}$  teljes oszloprangú,  $\mathbf{C}$  teljes sorrangú, ezért  $\mathbf{C}^+\mathbf{B}^+\mathbf{y} = \mathbf{C}^+\mathbf{w} = \mathbf{x}$ , vagyis a  $\mathbf{C}^+\mathbf{B}^+$  teljesíti a definíció a) feltételét.  $\mathbf{A}$  b) is teljesül, hisz  $\mathbf{A}$  és  $\mathbf{B}$  oszloptere megegyezik, így  $\mathbf{B}^+\mathbf{z} = \mathbf{0}$ , tehát  $\mathbf{C}^+\mathbf{B}^+\mathbf{z} = \mathbf{0}$  is fennáll. A (7.12) és a (7.13) képletek behelyettesítéssel, majd az  $(\mathbf{CC}^T)^{-1}(\mathbf{B}^T\mathbf{B})^{-1} = (\mathbf{B}^T\mathbf{BCC}^T)^{-1} = (\mathbf{B}^T\mathbf{AC}^T)^{-1}$  átalakítással azonnal adódnak.  $\square$

► A (7.10) képlet tökéletes összhangban van az **egyenletrendszer optimális megoldásáról** szóló 7.46. tétel állításával. Ott arra jutottunk, hogy az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszer optimális megoldásai megegyeznek az  $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$  egyenletrendszer megoldásaival. Ha pedig  $\mathbf{A}$  teljes oszloprangú, akkor  $\mathbf{A}^T\mathbf{A}$  invertálható, így az optimális megoldás  $\hat{\mathbf{x}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$ , azaz a (7.10) képlet szerint  $\hat{\mathbf{x}} = \mathbf{A}^+\mathbf{b}$ , ahogy azt célul tűztük ki e paragrafus elején.

► Általában nem igaz a pszeudoinverzre az  $(\mathbf{XY})^+ = \mathbf{Y}^+\mathbf{X}^+$  összefüggés. Csak azt bizonyítottuk, hogy fennáll, ha  $\mathbf{X}$  teljes oszloprangú, és  $\mathbf{Y}$  teljes sorrangú.

7.58. PÉLDA (A PSZEUDOINVERZ KISZÁMÍTÁSA). Számítsuk ki a

$$\mathbf{B} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}, \mathbf{C} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \text{ és } \mathbf{M} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix}$$

mátrixok pszeudoinverzét!

**MEGOLDÁS.** Mivel  $\mathbf{B}$  teljes oszloprangú, ezért a (7.10) képlet szerint

$$\begin{aligned}\mathbf{B}^+ &= (\mathbf{B}^\top \mathbf{B})^{-1} \mathbf{B}^\top = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} -1/3 & 1/3 & 2/3 \\ 2/3 & 1/3 & -1/3 \end{bmatrix}.\end{aligned}$$

A  $\mathbf{C}$  mátrix teljes sorrangú, így a (7.11) képlet szerint

$$\begin{aligned}\mathbf{C}^+ &= \mathbf{C}^\top (\mathbf{C} \mathbf{C}^\top)^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{bmatrix} = \begin{bmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \\ 1/3 & 1/3 \end{bmatrix}.\end{aligned}$$

Az  $\mathbf{M}^+$  kiszámításához első lépésként meghatározzuk  $\mathbf{M}$  bázisfelbontását. Mivel  $\text{rref}(\mathbf{M}) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ , és  $\mathbf{M}$  első két oszlopa bázisoszlop, ezért

$$\mathbf{M} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

A feladat első felében használt jelölésekkel  $\mathbf{M} = \mathbf{BC}$ , így a (7.12) képletel számolva – és fölhasználva az előbb kiszámolt pszeudoinverzeket

$$\mathbf{M}^+ = \mathbf{C}^+ \mathbf{B}^+ = \begin{bmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \\ 1/3 & 1/3 \end{bmatrix} \begin{bmatrix} -1/3 & 1/3 & 2/3 \\ 2/3 & 1/3 & -1/3 \end{bmatrix} = \frac{1}{9} \begin{bmatrix} -4 & 1 & 5 \\ 5 & 1 & -4 \\ 1 & 2 & 1 \end{bmatrix}$$

Számolhatunk közvetlenül a (7.13) képlettel is:

$$\begin{aligned}\mathbf{M}^+ &= \mathbf{C}^\top (\mathbf{B}^\top \mathbf{M} \mathbf{C}^\top)^{-1} \mathbf{B}^\top \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \left( \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \\ &= \frac{1}{9} \begin{bmatrix} -4 & 1 & 5 \\ 5 & 1 & -4 \\ 1 & 2 & 1 \end{bmatrix} \quad \square\end{aligned}$$

**A pszeudoinverz tulajdonságai** Az  $\mathbf{A}$  mátrix inverzét az  $\mathbf{AX} = \mathbf{I}$  egyenlőséggel definiáltuk. Hasonló egyenlőségeket keresünk a pszeudoinverhez is. Közben azt a fontos tényt is fölfedezzük, hogy  $\mathbf{A}^+ \mathbf{A}$  és  $\mathbf{AA}^+$  is egy-egy merőleges vetítés mátrixa.

Azt tudjuk garantálni, hogy az  $\mathbf{AA}^+$  és az  $\mathbf{A}^+ \mathbf{A}$  mátrixok az egységmátrixszal legyenek egyenlők, de legalább szimmetrikusak, és

az  $\mathbf{A}$ -val, illetve az  $\mathbf{A}^+$ -szal való szorzásra nézve egységmátrixként viselkednek, azaz  $\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}$  és  $\mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+$ . E feltételek már elegendők lesznek a pszeudoinverz algebrai leírásához.

7.59. TÉTEL (MOORE–PENROSE–TÉTEL). *A valós  $\mathbf{A}$  mátrixnak  $\mathbf{X}$  pontosan akkor pszeudoinverze, ha az alábbi négy feltétel mindegyike fennáll:*

$$a) \mathbf{A}\mathbf{X}\mathbf{A} = \mathbf{A}, \quad b) \mathbf{X}\mathbf{A}\mathbf{X} = \mathbf{X}, \quad c) (\mathbf{A}\mathbf{X})^\top = \mathbf{A}\mathbf{X}, \quad d) (\mathbf{X}\mathbf{A})^\top = \mathbf{X}\mathbf{A}.$$

**BIZONYÍTÁS.** Azt, hogy  $\mathbf{A}$  pszeudoinverze teljesíti e négy feltételt, egyszerű behelyettesítéssel ellenőrizhetjük.

$$\begin{aligned} \mathbf{A}\mathbf{A}^+\mathbf{A} &= \mathbf{A}\mathbf{R}^\top(\mathbf{R}\mathbf{R}^\top)^{-1}(\mathbf{B}^\top\mathbf{B})^{-1}\mathbf{B}^\top\mathbf{A} \\ &= \mathbf{B}\mathbf{R}\mathbf{R}^\top(\mathbf{R}\mathbf{R}^\top)^{-1}(\mathbf{B}^\top\mathbf{B})^{-1}\mathbf{B}^\top\mathbf{B}\mathbf{R} = \mathbf{B}\mathbf{R} = \mathbf{A} \\ \mathbf{A}^+\mathbf{A}\mathbf{A}^+ &= \mathbf{R}^\top(\mathbf{R}\mathbf{R}^\top)^{-1}(\mathbf{B}^\top\mathbf{B})^{-1}\mathbf{B}^\top\mathbf{B}\mathbf{R}\mathbf{R}^\top(\mathbf{R}\mathbf{R}^\top)^{-1}(\mathbf{B}^\top\mathbf{B})^{-1}\mathbf{B}^\top \\ &= \mathbf{R}^\top(\mathbf{R}\mathbf{R}^\top)^{-1}(\mathbf{B}^\top\mathbf{B})^{-1}\mathbf{B}^\top = \mathbf{A}^+ \end{aligned}$$

A  $c)$  és az  $d)$  ellenőrzéséhez egyszerűsítsük az  $\mathbf{A}^+\mathbf{A}$  és  $\mathbf{A}\mathbf{A}^+$  kifejezéseket:

$$\mathbf{A}^+\mathbf{A} = (\mathbf{R}^\top(\mathbf{R}\mathbf{R}^\top)^{-1}(\mathbf{B}^\top\mathbf{B})^{-1}\mathbf{B}^\top)(\mathbf{B}\mathbf{R}) = \mathbf{R}^\top(\mathbf{R}\mathbf{R}^\top)^{-1}\mathbf{R}, \quad (7.14)$$

$$\mathbf{A}\mathbf{A}^+ = (\mathbf{B}\mathbf{R})(\mathbf{R}^\top(\mathbf{R}\mathbf{R}^\top)^{-1}(\mathbf{B}^\top\mathbf{B})^{-1}\mathbf{B}^\top) = \mathbf{B}(\mathbf{B}^\top\mathbf{B})^{-1}\mathbf{B}^\top. \quad (7.15)$$

Ezeket fölhasználva kapjuk, hogy

$$\begin{aligned} (\mathbf{A}^+\mathbf{A})^\top &= (\mathbf{R}^\top(\mathbf{R}\mathbf{R}^\top)^{-1}\mathbf{R})^\top = \mathbf{R}^\top(\mathbf{R}\mathbf{R}^\top)^{-1}\mathbf{R} = \mathbf{A}^+\mathbf{A} \\ (\mathbf{A}\mathbf{A}^+)^\top &= (\mathbf{B}(\mathbf{B}^\top\mathbf{B})^{-1}\mathbf{B}^\top)^\top = \mathbf{B}(\mathbf{B}^\top\mathbf{B})^{-1}\mathbf{B}^\top = \mathbf{A}\mathbf{A}^+, \end{aligned}$$

ami bizonyítja az  $c)$  és az  $d)$  egyenlőségeket. Már csak azt kell bizonyítani, hogy ezeket az összefüggéseket legföljebb csak egy mátrix teljesíti. Tegyük fel, hogy  $\mathbf{X}$  és  $\mathbf{Y}$  is teljesíti a négy feltételt. Ekkor

$$\begin{aligned} \mathbf{A}\mathbf{Y} &\stackrel{a)}{=} \mathbf{A}\mathbf{X}\mathbf{A}\mathbf{Y} \stackrel{c)}{=} (\mathbf{A}\mathbf{X})^\top(\mathbf{A}\mathbf{Y})^\top = \mathbf{X}^\top\mathbf{A}^\top\mathbf{Y}^\top\mathbf{A}^\top \\ &= \mathbf{X}^\top(\mathbf{A}\mathbf{Y}\mathbf{A})^\top \stackrel{a)}{=} \mathbf{X}^\top\mathbf{A}^\top = (\mathbf{A}\mathbf{X})^\top \stackrel{c)}{=} \mathbf{A}\mathbf{X} \end{aligned} \quad (7.16)$$

$$\begin{aligned} \mathbf{Y}\mathbf{A} &\stackrel{a)}{=} \mathbf{Y}\mathbf{A}\mathbf{X}\mathbf{A} \stackrel{d)}{=} (\mathbf{Y}\mathbf{A})^\top(\mathbf{X}\mathbf{A})^\top = \mathbf{A}^\top\mathbf{Y}^\top\mathbf{A}^\top\mathbf{X}^\top \\ &= (\mathbf{A}\mathbf{Y}\mathbf{A})^\top\mathbf{X}^\top \stackrel{a)}{=} \mathbf{A}^\top\mathbf{X}^\top = (\mathbf{X}\mathbf{A})^\top \stackrel{d)}{=} \mathbf{X}\mathbf{A} \end{aligned} \quad (7.17)$$

$$\mathbf{Y} \stackrel{b)}{=} \mathbf{Y}\mathbf{A}\mathbf{Y} \stackrel{(7.16)}{=} \mathbf{Y}\mathbf{A}\mathbf{X} \stackrel{(7.17)}{=} \mathbf{X}\mathbf{A}\mathbf{X} \stackrel{b)}{=} \mathbf{X}.$$

Ezzel bizonyítottuk a tételt. □

7.60. KÖVETKEZMÉNY ( $\mathbf{A}^+\mathbf{A}$  és  $\mathbf{A}\mathbf{A}^+$  MERŐLEGES VETÍTÉS). *Tetszőleges  $\mathbf{A} \in \mathbb{R}^{m \times n}$  mátrix esetén*

$$\mathbf{A}^+\mathbf{A} = \text{proj}_{S(\mathbf{A})} \quad \text{és} \quad \mathbf{A}\mathbf{A}^+ = \text{proj}_{O(\mathbf{A})}.$$

Tehát  $\mathbf{A}^+\mathbf{A}$  az  $\mathbb{R}^n$  teret merőlegesen vetíti  $\mathbf{A}$  sorterére, míg  $\mathbf{A}\mathbf{A}^+$  az  $\mathbb{R}^m$  teret merőlegesen vetíti  $\mathbf{A}$  oszlopterére.



**BIZONYÍTÁS.** A (7.14) egyenlőség szerint  $\mathbf{A}^+ \mathbf{A} = \mathbf{R}^\top (\mathbf{R} \mathbf{R}^\top)^{-1} \mathbf{R}$ , ami az altérre való merőleges vetítés mátrixáról szóló 7.39. tétel szerint az  $\mathbf{R}^\top$  oszlopvektorai által kifeszített térre – azaz a sortérre – való merőleges vetítés mátrixa. Hasonlóképp a (7.15) egyenlet szerint  $\mathbf{A} \mathbf{A}^+ = \mathbf{B} (\mathbf{B}^\top \mathbf{B})^{-1} \mathbf{B}^\top$ , ami a  $\mathbf{B}$  oszlopvektorai által kifeszített térre – azaz az oszloptérre – való merőleges vetítés mátrixa.  $\square$

*A pszeudo inverz és a minimális abszolút értékű optimális megoldás* Megmutatjuk, hogy  $\mathbf{A}^+$  úgy használható egy tetszőleges együtthatómátrixú  $\mathbf{A} \mathbf{x} = \mathbf{b}$  egyenletrendszer optimális megoldásának meghatározására, ahogy  $\mathbf{A}^{-1}$  használható akkor, ha  $\mathbf{A}$  invertálható.

7.61. TÉTEL (OPTIMÁLIS MEGOLDÁS PSZEUDOINVERZZEL). *Legyen  $\mathbf{A}$  egy valós mátrix. Az  $\mathbf{A} \mathbf{x} = \mathbf{b}$  egyenletrendszernek az  $\hat{\mathbf{x}} = \mathbf{A}^+ \mathbf{b}$  a minimális abszolút értékű optimális megoldása.*

**BIZONYÍTÁS.** Először megmutatjuk, hogy  $\mathbf{A}^+ \mathbf{b}$  optimális megoldás, azaz megoldása az  $\mathbf{A}^\top \mathbf{A} \mathbf{x} = \mathbf{A}^\top \mathbf{b}$  normálegyenlet-rendszernek. Tehát igazolni kell, hogy  $\mathbf{A}^\top \mathbf{A} \mathbf{A}^+ \mathbf{b} = \mathbf{A}^\top \mathbf{b}$ . Ehhez elég belátni, hogy  $\mathbf{A}^\top \mathbf{A} \mathbf{A}^+ = \mathbf{A}^\top$ . Legyen  $\mathbf{A} = \mathbf{B} \mathbf{R}$  az  $\mathbf{A}$  mátrix bázisfelbontása. Ekkor

$$\begin{aligned} \mathbf{A}^\top \mathbf{A} \mathbf{A}^+ &= (\mathbf{R}^\top \mathbf{B}^\top) (\mathbf{B} (\mathbf{B}^\top \mathbf{B})^{-1} \mathbf{B}^\top) \\ &= \mathbf{R}^\top (\mathbf{B}^\top \mathbf{B}) (\mathbf{B}^\top \mathbf{B})^{-1} \mathbf{B}^\top = \mathbf{R}^\top \mathbf{B}^\top = \mathbf{A}^\top \end{aligned}$$

Mivel  $\mathbf{A}^+ \mathbf{b}$  a definícióból következőleg a sortérben van, és a sortérbe csak egyetlen optimális megoldás – a minimális abszolút értékű megoldás – esik, ezért  $\mathbf{A}^+ \mathbf{b}$  valóban a minimális abszolút értékű optimális megoldás.  $\square$

A gyakorlatban gyakran előfordul, hogy mért adatokból kell bizonyos változók értékét meghatározni. Ha az  $n$  ismeretlen értékre a mérési hibákat kiküszöbölendő  $n$ -nél több mérést végzünk, az egyenletrendszer könnyen ellentmondásossá válhat. Ehhez hasonló esetet mutat a következő példa.

7.62. PÉLDA (EGYENLETRENDSZER OPTIMÁLIS MEGOLDÁSA). *Az alábbi háromismeretlenes egyenletrendszer négy egyenletből áll:*

$$\begin{aligned} x + 3y + 6z &= 8 \\ x - y + 2z &= 2 \\ x + 3y + 2z &= 2 \\ x - y - 2z &= 0 \end{aligned}$$

*Bármelyik három egyértelműen megoldható egyenletrendszert ad, de a négy együtt ellentmondásos. Határozzuk meg az optimális megoldását!*

**MEGOLDÁS.** A 7.61. tétel szerint az optimális megoldás  $\mathbf{A}^+\mathbf{b}$ , ahol

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 6 \\ 1 & -1 & 2 \\ 1 & 3 & 2 \\ 1 & -1 & -2 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 8 \\ 2 \\ 2 \\ 0 \end{bmatrix}$$

Mivel  $\mathbf{A}$  teljes oszloprangú, ezért csak egyetlen optimális megoldása van, másrészt pszeudoinverze a (7.10) képlettel számolható, így

$$\mathbf{A}^+ = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top = \begin{bmatrix} 0 & 1/4 & 1/4 & 1/2 \\ 0 & -1/4 & 1/4 & 0 \\ 1/8 & 1/8 & -1/8 & -1/8 \end{bmatrix} \quad \mathbf{A}^+\mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Így az egyenletrendszer optimális megoldása  $(1, 0, 1)$ .  $\square$

A következő példa olyan egyenletrendszert vizsgál, melyben az együtthatómátrix rangja kisebb mind az egyenletek, mind az ismeretlenek számánál.

**7.63. PÉLDA (EGYENLETRENDSZER OPTIMÁLIS MEGOLDÁSA).** *Oldjuk meg az*

$$\begin{aligned} y + z &= 3 \\ x + y + 2z &= 2 \\ x + y &= 2 \end{aligned}$$

*egyenletrendszert. Ha nem oldható meg, akkor határozzuk meg minimális abszolút értékű optimális megoldását!*

**MEGOLDÁS.** Az egyenletrendszer nem oldható meg, ami bővített mátrixának redukált lépcsős alakjából leolvasható:

$$\left[ \begin{array}{ccc|c} 0 & 1 & 1 & 3 \\ 1 & 1 & 2 & 2 \\ 1 & 1 & 0 & 2 \end{array} \right] \Rightarrow \left[ \begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right]$$

Az együtthatómátrix pszeudoinverzét meghatároztuk a 10.10. példában. Ezt felhasználva a minimális abszolút értékű optimális megoldás

$$\hat{\mathbf{x}} = \mathbf{A}^+\mathbf{b} = \frac{1}{9} \begin{bmatrix} -4 & 1 & 5 \\ 5 & 1 & -4 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

$\square$

## Feladatok

## Pseudoinverz

Számítsuk ki az alábbi mátrixok pseudoinverzét!

$$7.18. \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

$$7.19. \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$7.20. \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$$

$$7.21. \begin{bmatrix} 0 & 2 \\ 0 & 2 \end{bmatrix}$$

$$7.22. \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$$7.23. \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 0 & 0 \end{bmatrix}$$

$$7.24. \begin{bmatrix} 1 & 0 \\ 2 & 2 \\ 0 & 1 \end{bmatrix}$$

$$7.25. \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$$

$$7.26. \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

Az alábbi feladatokban határozzuk meg a felbontásaikkal megadott mátrixok pseudoinverzét!

$$7.27. \mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

$$7.28. \mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

$$7.29. \mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}$$

7.30. **PSZEUDOINVERZ MÁTRIXA** Tetszőleges  $\mathbf{A} \in \mathbb{R}^{m \times n}$  mátrixnak létezik pseudoinverze és az egyértelmű.

7.31. Mutassuk meg, hogy ha  $\mathbf{A}$  egy merőleges vetítés mátrixa, azaz  $\mathbf{A}^T = \mathbf{A} = \mathbf{A}^2$ , akkor  $\mathbf{A}^+ = \mathbf{A}$ . Igaz-e az állítás megfordítása?

7.32. Mutassuk meg, hogy ha  $\mathbf{A} \in \mathbb{R}^{m \times n}$  és  $\mathbf{A}$  teljes oszlopprangú, akkor  $\mathbf{A}^+ \mathbf{A} = \mathbf{I}_m$ , ha pedig  $\mathbf{A}$  teljes sorrangú, akkor  $\mathbf{A} \mathbf{A}^+ = \mathbf{I}_n$ .

7.33. **1-RANGÚ MÁTRIXOK PSZEUDOINVERZE** Mutassuk meg, hogy ha  $r(\mathbf{A}) = 1$ , akkor

$$\mathbf{A}^+ = \frac{1}{\text{trace}(\mathbf{A}^T \mathbf{A})} \mathbf{A}^T,$$

ahol  $\text{trace}(\mathbf{A}^T \mathbf{A})$  az  $\mathbf{A}$  elemeinek négyzetösszege. Eszerint ha  $\mathbf{a} \neq \mathbf{0}$ , akkor

$$\mathbf{a}^+ = \frac{1}{\mathbf{a}^T \mathbf{a}} \mathbf{a}^T = \frac{1}{\mathbf{a} \cdot \mathbf{a}} \mathbf{a}^T.$$

E feladat eredményét fölhasználva ellenőrizzük a 7.18., 7.19., 7.20., 7.21., 7.23., 7.25., 7.26. feladatok eredményeit!

7.34. **BLOKKDIAGONÁLIS MÁTRIX PSZEUDOINVERZE** Igazoljuk, hogy blokkdiagonális mátrix esetén

$$\begin{bmatrix} \mathbf{A}_1 & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & \mathbf{A}_2 & \dots & \mathbf{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & \mathbf{O} & \dots & \mathbf{A}_k \end{bmatrix}^+ = \begin{bmatrix} \mathbf{A}_1^+ & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & \mathbf{A}_2^+ & \dots & \mathbf{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & \mathbf{O} & \dots & \mathbf{A}_k^+ \end{bmatrix}.$$

7.35. Számítsuk ki a

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

mátrix pseudoinverzét!

## Ortonormált bázis – ortogonális mátrix

Nem kell indokolni a merőlegesség fontosságát bizonyos természeti jelenségek leírásában. A lineáris algebrában is nélkülözhetetlen a fogalma. Egy altér ortonormált bázisának megkonstruálása, és azoknak a leképezéseknek az áttekintése, melyek ortonormált bázist ortonormáltba visznek alapvetően fontosak.

### Ortogonalis mátrixok

**Ortogonalis és ortonormált bázis** Segíti az alterek vizsgálatát, ha a bázisvektorok páronként merőlegesek egymásra, ekkor ugyanis a különböző bázisvektorok skaláris szorzata 0. További könnyítést jelenthet, ha a bázisvektorok egységvektorok, mert ekkor egy vektor velük vett skaláris szorzata a merőleges vetület hosszát adja.

Páronként merőleges vektorok egy rendszerét *ortogonalis rendszernek* nevezzük. Ortogonalis rendszernek lehetnek 0-vektor tagjai. Páronként merőleges egységvektorok egy rendszerét *ortonormált rendszernek* nevezzük. Ortonormált rendszerben *nincsenek* 0-vektorok. A **következő** tételből azonnal adódik, hogy zérusvektort nem tartalmazó ortogonalis rendszer vagy egy tetszőleges ortonormált rendszer mindig bázisa az általa kifeszített alternek. Ezt az altér *ortogonalis bázisának* (rövidítve OB), illetve *ortonormált bázisának* (rövidítve ONB) nevezzük. Ortogonalis bázisból mindig kaphatunk egy ortonormáltat, ha az OB minden bázisvektorát elosztjuk a hosszával. Ezt a vektor *normálásának* nevezzük.

**7.64. TÉTEL (ORTOGONÁLIS VEKTOROK FÜGGETLENSÉGE).** *Ha a nullvektortól különböző  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k$  vektorok páronként ortogonálisak, akkor függetlenek is.*

**BIZONYÍTÁS.** Tekintsük a  $c_1\mathbf{a}_1 + \dots + c_k\mathbf{a}_k = \mathbf{0}$  egyenletet. Be kell látnunk, hogy ez csak a  $c_1 = \dots = c_k = 0$  esetben áll fenn. Szorozzuk be az egyenlőség mindkét oldalát az  $\mathbf{a}_i$  vektorral ( $i = 1, 2, \dots, k$ ). Ekkor a jobb oldal 0, a bal oldalon pedig egy tag kivételével mindegyik 0 lesz:

$$\begin{aligned}(c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \dots + c_k\mathbf{a}_k) \cdot \mathbf{a}_i &= \mathbf{0} \cdot \mathbf{a}_i \\ c_i\mathbf{a}_i \cdot \mathbf{a}_i &= 0.\end{aligned}$$

Mivel  $\mathbf{a}_i \cdot \mathbf{a}_i \neq 0$ , ezért  $c_i = 0$ , és ez igaz minden  $i$ -re. □

Tudjuk, hogy a háromdimenziós térben bármely  $\mathbf{v} = (x, y, z)$  vektor koordinátáira igaz, hogy

$$x = \mathbf{v} \cdot \mathbf{i}, \quad y = \mathbf{v} \cdot \mathbf{j}, \quad z = \mathbf{v} \cdot \mathbf{k}.$$

Az is igaz, hogy az  $\mathbf{i}$  és  $\mathbf{j}$  által kifeszített síknak, azaz az  $xy$ -síknak a  $\mathbf{v}$  vektorhoz legközelebb fekvő pontja, illetve az oda mutató helyvektor  $\hat{\mathbf{v}} = (x, y, 0)$ , azaz

$$\hat{\mathbf{v}} = (\mathbf{v} \cdot \mathbf{i})\mathbf{i} + (\mathbf{v} \cdot \mathbf{j})\mathbf{j}$$

Azt is tudjuk, hogy a  $\mathbf{v}$ -hez legközelebbi pont épp  $\mathbf{v}$ -nek a síkra való merőleges vetülete.

A fenti nyilvánvaló összefüggések tetszőleges ONB esetén is használhatók, így igen értékesek.

**7.65. TÉTEL (LEGJOBB KÖZELÍTÉS ONB ESETÉN).** *Adva van az  $\mathbb{R}^n$  térben egy  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k\}$  ortonormált rendszer által kifeszített  $\mathcal{A}$  altér, valamint egy  $\mathbf{v}$  vektor. Ekkor a*

$$\hat{\mathbf{v}} = (\mathbf{v} \cdot \mathbf{e}_1)\mathbf{e}_1 + (\mathbf{v} \cdot \mathbf{e}_2)\mathbf{e}_2 + \dots + (\mathbf{v} \cdot \mathbf{e}_k)\mathbf{e}_k \quad (7.18)$$

*vektor az  $\mathcal{A}$  altér  $\mathbf{v}$ -hez legközelebb fekvő pontja, azaz  $\hat{\mathbf{v}} = \text{proj}_{\mathcal{A}} \mathbf{v}$ .*

**BIZONYÍTÁS.** Először megmutatjuk, hogy a (7.18) képlet szerinti pont van legközelebb  $\mathbf{v}$ -hez.  $\mathbf{v}$  és  $\hat{\mathbf{v}}$  távolságának négyzete

$$\begin{aligned} (\mathbf{v} - \hat{\mathbf{v}})^2 &= \left( \mathbf{v} - \sum_{i=1}^k (\mathbf{v} \cdot \mathbf{e}_i) \mathbf{e}_i \right)^2 \\ &= \mathbf{v}^2 - 2 \sum_{i=1}^k (\mathbf{v} \cdot \mathbf{e}_i)^2 + \sum_{i=1}^k (\mathbf{v} \cdot \mathbf{e}_i)^2 \\ &= \mathbf{v}^2 - \sum_{i=1}^k (\mathbf{v} \cdot \mathbf{e}_i)^2. \end{aligned}$$

$\mathbf{v}$  és az altér egy tetszőleges  $\mathbf{u}$  vektorának távolságnégyzete:

$$\begin{aligned} (\mathbf{v} - \mathbf{u})^2 &= \left( \mathbf{v} - \sum_{i=1}^k c_i \mathbf{e}_i \right)^2 \\ &= \mathbf{v}^2 - 2 \sum_{i=1}^k c_i (\mathbf{v} \cdot \mathbf{e}_i) + \sum_{i=1}^k c_i^2. \end{aligned}$$

Ha az utóbbiból kivonva az előbbi pozitív értéket kapunk, ez azt jelenti, hogy valóban  $\hat{\mathbf{v}}$  van  $\mathbf{v}$ -hez legközelebb:

$$\begin{aligned} (\mathbf{v} - \mathbf{u})^2 - (\mathbf{v} - \hat{\mathbf{v}})^2 &= \left( \mathbf{v}^2 - 2 \sum_{i=1}^k c_i (\mathbf{v} \cdot \mathbf{e}_i) + \sum_{i=1}^k c_i^2 \right) - \left( \mathbf{v}^2 - \sum_{i=1}^k (\mathbf{v} \cdot \mathbf{e}_i)^2 \right) \\ &= \sum_{i=1}^k c_i^2 - 2 \sum_{i=1}^k c_i (\mathbf{v} \cdot \mathbf{e}_i) + \sum_{i=1}^k (\mathbf{v} \cdot \mathbf{e}_i)^2 \\ &= \sum_{i=1}^k (c_i - \mathbf{v} \cdot \mathbf{e}_i)^2 \geq 0. \end{aligned}$$

Ebből a **legjobb közelítés tétele** szerint kapjuk, hogy  $\hat{\mathbf{v}} = \text{proj}_{\mathcal{A}} \mathbf{v}$ .  $\square$

7.66. PÉLDA (EGY PONT SÍKRA VALÓ MERŐLEGES VETÜLETE). Határozzuk meg a  $(3, 1, 2)$  pontnak az egymásra merőleges  $\mathbf{a} = \frac{1}{7}(2, 3, 6)$  és  $\mathbf{b} = \frac{1}{7}(3, -6, 2)$  vektorok által kifeszített síkra való merőleges vetületét!

**MEGOLDÁS.** Mivel  $\mathbf{a}$  és  $\mathbf{b}$  ortonormált bázisa az általuk kifeszített altérnek, ezért a  $\mathbf{v} = (3, 1, 2)$  vektornak e síkra eső merőleges vetülete

$$\begin{aligned}\hat{\mathbf{v}} &= (\mathbf{v} \cdot \mathbf{a})\mathbf{a} + (\mathbf{v} \cdot \mathbf{b})\mathbf{b} \\ &= ((3, 1, 2) \cdot (\frac{2}{7}, \frac{3}{7}, \frac{6}{7})) (\frac{2}{7}, \frac{3}{7}, \frac{6}{7}) + ((3, 1, 2) \cdot (\frac{3}{7}, -\frac{6}{7}, \frac{2}{7})) (\frac{3}{7}, -\frac{6}{7}, \frac{2}{7}) \\ &= 3(\frac{2}{7}, \frac{3}{7}, \frac{6}{7}) + 1(\frac{3}{7}, -\frac{6}{7}, \frac{2}{7}) \\ &= (\frac{9}{7}, \frac{3}{7}, \frac{20}{7}).\end{aligned}$$

Összehasonlításként: a standard elemi módszer az  $\mathbf{a} \times \mathbf{b}$  irányvektorú,  $(3, 1, 2)$  ponton átmenő egyenes és a sík metszéspontjának meghatározása lenne.  $\square$

**Ortogonalis mátrixok** Egy ortonormált vektorrendszerből képzett mátrixnak gyönyörű algebrai és geometriai tulajdonságai vannak.

7.67. DEFINÍCIÓ (ORTOGONÁLIS ÉS SZEMIORTOGONÁLIS MÁTRIX). Egy valós négyzetes mátrixot ortogonálisnak nevezzük, ha oszlopvektorai vagy sorvektorai ortonormált rendszert alkotnak. Ha nem kötjük ki, hogy a mátrix négyzetes legyen, szemiortogonális mátrixról beszélünk.

► Látni fogjuk, hogy az ortogonális mátrixok definíciójában elég csak az oszlopvektorok vagy csak a sorvektorok ortonormalitását kikötni, mert mindegyikből következik a másik. Ha azonban egy mátrix nem négyzetes, akkor vagy csak az oszlopvektorai, vagy csak a sorvektorai alkothatnak ONR-t. Ha például az  $n \times k$  méretű  $\mathbf{Q}$  mátrix szemiortogonális, akkor  $k \leq n$  pontosan akkor áll fenn, ha  $\mathbf{Q}$  oszlopvektorai alkotnak ONR-t. Az ugyanis, hogy valamely vektorok ONR-t alkotnak, maga után vonja lineáris függetlenségüket is.

► Nagyon szerencsétlen az ortogonális mátrix elnevezés, de annyira el van terjedve, hogy nem lehet eltérni tőle. Nyilván jobb lenne az ortonormált mátrix elnevezés.

► Minden ortogonális mátrix egyúttal szemiortogonális is.

7.68. PÉLDA (ORTOGONÁLIS MÁTRIXOK). Melyek ortogonálisak és melyek szemiortogonálisak az alábbi mátrixok közül?

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{C} = \frac{1}{7} \begin{bmatrix} 2 & 3 & 6 \\ 6 & 2 & -3 \\ 3 & -6 & 2 \end{bmatrix}.$$

**MEGOLDÁS.** Mindhárom mátrix szemiortogonális, hisz oszlopvektorai vagy sorvektorai ortonormált rendszert alkotnak (az  $\mathbf{A}$  mátrix sorai  $\mathbb{R}^4$  standard egységvektorai közül valók, a  $\mathbf{B}$  mátrix oszlopvektorai az  $\mathbb{R}^4$  első két standard egységvektorának  $xy$ -síkjában  $\alpha$  szöggel való elforgatásával kaphatók, a  $\mathbf{C}$  mátrix esetén az oszlopvektorok skaláris szorzatainak elvégzésével ellenőrizhetjük ortonormalitásukat). A három mátrix közül csak a  $\mathbf{C}$  négyzetes, így csak ez ortogonális.  $\square$

► Könnyen látható, hogy minden permutációmátrix, így az egység-mátrix is, ortogonális.

**7.69. TÉTEL (SZEMIORTOGONÁLIS MÁTRIXOK EKVIVALENS DEFINÍCIÓI).**

Legyen  $m \geq n$  és  $\mathbf{Q} \in \mathbb{R}^{m \times n}$ . Az alábbi állítások ekvivalensek:

- a)  $\mathbf{Q}$  szemiortogonális,
- b)  $\mathbf{Q}^T \mathbf{Q} = \mathbf{I}_n$ .

► Hasonlóképp  $m \leq n$  esetén  $\mathbf{Q}$  pontosan akkor szemiortogonális, ha  $\mathbf{Q} \mathbf{Q}^T = \mathbf{I}_m$ .

► A b) állítás algebrai nyelven azt mondja, hogy  $m \geq n$  esetén  $\mathbf{Q}$  pontosan akkor szemiortogonális, ha transzponáltja a bal oldali inverze.

**BIZONYÍTÁS.** a)  $\Rightarrow$  b): Ha  $\mathbf{Q}$  szemiortogonális és  $m \geq n$ , akkor  $\mathbf{Q}$  oszlopai alkotnak ONR-t. Legyen  $\mathbf{Q} = [\mathbf{q}_1 \ \mathbf{q}_2 \ \dots \ \mathbf{q}_n]$ . Ekkor  $[\mathbf{Q}^T \mathbf{Q}]_{ij} = \mathbf{q}_i^T \mathbf{q}_j = \mathbf{q}_i \cdot \mathbf{q}_j$ , de mivel a  $\{\mathbf{q}_i\}$  vektorrendszer ortonormált, ezért  $\mathbf{q}_i^2 = 1$  és  $\mathbf{q}_i \cdot \mathbf{q}_j = 0$ , ha  $i \neq j$ . Eszerint  $[\mathbf{Q}^T \mathbf{Q}]_{ii} = 1$ , és  $[\mathbf{Q}^T \mathbf{Q}]_{ij} = 0$ , ha  $i \neq j$  és  $i, j \leq n$ , vagyis  $\mathbf{Q}^T \mathbf{Q} = \mathbf{I}_n$ .

b)  $\Rightarrow$  a): A  $\mathbf{Q}^T \mathbf{Q} = \mathbf{I}_n$  összefüggésbeli mátrixszorzást sorvektorszor oszlopvektorként tekintve épp azt kapjuk, hogy  $\mathbf{q}_i^2 = 1$  és  $\mathbf{q}_i \cdot \mathbf{q}_j = 0$ , ha  $i \neq j$ , azaz a  $\{\mathbf{q}_i\}$  vektorrendszer ortonormált.  $\square$

**7.70. TÉTEL (ORTOGONÁLIS MÁTRIXOK EKVIVALENS DEFINÍCIÓI).**

Legyen  $\mathbf{Q} \in \mathbb{R}^{n \times n}$ . Az alábbi állítások ekvivalensek:

- a)  $\mathbf{Q}$  oszlopvektorai ortonormált rendszert alkotnak.
- b)  $\mathbf{Q}^T \mathbf{Q} = \mathbf{I}_n$ .
- c)  $\mathbf{Q}^{-1} = \mathbf{Q}^T$ .
- d)  $\mathbf{Q} \mathbf{Q}^T = \mathbf{I}_n$ .
- e)  $\mathbf{Q}$  sorvektorai ortonormált rendszert alkotnak.

**BIZONYÍTÁS.** Az a)  $\Leftrightarrow$  b) ekvivalenciát az előző állításban bizonyítottuk.

b)  $\Rightarrow$  c): Mivel  $\mathbf{Q}$  négyzetes, ezért a  $\mathbf{Q}^T \mathbf{Q} = \mathbf{I}$  összefüggés egyúttal azt is jelenti, hogy  $\mathbf{Q}$  invertálható, tehát  $\mathbf{Q}$  és  $\mathbf{Q}^T$  egymás inverzei, azaz  $\mathbf{Q}^{-1} = \mathbf{Q}^T$ .

c)  $\Rightarrow$  d): Mivel  $\mathbf{Q}^{-1} = \mathbf{Q}^T$ , ezért  $\mathbf{Q} \mathbf{Q}^T = \mathbf{I}_n$ .

d)  $\Rightarrow$  e): A  $\mathbf{Q} \mathbf{Q}^T = \mathbf{I}_n$  egyenletben a mátrixszorzásra sorvektorszor oszlopvektorként tekintve épp azt kapjuk, hogy  $\mathbf{Q}^T$  oszlopvektorai – és így  $\mathbf{Q}$  sorvektorai – ONB-t alkotnak.

$e) \Rightarrow a)$ : Az előzőekben beláttuk, hogy  $a)$ -ból következik  $e)$ , azaz ha  $\mathbf{Q}$  oszlopvektorai ONB-t alkotnak, akkor sorvektorai is. Ezt  $\mathbf{Q}^T$ -ra alkalmazva azt kapjuk, hogy ha  $\mathbf{Q}$  sorvektorai ONB-t alkotnak, akkor oszlopvektorai is.  $\square$

7.71. PÉLDA (ORTOGONÁLIS MÁTRIXOK INVERZE). Számítsuk ki az

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}, \quad \frac{1}{7} \begin{bmatrix} 2 & 3 & 6 \\ 6 & 2 & -3 \\ 3 & -6 & 2 \end{bmatrix}.$$

mátrixok inverzét!

**MEGOLDÁS.** Mindhárom mátrix ortogonális (az első permutációmátrix, a harmadik ortogonalitását a 7.68. példában ellenőriztük), így az előző tétel szerint inverzük megegyezik transzponáltjukkal, tehát az inverzek:

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix}, \quad \frac{1}{7} \begin{bmatrix} 2 & 6 & 3 \\ 3 & 2 & -6 \\ 6 & -3 & 2 \end{bmatrix}. \quad \square$$

**Ortogonalis mátrixok geometriája** Ortogonalis mátrixhoz tartozó mátrixleképezés ONB-t ONB-ba visz úgy, ahogy a síkban vagy térben a forgatás és a tükrözés.

7.72. TÉTEL (ORTOGONÁLIS MÁTRIXHOZ TARTOZÓ MÁTRIXLEKÉPEZÉS).

Legyen  $\mathbf{Q} \in \mathbb{R}^{n \times n}$ . Az alábbi állítások ekvivalensek:

- a)  $\mathbf{Q}$  ortogonalis.
- b)  $|\mathbf{Q}\mathbf{x}| = |\mathbf{x}|$  minden  $\mathbf{x} \in \mathbb{R}^n$  vektorra.
- c)  $\mathbf{Q}\mathbf{x} \cdot \mathbf{Q}\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$  minden  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  vektorra.

**BIZONYÍTÁS.**  $a) \Rightarrow b)$ : Ha  $\mathbf{Q}$  ortogonalis, akkor  $\mathbf{Q}^T \mathbf{Q} = \mathbf{I}$ , így tetszőleges  $\mathbf{x} \in \mathbb{R}^n$  vektorra

$$|\mathbf{Q}\mathbf{x}|^2 = \mathbf{Q}\mathbf{x} \cdot \mathbf{Q}\mathbf{x} = (\mathbf{Q}\mathbf{x})^T (\mathbf{Q}\mathbf{x}) = \mathbf{x}^T \mathbf{Q}^T \mathbf{Q} \mathbf{x} = \mathbf{x}^T \mathbf{x} = |\mathbf{x}|^2.$$

$b) \Rightarrow c)$ : A  $b)$ -ből következik, hogy

$$|\mathbf{Q}(\mathbf{x} + \mathbf{y})| = |\mathbf{x} + \mathbf{y}| \quad \text{és} \quad |\mathbf{Q}(\mathbf{x} - \mathbf{y})| = |\mathbf{x} - \mathbf{y}|.$$

Ezt, és a skalárszorítás és az abszolút érték közti kapcsolatot megadó



(1.8) egyenletet fölhasználva kapjuk, hogy minden  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  vektorra

$$\begin{aligned}\mathbf{Q}\mathbf{x} \cdot \mathbf{Q}\mathbf{y} &= \frac{1}{4} \left( |\mathbf{Q}\mathbf{x} + \mathbf{Q}\mathbf{y}|^2 - |\mathbf{Q}\mathbf{x} - \mathbf{Q}\mathbf{y}|^2 \right) \\ &= \frac{1}{4} \left( |\mathbf{Q}(\mathbf{x} + \mathbf{y})|^2 - |\mathbf{Q}(\mathbf{x} - \mathbf{y})|^2 \right) \\ &= \frac{1}{4} \left( |\mathbf{x} + \mathbf{y}|^2 - |\mathbf{x} - \mathbf{y}|^2 \right) \\ &= \mathbf{x} \cdot \mathbf{y}\end{aligned}$$

c)  $\Rightarrow$  a): A  $\mathbf{Q}$  mátrix  $i$ -edik oszlopát jelölje  $\mathbf{q}_i$ , azaz  $\mathbf{q}_i = \mathbf{Q}\mathbf{e}_i$ , ahol  $\mathbf{e}_i$  a standard bázis  $i$ -edik vektora. Ekkor

$$\mathbf{q}_i \cdot \mathbf{q}_j = \mathbf{Q}\mathbf{e}_i \cdot \mathbf{Q}\mathbf{e}_j = \mathbf{e}_i \cdot \mathbf{e}_j = \begin{cases} 0, & \text{ha } i \neq j, \\ 1, & \text{ha } i = j. \end{cases}$$

Tehát  $\mathbf{Q}$  oszlopvektorai ortonormált rendszert alkotnak, azaz  $\mathbf{Q}$  ortogonális.  $\square$

► A tétel egyik állítása úgy is kimondható, hogy egy  $\mathbf{Q}$  négyzetes mátrix pontosan akkor ortogonális, ha a  $Q : \mathbf{x} \mapsto \mathbf{Q}\mathbf{x}$  mátrixleképezés távolságtartó. A tétel másik állítása azt mondja, hogy  $\mathbf{Q}$  pontosan akkor ortogonális, ha a  $Q$  megtartja a skaláris szorzatot.

► Hasonló állítás igaz szemiortogonális mátrixokra is (ld. ?? feladat).

#### 7.73. TÉTEL (ORTOGONÁLIS MÁTRIXOK TULAJDONSÁGAI).

a) Ha  $\mathbf{Q}$  valós ortogonális mátrix, akkor  $|\det(\mathbf{Q})| = 1$ .

b) Az  $n \times n$ -es valós ortogonális mátrixok  $O(n)$  halmazából nem vezet ki a mátrixszorzás és invertálás művelete.

**BIZONYÍTÁS.** a) Mivel  $\mathbf{Q}^T \mathbf{Q} = \mathbf{I}$ , ezért  $\det(\mathbf{Q}^T) \det(\mathbf{Q}) = \det(\mathbf{I}) = 1$ , de  $\det(\mathbf{Q}^T) = \det(\mathbf{Q})$ , így  $\det(\mathbf{Q}) = 1$  vagy  $\det(\mathbf{Q}) = -1$ .

b) Ortogonális mátrix inverze megegyezik transzponáltjával, ami ugyancsak ortogonális, tehát inverze is az. Be kell még látni, hogy két ortogonális mátrix szorzata is ortogonális. Legyen  $\mathbf{Q}_1$  és  $\mathbf{Q}_2$  ortogonális. Ekkor

$$(\mathbf{Q}_1 \mathbf{Q}_2)^T \mathbf{Q}_1 \mathbf{Q}_2 = \mathbf{Q}_2^T \mathbf{Q}_1^T \mathbf{Q}_1 \mathbf{Q}_2 = \mathbf{Q}_2^T \mathbf{Q}_2 = \mathbf{I},$$

tehát  $\mathbf{Q}_1 \mathbf{Q}_2$  valóban ortogonális.  $\square$

► Az is azonnal látható, hogy az  $n \times n$ -es 1 determinánsú valós ortogonális mátrixok halmazából sem vezet ki a mátrixszorzás és invertálás művelete. E mátrixhalmazt  $SO(n)$  jelöli.

**A 2- és 3-dimenziós tér ortogonális transzformációi** Forgatások és tükrözések segítségével leírhatók az ortogonális mátrixok.

A valós ortogonális mátrixok  $O(n)$  halmaza a mátrixszorzás műveletével csoportot alkot. Ezt ortogonális csoportnak nevezik. A csoportokról a függelékben írunk. Az 1 determinánsú ortogonális mátrixok  $SO(n)$  csoportját speciális ortogonális csoportnak nevezik.  $O(10)$  fontos szerepet játszik a modern fizika húr-elméletében, mint a 10-dimenziós tér-idő szimmetriacsoportja.

**7.74. TÉTEL.** Minden  $O(2)$ -be eső ortogonális mátrix vagy egy  $\alpha$  szögű forgatás, vagy egy  $\alpha/2$  szögű egyenesre való tükrözés mátrixa.

**BIZONYÍTÁS.** Legyen  $\mathbf{Q} = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$ . Ha e mátrix ortogonális, akkor oszlopvektorai ortonormált rendszert alkotnak, azaz

$$\begin{aligned} a^2 + b^2 &= 1 \\ c^2 + d^2 &= 1 \\ ac + bd &= 0. \end{aligned}$$

Az utolsó egyenlet szerint  $a^2c^2 = b^2d^2$ , azaz  $a^2(1 - d^2) = (1 - a^2)d^2$ , amiből  $a^2 = d^2$ , és  $b^2 = c^2$  adódik. Végül kapjuk, hogy vagy  $d = a$  és  $c = -b$ , vagy  $d = -a$  és  $c = b$ . Az első esetben  $\det(\mathbf{Q}) = ad - bc = 1$ , a másodikban  $\det(\mathbf{Q}) = -1$ . Vegyük észre, hogy bármely megoldáshoz egyértelműen találunk egy olyan  $\alpha \in [0, 2\pi)$  valóst, hogy  $a = \cos \alpha$  és  $b = \sin \alpha$ . Vagyis az összes másodrendű ortogonális mátrix

$$\begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \text{ vagy } \begin{bmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{bmatrix}$$

alakba írható. Ha a determinánsa 1, akkor egy  $\alpha$  szögű forgatás, ha determinánsa  $-1$ , akkor egy  $\alpha/2$  szögű egyenesre való tükrözés mátrixa (ld. a 7.27. és a 7.32. pontokat).  $\square$

A 3-dimenziós eset kissé bonyolultabb. Számtalan klasszikus műszaki alkalmazásban – mindenek előtt a merev testek mozgásának leírásában – fontos szerepet játszanak  $SO(3)$  elemei, azaz az 1 determinánsú ortogonális mátrixok. Ezek itt és nagyobb dimenzió esetén is a forgások mátrixai. A forgás azonban csak a 3-dimenzióban írható le úgy, mint egy tengely körül  $\alpha$  szöggel való elfordulás. E tételt csak a sajátvektorok elméletének ismeretében fogjuk tudni bizonyítani (ld. ?? tétel).

**7.75. PÉLDA (FORGATÁS TENGELEJE ÉS SZÖGE).** Az 1 determinánsú ortogonális

$$\frac{1}{15} \begin{bmatrix} 14 & -5 & 2 \\ 5 & 10 & -10 \\ 2 & 10 & 11 \end{bmatrix}$$

mátrix milyen tengely körüli és mekkora szöggel való forgatás mátrixa?

**MEGOLDÁS.** Ha  $\mathbf{A}$  egy 0-tól különböző szöggel való forgatás mátrixa, és  $\mathbf{v}$  a tengely egy irányvektora, akkor csak  $\mathbf{v}$  skalárszorosai fogják kielégíteni az  $\mathbf{A}\mathbf{x} = \mathbf{x}$  egyenletet. Ez ekvivalens a homogén lineáris

$$(\mathbf{A} - \mathbf{I})\mathbf{x} = \mathbf{0}$$

egyenletrendszerrel, melynek alakja és megoldása esetünkben

$$\frac{1}{15} \begin{bmatrix} -1 & -5 & 2 \\ 5 & -5 & -10 \\ 2 & 10 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}.$$

Tehát a forgástengely egy irányvektora a  $\mathbf{v} = (2, 0, 1)$  vektor. A forgásszög, illetve a forgásszög koszinuszának meghatározásához elég egy olyan  $\mathbf{w}$  vektort találni, mely a tengelyre merőleges síkban van. Ilyen például a  $\mathbf{w} = (0, 1, 0)$  vektor. Ennek képe a forgatásnál

$$\frac{1}{15} \begin{bmatrix} 14 & -5 & 2 \\ 5 & 10 & -10 \\ 2 & 10 & 11 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/3 \\ 2/3 \\ 2/3 \end{bmatrix}.$$

A forgásszög megegyezik e két vektor szögével, tehát

$$\cos \alpha = \frac{\mathbf{w} \cdot \mathbf{Aw}}{|\mathbf{w}| |\mathbf{Aw}|} = \frac{2/3}{1 \cdot 1} = \frac{2}{3}.$$

Ez összhangban van a 7.27. feladat eredményével, ahol e forgatást a tengely és a szög ismeretében kellett megkonstruálni.  $\square$

A harmadrendű  $-1$  determinánsú ortogonális mátrixok, azaz  $O(3) - SO(3)$  elemei ugyan nem mind tükrözések, de egy tükrözés és egy forgatás egymás utáni alkalmazásával megkaphatók (ld. ?? tétel)!

**Givens-forgatás, Householder-tükrözés\*** Az  $n$ -dimenziós tér forgatásai és tükrözései közül kiválaszthatunk olyan egyszerű, ún. primitív ortogonális transzformációkat, melyek mátrixai szorzataként az összes ortogonális mátrix előállítható. E transzformációkat több hatékony numerikus matematikai módszer is használja.

Azt a forgatást, mely egy koordinátasík vektorain kívül minden más vektort helyben hagy, *Givens-forgatásnak* nevezzük. Az  $i$ -edik és  $j$ -edik koordinátatengely síkját érintő forgatás mátrixa

$$\mathbf{G} = \begin{bmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \dots & \cos \alpha & \dots & -\sin \alpha & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & \sin \alpha & \dots & \cos \alpha & \dots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix}$$

amit úgy kapunk meg, hogy az egységmátrix  $i$ -edik és  $j$ -edik sorának és oszlopának metszetében lévő négy helyre az  $\alpha$  szögű forgatás mátrixát tesszük.

E forgatással elérhető például, hogy egy  $\mathbf{x}$  vektort egy olyan vektorba forgassunk, melynek  $j$ -edik koordinátája 0. Csak az  $i$ -edik és  $j$ -edik sorokat és oszlopokat kiemelve

$$\begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}.$$

Ebből látható, hogy az

$$\begin{aligned} r &= \sqrt{a^2 + b^2} \\ \cos \alpha &= a/r \\ \sin \alpha &= -b/r \end{aligned} \quad (7.19)$$

egyenletek segítségével fölírható a forgatómátrix az  $a$  és  $b$  ismeretében. Ez használható mátrix háromszögalakra hozásában, például a következőkben vizsgált QR-felbontás Givens-forgatások segítségével is elvégezhető (ld. 7.83. példa). Ennek előnye a ritka mátrixok esetén mutatkoznak, és a számítások párhuzamosíthatóak is.

Egy adott  $\mathbf{a} \neq \mathbf{0}$  vektorra, vagy az  $\mathbf{e} = \mathbf{a}/|\mathbf{a}|$  egységvektorra merőleges hipersíkra való tükrözést *Householder-tükrözésnek* nevezzük. Mátrixa

$$\mathbf{H} = \mathbf{I} - \frac{2}{\mathbf{a}^T \mathbf{a}} \mathbf{a} \mathbf{a}^T$$

Feladatként az Olvasóra hagyjuk annak bizonyítását, hogy  $\mathbf{e}$  transzformáció valóban helyben hagyja az  $\mathbf{e}^\perp$  tér összes vektorát és  $-\mathbf{e}$ -be viszi az  $\mathbf{e}$  vektort (ld. ?? feladat). E tükrözés is használható egy mátrix háromszögalakra hozásához, QR-felbontásának megkonstruálásához. Ehhez szükség lesz az alábbi állításra.

**7.76. ÁLLÍTÁS (EGY VEKTOR TÜKRÖZÉSE EGY MÁSIKBA).** Ha  $\mathbf{a}$  és  $\mathbf{b}$  két különböző, de azonos hosszúságú vektor  $\mathbb{R}^n$ -ben, akkor az  $(\mathbf{a} - \mathbf{b})^\perp$  hipersíkra való Householder-tükrözés az  $\mathbf{a}$  vektort  $\mathbf{b}$ -be viszi és viszont.

**BIZONYÍTÁS.** Meg kell mutatnunk, hogy  $\mathbf{H}\mathbf{a} = \mathbf{b}$  és  $\mathbf{H}\mathbf{b} = \mathbf{a}$ , ahol

$$\mathbf{H} = \mathbf{I} - \frac{2}{(\mathbf{a} - \mathbf{b})^T (\mathbf{a} - \mathbf{b})} (\mathbf{a} - \mathbf{b})(\mathbf{a} - \mathbf{b})^T.$$

Kihasználjuk, hogy  $\mathbf{a}$  és  $\mathbf{b}$  azonos hosszúságúak, így  $\mathbf{a}^T \mathbf{a} = \mathbf{b}^T \mathbf{b}$ , és hogy a skaláris szorzás felcserélhető, azaz  $\mathbf{a}^T \mathbf{b} = \mathbf{b}^T \mathbf{a}$ . Így

$$(\mathbf{a} - \mathbf{b})^T (\mathbf{a} - \mathbf{b}) = \mathbf{a}^T \mathbf{a} - \mathbf{a}^T \mathbf{b} - \mathbf{b}^T \mathbf{a} + \mathbf{b}^T \mathbf{b} = 2(\mathbf{a}^T \mathbf{a} - \mathbf{b}^T \mathbf{a}) = 2(\mathbf{a} - \mathbf{b})^T \mathbf{a}.$$

Eszerint

$$\begin{aligned} \mathbf{H}\mathbf{a} &= \mathbf{a} - \frac{2}{(\mathbf{a} - \mathbf{b})^T (\mathbf{a} - \mathbf{b})} (\mathbf{a} - \mathbf{b})(\mathbf{a} - \mathbf{b})^T \mathbf{a} \\ &= \mathbf{a} - \frac{1}{(\mathbf{a} - \mathbf{b})^T \mathbf{a}} (\mathbf{a} - \mathbf{b})^T \mathbf{a} (\mathbf{a} - \mathbf{b}) \\ &= \mathbf{a} - (\mathbf{a} - \mathbf{b}) = \mathbf{b}. \end{aligned}$$

Mivel  $\mathbf{H}^{-1} = \mathbf{H}$ , ezért  $\mathbf{H}\mathbf{b} = \mathbf{H}^{-1}\mathbf{b} = \mathbf{a}$ . □

7.77. PÉLDA (HOUSEHOLDER-TÜKRÖZÉS). Határozzuk meg azt a  $\mathbf{H}$  mátrixot, mely az  $(1, -1, -1, 1)$  vektort olyan vektorba viszi, melynek az első kivéve minden koordinátája 0.

MEGOLDÁS.  $|(1, -1, -1, 1)| = 2$ , ezért a képvektor csak a  $\pm(2, 0, 0, 0)$  vektorok valamelyike lehet. Válasszuk a pozitív koordinátáját. Az  $(1, -1, -1, 1) - (2, 0, 0, 0) = (-1, -1, -1, 1)$  vektorra merőleges hipersíkra való tükrözés mátrixa

$$\begin{aligned}\mathbf{H} &= \mathbf{I} - \frac{2}{\mathbf{a}^\top \mathbf{a}} \mathbf{a} \mathbf{a}^\top = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} -1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & -1 & -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}\end{aligned}$$

Fejben számolva is könnyen ellenőrizhetjük, hogy  $\mathbf{H} \cdot (1, -1, -1, 1) = (2, 0, 0, 0)$ .  $\square$

### Ortogonalizáció

**Gram–Schmidt-ortogonalizáció\*** Nagy előnyökkel jár, ha egy altérnek nem csak egy bázisát, hanem egy ortogonális bázisát ismerjük. E paragrafusban megmutatjuk, hogy ilyen bázis létezik, és eljárást adunk a megkonstruálására. Ezt az eljárást Gram–Schmidt-ortogonalizációnak nevezzük.

7.78. TÉTEL (GRAM–SCHMIDT-ORTOGONALIZÁCIÓ). Ha  $\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k\}$  egy független vektorrendszer, akkor létezik olyan ortogonális  $\mathcal{V} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$  vektorrendszer, hogy minden  $i = 1, 2, \dots, k$  esetén

$$\text{span}(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_i) = \text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_i). \quad (7.20)$$

Az ortogonális  $\mathcal{V}$  rendszerből a vektorok normálásával kapott

$$\left\{ \frac{\mathbf{v}_1}{|\mathbf{v}_1|}, \frac{\mathbf{v}_2}{|\mathbf{v}_2|}, \dots, \frac{\mathbf{v}_k}{|\mathbf{v}_k|} \right\}$$

rendszer ortonormált.

**BIZONYÍTÁS.** A  $\text{span}(\mathbf{a}_1) = \text{span}(\mathbf{v}_1)$  összefüggés teljesül, ha

$$\mathbf{v}_1 = \mathbf{a}_1.$$

A  $\text{span}(\mathbf{a}_1, \mathbf{a}_2) = \text{span}(\mathbf{v}_1, \mathbf{v}_2)$  teljesülése érdekében olyan  $\mathbf{v}_2$  vektort kell választani, mely az  $\mathbf{a}_1$  és  $\mathbf{a}_2$  síkjában van, másrészt  $\mathbf{v}_2$ -nek merőlegesnek kell lennie  $\mathbf{v}_1$ -re. E feltételeket teljesíti az  $\mathbf{a}_2$ -nek a  $\mathbf{v}_1$  által

kifeszített altérre merőleges összetevője, azaz a

$$\mathbf{v}_2 = \mathbf{a}_2 - \left( \mathbf{a}_2 \cdot \frac{\mathbf{v}_1}{|\mathbf{v}_1|} \right) \frac{\mathbf{v}_1}{|\mathbf{v}_1|} = \mathbf{a}_2 - \frac{\mathbf{a}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1$$

vektor. Látható, hogy  $\mathbf{e}$  vektor nem lehet a 0-vektor, hisz  $\mathbf{v}_2 = 0$  esetén  $\mathbf{a}_2 = \frac{\mathbf{a}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 = \frac{\mathbf{a}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{a}_1$  lenne, azaz  $\mathbf{a}_1$  és  $\mathbf{a}_2$  nem lenne független, ami ellentmond annak, hogy  $\mathcal{A}$  független. Az előző képletekből látható, hogy  $\mathbf{v}_1$  és  $\mathbf{v}_2$  előállítható az  $\mathbf{a}_1$  és  $\mathbf{a}_2$  lineáris kombinációjaként, és viszont, így  $\text{span}(\mathbf{a}_1, \mathbf{a}_2) = \text{span}(\mathbf{v}_1, \mathbf{v}_2)$  fennáll. Az eljárás hasonlóképp folytatható. Ha már megkonstruáltuk  $\mathbf{v}_i$ -t, akkor a 7.65. tétel szerint kiszámoljuk az  $\mathbf{a}_{i+1}$  vektornak a  $\text{span}(\frac{\mathbf{v}_1}{|\mathbf{v}_1|}, \frac{\mathbf{v}_2}{|\mathbf{v}_2|}, \dots, \frac{\mathbf{v}_i}{|\mathbf{v}_i|})$  altérre merőleges összetevőjét, és ezt választjuk  $\mathbf{v}_{i+1}$ -nek, azaz

$$\mathbf{v}_{i+1} = \mathbf{a}_{i+1} - \frac{\mathbf{a}_{i+1} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 - \frac{\mathbf{a}_{i+1} \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2 - \dots - \frac{\mathbf{a}_{i+1} \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i} \mathbf{v}_i$$

Könnyen látható, hogy  $\mathbf{v}_{i+1} \neq \mathbf{0}$ , mert ellenkező esetben  $\mathcal{A}$  nem volna független. Látható az is, hogy  $\mathbf{v}_{i+1}$  kifejezhető az  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{i+1}$  vektorok lineáris kombinációjaként, és  $\mathbf{a}_{i+1}$  kifejezhető az  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{i+1}$  vektorok lineáris kombinációjaként, tehát a tétel kifeszített alterekre vonatkozó állítása is fennáll.  $\square$

**7.79. PÉLDA (GRAM-SCHMIDT-ORTOGONALIZÁCIÓ).** Keressünk ortonormált bázist az  $(1, 1, 1, 1)$ ,  $(3, -1, 3, -1)$ ,  $(6, 2, 2, -2)$  vektorok által kifeszített altérben.

**MEGOLDÁS.** Először keressünk egy ortogonális bázist:

$$\mathbf{v}_1 = (1, 1, 1, 1)$$

$$\mathbf{v}_2 = (3, -1, 3, -1) - \frac{(3, -1, 3, -1) \cdot (1, 1, 1, 1)}{(1, 1, 1, 1) \cdot (1, 1, 1, 1)} (1, 1, 1, 1) = (2, -2, 2, -2)$$

$$\begin{aligned} \mathbf{v}_3 &= (6, 2, 2, -2) - \frac{(6, 2, 2, -2) \cdot (1, 1, 1, 1)}{(1, 1, 1, 1) \cdot (1, 1, 1, 1)} (1, 1, 1, 1) \\ &\quad - \frac{(6, 2, 2, -2) \cdot (2, -2, 2, -2)}{(2, -2, 2, -2) \cdot (2, -2, 2, -2)} (2, -2, 2, -2) = (2, 2, -2, -2) \end{aligned}$$

Végül az ortonormált bázis:

$$\left\{ \left( \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2} \right), \left( \frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2} \right), \left( \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2} \right) \right\} \quad \square$$

► Könnyen igazolható, hogy a Gram-Schmidt-ortogonalizáció működik nem független vektorokból álló vektorrendszerre is, annyi változással, hogy pontosan akkor lesz  $\mathbf{v}_i = \mathbf{0}$ , ha  $\mathbf{a}_i$  nem független a kisebb indexű vektoroktól, azaz  $\mathbf{a}_i$  benne van a  $\text{span}(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{i-1})$  altérben.

**A QR-felbontás\*** Ahogyan egy mátrix elemi sorműveletekkel való háromszögalakra hozását tömör formában megőrizte az LU-felbontás, ugyanígy a QR-felbontás őrzi az ortogonalizációs eljárás eredményét. E felbontás mind a legkisebb négyzetek módszerében, mind a később tárgyalandó sajátértékprobléma megoldásában fontos szerephez jut.

**7.80. DEFINÍCIÓ (QR-FELBONTÁS).** Legyen  $\mathbf{A}$  egy teljes oszloprangú mátrix. Az  $\mathbf{A} = \mathbf{QR}$  felbontást QR-felbontásnak vagy redukált QR-felbontásnak nevezzük, ha  $\mathbf{Q}$  az  $\mathbf{A}$ -val azonos méretű szemiorтогоnalis mátrix, és  $\mathbf{R}$  négyzetes felső háromszögmátrix, főátlójában pozitív elemekkel.

► Ha a  $\mathbf{Q}$  mátrixot ortonormált oszlopvektorok hozzávételével kiegészítjük egy ortogonális mátrixszá, az  $\mathbf{R}$  mátrixot pedig zérussorok hozzávételével egy  $m \times n$ -es felső háromszögmátrixszá, akkor e mátrixok szorzata is  $\mathbf{A}$ , ugyanis

$$\mathbf{A} = \begin{bmatrix} \mathbf{Q} & \hat{\mathbf{Q}} \end{bmatrix} \begin{bmatrix} \mathbf{R} \\ \mathbf{O} \end{bmatrix} = \mathbf{QR} + \hat{\mathbf{Q}}\mathbf{O} = \mathbf{QR}$$

Ezt a bővebb felbontást is szokás QR-felbontásnak nevezni. Mi erre inkább a *teljes QR-felbontás* elnevezést használjuk. Ekkor tehát az  $\mathbf{A}$  mátrixot egy ortogonális mátrix, és egy  $\mathbf{A}$ -val azonos méretű felső háromszögmátrix szorzatára bontjuk.

► Vannak művek és programok, amelyek QR-felbontásnak tekintik azt is, ha  $\mathbf{Q}$  semiortogonális,  $\mathbf{R}$  felső háromszögmátrix, de a főátló elemeiről nincs kikötve, hogy pozitív legyen! (Pl. a mátrixalapú nyelvek is ilyen alakú felbontást adnak.) Ezeket könnyen átkonvertálhatjuk pozitív főátlójú felbontásra. Ha valamely  $i$ -re  $r_{ii} < 0$  lenne, akkor szorozzuk be az  $\mathbf{R}$  mátrix  $i$ -edik sorvektorát és a  $\mathbf{Q}$  mátrix  $i$ -edik oszlopát, azaz a  $\mathbf{q}_i$  vektort  $-1$ -gyel. Ez a szorzaton nem változtat. Így elérhetjük, hogy  $\mathbf{R}$  minden főátlón lévő eleme pozitív legyen.

A Gram–Schmidt-ortogonalizációs eljárásból könnyen előállítható egy mátrix QR-felbontása. Legyen  $\mathbf{A} = [\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_k] \in \mathbb{R}^{n \times k}$ . Mivel  $\mathbf{A}$  teljes oszloprangú, azaz oszlopai függetlenek, ezért  $k \leq n$ . Az ortogonalizációs eljárás végén kapott egységvektorokat jelölje  $\mathbf{q}_i$ , azaz  $\mathbf{q}_i = \frac{\mathbf{v}_i}{\|\mathbf{v}_i\|}$  ( $i = 1, 2, \dots, k$ ). Mivel a [Gram–Schmidt-ortogonalizációs tétel](#) szerint  $\text{span}(\mathbf{a}_1, \dots, \mathbf{a}_i) = \text{span}(\mathbf{q}_1, \dots, \mathbf{q}_i)$  minden  $i = 1, 2, \dots, k$  értékre, ezért léteznek olyan  $r_{ij}$  skalárok, hogy

$$\begin{aligned} \mathbf{a}_1 &= r_{11}\mathbf{q}_1 \\ \mathbf{a}_2 &= r_{12}\mathbf{q}_1 + r_{22}\mathbf{q}_2 \\ &\vdots \\ \mathbf{a}_k &= r_{1k}\mathbf{q}_1 + r_{2k}\mathbf{q}_2 + \dots + r_{kk}\mathbf{q}_k. \end{aligned} \tag{7.21}$$

Ezt mátrixszorzat-alakba írva épp a kívánt felbontást kapjuk:

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_k \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \dots & \mathbf{q}_k \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1k} \\ 0 & r_{22} & \dots & r_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_{kk} \end{bmatrix} = \mathbf{Q}\mathbf{R}.$$

A Gram–Schmidt-eljárásból az is látható, hogy  $r_{ii} = |\mathbf{v}_i|$ , tehát  $r_{ii} > 0$ . Ezzel bizonyítottuk a QR-felbontás létezését.

A Gram-Schmidt-eljárásból megkaphatjuk a  $\mathbf{Q}$  mátrixot, azonban kérdés, hogy az  $\mathbf{R}$  hogyan számítható ki egyszerűen. A 7.69. állítás egyszerű megoldást ad. Ha  $\mathbf{A} = \mathbf{Q}\mathbf{R}$ , akkor az egyenlőség mindkét oldalát  $\mathbf{Q}^\top$ -tal szorozva kapjuk, hogy  $\mathbf{Q}^\top\mathbf{A} = \mathbf{Q}^\top\mathbf{Q}\mathbf{R} = \mathbf{I}_k\mathbf{R} = \mathbf{R}$ , tehát

$$\mathbf{R} = \mathbf{Q}^\top\mathbf{A}.$$

7.81. PÉLDA (QR-FELBONTÁS KISZÁMÍTÁSA). *Határozzuk meg az*

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 6 \\ 1 & -1 & 2 \\ 1 & 3 & 2 \\ 1 & -1 & -2 \end{bmatrix}$$

*mátrix QR-felbontását.*

**MEGOLDÁS.** A 7.79. példában épp az  $\mathbf{A}$  mátrix oszlopvektoraiból álló vektorrendszert ortogonalizáltuk. Mivel a három vektor lineárisan független, ezért  $\mathbf{A}$  teljes oszloprangú. A 7.79. példa megoldása alapján az  $\mathbf{A}$  oszlopvektorainak ortogonalizálásával kapott vektorokkal fölírható a  $\mathbf{Q}$  mátrix:

$$\mathbf{Q} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \end{bmatrix}$$

Innen

$$\mathbf{R} = \mathbf{Q}^\top\mathbf{A} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 6 \\ 1 & -1 & 2 \\ 1 & 3 & 2 \\ 1 & -1 & -2 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 4 \\ 0 & 4 & 4 \\ 0 & 0 & 4 \end{bmatrix}$$

Valóban,

$$\begin{bmatrix} 1 & 3 & 6 \\ 1 & -1 & 2 \\ 1 & 3 & 2 \\ 1 & -1 & -2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 2 & 4 \\ 0 & 4 & 4 \\ 0 & 0 & 4 \end{bmatrix}.$$

□



**7.82. TÉTEL (QR-FELBONTÁS LÉTEZÉSE ÉS EGYÉRTELMŰSÉGE).** *Bármely valós, teljes oszloprangú  $\mathbf{A}$  mátrixnak létezik QR-felbontása, azaz létezik egy szemiortogonális  $\mathbf{Q}$  mátrix és egy  $\mathbf{R}$  felső háromszögmátrix pozitív főátlóbeli elemekkel, hogy  $\mathbf{A} = \mathbf{QR}$ . Az így kapott felbontás egyértelmű.*

**BIZONYÍTÁS.** A felbontás létezését a Gram–Schmidt-ortogonalizációra alapozva az előzőekben megmutattuk.

Tegyük fel, hogy létezik két felbontás is, azaz  $\mathbf{A} = \mathbf{QR} = \hat{\mathbf{Q}}\hat{\mathbf{R}}$ , ahol  $\mathbf{Q}$  és  $\hat{\mathbf{Q}}$  szemiortogonális, azaz  $\mathbf{Q}^T\mathbf{Q} = \hat{\mathbf{Q}}^T\hat{\mathbf{Q}} = \mathbf{I}$ . Ekkor  $\mathbf{R}^T\mathbf{R} = \hat{\mathbf{R}}^T\hat{\mathbf{R}}$ , ugyanis

$$\mathbf{A}^T\mathbf{A} = (\mathbf{QR})^T\mathbf{QR} = \mathbf{R}^T\mathbf{Q}^T\mathbf{QR} = \mathbf{R}^T\mathbf{R},$$

$$\mathbf{A}^T\mathbf{A} = (\hat{\mathbf{Q}}\hat{\mathbf{R}})^T\hat{\mathbf{Q}}\hat{\mathbf{R}} = \hat{\mathbf{R}}^T\hat{\mathbf{Q}}^T\hat{\mathbf{Q}}\hat{\mathbf{R}} = \hat{\mathbf{R}}^T\hat{\mathbf{R}}.$$

Innen  $(\hat{\mathbf{R}}^{-1})^T\mathbf{R}^T = \hat{\mathbf{R}}\mathbf{R}^{-1}$ . Itt a bal oldalon egy alsó, a jobb oldalon egy felső háromszögmátrix áll. Ez csak úgy lehetséges, ha mindkét szorzat diagonális. Jelölje  $\mathbf{R}$ , illetve  $\hat{\mathbf{R}}$  átlójának elemeit  $r_i$ , illetve  $\hat{r}_i$  ( $i = 1, \dots, n$ ). A főátló elemei tehát

$$\frac{r_i}{\hat{r}_i} = \frac{\hat{r}_i}{r_i},$$

ahonnan  $r_i > 0$  és  $\hat{r}_i > 0$  miatt  $r_i = \hat{r}_i$  következik. Eszerint  $(\hat{\mathbf{R}}^{-1})^T\mathbf{R}^T = \hat{\mathbf{R}}\mathbf{R}^{-1} = \mathbf{I}$ , amiből  $\mathbf{R} = \hat{\mathbf{R}}$  adódik, majd abból  $\mathbf{A} = \mathbf{QR} = \hat{\mathbf{Q}}\hat{\mathbf{R}}$  miatt  $\mathbf{Q} = \hat{\mathbf{Q}}$ .  $\square$

**QR-felbontás primitív ortogonális transzformációkkal\*** A QR-felbontás a Gram–Schmidt-ortogonalizáció helyett más technikákkal, így például a primitív ortogonális transzformációkkal is kiszámolható. Hasonlóan az Gauss-eliminációhoz itt is a háromszögalakra hozást eliminációval valósítjuk meg, de most nem elemi mátrixokkal, hanem ortogonálisokkal.

**7.83. PÉLDA (QR-FELBONTÁS GIVENS-FORGATÁSOKKAL).** *Határozzuk meg az*

$$\mathbf{A} = \begin{bmatrix} 4 & 5 & 8 \\ 3 & 10 & 6 \\ 0 & 12 & 13 \end{bmatrix}$$

*mátrix QR-felbontását Givens-forgatások segítségével!*

**MEGOLDÁS.** Először az első és második sorokat és oszlopokat figyelve elimináljuk a második sor első elemét. Itt a (7.19) egyenletekben is használt jelölésekkel  $a = 4$ ,  $b = 3$ , tehát  $r = \sqrt{3^2 + 4^2} = 5$ ,  $\cos \alpha = 4/5$ ,  $\sin \alpha = -3/5$ . Így első lépésben a következő mátrixszorzással eliminálhatunk:

$$\mathbf{Q}_1 = \begin{bmatrix} 4/5 & 3/5 & 0 \\ -3/5 & 4/5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathbf{Q}_1\mathbf{A} = \begin{bmatrix} 5 & 10 & 10 \\ 0 & 5 & 0 \\ 0 & 12 & 13 \end{bmatrix}.$$

Következő lépésben a  $\mathbf{Q}_1\mathbf{A}$  mátrix harmadik sorának második elemét elimináljuk:

$$\mathbf{Q}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5/13 & 12/13 \\ 0 & -12/13 & 5/13 \end{bmatrix}, \quad \mathbf{R} = \mathbf{Q}_2\mathbf{Q}_1\mathbf{A} = \begin{bmatrix} 5 & 10 & 10 \\ 0 & 13 & 12 \\ 0 & 0 & 5 \end{bmatrix}.$$

és innen

$$\mathbf{Q} = (\mathbf{Q}_2\mathbf{Q}_1)^{-1} = \mathbf{Q}_1^T\mathbf{Q}_2^T = \begin{bmatrix} 4/5 & -3/13 & 36/65 \\ 3/5 & 4/13 & -48/65 \\ 0 & 12/13 & 5/13 \end{bmatrix},$$

amely mátrixokkal  $\mathbf{A} = \mathbf{QR}$  valóban fennáll.  $\square$

A Householder-tükrözést alkalmazva a QR-felbontásra egy további módszert kaphatunk. Az ötlet lényege, hogy a 7.77. példában látott módon először az első oszlopban elimináljuk az első elem alattiakat, majd olyan transzformációt választunk, mely az első sort és oszlopot nem változtatja, de a második sor második eleme alattiakat eliminálja, és így tovább. Egy  $4 \times 4$ -es mátrixon szemléltetjük az eljárást.

$$\mathbf{A} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} \rightarrow \mathbf{Q}_1\mathbf{A} = \begin{bmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{bmatrix} \rightarrow \mathbf{Q}_2\mathbf{Q}_1\mathbf{A} = \begin{bmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & * & * \end{bmatrix} \rightarrow \mathbf{Q}_3\mathbf{Q}_2\mathbf{Q}_1\mathbf{A} = \begin{bmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \end{bmatrix}$$

$$\mathbf{Q}_1 = \mathbf{H}_1 \quad \mathbf{Q}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & \mathbf{H}_2 & \\ 0 & & & \end{bmatrix} \quad \mathbf{Q}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & & \\ 0 & 0 & & \mathbf{H}_3 \end{bmatrix}$$

Az első lépésben az  $\mathbf{A}$  mátrix első oszlopához ( $\mathbf{a}_1$ ) keresünk egy  $\mathbf{b}_1$  vektort, mely vele egyenlő hosszú, és csak az első koordinátája nem 0. Ezután az  $\mathbf{a}_1 - \mathbf{b}_1$  vektorhoz megkonstruáljuk a  $\mathbf{Q}_1 = \mathbf{H}_1$  Householder-mátrixot. Így  $\mathbf{Q}_1\mathbf{A}$  első oszlopában kinulláztuk az első sor alatti elemeket. Ezután elhagyjuk az első sort és oszlopot, és az így kapott mátrix első oszlopvektorával ( $\mathbf{a}_2$ ) és a vele egyenlő hosszú, és az első koordinátát kivéve 0 koordinátájú  $\mathbf{b}_2$  vektorral megkonstruáljuk a  $\mathbf{H}_2$  Householder-mátrixot, melyet kiegészítünk egy sorral és oszloppal úgy, hogy a  $\mathbf{Q}_1\mathbf{A}$  mátrixszal szorozva annak első sorát és oszlopát ne változtassa. Ez lesz a  $\mathbf{Q}_2$  mátrix. Hasonlóan folytatva végül egy  $\mathbf{R} = \mathbf{Q}_{n-1} \dots \mathbf{Q}_2\mathbf{Q}_1\mathbf{A}$  felső háromszögmátrixhoz jutunk (a fenti mintán  $n = 4$ ). Mivel a  $\mathbf{Q}_i$  mátrixok mindegyike ortogonális, ezért szoroztuk inverze is az lesz. Így a QR-felbontásban  $\mathbf{Q} = \mathbf{Q}_1^T\mathbf{Q}_2^T \dots \mathbf{Q}_{n-1}^T$ . A QR-felbontás ilyen módon való meghatározását nevezzük *Householder-módszernek*.

7.84. PÉLDA (QR-FELBONTÁS HOUSEHOLDER-TÜKRÖZÉSEL). *Határozzuk meg az*

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & -3 \\ -2 & 5 & -7 \end{bmatrix}$$

*mátrix QR-felbontását Householder-módszerrel!*

**MEGOLDÁS.** Az  $(1, 2, -2) \mapsto (3, 0, 0)$  transzformációhoz az

$$\mathbf{a} = (1, 2, -2) - (3, 0, 0) = (-2, 2, -2)$$

vektorral Householder-tükrözést végzünk:

$$\begin{aligned} \mathbf{Q}_1 &= \mathbf{I}_3 - \frac{2}{\mathbf{a}^\top \mathbf{a}} \mathbf{a} \mathbf{a}^\top = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{1}{6} \begin{bmatrix} 4 & -4 & 4 \\ -4 & 4 & -4 \\ 4 & -4 & 4 \end{bmatrix} \\ &= \frac{1}{3} \begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix} \\ \mathbf{Q}_1 \mathbf{A} &= \frac{1}{3} \begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & -3 \\ -2 & 5 & -7 \end{bmatrix} = \begin{bmatrix} 3 & -2 & 3 \\ 0 & 4 & -5 \\ 0 & 3 & -5 \end{bmatrix} \end{aligned}$$

Ezután a  $\mathbf{Q}_1 \mathbf{A}$  mátrixból képzeletben elhagyva az első sort és oszlopot a  $(4, 3) \mapsto (5, 0)$  transzformációhoz kell az  $\mathbf{a} = (4, 3) - (5, 0) = (-1, 3)$  vektorral Householder-tükrözést végezni:

$$\begin{aligned} \mathbf{H}_2 &= \mathbf{I}_2 - \frac{2}{\mathbf{a}^\top \mathbf{a}} \mathbf{a} \mathbf{a}^\top = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{1}{5} \begin{bmatrix} 1 & -3 \\ -3 & 9 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 4 & 3 \\ 3 & -4 \end{bmatrix} \\ \mathbf{Q}_2 &= \left[ \begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & 4/5 & 3/5 \\ 0 & 3/5 & -4/5 \end{array} \right], \quad \mathbf{R} = \mathbf{Q}_2 \mathbf{Q}_1 \mathbf{A} = \begin{bmatrix} 3 & -2 & 3 \\ 0 & 5 & -7 \\ 0 & 0 & 1 \end{bmatrix} \\ \mathbf{Q} &= (\mathbf{Q}_2 \mathbf{Q}_1)^{-1} = \mathbf{Q}_1^\top \mathbf{Q}_2^\top = \frac{1}{15} \begin{bmatrix} 5 & 2 & 14 \\ 10 & 10 & -5 \\ -10 & 11 & 2 \end{bmatrix}. \end{aligned}$$

Az  $\mathbf{A} = \mathbf{Q} \mathbf{R}$  egyenlőség fennállásának ellenőrzését az Olvasóra hagyjuk.  $\square$

**Egyenletrendszer optimális megoldása QR-felbontással\*** Ha egy egyenletrendszer ellentmondásos, az optimális megoldás megtalálásához fölírt normálegyenlet gyakran rosszul kondicionált, ezért érdemes olyan megoldási technikát keresni, mely hatékonyabb a számítási hibák kezelésében. Egy ilyen technikát ismertetünk.

**7.85. TÉTEL (LEGKISEBB NÉGYZETEK QR-FELBONTÁSSAL).** Legyen  $\mathbf{A}$  egy teljes oszloprangú  $m \times n$ -es valós mátrix,  $\mathbf{A} = \mathbf{QR}$  egy QR-felbontása, és legyen  $\mathbf{b}$  egy  $\mathbb{R}^m$ -beli vektor. Ekkor az  $\mathbf{Ax} = \mathbf{b}$  egyenletrendszer egyetlen optimális megoldása  $\hat{\mathbf{x}} = \mathbf{R}^{-1}\mathbf{Q}^T\mathbf{b}$ , ami megkapható az

$$\mathbf{R}\hat{\mathbf{x}} = \mathbf{Q}^T\mathbf{b}$$

egyenletrendszerből egyszerű visszahelyettesítéssel is.

**BIZONYÍTÁS.** Az egyenletrendszer optimális megoldásáról szóló 7.46. tétel szerint az optimális megoldás a normálegyenletből megkapható. Eszerint

$$\begin{aligned} \mathbf{A}^T\mathbf{A}\hat{\mathbf{x}} &= \mathbf{A}^T\mathbf{b} & \mathbf{A} = \mathbf{QR} \text{ behelyettesítése után} \\ (\mathbf{QR})^T\mathbf{QR}\hat{\mathbf{x}} &= (\mathbf{QR})^T\mathbf{b} \\ \mathbf{R}^T\mathbf{Q}^T\mathbf{QR}\hat{\mathbf{x}} &= \mathbf{R}^T\mathbf{Q}^T\mathbf{b} & \mathbf{Q}^T\mathbf{Q} = \mathbf{I} \\ \mathbf{R}^T\mathbf{R}\hat{\mathbf{x}} &= \mathbf{R}^T\mathbf{Q}^T\mathbf{b} & \text{balról szorzás az } (\mathbf{R}^T)^{-1} \text{ mátrixszal} \\ \mathbf{R}\hat{\mathbf{x}} &= \mathbf{Q}^T\mathbf{b}. \end{aligned}$$

Az utolsó egyenlet visszahelyettesítésekkel is megoldható, mivel  $\mathbf{R}$  felső háromszögmátrix. Mivel  $\mathbf{R}$  főátlójában nincsenek zéruselemek, ezért  $\mathbf{R}$  invertálható (ezt kihasználtuk, amikor  $(\mathbf{R}^T)^{-1}$ -gyel szoroztunk), tehát az egyenletből  $\hat{\mathbf{x}}$  kifejezhető:  $\hat{\mathbf{x}} = \mathbf{R}^{-1}\mathbf{Q}^T\mathbf{b}$ .  $\square$

**7.86. PÉLDA (EGYENLETRENDSZER OPTIMÁLIS MEGOLDÁSA).** Az alábbi háromismeretlenes egyenletrendszert megoldottuk a 7.62. példában:

$$\begin{aligned} x + 3y + 6z &= 8 \\ x - y + 2z &= 2 \\ x + 3y + 2z &= 2 \\ x - y - 2z &= 0 \end{aligned}$$

Adjunk rá új, a QR-felbontást használó megoldást!

**MEGOLDÁS.** Az egyenletrendszer együtthatómátrixának QR-felbontását meghatároztuk a 7.81. példában. Eszerint

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 6 \\ 1 & -1 & 2 \\ 1 & 3 & 2 \\ 1 & -1 & -2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 2 & 4 \\ 0 & 4 & 4 \\ 0 & 0 & 4 \end{bmatrix}.$$

Egyik lehetőség, hogy fölírjuk az  $\mathbf{R}\hat{\mathbf{x}} = \mathbf{Q}^T\mathbf{b}$  mátrixegyenletet:

$$\begin{bmatrix} 2 & 2 & 4 \\ 0 & 4 & 4 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 8 \\ 2 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \\ 4 \end{bmatrix}$$

Ezt az egyenletrendszert fejben is meg tudjuk oldani visszahelyettesítéssel:  $(\hat{x}_1, \hat{x}_2, \hat{x}_3) = (1, 0, 1)$ . Természetesen ha már kiszámoltuk az  $\mathbf{R}^{-1}$  mátrixot, akkor segítségével is megkapható az optimális megoldás:

$$\hat{\mathbf{x}} = \mathbf{R}^{-1} \mathbf{Q}^T \mathbf{b} = \frac{1}{4} \begin{bmatrix} 2 & -1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 8 \\ 2 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

□

**Feladatok**

7.36.  $\mathbb{R}^4$  egy 2-dimenziós alterének bázisvektora az  $(1, 1, -1, -1)$  és a  $(9, 3, -1, 5)$  vektorok. A Gram–Schmidt-eljárással adjuk meg az alter egy ortonormált bázisát.

7.37.  $\mathbb{R}^4$  egy 2-dimenziós alterének bázisvektora az  $(1, -1, 1, -1)$  és a  $(8, 6, 2, 0)$  vektorok. A Gram–Schmidt-eljárással adjuk meg az alter egy ortonormált bázisát.

Adjuk meg az alábbi mátrixok QR-felbontását a Gram–Schmidt-eljárás alkalmazásával!

7.38. 
$$\begin{bmatrix} 1 & 2 \\ 8 & -20 \\ 4 & -1 \end{bmatrix}$$

7.39. 
$$\begin{bmatrix} 1 & 8 \\ -1 & 6 \\ 1 & 2 \\ -1 & 0 \end{bmatrix}$$

7.40. Adjuk meg az alábbi mátrix QR-felbontását Givens-forgatással!

$$\begin{bmatrix} 8 & 3 & -4 \\ 0 & 4 & 2 \\ 15 & 12 & 1 \end{bmatrix}$$

7.41. Adjuk meg az alábbi mátrix QR-felbontását Householder-tükrözésekkel!

$$\begin{bmatrix} 1 & -1 & 2 & 4 \\ 1 & 4 & 1 & 1 \\ 1 & 4 & -1 & 2 \\ 1 & -1 & 4 & 1 \end{bmatrix}$$

7.42. Milyen geometriai interpretáció adható az  $\mathbf{A}$  mátrix oszlopvektoraival kifejezve az  $\mathbf{A}$  QR-felbontásában szereplő  $\mathbf{R}$  mátrix főátlóbeli elemeinek?

7.43. Számítsuk ki a

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

mátrix QR-felbontását!

7.44. Igazoljuk, hogy ha  $\mathbf{e}$  az  $\mathbb{R}^n$  egy egységvektora, akkor  $\mathbf{H} = \mathbf{I} - 2\mathbf{e}\mathbf{e}^T$  lesz a Householder-transzformáció mátrixa, mely helyben hagyja az  $\mathbf{e}^\perp$  altér összes vektorát és  $\mathbf{H}\mathbf{e} = -\mathbf{e}$ .

## Komplex és véges test feletti terek\*

A következőkben egyre többször szembesülünk azzal, hogy valós számokkal megfogalmazható problémák megválaszolásához is szükség van a komplex számokra. E fejezetben a geometriai szemléletmódot is kiterjesztjük a komplex terekre. Meglepő módon a geometriai analógiák még a véges test feletti terek esetén is jól használhatók.

### Komplex vektorok és terek

**Komplex vektorok skaláris szorzata** Ha  $\mathbb{C}^n$ -beli vektorok skaláris szorzatát úgy értelmeznénk, mint a valós vektorok esetén, fura dolgok történnének.

Vegyük például a  $(1, i)$  és az  $(i, i)$  vektorokat. Önmagával vett skaláris szorzata e két vektornak ez lenne:

$$\begin{aligned}(1, i) \cdot (1, i) &\stackrel{?}{=} 1 - 1 = 0 \\ (i, i) \cdot (i, i) &\stackrel{?}{=} -1 - 1 = -2\end{aligned}$$

Ez azt mutatja, hogy ha a komplex vektorok abszolút értékét (hosszát), a valósban használt skaláris szorzattal definiálnánk, az abszolút érték legfontosabb tulajdonságai nem maradnának igazak! Kérdés, kiterjeszthető-e a valós vektorok skaláris szorzatának definíciója a komplex vektorokra úgy, hogy a fontosabb tulajdonságok érvényben maradjanak? Az ötletet a komplex számok – mint egydimenziós vektorok – abszolút értéke adja. A  $z = a + ib$  szám abszolút értékének négyzete  $z\bar{z}$ , és nem  $z^2$ ! Eszerint az egydimenziós  $z$  vektor önmagával vett skaláris szorzatának  $z\bar{z}$ -t vagy  $\bar{z}z$ -t kell adnia. Ennek megfelelően a  $\mathbf{z} = (z_1, z_2, \dots, z_n)$  és a  $\mathbf{w} = (w_1, w_2, \dots, w_n)$  vektorok skaláris szorzatának egy lehetséges definíciója

$$\begin{aligned}\mathbf{z} \cdot \mathbf{w} &= z_1 \bar{w}_1 + z_2 \bar{w}_2 + \dots + z_n \bar{w}_n, \text{ vagy} \\ \mathbf{z} \cdot \mathbf{w} &= \bar{z}_1 w_1 + \bar{z}_2 w_2 + \dots + \bar{z}_n w_n.\end{aligned}$$

Mindkét fenti képlet használható, ízlés kérdése melyiket választjuk (könyvenként változik). Mi az utóbbit fogjuk használni a skaláris szorzat mátrixszorzatos alakjának egyszerűbb volta miatt (ld. majd a 7.88. definícióban). Mindenek előtt egy elnevezés:

**7.87. DEFINÍCIÓ (KOMPLEX MÁTRIX ADJUNGÁLTJA).** Az  $\mathbf{A}$  komplex mátrix adjungáltján (vagy Hermite-féle transzponáltján) elemenkénti konjugáltjának transzponáltját értjük. Az  $\mathbf{A}$  adjungáltját  $\mathbf{A}^*$ , vagy Hermite neve után  $\mathbf{A}^H$  jelöli, tehát  $\mathbf{A}^H = \overline{\mathbf{A}}^T$ .

Például  $\begin{bmatrix} 1 & 1+i \\ -i & 2 \end{bmatrix}^H = \begin{bmatrix} -i & 1 \\ 1-i & 2 \end{bmatrix}$ , míg  $[1-i \ i]^H = \begin{bmatrix} 1+i \\ -i \end{bmatrix}$ .

**7.88. DEFINÍCIÓ (KOMPLEX VEKTOROK SKALÁRIS SZORZATA).** A  $\mathbb{C}^n$ -beli  $\mathbf{z} = (z_1, z_2, \dots, z_n)$  és  $\mathbf{w} = (w_1, w_2, \dots, w_n)$  vektorok skaláris szorzatán a

$$\mathbf{z} \cdot \mathbf{w} = \bar{z}_1 w_1 + \bar{z}_2 w_2 + \dots + \bar{z}_n w_n$$

komplex skalárt értjük. Ennek mátrixszorzatos alakja  $\mathbf{z} \cdot \mathbf{w} = \mathbf{z}^H \mathbf{w}$ .

Így a fent említett  $(1, i)$  és az  $(i, i)$  önmagukkal és egymással vett skaláris szorzatai:

$$\begin{aligned} (1, i) \cdot (1, i) &= \begin{bmatrix} 1 \\ i \end{bmatrix}^H \begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} 1 & -i \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = 1 - i^2 = 2, \\ (i, i) \cdot (i, i) &= \begin{bmatrix} i \\ i \end{bmatrix}^H \begin{bmatrix} i \\ i \end{bmatrix} = \begin{bmatrix} -i & -i \end{bmatrix} \begin{bmatrix} i \\ i \end{bmatrix} = -i^2 - i^2 = 2, \\ (1, i) \cdot (i, i) &= \begin{bmatrix} 1 \\ i \end{bmatrix}^H \begin{bmatrix} i \\ i \end{bmatrix} = \begin{bmatrix} 1 & -i \end{bmatrix} \begin{bmatrix} i \\ i \end{bmatrix} = i - i^2 = 1 + i, \\ (i, i) \cdot (1, i) &= \begin{bmatrix} i \\ i \end{bmatrix}^H \begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} -i & -i \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = -i - i^2 = 1 - i. \end{aligned}$$

► Világos, hogy két valós vektor skaláris szorzata az eredeti és e definíció szerint is ugyanazt az eredményt adja, ugyanis minden valós  $r$  számra  $\bar{r} = r$ , tehát valós  $\mathbf{u}$  és  $\mathbf{v}$  vektorok esetén  $\mathbf{u}^H = \mathbf{u}^T$ , így  $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^H \mathbf{v} = \mathbf{u}^T \mathbf{v}$ . Tehát a fenti definíció kiterjesztése a valósban használt definíciónak.

► E definícióval a vektorok hosszára vonatkozó tulajdonságok is érvényben maradnak, amit hamarosan belátunk (ld. ??? tétel).

Az adjungált tulajdonságai kiterjesztései a valós mátrixok transzponáltja tulajdonságainak, hisz valós mátrix konjugáltja megegyezik önmagával. Ez azonnal bizonyítja az alábbi tételt:

**7.89. TÉTEL (AZ ADJUNGÁLT TULAJDONSÁGAI).** Legyenek  $\mathbf{A}$  és  $\mathbf{B}$  komplex mátrixok,  $c$  komplex szám. Ekkor

- $(\mathbf{A}^H)^H = \mathbf{A}$ ,
- $(\mathbf{A} + \mathbf{B})^H = \mathbf{A}^H + \mathbf{B}^H$ ,
- $(c\mathbf{A})^H = \bar{c}\mathbf{A}^H$
- $(\mathbf{AB})^H = \mathbf{B}^H \mathbf{A}^H$ .

Az adjungált tulajdonságaiból azonnal következik a következő tétel:

**7.90. TÉTEL (A KOMPLEX SKALÁRIS SZORZÁS TULAJDONSÁGAI).** Legyen  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{C}^n$ , és legyen  $c \in \mathbb{C}$ . Ekkor

- $\mathbf{u} \cdot \mathbf{v} = \overline{\mathbf{v} \cdot \mathbf{u}}$

Legyünk óvatosak az *adjungált* kifejezéssel: a könyvünk determinánsokról szóló fejezetében *klasszikus adjungáltnak* nevezett fogalmat ne keverjük össze ezzel az adjungálttal, nincs közük egymáshoz!



- b)  $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$ ,  
 c)  $(c\mathbf{u}) \cdot \mathbf{v} = \overline{c}(\mathbf{u} \cdot \mathbf{v})$  és  $\mathbf{u} \cdot (c\mathbf{v}) = c(\mathbf{u} \cdot \mathbf{v})$ ,  
 d)  $\mathbf{u} \cdot \mathbf{u} > 0$ , ha  $\mathbf{u} \neq \mathbf{0}$ , és  $\mathbf{u} \cdot \mathbf{u} = 0$ , ha  $\mathbf{u} = \mathbf{0}$ .

► Könnyen látható, hogy e tétel kiterjesztése a valós térbeli vektorokra kimondott 1.18. tételnek, bár első pillanatra úgy tűnhet, hogy ellentmond neki. Például valósban a skaláris szorzás kommutatív, itt nem, de a most kimondott változat érvényes valós vektorokra is, hisz valós vektor konjugáltja megegyezik önmagával. Hasonló állítható a c) tulajdonságról is.

► A c)-beli két tulajdonság bármelyike következik a másikkól az a) alkalmazásával. Ha a skaláris szorzatot az  $\mathbf{u} \cdot \mathbf{v} = \mathbf{v}^H \mathbf{u}$  képlettel definiáltuk volna, akkor a természetesebbnek ható  $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v})$  összefüggés volna igaz.

► d)-ben az is az állítás része, hogy egy komplex vektor önmagával vett skaláris szorzata egyáltalán valós szám.

► A d) úgy is megfogalmazható, hogy  $\mathbf{u} \cdot \mathbf{u} \geq 0$ , és  $\mathbf{u} \cdot \mathbf{u} = 0$  pontosan akkor áll fenn, ha  $\mathbf{u} = \mathbf{0}$ .

**BIZONYÍTÁS.** A bizonyítás a skaláris szorzás mátrixszorzatos alakjából, és a konjugált tulajdonságaiból azonnal adódik. Példaként megmutatjuk az a) bizonyítását:

$$\begin{aligned} \overline{\mathbf{v} \cdot \mathbf{u}} &= \overline{\mathbf{v}^H \mathbf{u}} = \overline{\mathbf{v}^T \mathbf{u}} = \mathbf{v}^T \overline{\mathbf{u}} = \overline{\mathbf{u}}^T \mathbf{v} = \mathbf{u}^H \mathbf{v} \\ &= \mathbf{u} \cdot \mathbf{v} \end{aligned}$$

A többi állítás hasonlóan bizonyítható. □

**$\mathbb{C}^n$  kitüntetett alterei** Mivel a komplex mátrixszorzásban nem az egyik mátrix sorának és a másik egy oszlopának skaláris szorzatát számítjuk ki, megváltoznak a kitüntetett alterek.

Egy  $\mathbf{A} \in \mathbb{C}^{m \times n}$  esetén az  $\mathbf{A}\mathbf{x}$  szorzatban nem  $\mathbf{A}$  sortérének egy  $\mathbf{a}_i$  vektorát szorozzuk az  $\mathbf{x}$  vektorral, hisz  $\mathbf{a}_i \cdot \mathbf{x} = \overline{\mathbf{a}_i}^T \mathbf{x}$ . Így az  $\mathbf{A}\mathbf{x} = \mathbf{0}$  egyenlet azt jelenti, hogy a sortér vektorainak konjugáltjaiból álló tér merőleges a nulltérre, így az is adódik, hogy  $\mathbb{C}^n = \mathcal{S}(\overline{\mathbf{A}}) \oplus \mathcal{N}(\mathbf{A})$  és hasonlóképp  $\mathbb{C}^m = \mathcal{O}(\overline{\mathbf{A}}) \oplus \mathcal{N}(\mathbf{A}^T)$ . Az  $\overline{\mathbf{A}}$  sortere nyilván megegyezik az  $\mathbf{A}$  sortere konjugáltjával, azaz  $\overline{\mathcal{S}(\mathbf{A})}$ -val. Az  $\mathbf{A}$  komplex mátrix kitüntetett alterein tehát az  $\overline{\mathcal{S}(\mathbf{A})} = \mathcal{O}(\mathbf{A}^H)$ ,  $\mathcal{N}(\mathbf{A})$ ,  $\mathcal{O}(\mathbf{A})$ ,  $\mathcal{N}(\mathbf{A}^H)$  tereket értjük. Innen adódik a következő tétel:

**7.91. TÉTEL (KOMPLEX MÁTRIX KITÜNTETETT ALTEREI).** Az  $\mathbf{A} \in \mathbb{C}^{m \times n}$  mátrix kitüntetett altereire igazak a következő állítások:

- a)  $\mathbb{C}^n = \mathcal{O}(\mathbf{A}^H) \oplus \mathcal{N}(\mathbf{A})$ ,  $\mathbb{C}^m = \mathcal{O}(\mathbf{A}) \oplus \mathcal{N}(\mathbf{A}^H)$ ,  
 b)  $\mathcal{O}(\mathbf{A}^H) \perp \mathcal{N}(\mathbf{A})$ ,  $\mathcal{O}(\mathbf{A}) \perp \mathcal{N}(\mathbf{A}^H)$ ,  
 c) az  $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$  mátrixleképezés kölcsönösen egyértelmű  $\mathcal{O}(\mathbf{A}^H)$  és  $\mathcal{O}(\mathbf{A})$  között.

**Önadjungált mátrixok** Ahogy a transzponált fogalmának a – komplex skaláris szorzatot figyelembe vevő – kiterjesztése az adjungált, ugyanúgy a szimmetrikus mátrix fogalmának kiterjesztése az önadjungált mátrix. Szimmetrikus mátrix az, amelyik megegyezik saját transzponáltjával, önadjungált az, amelyik megegyezik saját adjungáltjával.

Az  $\mathbf{A}$  komplex mátrix *önadjungált*, ha

$$\mathbf{A}^H = \mathbf{A}. \quad (7.22)$$

- Az önadjungált mátrixokat *Hermite-féle mátrixnak* is nevezik.
- Világos, hogy önadjungált mátrix főátlójában csak valósok állhatnak, mert csak azok egyeznek meg saját konjugáltjukkal.
- Minden valós szimmetrikus mátrix önadjungált, hisz a valós számok megegyeznek saját konjugáltjukkal. Sőt, mivel a nem valós komplex számok nem egyeznek meg saját konjugáltjukkal, ezért a komplex szimmetrikus mátrixok pontosan akkor önadjungáltak, ha minden elemük valós.

7.92. PÉLDA (ÖNADJUNGÁLT MÁTRIXOK). Az

$$\begin{bmatrix} 1 & i & 1+i \\ -i & 2 & 2-3i \\ 1-i & 2+3i & 3 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, \quad \begin{bmatrix} i & 1+i \\ 1-i & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1+i \\ 1+i & 2 \end{bmatrix}$$

mátrixok közül az első kettő önadjungált, a harmadik nem, mert főátlójában nem minden szám valós, a negyedik sem, az viszont komplex szimmetrikus mátrix!

**Távolság és a merőleges vetítés komplex terekben** A komplex skaláris szorzás segítségével – a valós esethez hasonlóan – definiálható a komplex vektorok távolsága és szöge, és így a merőlegessége is.

A komplex  $\mathbf{u} \in \mathbb{C}^n$  vektor *hossza*, vagy *abszolút értéke*  $|\mathbf{u}| = \sqrt{\mathbf{u} \cdot \mathbf{u}}$ , két vektor *távolsága* megegyezik különbségük hosszával, azaz  $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$  vektorok esetén  $d(\mathbf{u}, \mathbf{v}) = |\mathbf{u} - \mathbf{v}|$ . Két vektor *szögének* koszinusza, ahogy azt az (1.4) képlettel definiáltuk,  $\cos(\mathbf{u}, \mathbf{v})_{\angle} := \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}||\mathbf{v}|}$ . Így két vektor *merőleges*, ha skaláris szorzatuk 0.

**Unitér mátrixok** Az ortogonális mátrixok komplex analogonjai az unitér mátrixok.

7.93. DEFINÍCIÓ (UNITÉR MÁTRIX). Egy komplex négyzetes  $\mathbf{U}$  mátrix unitér, ha  $\mathbf{U}^H \mathbf{U} = \mathbf{I}$ .

- Az ortogonális mátrixokhoz hasonlóan bizonyítható, hogy egy  $\mathbf{U} \in \mathbb{C}^{n \times n}$  mátrix pontosan akkor unitér, ha az alábbiak bármelyike teljesül:
  - $\mathbf{U} \mathbf{U}^H = \mathbf{I}$ ,

- $U^{-1} = U^H$ ,
- $U$  oszlopvektorai ortonormált bázist alkotnak a komplex skalárszorzásra nézve,
- $U$  sorvektorai ortonormált bázist alkotnak a komplex skalárszorzásra nézve,
- $|Ux| = |x|$  minden  $x \in \mathbb{C}^n$  vektorra,
- $Ux \cdot Uy = x \cdot y$ .

### Diszkrét Fourier-transzformált

**Fourier-mátrixok** Az  $N$ -edik komplex egységgyök hatványaiból képzett Vandermonde-mátrix kiemelkedően fontos szerepet kapott a modern műszaki alkalmazásokban. E mátrix alaptulajdonságainak megismeréséhez a Fourier-összegek együttthatói és helyettesítési értékei közti kapcsolaton keresztül közelítünk.

A Fourier-sorok komplex

$$\sum_{n=-\infty}^{\infty} c_n e^{nit}$$

alakja, és ezek

$$\sum_{n=0}^{N-1} c_n e^{nit} = c_0 + c_1 e^{it} + c_1 e^{2it} + \dots + c_{N-1} e^{(N-1)it} \quad (7.23)$$

alakú részletösszegei kulcsszerepet játszanak a periodikus, illetve a korlátos tartományon értelmezett függvények leírásában. A (7.23) összeget (diszkrét) *Fourier-összegnek* nevezzük.

**7.94. ÁLLÍTÁS (FOURIER-ÖSSZEG HELYETTESÍTÉSI ÉRTÉKEI).** A (7.23) Fourier-összeg együttthatóihoz a Fourier-összegnek a  $[0, 2\pi]$  intervallumot  $N$  egyenlő részre osztó  $0, \frac{2\pi}{N}, \frac{4\pi}{N}, \dots, \frac{2(N-1)\pi}{N}$  pontokban vett helyettesítési értékeit rendelő leképezés lineáris, melynek mátrixa  $[e^{\frac{2\pi i}{N} mn}]$  ( $0 \leq m, n < N$ ).

**BIZONYÍTÁS.** Először vizsgáljuk meg az  $N = 3$  esetet. Az osztópontok:  $t_0 = 0, t_1 = 2\pi/3, t_2 = 4\pi/3$ . A Fourier-összeg  $c_0 + c_1 e^{it} + c_2 e^{2it}$ , ennek  $t_k$ -beli helyettesítési értékét jelölje  $y_k$ . Tehát

$$\begin{aligned} y_0 &= c_0 + c_1 e^{i0} + c_2 e^{2i0} = c_0 + c_1 + c_2 \\ y_1 &= c_0 + c_1 e^{\frac{2\pi i}{3}} + c_2 e^{\frac{4\pi i}{3}} = c_0 + c_1 \varepsilon + c_2 \varepsilon^2 \\ y_2 &= c_0 + c_1 e^{\frac{4\pi i}{3}} + c_2 e^{\frac{8\pi i}{3}} = c_0 + c_1 \varepsilon^2 + c_2 \varepsilon^4 \end{aligned}$$

ahol  $\varepsilon = e^{\frac{2\pi i}{3}}$  a legkisebb pozitív argumentumú harmadik komplex egységgyököt jelöli. Világos, hogy a  $(c_0, c_1, c_2) \mapsto (y_0, y_1, y_2)$  leképezés

lineáris, melynek mátrixszorzat-alakja

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \varepsilon & \varepsilon^2 \\ 1 & \varepsilon^2 & \varepsilon^4 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix}$$

Hasonlóan egyszerű az általános eset is, azonban még tekintsük át az  $N = 2$  és az  $N = 4$  eset is.  $N = 2$  esetén  $\varepsilon = e^{\frac{2\pi i}{2}} = -1$  a primitív egységyökök, így

$$\begin{bmatrix} y_0 \\ y_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & \varepsilon \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix},$$

míg  $N = 4$  esetén  $\varepsilon = e^{\frac{2\pi i}{4}} = i$ , tehát

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \varepsilon & \varepsilon^2 & \varepsilon^3 \\ 1 & \varepsilon^2 & \varepsilon^4 & \varepsilon^6 \\ 1 & \varepsilon^3 & \varepsilon^6 & \varepsilon^9 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix}.$$

Általános esetben az  $n$ -edik osztópont  $\frac{2n\pi}{N}$  ( $n = 0, 1, 2, \dots, N-1$ ), a Fourier-összeg e pontbeli helyettesítési értékét  $y_n$ -nel jelölve

$$\begin{aligned} y_0 &= c_0 + c_1 e^{i0} + c_2 e^{2i0} + \dots + c_{N-1} e^{(N-1)i0} = c_0 + c_1 + \dots + c_{N-1} \\ y_1 &= c_0 + c_1 e^{\frac{2\pi i}{N}} + c_2 e^{\frac{4\pi i}{N}} + \dots + c_{N-1} e^{\frac{2(N-1)\pi i}{N}} \\ y_2 &= c_0 + c_1 e^{\frac{4\pi i}{N}} + c_2 e^{\frac{8\pi i}{N}} + \dots + c_{N-1} e^{\frac{4(N-1)\pi i}{N}} \\ &\vdots \\ y_{N-1} &= c_0 + c_1 e^{\frac{2\pi i(N-1)}{N}} + c_2 e^{\frac{4\pi i(N-1)}{N}} + \dots + c_{N-1} e^{\frac{2\pi i(N-1)^2}{N}} \end{aligned}$$

Az  $\varepsilon = e^{2\pi i/N}$  jelöléssel mátrixszorzat-alakban

$$\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{N-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & \varepsilon & \varepsilon^2 & \varepsilon^3 & \dots & \varepsilon^{N-1} \\ 1 & \varepsilon^2 & \varepsilon^4 & \varepsilon^6 & \dots & \varepsilon^{2(N-1)} \\ 1 & \varepsilon^3 & \varepsilon^6 & \varepsilon^9 & \dots & \varepsilon^{3(N-1)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \varepsilon^{N-1} & \varepsilon^{2(N-1)} & \varepsilon^{3(N-1)} & \dots & \varepsilon^{(N-1)^2} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{N-1} \end{bmatrix} \quad \square$$

E példában szereplő együtthatómátrix egy Vandermonde-mátrix, mégpedig a  $\mathbf{V}_N(1, \varepsilon, \varepsilon^2, \dots, \varepsilon^{N-1})$  mátrix, melyet  $\Phi_{N,\varepsilon}$ -nal jelölünk. Fontos lesz még e mátrix konjugáltja, mely az  $\omega = \bar{\varepsilon} = e^{-2\pi i/N}$  egységyökhöz tartozó Vandermonde-mátrix. E két mátrixot *Fourier-mátrixnak* is nevezik. Tehát  $[\Phi_{N,\varepsilon}]_{kn} = \varepsilon^{kn}$ ,  $[\Phi_{N,\omega}]_{kn} = \omega^{kn}$  ( $0 \leq k, n < N$ ), azaz

részletesen

$$\Phi_{N,\varepsilon} = \mathbf{V}_N(1, \varepsilon, \varepsilon^2, \dots, \varepsilon^{N-1}) = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & \varepsilon & \dots & \varepsilon^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \varepsilon^{N-1} & \dots & \varepsilon^{(N-1)^2} \end{bmatrix} \quad (7.24)$$

$$\Phi_{N,\omega} = \mathbf{V}_N(1, \omega, \dots, \omega^{N-1}) = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & \omega & \dots & \omega^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \dots & \omega^{(N-1)^2} \end{bmatrix} \quad (7.25)$$

Az előző példában a Fourier-együtthatók ismeretében meghatároztuk a függvény helyettesítési értékeit. A gyakorlati alkalmazásokban főként a fordított sorrend érdekes, vannak  $y_k$  mért adataink, és keressük a  $c_k$  együtthatókat. Ehhez nyújt alapismereteket a következő tétel.

**7.95. TÉTEL (A FOURIER-MÁTRIXOK TULAJDONSÁGAI).** Legyen  $N$  pozitív egész szám,  $\varepsilon = e^{2\pi i/N}$ ,  $\omega = \bar{\varepsilon} = e^{-2\pi i/N}$ . Az  $\Phi_{N,\varepsilon}$  és  $\Phi_{N,\omega}$  Fourier-mátrixok a következő tulajdonságokkal rendelkeznek:

- a) Bármelyik Fourier-mátrix  $k$ -adik és  $N - k$ -adik sora egymás konjugáltja, páros  $N$  esetén pedig az  $N/2$ -edik sorvektor  $(1, -1, 1, -1, \dots)$ .
- b) A két Fourier-mátrix egymás konjugáltja és egyúttal egymás adjungáltja is, azaz  $\Phi_{N,\omega} = \overline{\Phi_{N,\varepsilon}} = \Phi_{N,\varepsilon}^H$  és  $\Phi_{N,\varepsilon} = \overline{\Phi_{N,\omega}} = \Phi_{N,\omega}^H$ .
- c)  $\Phi_{N,\varepsilon} \Phi_{N,\omega} = N\mathbf{I}_N$ , így  $\Phi_{N,\varepsilon}$  és  $\Phi_{N,\omega}$  invertálható,

$$\Phi_{N,\varepsilon}^{-1} = \frac{1}{N} \Phi_{N,\omega}, \quad \Phi_{N,\omega}^{-1} = \frac{1}{N} \Phi_{N,\varepsilon},$$

továbbá  $\frac{1}{\sqrt{N}} \Phi_{N,\varepsilon}$  és  $\frac{1}{\sqrt{N}} \Phi_{N,\omega}$  unitér.

**BIZONYÍTÁS.** a) Az  $\Phi_{N,\varepsilon}$  mátrix  $k$ -adik, illetve  $N - k$ -adik sorának  $n$ -edik eleme  $\varepsilon^{kn}$ , illetve  $\varepsilon^{(N-k)n}$ . Ez utóbbit átalakítva kapjuk, hogy

$$\varepsilon^{(N-k)n} = \varepsilon^{Nn} \varepsilon^{-kn} = (\varepsilon^{-1})^{kn} = \bar{\varepsilon}^{kn}.$$

Mivel minden pozitív páros  $N$ -re  $\varepsilon^{\frac{N}{2}} = -1$ , ezért az  $N/2$ -edik sorban  $-1$  hatványai szerepelnek.

b) Mivel  $\omega = \bar{\varepsilon}$ , ezért  $\omega^s = \bar{\varepsilon}^s$ , tehát  $\Phi_{N,\omega} = \overline{\Phi_{N,\varepsilon}}$ . Másrészt  $\Phi_{N,\varepsilon}$  szimmetrikus, következésképp  $\overline{\Phi_{N,\varepsilon}} = \Phi_{N,\varepsilon}^T = \Phi_{N,\varepsilon}^H$ .

c) Számítsuk ki a  $\Phi_{N,\varepsilon} \Phi_{N,\omega}$  mátrixot! A szorzat  $k$ -adik sorának  $n$ -edik oszlopában a

$$\sum_{m=0}^{N-1} \varepsilon^{km} \omega^{mn} = \sum_{m=0}^{N-1} \varepsilon^{m(k-n)} = \sum_{m=0}^{N-1} (\varepsilon^{k-n})^m$$

összeg szerepel. Ha  $k = n$ , azaz  $\varepsilon^{k-n} = 1$ , akkor ez az összeg  $N$ , minden más esetben 0 (ld. a ?? állítást a komplex számokról szóló

függelékben). Mindezek egyik következménye, hogy  $\Phi_{N,\varepsilon}$  és  $\Phi_{N,\omega}$  invertálhatók,  $\Phi_{N,\varepsilon}$  inverze  $\frac{1}{N}\Phi_{N,\omega}$ ,  $\Phi_{N,\omega}$  inverze  $\frac{1}{N}\Phi_{N,\varepsilon}$ . A másik következmény, hogy

$$\left(\frac{1}{\sqrt{N}}\Phi_{N,\varepsilon}\right)\left(\frac{1}{\sqrt{N}}\Phi_{N,\omega}\right) = \mathbf{I}_N,$$

ami a b)-t is figyelembe véve épp azt jelenti, hogy

$$\left(\frac{1}{\sqrt{N}}\Phi_{N,\varepsilon}\right)\left(\frac{1}{\sqrt{N}}\Phi_{N,\varepsilon}\right)^H = \mathbf{I}_N,$$

azaz hogy  $\frac{1}{\sqrt{N}}\Phi_{N,\varepsilon}$  unitér. □

**Diszkrét Fourier-transzformáció** A diszkrét Fourier-transzformációra úgy gondolhatunk, mint egy – általában komplex – függvény helyettesítési értékeinek vektorához a függvény trigonometrikus összetevői együtthatóinak vektorát rendelő lineáris  $\mathbb{C}^N \rightarrow \mathbb{C}^N$  leképezésre.

A 7.94. példában egy Fourier-összeg együtthatóival kifejeztük a függvény megadott helyeken fölvelt értékeit. A fordított irány sokkal érdekesebb: ismerjük egy  $f$  függvény  $N$  különböző megadott helyen fölvelt értékét, és meg van adva  $N$  lineárisan független függvény. Olyan lineáris kombinációjuk együtthatóit keressük e függvényeknek, mely lineáris kombináció a megadott helyeken megegyezik  $f$ -fel. Mi a következőkben definiálandó diszkrét Fourier-transzformáció esetén az

$$f(t) = \frac{1}{N} \sum_{n=0}^{N-1} c_n e^{nit}$$

függvényből indulunk ki, a megadott helyek a  $[0, 2\pi]$  intervallumot  $N$  részre osztó  $2k\pi/N$  ( $k = 0, 1, \dots, N-1$ ) pontok. A  $(c_0, c_1, \dots, c_{N-1}) \mapsto (y_0, y_1, \dots, y_{N-1})$  leképezés inverzét fogjuk diszkrét Fourier-transzformálnak nevezni. Ennek mátrixa  $\Phi_{N,\omega}$ , amelyre a továbbiakban az  $\mathbf{F}_N$  jelölést is használjuk. E megközelítésből az  $f$  függvény teljesen elhagyható, hisz a lényeg az, hogy egy szám- $N$ -eshez hozzárendelünk egy másikat!

**7.96. DEFINÍCIÓ (DISZKRÉT FOURIER-TRANSZFORMÁCIÓ (DFT)).** Az  $F_N : \mathbb{C}^N \rightarrow \mathbb{C}^N : \mathbf{x} \mapsto \mathbf{X} = \mathbf{F}_N \mathbf{x}$  leképezést diszkrét Fourier-transzformációnak nevezzük.

- A diszkrét Fourier-transzformáció tehát a (7.25) képlettel megadott  $\mathbf{F}_N = \Phi_{N,\omega}$  mátrixhoz tartozó mátrixleképezés.
- A leképezést kifejtve koordinátánként:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} kn} = \sum_{n=0}^{N-1} x_n \omega^{kn} \quad (\omega = e^{-\frac{2\pi i}{N}}). \quad (7.26)$$

► Az  $F_N$  transzformáció mátrixszorzatos alakja

$$F_N : \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{bmatrix} \mapsto \begin{bmatrix} X_0 \\ X_1 \\ \vdots \\ X_{N-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & \omega & \dots & \omega^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \dots & \omega^{(N-1)^2} \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{bmatrix}.$$

► E témában elterjedt jelölések: a transzformálandó vektor dimenzióját nagy  $N$  jelöli, a képvektort a transzformálandó vektor nagybetűs változata jelöli, azaz  $\mathbf{x}$  képe  $\mathbf{X}$ ,  $\mathbf{y}$  képe  $\mathbf{Y}$ , stb., a vektorok koordinátái 0-tól  $N-1$ -ig vannak indexelve.

► A diszkrét Fourier-transzformációt gyakran a Fourier-mátrixok valamelyikének egy másik konstansszorosával definiálják. Előfordul az unitér  $\frac{1}{\sqrt{N}}F_N$ , az  $\frac{1}{N}F_N$  vagy a  $\Phi_{N,\varepsilon}$  mátrix is a transzformáció mátrixaként, sőt van aki minden olyan  $\Phi_{N,\varepsilon}$  mátrixot egy DFT mátrixának tekint, ahol  $\varepsilon$  primitív  $N$ -edik egységgyök.

► Az általunk adott definíció a legelterjedtebb, a legtöbb ismert szoftver is ezt használja. Ennek oka e definíciónak a folytonos Fourier-transzformálttal való szorosabb kapcsolata, és a jelfeldolgozásban is ezt használják leginkább. Más alkalmazásokhoz viszont megfelelőbb lehet valamelyik fent említett másik definíció.

► Konkrétan az  $F_1$ ,  $F_2$ ,  $F_4$  és  $F_8$  mátrixok:

$$F_1 = [1], \quad F_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad F_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{bmatrix},$$

$$F_8 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \frac{1-i}{\sqrt{2}} & -i & \frac{-1-i}{\sqrt{2}} & -1 & \frac{-1+i}{\sqrt{2}} & i & \frac{1+i}{\sqrt{2}} \\ 1 & -i & -1 & i & 1 & -i & -1 & i \\ 1 & \frac{-1-i}{\sqrt{2}} & i & \frac{1-i}{\sqrt{2}} & -1 & \frac{1+i}{\sqrt{2}} & -i & \frac{-1+i}{\sqrt{2}} \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & \frac{-1+i}{\sqrt{2}} & -i & \frac{1+i}{\sqrt{2}} & -1 & \frac{1-i}{\sqrt{2}} & i & \frac{-1-i}{\sqrt{2}} \\ 1 & i & -1 & i & 1 & i & -1 & i \\ 1 & \frac{1+i}{\sqrt{2}} & i & \frac{-1+i}{\sqrt{2}} & -1 & \frac{-1-i}{\sqrt{2}} & -i & \frac{1-i}{\sqrt{2}} \end{bmatrix}$$

**7.97. TÉTEL (A DFT TULAJDONSÁGAI).** Tekintsük a diszkrét  $F_N$  Fourier-transzformációt, és legyen az  $\mathbf{x} = (x_0, x_1, \dots, x_{N-1})$  vektor képe  $\mathbf{X} = (X_0, X_1, \dots, X_{N-1})$ . Ekkor a következők igazak:

a) Konstans vektor képe impulzusvektor (melynek a nulladikat kivéve mindegyik koordinátája 0), és fordítva, konkrétan

$$F_N(c, c, \dots, c) = (Nc, 0, \dots, 0), \quad F_N(c, 0, \dots, 0) = (c, c, \dots, c).$$

ahol  $c \in \mathbb{C}$  tetszőleges konstans.

- b) Ha  $\mathbf{x}$  valós vektor, akkor  $X_{N-k} = \overline{X}_k$ .  
 c) Az  $F_N$  transzformáció invertálható, inverze (IDFT) többféle felírásban:

$$\mathbf{x} = F_N^{-1} \mathbf{X} = \frac{1}{N} \Phi_{N,\varepsilon} \mathbf{X}, \quad x_k = \frac{1}{N} \sum_{n=0}^{N-1} X_n \varepsilon^{kn} = \frac{1}{N} \sum_{n=0}^{N-1} X_n e^{\frac{2\pi i}{N} kn}.$$

**BIZONYÍTÁS.** a) Az állítás első részének bizonyítása közvetlenül leolvasható az  $\Phi_{N,\varepsilon}(c\Phi_{N,\omega}) = cN\mathbf{I}_N$  szorzat első oszlopából, második része a  $c\Phi_{N,\varepsilon}$  mátrix első oszlopából. De a 7.95. tétel bizonyításában is használt ?? állításra hivatkozva közvetlenül is azonnal adódik.

b) A (7.26) képletet használva

$$\begin{aligned} X_{N-k} &= \sum_{n=0}^{N-1} x_n \omega^{(N-k)n} = \sum_{n=0}^{N-1} x_n \omega^{-kn} \\ &= \sum_{n=0}^{N-1} x_n \overline{\omega}^{kn} = \overline{\sum_{n=0}^{N-1} x_n \omega^{kn}} = \overline{X_k} \end{aligned}$$

- c) Az invertálhatóság azonnal adódik abból, hogy a Fourier-mátrixok Vandermonde-mátrixok is egyúttal, melynek determinánsa nem 0. A tételbeli mindegyik összefüggés azonnali következménye az  $F_N^{-1} = \Phi_{N,\omega}^{-1} = \frac{1}{N} \Phi_{N,\varepsilon}$  képletnek.  $\square$

7.98. PÉLDA (DFT KISZÁMÍTÁSA). Határozzuk meg az  $\mathbf{x} = (1, i, i, 2)$  vektor diszkrét Fourier-transzformáltját!

**MEGOLDÁS.**  $N = 4$ , így

$$\mathbf{X} = F_4 \mathbf{x} = \mathbf{F}_4 \mathbf{x} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{bmatrix} \begin{bmatrix} 1 \\ i \\ i \\ 2 \end{bmatrix} = \begin{bmatrix} 3 + 2i \\ 2 + i \\ -1 \\ -3i \end{bmatrix}. \quad \square$$

**Periodikus összetevők szűrése** Műszaki alkalmazásokban gyakran előfordul, hogy egy periodikus függvénnyel leírható jelhez magasabb frekvenciájú zaj adódik, amit utólag ki szeretnénk „szűrni”. Ez egy DFT-IDFT párral könnyen elvégezhető.

A szűrés általános modellje három lépésből áll, melyet az alábbi séma szemléltet:

$$\mathbf{x} \xrightarrow{\text{DFT}} \mathbf{X} \xrightarrow{\text{„szűrés”}} \hat{\mathbf{X}} \xrightarrow{\text{IDFT}} \hat{\mathbf{x}}$$

A műszaki gyakorlatban „szűrésen” sokféle transzformációt értenek, mely az  $\mathbf{X}$  vektort az  $\hat{\mathbf{X}}$ -ba képz. Mi csak a legegyszerűbb esettel foglalkozunk,  $\mathbf{X}$  bizonyos koordinátáinak elhagyásával (kiszűréssel).



A következőkben egy mesterkélten leegyszerűsített, fejben számolva is követhető példát mutatunk a DFT e tipikus alkalmazására.

**7.99. PÉLDA** (MAGAS FREKVENCIÁJÚ ÖSSZETEVŐK SZŰRÉSE). *Adva van egy  $p$  szerint periodikus függvény  $t_k = kp/6$  ( $k = 0, 1, \dots, 5$ ) helyeken fölvevett függvényértékeinek  $\mathbf{x} = (4, 1, -2, -2, -2, 1)$  vektora. Bontsuk fel e függvényt egy  $p$  szerint periodikus trigonometrikus függvény, és  $p/m$  periódusú függvények (zaj) összegére ( $m > 1$  egész).*

**MEGOLDÁS.** Az  $\mathbf{x}$  vektor diszkrét Fourier-transzformáltja

$$\mathbf{X} = F_6 \mathbf{x} = \mathbf{F}_{6,\omega} \mathbf{x} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \omega & \omega^2 & -1 & \omega^4 & \omega^5 \\ 1 & \omega^2 & \omega^4 & 1 & \omega^2 & \omega^4 \\ 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & \omega^4 & \omega^2 & 1 & \omega^4 & \omega^2 \\ 1 & \omega^5 & \omega^4 & -1 & \omega^2 & \omega \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ -2 \\ -2 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 9 \\ 3 \\ 0 \\ 3 \\ 9 \end{bmatrix}$$

ahol kihasználtuk, hogy  $\omega^6 = 1$ ,  $\omega + \omega^5 = 1$ ,  $\omega^2 + \omega^4 = -1$ . Például

$$X_2 = [\mathbf{F}_{6,\omega}]_{2*} \mathbf{x} = 4 + \omega^2 - 2\omega^4 - 2 - 2\omega^2 + \omega^4 = 2 - \omega^2 - \omega^4 = 3.$$

Mivel  $\mathbf{x}$  valós, ezért a 7.97. tétel b) pontja szerint  $X_{N-k} = \overline{X_k}$ , amit azonnal ellenőrizhetünk is. Az  $\mathbf{x}$  vektor „mögött” lévő  $p$  szerint periodikus függvény e modellben

$$x(t) = \frac{1}{6} \sum_{n=0}^5 X_n e^{i n \frac{2\pi}{p} t} = \frac{1}{6} \left( 9e^{i \frac{2\pi}{p} t} + 3e^{2i \frac{2\pi}{p} t} + 3e^{4i \frac{2\pi}{p} t} + 9e^{5i \frac{2\pi}{p} t} \right)$$

A  $p$  értékének valójában semmi szerepe, mert e függvényt csak a  $k \frac{p}{n}$  ( $k = 0, 1, \dots, 5$ ) pontokban értékeljük ki, így a fenti összegben csak a

$$e^{i \frac{2\pi}{p} \frac{pn}{6}} = (e^{i \frac{2\pi}{6}})^n = \varepsilon^n$$

értékek szerepelnek. Ezekre használható a

$$\begin{aligned} \varepsilon^n + \varepsilon^{6-n} &= (e^{i \frac{2\pi}{6}})^n + (e^{-i \frac{2\pi}{6}})^n = (e^{i \frac{2\pi n}{6}}) + (e^{-i \frac{2\pi n}{6}}) \\ &= 2 \cos n \frac{\pi}{3} = 2 \cos \frac{2\pi}{p} \frac{pn}{6} \end{aligned}$$

összefüggés, így

$$x(t) = \frac{1}{6} \left( 18 \cos \frac{2\pi}{p} t + 6 \cos 2 \frac{2\pi}{p} t \right) = 3 \cos \frac{2\pi}{p} t + \cos 2 \frac{2\pi}{p} t.$$

A függvény tehát  $3 \cos \frac{2\pi}{p} t$ , a „zaj”  $\cos 2 \frac{2\pi}{p} t$ . □

**Gyors Fourier-transzformáció** A diszkrét Fourier-transzformáció gyors kiszámítására konstruált algoritmusoknak döntő szerepük van a mai kultúránk alapját jelentő digitális technika fejlődésében.

A diszkrét Fourier-transzformált kiszámításához, azaz az  $n$ -edrendű Fourier-mátrixszal való szorzás kiszámításához  $n^2$  szorzás elvégzésére van szükség. Bármely olyan algoritmust, mely e transzformáció eredményét  $O(n \log n)$ , azaz konstansszor  $n \log n$  lépésben elvégzi, *gyors Fourier-transzformációnak* nevezzük. Sok változata létezik, mi csak az elsőként publikált, legismertebbet ismertetjük.

A transzformáció gyorsaságának becsléséhez most minden aritmetikai művelet elvégzésének idejét tekintünk azonosnak. A DFT kiszámítására, azaz a Fourier-mátrixszal való szorzáshoz minden sorban  $N$  szorzás,  $N - 1$  összeadás kell, és  $N$  sor van, így a szükséges műveletek száma  $N(2N - 1)$ .

Az egyszerűség kedvéért legyen a továbbiakban  $N$  kettőhatvány, és csoportosítsuk az  $X_k$ -t megadó összeget az indexek paritása szerint, azaz külön adjuk össze a páros és külön a páratlan indexűket. Vegyük észre, hogy ez az összeg két fele akkora méretű Fourier-transzformációból megkapható:

$$\begin{aligned}
 X_k &= \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} kn} = \sum_{n=0}^{N-1} x_n \omega_N^{kn} \\
 &= \sum_{n=0}^{N/2-1} x_{2n} e^{-\frac{2\pi i}{N} 2nk} + \sum_{n=0}^{N/2-1} x_{2n+1} e^{-\frac{2\pi i}{N} (2n+1)k} \\
 &= \sum_{n=0}^{N/2-1} x_{2n} e^{-\frac{2\pi i}{N/2} nk} + e^{-\frac{2\pi i}{N} k} \sum_{n=0}^{N/2-1} x_{2n+1} e^{-\frac{2\pi i}{N/2} nk} \\
 &= \sum_{n=0}^{N/2-1} x_{2n} \omega_{N/2}^{nk} + \omega_N^k \sum_{n=0}^{N/2-1} x_{2n+1} \omega_{N/2}^{nk} \\
 &= E_k + \omega_N^k O_k.
 \end{aligned}$$

Hogy különbséget tegyünk az  $N$  és  $N/2$  dimenziós vektorok transzformációi közt, az  $N$ -edik és  $N/2$ -edik egységgyököket jelölje

$$\omega_N = e^{-\frac{2\pi i}{N}}, \text{ és } \omega_{N/2} = e^{-\frac{2\pi i}{N/2}}.$$

És ezen a ponton azzal tudjuk csökkenteni a számításokat, hogy mivel az  $E_k$  és  $O_k$  összegek  $N/2$  szerint periodikusak, így  $k \geq N/2$  esetén  $E_k$  és  $O_k$  értékét már nem kell újra számolni, ugyanis

$$E_{k+N/2} = E_k, \quad O_{k+N/2} = O_k,$$

és az  $\omega_N^k$  együttható is újrahasznosítható:

$$\omega_N^{k+N/2} = e^{-\frac{2\pi i}{N}(k+N/2)} = e^{-\frac{2\pi i}{N}(k)} e^{-\frac{2\pi i}{N} \frac{N}{2}} = -e^{-\frac{2\pi i}{N}(k)} = -\omega_N^k.$$

Ezeket összevetve tehát  $k < N/2$  esetén

$$\begin{aligned} X_k &= E_k + \omega_N^k O_k, \\ X_{k+N/2} &= E_k - \omega_N^k O_k. \end{aligned}$$

Így, ha  $E_k$  és  $O_k$  már ki van számolva,  $X_k$  és  $X_{k+N/2}$  kiszámításához csak egy szorzásra, egy összeadásra és egy kivonásra van szükség, azaz az  $X_k$  ( $0 \leq k < N$ ) együtthatók  $3N/2$  művelettel megkaphatók. Ezután rekurzív módon  $E_k$  és  $O_k$  kiszámítását is ugyanígy végezzük: mivel fele akkora a vektor, de kettő van belőle, itt is  $3N/2$  műveletre lesz szükség. Mivel  $N$  kettőhatvány, például  $N = 2^s$ , így  $s = \log_2 N$ -szer kell megismételnünk ezt a lépést, vagyis a teljes transzformáció műveletigénye  $\frac{3}{2}N \log_2 N$ . Konkrétan néhány  $N$  esetén:

| $N$      | $2^4 = 16$ | $2^8 = 256$ | $2^{10} = 1024$ | $2^{16} = 65536$ |
|----------|------------|-------------|-----------------|------------------|
| DFT      | 496        | 130816      | 2096128         | 8589869056       |
| FFT      | 96         | 3072        | 15360           | 1572864          |
| hányados | $> 5$      | $> 42$      | $> 136$         | $> 5461$         |

A műveletigény fenti számításában nem vettük figyelembe  $\omega_N^k$  kiszámításának költségeit. Ha e hatvány kiszámítása  $C$  aritmetikai művelettel egyenértékű, akkor is csak  $\frac{C+3}{2}N \log_2 N$  műveletre van szükségünk. Ezzel tehát bizonyítottuk a következő tételt:

**7.100. TÉTEL (GYORS FOURIER-TRANSZFORMÁCIÓ).** *Létezik olyan algoritmus, mely egy  $N$ -dimenziós vektor diszkrét Fourier-transzformáltját legfeljebb  $O(N \log_2 N)$  aritmetikai művelet elvégzésével kiszámolja.*

A fenti bizonyításban szereplő algoritmus pszeudokódja:

Mivel e transzformáció is lineáris leképezésekből áll, a gyors Fourier-transzformáció mátrixszorzat-alakba is fölírható:

$$\mathbf{F}_N = \Delta_N \begin{bmatrix} \mathbf{F}_{N/2} & \mathbf{O} \\ \mathbf{O} & \mathbf{F}_{N/2} \end{bmatrix} \Pi_N,$$

ahol  $\Pi_N$  az a permutációs mátrix, mely előre veszi a páros indexű elemeket,  $\Delta_N$  pedig a „fél” transzformáltakat összeadó, és a páratlan indexűeket egy  $\omega$ -hatvánnyal beszorzó mátrix. Ezek kisebb indexű

A tételbeli eljárást Gauss már ismerte és 1805-ben használta a másodiknak fölfedezett Pallas és a harmadiknak fölfedezett Juno nevű kisbolygó pályájának kiszámításához. A felező eljárást Danielson és Lánzos 1942-ben újra fölfedezték, de ők sem vizsgálták az algoritmus sebességét. Az FFT ismertté és népszerűvé Cooley és Tukey 1965-ben megjelent cikke után vált.

```

function FFT(x)
  N ← dim(x)
  X legyen N-dimenziós vektor
  if N = 1 then
    | X0 ← x0
  else
    y ← x páros indexű elemei
    z ← x páratlan indexű elemei
    Y ← FFT(y)
    Z ← FFT(z)
    for k ← 0 to N/2 - 1 do
      | E ← Yk
      | O ← e-2πi/N k Zk
      | Xk ← E + O
      | Xk+N/2 ← E - O
  return X

```

7.26. ábra: FFT algoritmus. A rekurzív függvény bemenete egy tetszőleges komplex  $x$  vektor, kimenete a diszkrét Fourier-transzformált  $X$  vektor.

példányai:

$$\Pi_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \Pi_8 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\Delta_4 = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -i \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & i \end{bmatrix} = \begin{bmatrix} \mathbf{I}_2 & \mathbf{D}_2 \\ \mathbf{I}_2 & -\mathbf{D}_2 \end{bmatrix} \quad \Delta_8 = \begin{bmatrix} \mathbf{I}_4 & \mathbf{D}_4 \\ \mathbf{I}_4 & -\mathbf{D}_4 \end{bmatrix}$$

A  $\Delta$  mátrixokban szereplő diagonális mátrixok az egységmátrixok, és az  $\omega$  hatványait tartalmazó  $\mathbf{D}$  mátrixok, ahol  $\mathbf{D}_k = \text{diag}(1, \omega, \omega^2, \dots, \omega^{k-1})$ .

Tehát például

$$\begin{aligned} \mathbf{F}_8 &= \Delta_8 \begin{bmatrix} \mathbf{F}_4 & \mathbf{O} \\ \mathbf{O} & \mathbf{F}_4 \end{bmatrix} \Pi_8 \\ &= \Delta_8 \begin{bmatrix} \Delta_4 & \mathbf{O} \\ \mathbf{O} & \Delta_4 \end{bmatrix} \begin{bmatrix} \mathbf{F}_2 & \mathbf{O} & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{F}_2 & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{F}_2 & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{O} & \mathbf{F}_2 \end{bmatrix} \begin{bmatrix} \Pi_4 & \mathbf{O} \\ \mathbf{O} & \Pi_4 \end{bmatrix} \Pi_8. \end{aligned}$$

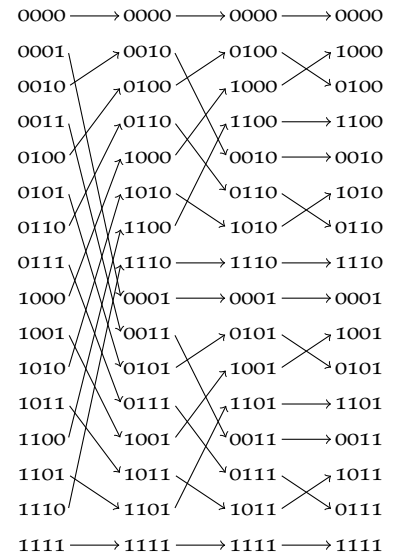
Látjuk, hogy a rekurzió következtében a transzformálandó vektort elő-

ször a  $\Pi$ -mátrixokból álló blokkmátrixokkal kell szorozni. E mátrixok szorzata is permutációmátrix. Hatását e konkrét esetben kiszámoljuk a fent megadott  $\Pi_4$  és  $\Pi_8$  mátrixok behelyettesítésével:

$$\begin{bmatrix} \Pi_4 & \mathbf{O} \\ \mathbf{O} & \Pi_4 \end{bmatrix} \Pi_8 \mathbf{x} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = \begin{bmatrix} x_0 \\ x_1 \\ x_4 \\ x_2 \\ x_6 \\ x_5 \\ x_3 \\ x_7 \end{bmatrix}.$$

Ez első pillanatban áttekinthetetlen permutációnak tűnik, de valójában egy igen egyszerűen leírható transzformációt kapunk: a transzformálandó  $\mathbf{x}$  vektor  $k$ -adik koordinátáját ( $k = 0, 1, \dots, N-1$ ) a  $j$ -edik helyébe viszi, ha  $j$  bináris alakja éppen fordítottja  $k$  bináris alakjának. Például ha  $N = 16$ , és  $k = 6$ , akkor  $x_{12}$  a harmadik koordináta helyére kerül a permutáció során, mivel  $12 = 1100_2$ , és ennek fordítottja  $0011_2 = 3$ . Ennek igazolása rendkívül egyszerű, ha észrevesszük, hogy az  $i$ -edik  $\Delta$ -mátrixszal való szorzás épp jobbról az első  $i$  koordináta szerinti lexicografikus sorrendbe rendezi az elemeket. Ennek szemléltetésére elég az a 7.27 ábrán bemutatott  $N = 16$  eset vizsgálata.

**Vektorok konvolúciója** Vektorok konvolúciója igen sok helyen felmerül: a polinomok szorzásától kezdve az olyan transzformációkig, ahol egy koordinátát szomszédainak egy rögzített lineáris kombinációjával kell helyettesíteni. A gyors Fourier-transzformációval hatékonyan számolható.



7.27. ábra: Az  $\mathbf{x}$  vektor koordinátáinak indexeit binárisan fölírva, jól követhető azok mozgása a permutációk során.

## Megoldások

7.1. Igen,  $A : \mathbb{R}^3 \rightarrow \mathbb{R}$ , mátrixa  $\mathbf{A} = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}$ .

7.2. Nem, az  $A : \mathbf{x} \mapsto \mathbf{a} + \mathbf{x}$  leképezés a  $\mathbf{0}$  vektort nem a  $\mathbf{0}$ -ba képi, így nem lehet mátrixleképezés!

7.4. Igen, a mátrix

$$\mathbf{A} = \mathbf{a}(\mathbf{a} \cdot \mathbf{x}) = \mathbf{a}(\mathbf{a}^\top \mathbf{x}) = (\mathbf{a}\mathbf{a}^\top) \mathbf{x} = \begin{bmatrix} a_1^2 & a_1 a_2 & a_1 a_3 \\ a_2 a_1 & a_2^2 & a_2 a_3 \\ a_3 a_1 & a_3 a_2 & a_3^2 \end{bmatrix}$$

7.5. A bizonyítások a mátrixműveletek tulajdonságaiból következnek. Ott, ahol valamelyik mátrixazonosságot használjuk, az ekvivalenciát kimondó nyíl fölé M-betűt írunk, ahol pedig a függvények közti műveleti tulajdon-

ságokat használjuk, ott egy F-betűt:

- a)  $(A+B)(\mathbf{x}) = C(\mathbf{x}) \xleftrightarrow{F} A(\mathbf{x}) + B(\mathbf{x}) = C(\mathbf{x}) \iff \mathbf{Ax} + \mathbf{Bx} = \mathbf{Cx} \xleftrightarrow{M} (\mathbf{A} + \mathbf{B})\mathbf{x} = \mathbf{Cx}$   
b)  $(cA)(\mathbf{x}) = C(\mathbf{x}) \xleftrightarrow{F} cA(\mathbf{x}) = C(\mathbf{x}) \iff c\mathbf{Ax} = \mathbf{Cx} \xleftrightarrow{M} (c\mathbf{A})\mathbf{x} = \mathbf{Cx}$   
c)  $(X \circ Y)(\mathbf{x}) = Z(\mathbf{x}) \xleftrightarrow{F} X(Y(\mathbf{x})) = Z(\mathbf{x}) \iff \mathbf{X}(\mathbf{Yx}) = \mathbf{Zx} \xleftrightarrow{M} (\mathbf{XY})\mathbf{x} = \mathbf{Zx}$ .

7.6. A 7.2. tételből, illetve a 7.5. feladatból tudjuk, hogy  $(A \circ B)(\mathbf{x}) = \mathbf{ABx}$  és  $(B \circ A)(\mathbf{x}) = \mathbf{BAx}$ . Így ha  $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$ , azaz  $\mathbf{A}$  inverze  $\mathbf{B}$ , akkor  $(A \circ B)(\mathbf{x}) = \mathbf{ABx} = \mathbf{Ix} = \mathbf{x}$ , és hasonlóan  $(B \circ A)(\mathbf{x}) = \mathbf{BAx} = \mathbf{Ix} = \mathbf{x}$ , azaz  $A \circ B$  és  $B \circ A$  az identikus leképezés. Hasonlóképp, ha  $A \circ B$  és  $B \circ A$  az identikus leképezés, akkor  $(\mathbf{AB})\mathbf{x} = (A \circ B)(\mathbf{x}) = \mathbf{x}$  és  $(\mathbf{BA})\mathbf{x} = (B \circ A)(\mathbf{x}) = \mathbf{x}$ , így  $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$  (ld. ?? feladat), vagyis  $\mathbf{A}$  és  $\mathbf{B}$  egymás inverzei.

7.11. Ha  $\mathbf{A} = \mathbf{C}^{-1}\mathbf{B}\mathbf{C}$ , akkor  $C_{ij}$ -vel jelölve a  $\mathbf{C}$  mátrix  $c_{ij}$ -hez tartozó előjeles aldeterminánsát kapjuk, hogy

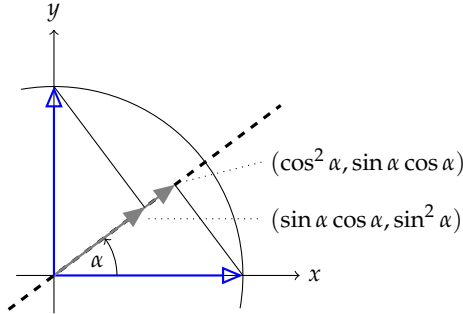
$$\begin{aligned}\text{trace}(\mathbf{A}) &= \sum_i a_{ii} = \sum_i \frac{1}{\det \mathbf{C}} \sum_{j,k} C_{ij} b_{jk} c_{ki} \\ &= \frac{1}{\det \mathbf{C}} \left( \sum_{j \neq k} b_{jk} \sum_i C_{ij} c_{ki} + \sum_j b_{jj} \sum_i C_{ij} c_{ji} \right) = \dots\end{aligned}$$

7.14. Az új egyenletrendszer sorterét az eredeti egyenletrendszer megoldásvektorai feszítik ki, ezért ennek nullterére megegyezik az eredeti sorterével, melyet a sorvektorok feszítenek ki. (Természetesen elég a sorvektorok közül a függetleneket kiválasztani. Esetünkben tehát a megadott egyenletrendszer nullterét kifeszítik az  $(1, 2, 1, 2, 1)$  és az  $(1, 2, 3, 3, 1)$  vektorok.)

7.16. a) A bizonyítandó képlet:

$$\mathbf{P} = \begin{bmatrix} \cos^2 \alpha & \sin \alpha \cos \alpha \\ \sin \alpha \cos \alpha & \sin^2 \alpha \end{bmatrix}.$$

Ez leolvasható a következő ábráról:



Az itt látható két derékszögű háromszög befogóinak hossza  $\cos \alpha$ , illetve  $\sin \alpha$ , és így például a  $\cos \alpha$  hosszú szakasz két tengelyvetülete  $\cos^2 \alpha$  és  $\cos \alpha \sin \alpha$  hosszú, tehát  $\mathbf{i}$  képe  $(\cos^2 \alpha, \cos \alpha \sin \alpha)$ . Hasonlóan  $\mathbf{j}$  képe  $(\sin \alpha \cos \alpha, \sin^2 \alpha)$ . E két oszlopvektorból álló mátrix pedig valóban megegyezik a fent megadottal.

b) Az  $\mathbf{i}$  és  $\mathbf{j}$  vetülete az egyenesre

$$\text{proj}_{\mathbf{b}} \mathbf{i} = \frac{\mathbf{i} \cdot \mathbf{b}}{\mathbf{b} \cdot \mathbf{b}} \mathbf{b} = \frac{1}{b_1^2 + b_2^2} \begin{bmatrix} b_1^2 \\ b_1 b_2 \end{bmatrix} \quad \text{és}$$

$$\text{proj}_{\mathbf{b}} \mathbf{j} = \frac{\mathbf{j} \cdot \mathbf{b}}{\mathbf{b} \cdot \mathbf{b}} \mathbf{b} = \frac{1}{b_1^2 + b_2^2} \begin{bmatrix} b_1 b_2 \\ b_2^2 \end{bmatrix}.$$

E két vektor egymás mellé írásával kapott mátrix lesz a leképezés mátrixa:

$$\frac{1}{b_1^2 + b_2^2} \begin{bmatrix} b_1^2 & b_1 b_2 \\ b_1 b_2 & b_2^2 \end{bmatrix}.$$

Ez megegyezik a (7.8) képlettel, azaz

$$\frac{1}{b_1^2 + b_2^2} \begin{bmatrix} b_1^2 & b_1 b_2 \\ b_1 b_2 & b_2^2 \end{bmatrix} = \begin{bmatrix} \cos^2 \alpha & \sin \alpha \cos \alpha \\ \sin \alpha \cos \alpha & \sin^2 \alpha \end{bmatrix},$$

ugyanis ha a  $\mathbf{b}$  vektor  $x$ -tengellyel bezárt szöge  $\alpha$ , akkor  $\cos \alpha = b_1 / \sqrt{b_1^2 + b_2^2}$ , és  $\sin \alpha = b_2 / \sqrt{b_1^2 + b_2^2}$ .

7.17. Ha  $P$  merőleges vetítés, akkor bármely  $\mathbf{v}$  vektorra  $\mathbf{v} - P\mathbf{v} \perp \mathbf{v}$ , tehát a Pitagorasz-tétel szerint  $|\mathbf{v}|^2 = |P\mathbf{v}|^2 + |\mathbf{v} - P\mathbf{v}|^2 \geq |P\mathbf{v}|^2$ .

Fordítva, tegyük fel, hogy bár minden  $\mathbf{v}$ -re  $|P\mathbf{v}| \leq |\mathbf{v}|$ , de indirekt módon van olyan  $\mathbf{v} \in \text{Im } P$ , hogy  $\mathbf{v}$  nem merőleges  $\text{Ker } P$ -re, azaz  $P$  nem merőleges vetítés. Legyen  $\mathbf{v}$ -nek  $\text{Ker } P$ -re való merőleges vetülete  $\mathbf{w}$ . Ekkor  $P(\mathbf{v} - \mathbf{w}) = P\mathbf{v} = \mathbf{v}$ , és  $|\mathbf{v} - \mathbf{w}| < |\mathbf{v}|$ , így  $|\mathbf{v} - \mathbf{w}| < |P(\mathbf{v} - \mathbf{w})|$ , ami ellentmond feltevésünknek.

7.18.  $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix}$ , így a (7.13) képlet szerint

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}^+ = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \left( \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & 0 \end{bmatrix}$$

7.19.  $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix}$ , így a (7.13) képlet szerint

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}^+ = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \left( \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{bmatrix}$$

$$7.20. \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}^+ = \begin{bmatrix} 1/10 & 2/10 \\ 1/10 & 2/10 \end{bmatrix}$$

$$7.21. \begin{bmatrix} 0 & 2 \\ 0 & 2 \end{bmatrix}^+ = \begin{bmatrix} 0 & 0 \\ 1/4 & 1/4 \end{bmatrix}$$

$$7.22. \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}^+ = \begin{bmatrix} 2/3 & -1/3 \\ 1/3 & 1/3 \\ -1/3 & 2/3 \end{bmatrix}$$

$$7.23. \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 0 & 0 \end{bmatrix}^+ = \begin{bmatrix} 1/10 & 2/10 & 0 \\ 1/10 & 2/10 & 0 \end{bmatrix}$$

$$7.24. \begin{bmatrix} 1 & 0 \\ 2 & 2 \\ 0 & 1 \end{bmatrix}^+ = \begin{bmatrix} 5/9 & 2/9 & -4/9 \\ -4/9 & 2/9 & 5/9 \end{bmatrix}$$

$$7.25. \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}^+ = \begin{bmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}$$

$$7.26. \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}^+ = 1/30 \begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}$$

7.27. Mivel az  $\mathbf{A}$  mátrix egy olyan  $\mathbf{XY}$  felbontásával van megadva, melyben  $\mathbf{X}$  teljes oszlop-,  $\mathbf{Y}$  teljes sorrangú, ezért használható a ?? feladat eredménye. Így

$$\mathbf{A}^+ = \mathbf{Y}^T(\mathbf{Y}\mathbf{Y}^T)^{-1}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T = \frac{1}{3} \begin{bmatrix} -1 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 2 & -1 & 0 & 1 \end{bmatrix}$$

7.28. Az előző feladathoz hasonlóan a ?? feladat szerint

$$\mathbf{A}^+ = \mathbf{Y}^T(\mathbf{Y}\mathbf{Y}^T)^{-1}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T = \frac{1}{15} \begin{bmatrix} -4 & 5 & 1 \\ -4 & 5 & 1 \\ 3 & 0 & 3 \\ 7 & -5 & 2 \end{bmatrix}$$

7.29. A 7.30. tétel feladat szerint

$$\mathbf{A}^+ = \mathbf{Y}^T(\mathbf{Y}\mathbf{Y}^T)^{-1}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T = \frac{1}{132} \begin{bmatrix} -32 & 34 & 2 & 4 \\ 1 & 1 & 2 & 4 \\ 34 & -32 & 2 & 4 \end{bmatrix}$$

7.30. A létezés bizonyítható a pszeudoinverz megkonstruálása nélkül is, de mi inkább konstruktív bizonyítást adunk. Ha  $\mathbf{A} \in \mathbb{R}^{m \times n}$ , akkor megadunk egy olyan  $\mathbf{A}^+ \in \mathbb{R}^{n \times m}$  mátrixot, melyre  $\mathbf{A}^+(\mathbf{Ax}) = \mathbf{x}$ , ha  $\mathbf{x} \in \mathcal{S}(\mathbf{A})$ , és  $\mathbf{A}^+\mathbf{z} = \mathbf{0}$ , ha  $\mathbf{z} \perp \mathcal{O}(\mathbf{A})$  ( $\mathbf{z} \in \mathcal{N}(\mathbf{A}^T)$ ).

Legyen  $\mathbf{A} = \mathbf{BR}$  az  $\mathbf{A}$  bázisfelbontása, ahol  $\mathbf{R} = \text{rref}(\mathbf{A})$  egy  $r \times n$ -es mátrix,  $\mathbf{B}$  az  $\mathbf{A}$  főoszlopaiból álló  $m \times r$ -es mátrix, és  $r = \text{r}(\mathbf{A})$ .  $\mathbf{R}$  sorvektorai a sortér egy  $\mathcal{R}$  bázisát adják. Jelölje egy tetszőleges  $\mathbf{x}$  sortérbeli vektor  $\mathbf{e}$  bázisra vonatkozó koordinátás alakját  $[\mathbf{x}]_{\mathcal{R}}$ , így  $\mathbf{x} = \mathbf{R}^T[\mathbf{x}]_{\mathcal{R}}$  és  $\mathbf{y} = \mathbf{Ax} = \mathbf{AR}^T[\mathbf{x}]_{\mathcal{R}}$ . Olyan mátrixot keresünk tehát, melyre  $\mathbf{A}^+\mathbf{y} = \mathbf{x}$ , azaz  $\mathbf{A}^+\mathbf{AR}^T[\mathbf{x}]_{\mathcal{R}} = \mathbf{R}^T[\mathbf{x}]_{\mathcal{R}}$ . Kihasználjuk, hogy  $\mathbf{B}$  és  $\mathbf{R}^T$  teljes oszloprangú, így  $\mathbf{B}^T\mathbf{B}$  és  $\mathbf{RR}^T$  invertálható. Kiindulva az  $\mathbf{y} = \mathbf{Ax}$  egyenletből, annak mindkét oldalát  $(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T$ -tal, majd  $\mathbf{R}^T(\mathbf{RR}^T)^{-1}$ -zel szorozva kapjuk, hogy

$$\begin{aligned} \mathbf{y} &= \mathbf{AR}^T[\mathbf{x}]_{\mathcal{R}} = \mathbf{BRR}^T[\mathbf{x}]_{\mathcal{R}} \\ (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{y} &= (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{BRR}^T[\mathbf{x}]_{\mathcal{R}} = \mathbf{RR}^T[\mathbf{x}]_{\mathcal{R}} \\ \mathbf{R}^T(\mathbf{RR}^T)^{-1}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{y} &= \mathbf{R}^T[\mathbf{x}]_{\mathcal{R}} = \mathbf{x}. \end{aligned}$$

Eszerint  $\mathbf{A}^+ = \mathbf{R}^T(\mathbf{RR}^T)^{-1}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T$  jó jelölt. Megmutatjuk, hogy az így definiált  $\mathbf{A}^+$  mátrix eleget tesz a pszeudoinverz definíciójának. Ha  $\mathbf{x} \in \mathcal{S}(\mathbf{A})$ , és  $\mathbf{y} = \mathbf{Ax}$ , akkor épp  $\mathbf{A}^+$  származtatásából adódóan  $\mathbf{A}^+\mathbf{y} = \mathbf{x}$ . Legyen tehát  $\mathbf{z} \perp \mathcal{O}(\mathbf{A})$ , azaz  $\mathbf{z} \in \mathcal{N}(\mathbf{A}^T)$ . Ekkor, mivel  $\mathbf{B}$  oszlopai

kifeszítik az  $\mathcal{O}(\mathbf{A})$  alteret,  $\mathbf{B}^T\mathbf{z} = \mathbf{0}$ . Így

$$\mathbf{A}^+\mathbf{z} = \mathbf{R}^T(\mathbf{RR}^T)^{-1}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{z} = \mathbf{0}.$$

Képezzük az  $m \times m$ -es  $\mathbf{C} = [\mathbf{AR}^T|\mathbf{Z}]$  mátrixot, ahol  $\mathbf{Z}$  oszlopvektorai alkossák  $\mathcal{N}(\mathbf{A}^T)$  egy bázisát. Mivel  $\mathbf{AR}^T$  oszlopai az  $\mathcal{O}(\mathbf{A})$ -ban bázist alkotnak, ezért  $\mathbf{C}$  oszlopai függetlenek, így  $\mathbf{C}$  invertálható. Ha  $\mathbf{X}$  egy pszeudoinverz, akkor annak definíciója szerint  $\mathbf{XC} = [\mathbf{R}^T|\mathbf{O}]$ , így  $\mathbf{A}^+\mathbf{C} = \mathbf{XC}$ , amiből  $\mathbf{X} = \mathbf{A}^+$ . Ezzel az egyértelműséget is igazoltuk.

7.31. 1. megoldás: Ha  $\mathbf{A}$  merőleges vetítés mátrixa, akkor  $\mathbf{A}^T = \mathbf{A}$  okán  $\mathcal{S}(\mathbf{A}) = \mathcal{O}(\mathbf{A})$ , és a sortér minden  $\mathbf{x}$  vektorára  $\mathbf{Ax} = \mathbf{x}$ , így  $\mathbf{A}^+\mathbf{x} = \mathbf{x}$  is igaz lesz. Másrészt  $\mathbf{A}^T = \mathbf{A}$  miatt  $\mathcal{N}(\mathbf{A}^T) = \mathcal{N}(\mathbf{A})$ , tehát minden  $\mathbf{z} \in \mathcal{N}(\mathbf{A}^T)$  esetén  $\mathbf{Az} = \mathbf{0}$ , és definíció szerint  $\mathbf{A}^+\mathbf{z} = \mathbf{0}$  is fennáll, tehát  $\mathbf{A}$  és  $\mathbf{A}^+$  hatása az  $\mathcal{O}(\mathbf{A})$  és annak merőleges kiegészítő alterén is megegyezik, így a két mátrix azonos.

2. megoldás: A Penrose-tétel alapján azt kell ellenőriznünk, hogy  $\mathbf{A}$  a négy feltétel mindegyikét teljesíti, mint pszeudoinverz. Ez igaz, hisz  $\mathbf{A}^T = \mathbf{A} = \mathbf{A}^2$  miatt  $\mathbf{A}^3 = \mathbf{A}$ , és  $(\mathbf{A}^2)^T = \mathbf{A}^2$ .

Az állítás megfordítása nem igaz. Például az  $\mathbf{A} = -\mathbf{I}$  mátrixra  $\mathbf{A}^+ = \mathbf{A}$ , de a  $-\mathbf{I}$  mátrixra  $(-\mathbf{I})^2 \neq -\mathbf{I}$ , tehát  $-\mathbf{I}$  nem vetítés mátrixa.

7.32. Az első állítás azonnal következik a (7.10) egyenlőségből, ugyanis ha  $\mathbf{A}$  teljes oszloprangú, akkor

$$\begin{aligned} \mathbf{A}^+\mathbf{A} &= ((\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T)\mathbf{A} \\ &= (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{A} = \mathbf{I} \end{aligned}$$

A másik állítás hasonlóan adódik a (7.11) képletből.

7.33. Ha  $\text{r}(\mathbf{A}) = 1$ , akkor létezik olyan  $\mathbf{a}$  és  $\mathbf{b}$  vektor, hogy  $\mathbf{A} = \mathbf{ab}^T$ . Ekkor a pszeudoinverz kiszámításáról szóló ?? feladat (??) képlete szerint

$$\begin{aligned} \mathbf{A}^+ &= \mathbf{b}(\mathbf{b}^T\mathbf{b})^{-1}(\mathbf{a}^T\mathbf{a})^{-1}\mathbf{a}^T \\ &= \frac{1}{\mathbf{a}^T\mathbf{ab}^T\mathbf{b}}\mathbf{ba}^T \\ &= \frac{1}{\text{trace}(\mathbf{A}^T\mathbf{A})}\mathbf{A}^T. \end{aligned}$$

Az utóbbi egyenlőség azon múlik, hogy mind  $\mathbf{a}^T\mathbf{ab}^T\mathbf{b}$ , mind  $\text{trace}(\mathbf{A}^T\mathbf{A})$  az összes  $a_i^2b_j^2$  alakú elem összege, ahol  $\mathbf{a} = (a_1, \dots, a_m)$ ,  $\mathbf{b} = (b_1, \dots, b_n)$ . A vektorokra vonatkozó állítás ennek speciális esete.

7.34. Az állítás igazolásához elég csak a Penrose-tétel négy feltételét ellenőrizni.

7.35. A 7.34. feladat alapján az alábbi blokkosítással és 1-rangú mátrixok pszeudoinverzére vonatkozó állításból

azonnal adódik a válasz:

$$\left[ \begin{array}{ccc|ccc} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{array} \right]^+ = \left[ \begin{array}{ccc|ccc} 1/3 & 0 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 0 \end{array} \right]$$

7.36.  $\mathbf{v}_2 = (9, 3, -1, 5) - \frac{(9, 3, -1, 5)(1, 1, -1, -1)}{|(1, 1, -1, -1)|^2} = (7, 1, 1, 7)$ ,  
így az ortonormált bázis vektorai:  $\frac{1}{2}(1, 1, -1, -1)$ ,  
 $\frac{1}{10}(7, 1, 1, 7)$

7.37.  $\mathbf{v}_2 = (8, 6, 2, 0) - \frac{(8, 6, 2, 0)(1, -1, 1, -1)}{|(1, -1, 1, -1)|^2} = (7, 7, 1, 1)$ ,  
így az ortonormált bázis vektorai:  
 $\frac{1}{2}(1, -1, 1, -1)$ ,  $\frac{1}{10}(7, 7, 1, 1)$

7.38.  $(2, -1, -20) - \frac{(1, 4, 8)(2, -1, -20)}{(1, 4, 8)(1, 4, 8)}(1, 4, 8) = (4, 7, -4)$ .

$$\mathbf{Q} = \frac{1}{9} \begin{bmatrix} 1 & 4 \\ 4 & 7 \\ 8 & -4 \end{bmatrix}, \text{ és } \mathbf{R} = \mathbf{Q}^T \mathbf{A} = \begin{bmatrix} 1 & -18 \\ 0 & 9 \end{bmatrix}.$$

7.39. A 7.37.. feladat eredményét alkalmazva

$$\mathbf{A} = \mathbf{QR} = \begin{bmatrix} 1/2 & 7/10 \\ -1/2 & 7/10 \\ 1/2 & 1/10 \\ -1/2 & 1/10 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 0 & 10 \end{bmatrix}.$$

7.40. Először a harmadik sor első elemét elimináljuk. Ekkor  $a = 8$ ,  $b = 15$ , így  $r = \sqrt{8^2 + 15^2} = 17$ . Eszerint az első és harmadik sorok és oszlop kereszteződéseiben lévő  $\begin{bmatrix} 8 & -4 \\ 15 & 1 \end{bmatrix}$  részmátrix első oszlopvektorát, azaz a  $(8, 15)$  vektort beforgatjuk a  $(17, 0)$  vektorba. A forgatás  $\alpha$  szögére  $\cos \alpha = 8/17$ ,  $\sin \alpha = -15/17$ . Ebből:

$$\mathbf{Q}_1 = \begin{bmatrix} 8/17 & 0 & 15/17 \\ 0 & 1 & 0 \\ -15/17 & 0 & 8/17 \end{bmatrix} \quad \mathbf{Q}_1 \mathbf{A} = \begin{bmatrix} 17 & 12 & -1 \\ 0 & 4 & 2 \\ 0 & 3 & 4 \end{bmatrix}$$

A  $\mathbf{Q}_1 \mathbf{A}$  mátrix harmadik sor második elemének eliminálásához  $a = 4$ ,  $b = 3$ ,  $r = \sqrt{4^2 + 3^2} = 5$ ,  $\cos \alpha = 4/5$ ,  
 $-\sin \alpha = 3/5$ :

$$\mathbf{Q}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4/5 & 3/5 \\ 0 & -3/5 & 4/5 \end{bmatrix} \quad \mathbf{R} = \mathbf{Q}_2 \mathbf{Q}_1 \mathbf{A} = \begin{bmatrix} 17 & 12 & -1 \\ 0 & 5 & 4 \\ 0 & 0 & 2 \end{bmatrix}$$

Végül

$$\mathbf{Q} = (\mathbf{Q}_2 \mathbf{Q}_1)^{-1} = \mathbf{Q}_1^T \mathbf{Q}_2^T = \begin{bmatrix} 8/17 & -9/17 & -12/17 \\ 0 & 4/5 & -3/5 \\ 15/17 & 24/85 & 32/85 \end{bmatrix}.$$

7.41. Az első oszlop főátló alatti elemeinek eliminálását az  $(1, 1, 1, 1) \mapsto (2, 0, 0, 0)$  hozzárendelést eredményező, az

$$\mathbf{a} = (1, 1, 1, 1) - (2, 0, 0, 0) = (-1, 1, 1, 1)$$

vektorra merőleges hipersíkra való tükrözéssel valósítjuk meg. E tükrözés mátrixa:

$$\begin{aligned} \mathbf{Q}_1 &= \mathbf{I}_4 - \frac{2}{\mathbf{a}^T \mathbf{a}} \mathbf{a} \mathbf{a}^T \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \\ \mathbf{Q}_1 \mathbf{A} &= \begin{bmatrix} 2 & 3 & 3 & 4 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -2 & 2 \\ 0 & -5 & 3 & 1 \end{bmatrix} \end{aligned}$$

A  $\mathbf{Q}_1 \mathbf{A}$  második oszlopában a főátló alatti elemek eliminálása a  $(0, 0, -5)$  vektort az  $(5, 0, 0)$ -ba képző tükrözéssel valósítható meg. Ez az  $\mathbf{a} = (0, 0, -5) - (5, 0, 0)$  normálvektorú síkra való tükrözés, melynek mátrixa:

$$\begin{aligned} \mathbf{H}_2 &= \mathbf{I}_3 - \frac{2}{\mathbf{a}^T \mathbf{a}} \mathbf{a} \mathbf{a}^T \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix} \end{aligned}$$

Innen

$$\begin{aligned} \mathbf{Q}_2 &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \quad \hat{\mathbf{R}} = \mathbf{Q}_2 \mathbf{Q}_1 \mathbf{A} = \begin{bmatrix} 2 & 3 & 3 & 4 \\ 0 & 5 & -3 & -1 \\ 0 & 0 & -2 & 2 \\ 0 & 0 & 0 & -1 \end{bmatrix} \\ \hat{\mathbf{Q}} &= (\mathbf{Q}_2 \mathbf{Q}_1)^{-1} = \mathbf{Q}_1^T \mathbf{Q}_2^T = \mathbf{Q}_1 \mathbf{Q}_2 = \frac{1}{2} \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix}. \end{aligned}$$

E mátrixokra  $\mathbf{A} = \hat{\mathbf{Q}} \hat{\mathbf{R}}$ , amit szokás QR-felbontásnak tekinteni, de az általunk adott definíciónak nem felel meg, mert  $\hat{\mathbf{R}}$  főátlójában nem csak pozitív elemek szerepelnek. Az  $\hat{\mathbf{R}}$  harmadik és negyedik sorának, valamint a  $\hat{\mathbf{Q}}$  harmadik és negyedik oszlopának  $-1$ -gyel szorzása nem változtat a



szorzatukon, így a QR-felbontás mátrixai:

$$\mathbf{Q} = \frac{1}{2} \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} 2 & 3 & 3 & 4 \\ 0 & 5 & -3 & -1 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

7.42.  $r_{ii}$  az  $\mathbf{a}_i$  távolsága az  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{i-1}$  vektorok által kifeszített altértől!

$$7.43. \quad \mathbf{Q} = \begin{bmatrix} 0.70711 & -0.40825 & 0.28868 & -0.50000 \\ 0.70711 & 0.40825 & -0.28868 & 0.50000 \\ 0.00000 & 0.81650 & 0.28868 & -0.50000 \\ 0.00000 & 0.00000 & 0.86603 & 0.50000 \end{bmatrix},$$

$$\mathbf{R} = \begin{bmatrix} 1.41421 & 2.12132 & 0.70711 & 0.00000 \\ 0.00000 & 1.22474 & 2.04124 & 0.81650 \\ 0.00000 & 0.00000 & 1.15470 & 2.02073 \\ 0.00000 & 0.00000 & 0.00000 & 0.50000 \end{bmatrix}$$

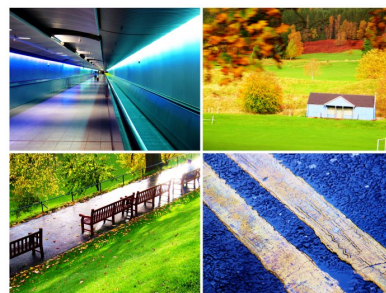


### **III. rész**

## **Mátrixok sajátosságai**



E rész a mátrixok legfontosabb tulajdonságait, sajátságait vizsgálja. A *sajátságokra* való utalás egyúttal egy szójáték, mely a mátrix *sajátértékének*, *sajátvektorának* és *sajátalterének* igen fontos fogalmára, valamint a születésekor ugyancsak *sajátértéknek* nevezett szinguláris értékre utal. Vizsgálatainkat a négyzetes mátrixok lehető legegyszerűbb – diagonális – alakra hozásával kezdjük, amihez a sajátértékek meghatározása vezet. Az alkalmazásokban az utóbbi időben különösen fontossá vált, és minden mátrixra működő másik diagonalizáló technika a szinguláris értékekhez kapcsolódik. Ennek tárgyalását a diagonalizálhatósághoz kapcsolódó kérdések tisztázása, a „majdnem diagonális alak”, a Jordan-féle normálalak leírása követi, végül e részt az alkalmazásokban különösen fontos nemnegatív mátrixok vizsgálata zárja.



green&blue (CC) by Joós Andi



## 8

### Sajátérték, diagonalizálás

Egy mátrix jellemzésének különösen hatékony eszköze azoknak a nullvektortól különböző  $\mathbf{x}$  vektoroknak a meghatározása, amelyeket a mátrixszal való szorzás önmagukkal párhuzamos vektorokba visz, azaz amelyekre  $\mathbf{Ax} = \lambda\mathbf{x}$ . E vektorok ismerete olyan bázis megtalálásához is hozzásegít, amelyben e mátrix lényegesen egyszerűbb – például diagonális – alakot ölt.

#### Sajátérték, sajátvektor, sajátaltér

**A sajátérték és a sajátvektor fogalma** Kezdjük egy egyszerű feladattal, melyből kiolvasható annak lényege, amiről e fejezetben szó lesz.

**8.1. PÉLDA (JÓ BÁZIS TÜKRÖZÉSHEZ).** *Tükrözzük a 3-dimenziós tér vektorait a tér egy megadott síkjára! Geometriai szemléletünkre hagyatkozva válasszunk e lineáris leképezés leírásához egy megfelelő bázist, majd írjuk fel a tükrözés e bázisra vonatkozó mátrixát!*

**MEGOLDÁS.** A síkra való tükrözés a síkra merőleges vektorokat ellentettjükbe viszi, míg a sík vektorait helyben hagyja. A tér minden vektorra egyértelműen előáll egy síkba eső és egy rá merőleges vektor összegként. A mellékelt ábra szemlélteti, hogy a tér mindegyik vektorának tükörképe úgy kapható meg, hogy a síkba eső összetevőjét meghagyjuk, és ahhoz adjuk a síkra merőleges összetevő ellentettjét. Természetes módon adódik az ötlet, hogy válasszuk ki a sík egy tetszőleges bázisát (álljon ez az  $\mathbf{a}$  és  $\mathbf{b}$  vektorokból), és e két vektorhoz vegyünk hozzá egy a síkra merőleges  $\mathbf{c}$  vektort harmadik bázisvektornak. Ekkor a tükröző  $T$  leképezés hatása e vektorokon:  $T\mathbf{a} = \mathbf{a}$ ,  $T\mathbf{b} = \mathbf{b}$  és  $T\mathbf{c} = -\mathbf{c}$ . Az  $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$  bázisban  $T$  mátrixa

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Így e bázisban egy tetszőleges  $(x, y, z)$  vektor tükörképe  $(x, y, -z)$ .  $\square$

E példában úgy választottunk bázist, hogy olyan vektorokat kerestünk, melyek önmaguk skalárszorosaiba mennek, azaz amelyek kielégítenek egy  $T\mathbf{x} = \lambda\mathbf{x}$  alakú egyenletet. Ez a következő definícióhoz vezet, melyet először csak mátrixokra mondunk ki.

**8.2. DEFINÍCIÓ (SAJÁTÉRTÉK, SAJÁTVEKTOR).** Azt mondjuk, hogy a  $\lambda$  szám az  $\mathbf{A}$  mátrix sajátértéke, ha létezik olyan nemnulla  $\mathbf{x}$  vektor, melyre  $\mathbf{Ax} = \lambda\mathbf{x}$ . Az ilyen  $\mathbf{x}$  vektorokat az  $\mathbf{A}$  mátrix  $\lambda$  sajátértékéhez tartozó sajátvektorainak, a  $(\lambda, \mathbf{x})$  párokat pedig az  $\mathbf{A}$  sajátpárjainak nevezzük.

**8.3. PÉLDA (SAJÁTÉRTÉK, SAJÁTVEKTOR).** Igazoljuk, hogy az  $\mathbf{A} = \begin{bmatrix} -2 & 2 \\ -2 & 3 \end{bmatrix}$  mátrixnak  $-1$  egy sajátértéke, és  $(2, 1)$  az egyik hozzátartozó sajátvektora, azaz  $(-1, (2, 1))$  egy sajátpár. Mutassuk meg, hogy a  $(2, (1, 2))$  pár egy másik sajátpár!

**MEGOLDÁS.** Valóban,

$$\begin{bmatrix} -2 & 2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \end{bmatrix}, \quad \text{azaz} \quad \begin{bmatrix} -2 & 2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = (-1) \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

E mátrix egy másik sajátértéke 2, ugyanis

$$\begin{bmatrix} -2 & 2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \quad \text{azaz} \quad \begin{bmatrix} -2 & 2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 2 \end{bmatrix}. \quad \square$$

Ha  $\mathbf{x}$  egy sajátvektor, akkor minden nemnulla konstansszorosa is az, ugyanis

$$\mathbf{A}(c\mathbf{x}) = c\mathbf{Ax} = c\lambda\mathbf{x} = \lambda(c\mathbf{x}),$$

azaz  $\mathbf{A}(c\mathbf{x}) = \lambda(c\mathbf{x})$ . Ennél több is igaz:

**8.4. ÁLLÍTÁS (A SAJÁTVEKTOROK ALTEREI).** Ha az  $\mathbf{A}$  mátrixnak  $\lambda$  egy sajátértéke, akkor a  $\lambda$ -hoz tartozó sajátvektorok a nullvektorral együtt alteret alkotnak, mely megegyezik  $\mathbf{A} - \lambda\mathbf{I}$  nullterével.

**BIZONYÍTÁS.** A nem nullvektor  $\mathbf{x}$  pontosan akkor egy  $\lambda$  sajátértékhez tartozó sajátvektor, ha kielégíti az  $\mathbf{Ax} = \lambda\mathbf{x}$  egyenletet, azaz az  $\mathbf{Ax} - \lambda\mathbf{x} = \mathbf{0}$  egyenletet, vagyis ha megoldása a homogén lineáris  $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$  egyenletnek. Ez pedig épp azt jelenti, hogy  $\mathbf{x}$  eleme  $\mathbf{A} - \lambda\mathbf{I}$  nullterének.  $\square$

**8.5. DEFINÍCIÓ (SAJÁTALTÉR).** A négyzetes  $\mathbf{A}$  mátrix  $\lambda$  sajátértékéhez tartozó sajátvektorai és a nullvektor által alkotott alteret a  $\lambda$  sajátértékhez tartozó sajátaltérnek nevezzük.

**8.6. PÉLDA (SAJÁTALTÉR BÁZISÁNAK MEGHATÁROZÁSA).** Adjuk meg az

$$\mathbf{A} = \begin{bmatrix} 3 & 6 & 1 \\ 1 & 8 & 1 \\ 1 & 6 & 3 \end{bmatrix}$$



mátrix 2-höz, mint sajátértékhez tartozó sajátalterét úgy, hogy megadjuk egy bázisát! Tegyük meg ugyanezt a 10-hez tartozó sajátaltérrel is.

**MEGOLDÁS.** Először ellenőrizzük, hogy a 2 sajátérték! Ehhez meg kell mutatni, hogy az  $(\mathbf{A} - 2\mathbf{I})\mathbf{x} = \mathbf{0}$  egyenletrendszernek van nemtriviális megoldása. Hozzuk az együtthatómátrixot redukált lépcsős alakra:

$$\mathbf{A} - 2\mathbf{I} = \begin{bmatrix} 1 & 6 & 1 \\ 1 & 6 & 1 \\ 1 & 6 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 6 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Mivel  $r(\mathbf{A} - 2\mathbf{I}) = 1$ , ezért az  $(\mathbf{A} - 2\mathbf{I})\mathbf{x} = \mathbf{0}$  egyenletrendszer szabad változóinak száma 2, és megoldása

$$\mathbf{x} = \begin{bmatrix} -6s - t \\ s \\ t \end{bmatrix} = s \begin{bmatrix} -6 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Tehát a sajátaltér egy bázisa a  $(-6, 1, 0)$  és  $(-1, 0, 1)$  vektorokból áll.

A 10 is sajátérték, mivel az  $(\mathbf{A} - 10\mathbf{I})\mathbf{x} = \mathbf{0}$  egyenletrendszernek van nemtriviális megoldása, ugyanis

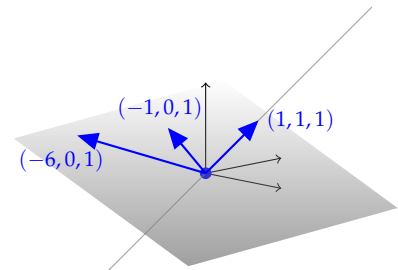
$$\mathbf{A} - 10\mathbf{I} = \begin{bmatrix} -7 & 6 & 1 \\ 1 & -2 & 1 \\ 1 & 6 & -7 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix},$$

tehát a megoldás

$$\mathbf{x} = \begin{bmatrix} t \\ t \\ t \end{bmatrix} = t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Így a sajátalteret az  $(1, 1, 1)$  vektor feszíti ki.

A két sajátaltér egyike 2-, másika 1-dimenziós altér. Ezt szemlélteti a 8.1. ábra. □



8.1. ábra: A 8.6. feladatbeli  $\mathbf{A}$  mátrix sajátalterei: a 2-höz tartozó 2-dimenziós altér, melyet a  $(-6, 1, 0)$  és  $(-1, 0, 1)$  vektorok feszítenek ki, és a 10-hez tartozó 1-dimenziós altér, melyet a  $(1, 1, 1)$  vektor feszít ki.

**Karakterisztikus polinom** Láttuk, hogy az  $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$  egyenletnek pontosan akkor van a zérusvektortól különböző megoldása, ha a homogén lineáris  $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$  egyenletrendszernek van nemtriviális megoldása. Ez a 6.3. tétel szerint pontosan akkor igaz, ha

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0. \quad (8.1)$$

Ez tehát azt jelenti, hogy  $\lambda$  pontosan akkor sajátérték, ha kielégíti a (8.1) egyenletet. Ezt az egyenletet az  $\mathbf{A}$  mátrix *karakterisztikus egyenletének* nevezzük. Ha  $\mathbf{A}$  egy  $n \times n$ -es mátrix, akkor az egyenlet bal oldala a determináns kifejtése után egy  $n$ -edfokú polinom, melyet *karakterisztikus polinomnak* nevezünk. Az  $n$ -edrendű  $\mathbf{A}$  mátrix karakterisztikus polinomját  $\chi_{\mathbf{A}}$ -val jelöljük, ennek általános alakja tehát

$$\chi_{\mathbf{A}}(\lambda) = (-1)^n \lambda^n + p_{n-1} \lambda^{n-1} + \dots + p_1 \lambda + p_0. \quad (8.2)$$

8.7. PÉLDA (KARAKTERISZTIKUS POLINOM FELÍRÁSA). *Határozzuk meg az*

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} a & b & c \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

*mátrixok karakterisztikus polinomját és ahol lehet, próbáljunk meg általánosabb érvényű állításokat megsejteni az eredmény alapján!*

**MEGOLDÁS.** Ki kell számítanunk a  $\det(\mathbf{A} - \lambda \mathbf{I})$  determináns értékét:

$$\begin{aligned} \det(\mathbf{A} - \lambda \mathbf{I}) &= \begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} = (a - \lambda)(d - \lambda) - bc \\ &= \lambda^2 - (a + d)\lambda + (ad - bc) \\ &= \lambda^2 - \text{trace}(\mathbf{A})\lambda + \det \mathbf{A}. \end{aligned} \quad (8.3)$$

Kimondható a következtetés:  $2 \times 2$ -es mátrixok karakterisztikus polinomját a mátrix nyomával és determinánsával is ki tudjuk fejezni.

A  $\mathbf{B}$  mátrix karakterisztikus polinomja

$$\det(\mathbf{B} - \lambda \mathbf{I}) = \begin{vmatrix} 1 - \lambda & a & b \\ 0 & 1 - \lambda & c \\ 0 & 0 & 1 - \lambda \end{vmatrix} = (1 - \lambda)^3.$$

Ebből leolvasható, hogy a háromszögmátrixok karakterisztikus polinomjának alakját nem befolyásolják a főátlón kívüli elemek (ld. a 8.8. tételt).

A  $\mathbf{C}$  mátrix karakterisztikus polinomja

$$\begin{aligned} \det(\mathbf{C} - \lambda \mathbf{I}) &= \begin{vmatrix} a - \lambda & b & c \\ 1 & -\lambda & 0 \\ 0 & 1 & -\lambda \end{vmatrix} \\ &= (a - \lambda)\lambda^2 + b\lambda + c \\ &= -\lambda^3 + a\lambda^2 + b\lambda + c. \end{aligned}$$

Ez azt sejteti, hogy minden polinomhoz könnyen konstruálható olyan mátrix, melynek az a karakterisztikus polinomja (ld. a ?? feladatot).  $\square$

Az előző feladat tanulságait külön állításokban is megfogalmazzuk:

8.8. ÁLLÍTÁS (HÁROMSZÖGMÁTRIXOK SAJÁTÉRTÉKEI). *A háromszögmátrixok és így a diagonális mátrixok sajátértékei megegyeznek a főátló elemeivel.*

**BIZONYÍTÁS.** Ha  $\mathbf{A}$  háromszögmátrix, akkor  $\mathbf{A} - \lambda \mathbf{I}$  is, és egy háromszögmátrix determinánsa megegyezik főátlóbeli elemeinek szorzatával. Eszerint az  $\mathbf{A} = [a_{ij}]$  háromszögmátrix karakterisztikus egyenlete

$$(a_{11} - \lambda)(a_{22} - \lambda) \dots (a_{nn} - \lambda) = 0,$$

A karakterisztikus polinomot a

$$\det(\lambda \mathbf{I} - \mathbf{A})$$

determinánssal is szokás definiálni. Előnye, hogy ekkor a polinom főegyütthatója mindig 1, míg az általunk használt definíció szerint a páratlan rendű mátrixok karakterisztikus polinomjának  $-1$  a főegyütthatója. Hátránya viszont az, hogy a konstans tag nem mindig a determináns, másrészt a kézzel való számolás is nehezebb, ezért az elemi feladatok egyszerűbb számolhatósága érdekében is hasznosabb a  $\det(\mathbf{A} - \lambda \mathbf{I})$  alakot választani.

aminek a gyökei  $a_{ii}$  ( $i = 1, \dots, n$ ). Így ezek az  $\mathbf{A}$  sajátértékei.  $\square$

**8.9. ÁLLÍTÁS (DETERMINÁNS, NYOM ÉS A SAJÁTÉRTÉKEK).** Ha az  $n$ -edrendű  $\mathbf{A}$  mátrix sajátértékei  $\lambda_1, \dots, \lambda_n$ , akkor

$$\det(\mathbf{A}) = \lambda_1 \lambda_2 \dots \lambda_n$$

$$\text{trace}(\mathbf{A}) = \lambda_1 + \lambda_2 + \dots + \lambda_n$$

Ezek az értékek megjelennek a karakterisztikus polinomban: a determináns a konstans tag, a nyom a  $(-\lambda)^{n-1}$  együtthatója.

**BIZONYÍTÁS.** A karakterisztikus polinom gyöktényezős alakja:

$$\det(\mathbf{A} - \lambda \mathbf{I}) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \dots (\lambda_n - \lambda)$$

$\lambda = 0$  behelyettesítése után kapjuk, hogy

$$\det(\mathbf{A}) = \lambda_1 \lambda_2 \dots \lambda_n.$$

Az állítás nyomra vonatkozó részének bizonyítását feladatként tűzzük ki.  $\square$

*A valós  $2 \times 2$ -es mátrixok sajátaltérjeinek jellemzése* Olyan eredményekkel fogunk megismerkedni e paragrafusban, melyek általánosíthatóak lesznek magasabb dimenzióra, de 2-dimenzió esetén egyszerűbb a szemléltetésük.

Láttuk, hogy ha  $\mathbf{x}$  sajátvektor, akkor bármely konstansszorosa is az. Így egy egyenessel párhuzamos vektorok közül elég csak egy vektor képét vizsgálni, mondjuk az egységvektorét. Hasznos lesz tehát a lineáris leképezések korábban megismert egységkör-ábrázolása (ld. 7.7 ábra).

**8.10. PÉLDA ( $2 \times 2$ -ES MÁTRIXOK SAJÁTVEKTORAINAK ÁBRÁZOLÁSA).** Határozzuk meg a 7.11. és a ?? példákban is szereplő

$$\mathbf{A} = \begin{bmatrix} 5 & 3 \\ 4 & 4 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 3 & 5 \\ 4 & 4 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} -5 & 3 \\ -4 & 4 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} -3 & 5 \\ -4 & 4 \end{bmatrix}.$$

mátrixok sajátértékeit és sajátvektorait. Szemléltessük ezeket az egységkör-ábrákban.

**MEGOLDÁS.** Egyszerű számolással meghatározható mind a négy mátrix karakterisztikus egyenlete, sajátértékei és sajátvektorai, bár a  $\mathbf{D}$  mátrix esetén ezek komplex számokat is tartalmaznak. A karakterisztikus polinomot jelölje  $p$ , indexében a mátrix jelével. Ezután megadjuk

a sajátértékeket, majd a sajátvektorokat:

$$\chi_A(\lambda) = \lambda^2 - \frac{5}{2}\lambda + 1, \quad \lambda_1 = 2, \quad \lambda_2 = \frac{1}{2}, \quad \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

$$\chi_B(\lambda) = \lambda^2 - \frac{3}{2}\lambda - 1, \quad \lambda_1 = 2, \quad \lambda_2 = -\frac{1}{2}, \quad \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

$$\chi_C(\lambda) = \lambda^2 - 1, \quad \lambda_1 = 1, \quad \lambda_2 = -1, \quad \mathbf{x}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

$$\chi_D(\lambda) = \lambda^2 + 1, \quad \lambda_1 = i, \quad \lambda_2 = -i, \quad \mathbf{x}_1 = \begin{bmatrix} \frac{3}{5} - \frac{4}{5}i \\ 1 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} \frac{3}{5} + \frac{4}{5}i \\ 1 \end{bmatrix}.$$

Némi nehézséget a **D** mátrix okoz, ezért az ahhoz tartozó számításokat részletezzük:

$$|\mathbf{D} - \lambda \mathbf{I}| = \begin{vmatrix} -\frac{3}{4} - \lambda & \frac{5}{4} \\ -\frac{5}{4} & \frac{3}{4} - \lambda \end{vmatrix} = \lambda^2 + 1$$

$$\lambda_1 = i: \begin{bmatrix} -\frac{3}{4} - i & \frac{5}{4} \\ -\frac{5}{4} & \frac{3}{4} - i \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -\frac{3}{5} + \frac{4}{5}i \\ 0 & 0 \end{bmatrix} \text{ amiből } \mathbf{x}_1 = \begin{bmatrix} \frac{3}{5} - \frac{4}{5}i \\ 1 \end{bmatrix},$$

$$\lambda_2 = -i: \begin{bmatrix} -\frac{3}{4} + i & \frac{5}{4} \\ -\frac{5}{4} & \frac{3}{4} + i \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -\frac{3}{5} - \frac{4}{5}i \\ 0 & 0 \end{bmatrix} \text{ amiből } \mathbf{x}_2 = \begin{bmatrix} \frac{3}{5} + \frac{4}{5}i \\ 1 \end{bmatrix}.$$

A négy mátrixhoz tartozó egységkörábra a 8.2 ábrán látható.  $\square$

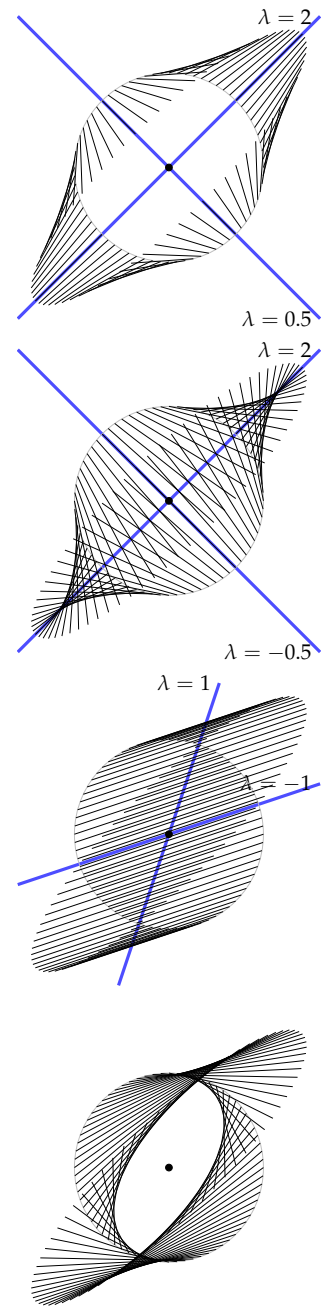
**8.11. TÉTEL** ( $2 \times 2$ -ES SZIMMETRIKUS MÁTRIXOK SAJÁTALTEREI). Legyen **A** egy  $2 \times 2$ -es szimmetrikus mátrix. Ekkor

- A** minden sajátértéke valós,
- A**-nak pontosan akkor van két azonos sajátértéke, ha  $a\mathbf{I}$  alakú, ekkor a sík összes vektora sajátvektor,
- minden egyéb esetben **A**-nak két különböző sajátértéke van, és ekkor sajátalterei merőlegesek egymásra.

**BIZONYÍTÁS.** A  $2 \times 2$ -es szimmetrikus valós mátrix általános alakja  $\mathbf{A} = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$ , ahol  $a, b, d \in \mathbb{R}$ . Ennek karakterisztikus egyenlete a (8.3) szerint  $\lambda^2 - (a+d)\lambda + (ad - b^2)$ . Az egyenlet diszkriminánsa  $D = (a+d)^2 - 4(ad - b^2) = (a-d)^2 + 4b^2 \geq 0$ . Tehát a gyökök, vagyis a sajátértékek valósak. Ez bizonyítja a)-t. A két sajátérték pontosan akkor egyezik meg, ha  $D = 0$ , ez viszont csak  $a = d$  és  $b = 0$  esetén lehetséges, ami bizonyítja b)-t. A c) állítás igazolását a feladatok közt tűzzük ki.  $\square$

**Mátrix összes sajátértékének és sajátvektorának meghatározása** Az előző paragrafusokban leírtak alapján egy mátrix sajátértékeinek és sajátvektorainak meghatározása két lépésben elvégezhető:

- megoldjuk a  $\det(\mathbf{A} - \lambda \mathbf{I}) = 0$  karakterisztikus egyenletet, ennek gyökei a sajátértékek,



8.2. ábra: A négy leképezés sajátirányai.

2. minden  $\lambda$  sajátértékhez meghatározzuk az  $\mathbf{A} - \lambda \mathbf{I}$  nullterének egy bázisát, az általa kifeszített altér nemzérus vektorai a  $\lambda$ -hoz tartozó sajátvektorok.

8.12. PÉLDA (AZ ÖSSZES SAJÁTÉRTÉK ÉS SAJÁTVEKTOR MEGHATÁROZÁSA). Határozzuk meg a

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

mátrix sajátértékeit és sajátvektorait!

**MEGOLDÁS.** Az első lépés a karakterisztikus egyenletet felírása és megoldása. A kiszámítandó determináns háromszög alakú, így értéke a főátlóbeli elemek szorzata:

$$\det(\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} 0 - \lambda & 1 & 1 \\ 0 & 2 - \lambda & 0 \\ 0 & 0 & 2 - \lambda \end{vmatrix} = -\lambda(2 - \lambda)^2$$

A karakterisztikus egyenlet gyökei és így az  $\mathbf{A}$  mátrix sajátértékei  $\lambda_1 = 0$ ,  $\lambda_2 = \lambda_3 = 2$ .

Tekintsük először a  $\lambda_1 = 0$  esetet.  $\mathbf{A} - \lambda_1 \mathbf{I}$  nullterének meghatározásához redukált lépcsős alakra hozzuk az  $\mathbf{A} - \lambda_1 \mathbf{I}$  mátrixot:

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{matrix} x_2 = 0 \\ x_3 = 0. \end{matrix}$$

Ennek megoldása  $x_1 = t$ , azaz az összes megoldás

$$\begin{bmatrix} t \\ 0 \\ 0 \end{bmatrix} = t \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Tehát a  $\lambda_1 = 0$  sajátértékhez tartozó sajátaltér az  $(1, 0, 0)$  vektor által kifeszített altér.

Tekintsük ezután a  $\lambda_2 = \lambda_3 = 2$  esetet. Meghatározzuk az  $\mathbf{A} - 2\mathbf{I}$  mátrix nullterét.

$$\begin{bmatrix} -2 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 2 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow 2x_1 - x_2 - x_3 = 0$$

Ennek az (egy egyenletből álló) egyenletrendszernek a megoldása:  $x_2 = s$ ,  $x_3 = t$ ,  $x_1 = (s + t)/2$ , azaz

$$\begin{bmatrix} (s + t)/2 \\ s \\ t \end{bmatrix} = s \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1/2 \\ 0 \\ 1 \end{bmatrix}.$$

Tehát a  $\lambda_2 = \lambda_3 = 2$  sajátértékhez tartozó sajátaltér az  $(\frac{1}{2}, 1, 0)$  és az  $(\frac{1}{2}, 0, 1)$  vektorok által kifeszített altér.  $\square$

Az  $n \times n$ -es mátrixok karakterisztikus egyenlete  $n$ -edfokú. Egy ilyen egyenlet megoldására  $n \leq 4$  esetén van megoldóképlet, ezért ezeket az egyenleteket – például egy komputer algebra program segítségével – meg tudjuk oldani. Egyébként vagy szerencsénk van, és az egyenlet olyan alakú, amilyenhez vannak gyors megoldási lehetőségek, vagy csak közelítő megoldás megtalálására van esély.

8.13. PÉLDA (MAGASABBFOKÚ KARAKTERISZTIKUS EGYENLET). *Határozzuk meg az*

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 3 & 3 & 2 \end{bmatrix}$$

*mátrix sajátértékeit és sajátvektorait!*

**MEGOLDÁS.** A karakterisztikus egyenlet:

$$\begin{aligned} \mathbf{A} - \lambda \mathbf{I} &= \begin{bmatrix} 1-\lambda & 2 & 2 \\ 2 & 1-\lambda & 2 \\ 3 & 3 & 2-\lambda \end{bmatrix} \\ &= (1-\lambda)^2(2-\lambda) + 24 - 12(1-\lambda) - 4(2-\lambda) \\ &= -(\lambda^3 - 4\lambda^2 - 11\lambda - 6) \end{aligned}$$

E harmadfokú egyenlet megoldására használhatunk számológépet, megoldóképletet, vagy például a függelékben megtalálható racionálisgyöktételt. Eszerint a karakterisztikus egyenlet  $-(\lambda + 1)^2(\lambda - 6) = 0$ , így gyökei  $\lambda_1 = \lambda_2 = -1$  és  $\lambda_3 = 6$ .

A  $\lambda_1 = \lambda_2 = -1$  esetben

$$\mathbf{A} + \mathbf{I} = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow x_1 + x_2 + x_3 = 0.$$

Ennek megoldása

$$\begin{bmatrix} -s-t \\ s \\ t \end{bmatrix} = s \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix},$$

azaz a  $-1$  sajátértékhez tartozó sajátalteret a  $(-1, 1, 0)$  és a  $(-1, 0, 1)$  vektorok feszítik ki.

A  $\lambda_3 = 6$  esetben

$$\mathbf{A} - 6\mathbf{I} = \begin{bmatrix} -5 & 2 & 2 \\ 2 & -5 & 2 \\ 3 & 3 & -4 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & -2/3 \\ 0 & 1 & -2/3 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{aligned} x_1 - \frac{2}{3}x_3 &= 0 \\ x_2 - \frac{2}{3}x_3 &= 0. \end{aligned}$$

Ennek megoldása a törtek alkalmazását elkerülő  $x_3 = 3t$  paraméterválasztással

$$\begin{bmatrix} 2t \\ 2t \\ 3t \end{bmatrix} = t \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}.$$

Tehát a  $\lambda_3 = 6$  sajátértékhez tartozó sajátalteret a  $(2, 2, 3)$  vektor feszíti ki.  $\square$

**A karakterisztikus egyenlet komplex gyökei** Ha valóselemű mátrixot vizsgálunk, megeshet, hogy a karakterisztikus egyenletnek vannak nem valós gyökei. Mivel a valós számok egyúttal komplexek is, a valós elemű mátrixot tekinthetjük komplex eleműnek is, ekkor viszont a karakterisztikus egyenlet komplex gyökeit is sajátértéknek tekinthetjük. Ebben az esetben a komplex sajátértékhez komplex elemű sajátvektor fog tartozni, amint azt már láthattuk a 8.10. példában.

**8.14. PÉLDA (KOMPLEX SAJÁTÉRTÉKEK ÉS KOMPLEX ELEMŰ SAJÁTVEKTOROK).** Határozzuk meg a komplex elemű

$$\mathbf{A} = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$

mátrix sajátértékeit és sajátvektorait!

**MEGOLDÁS.** A karakterisztikus egyenlet

$$\begin{vmatrix} \frac{1}{2} - \lambda & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} - \lambda \end{vmatrix} = \left(\frac{1}{2} - \lambda\right)^2 + \left(\frac{\sqrt{3}}{2}\right)^2 = \lambda^2 - \lambda + 1.$$

A  $\lambda^2 - \lambda + 1 = 0$  egyenlet gyökei  $\frac{1}{2} \pm \frac{\sqrt{3}}{2}i$ .

Először vizsgáljuk az  $\frac{1}{2} + \frac{\sqrt{3}}{2}i$  sajátértéket:

$$\mathbf{A} - \left(\frac{1}{2} + \frac{\sqrt{3}}{2}i\right) \mathbf{I} = \begin{bmatrix} -\frac{\sqrt{3}}{2}i & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2}i \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -i \\ 0 & 0 \end{bmatrix} \Rightarrow x - iy = 0.$$

Ennek az egyenlet(rendszer)nek a megoldása az  $y = t$  paraméterválasztással

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} it \\ t \end{bmatrix} = t \begin{bmatrix} i \\ 1 \end{bmatrix}.$$

Tehát az  $\frac{1}{2} + \frac{\sqrt{3}}{2}i$  sajátértékhez tartozó sajátaltér egy bázisa az  $(i, 1)$  vektorból áll.

Az  $\frac{1}{2} - \frac{\sqrt{3}}{2}i$  sajátérték esetén

$$\mathbf{A} - \left(\frac{1}{2} - \frac{\sqrt{3}}{2}i\right) \mathbf{I} = \begin{bmatrix} \frac{\sqrt{3}}{2}i & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2}i \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix} \Rightarrow x + iy = 0.$$

Ennek az egyenlet(rendszer)nek a megoldása az  $y = t$  paraméterválasztással

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -it \\ t \end{bmatrix} = t \begin{bmatrix} -i \\ 1 \end{bmatrix}.$$

Tehát az  $\frac{1}{2} - \frac{\sqrt{3}}{2}i$  sajátértékhez tartozó sajátalteret a  $(-i, 1)$  sajátvektor feszíti ki.  $\square$

*A karakterisztikus egyenlet többszörös gyökei: az algebrai és a geometriai multiplicitás* Ha  $\lambda$  a karakterisztikus egyenlet  $k$ -szoros gyöke, vagy más szóval  $\lambda$  multiplicitása vagy *algebrai multiplicitása*  $k$ , akkor a  $\lambda$ -hoz tartozó sajátalter  $d$  dimenziójára  $1 \leq d \leq k$ . Ezt az állítást később bebizonyítjuk. A sajátalter dimenzióját szokták a  $\lambda$  sajátérték *geometriai multiplicitásának* nevezni. A 8.12. és a 8.13. példák olyan eseteket mutattak, amikor a sajátértékek algebrai és geometriai multiplicitása azonos, azaz minden sajátalter épp annyi dimenziós, amennyi a gyök (algebrai) multiplicitása. A következő feladat azt mutatja, hogy a sajátalter dimenziója kisebb is lehet.

8.15. PÉLDA (SAJÁTÉRTÉK ALGEBRAI ÉS GEOMETRIAI MULTIPLICITÁSA).  
Határozzuk meg az

$$\mathbf{A} = \begin{bmatrix} 4 & 1 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{bmatrix} \text{ és a } \mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

*mátrix sajátértékeit és azok algebrai és geometriai multiplicitását!*

**MEGOLDÁS.** Mivel  $\mathbf{A}$  háromszögmátrix, ezért karakterisztikus polinomja  $(4 - \lambda)^3$ , a 4 tehát háromszoros gyök, azaz algebrai multiplicitása 3. Mivel

$$\mathbf{A} - 4\mathbf{I} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

ezért az  $(\mathbf{A} - 4\mathbf{I})\mathbf{x} = \mathbf{0}$  egyenletrendszer az  $y = 0, z = 0$  alakot ölti, aminek megoldása

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} t \\ 0 \\ 0 \end{bmatrix} = t \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Eszerint  $\mathbf{A}$  sajátaltere 1-dimenziós, melyet az  $(1, 0, 0)$  vektor feszít ki. A  $\lambda = 4$  sajátérték geometriai multiplicitása tehát 1.

A  $\mathbf{B}$  mátrix karakterisztikus polinomja  $(1 - \lambda)^2(2 - \lambda)^2$ , ennek gyökei 1 és 2, és mindegyiknek kettő az algebrai multiplicitása. Meghatá-



rozzuk sajátaltérüket.  $\lambda = 1$  esetén

$$\mathbf{B} - \lambda \mathbf{I} = \mathbf{B} - \mathbf{I} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Az ehhez tartozó homogén egyenletrendszer megoldása:

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} s \\ t \\ 0 \\ 0 \end{bmatrix} = s \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

Tehát az altér dimenziója 2, vagyis a geometriai multiplicitás megegyezik az algebraival. Ha  $\lambda = 2$ , akkor

$$\mathbf{B} - \lambda \mathbf{I} = \mathbf{B} - 2\mathbf{I} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Az ehhez tartozó homogén egyenletrendszer megoldása:

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ t \\ 0 \end{bmatrix} = t \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

Tehát az altér dimenziója most 1, vagyis a geometriai multiplicitás kisebb, mint az algebrai.  $\square$

**Sajátértékek és a mátrix hatványai** A mátrixok függvényeinek számolása szoros kapcsolatban van a sajátértékekkel. E témában első lépés a mátrixhatványok sajátértékeinek és sajátvektorainak meghatározása.

**8.16. TÉTEL (MÁTRIX INVERTÁLHATÓSÁGA ÉS A 0 SAJÁTÉRTÉK).** Az  $\mathbf{A}$  mátrix pontosan akkor invertálható, ha a 0 nem sajátértéke.

**BIZONYÍTÁS.**  $\mathbf{A}$  pontosan akkor invertálható, ha  $\det(\mathbf{A}) \neq 0$ , de ez ekvivalens azzal, hogy  $\det(\mathbf{A} - 0\mathbf{I}) \neq 0$ , azaz 0 nem sajátértéke  $\mathbf{A}$ -nak.  $\square$

**8.17. TÉTEL (MÁTRIX HATVÁNYAINAK SAJÁTÉRTÉKEI ÉS SAJÁTVEKTORAI).** Ha  $\lambda$  az  $\mathbf{A}$  mátrix egy sajátértéke és  $\mathbf{x}$  egy hozzá tartozó sajátvektor, akkor bármely egész  $n$  esetén  $\lambda^n$  sajátértéke az  $\mathbf{A}^n$  mátrixnak és  $\mathbf{x}$  egy hozzá tartozó sajátvektor, amennyiben  $\lambda^n$  és  $\mathbf{A}^n$  is értelmezve van.

**BIZONYÍTÁS.**  $n = 0$  esetén  $\lambda^0 = 1$  és  $\mathbf{A}^0 = \mathbf{I}$ , és ekkor minden vektor az 1 sajátértékhez tartozó sajátvektor, tehát ekkor az állítás igaz.

Pozitív  $n$ -re indukcióval igazoljuk az állítást:  $n = 1$  esetén nyilván igaz,  $n = 2$  esetén:

$$\mathbf{A}^2 \mathbf{x} = \mathbf{A}(\mathbf{A} \mathbf{x}) = \mathbf{A}(\lambda \mathbf{x}) = \lambda(\mathbf{A} \mathbf{x}) = \lambda(\lambda \mathbf{x}) = \lambda^2 \mathbf{x}.$$

Hasonlóan kapjuk, hogy ha  $n = k - 1$  esetén már igaz az állítás, akkor  $n = k$  esetén is:

$$\mathbf{A}^k \mathbf{x} = \mathbf{A}(\mathbf{A}^{k-1} \mathbf{x}) = \mathbf{A}(\lambda^{k-1} \mathbf{x}) = \lambda^{k-1}(\mathbf{A} \mathbf{x}) = \lambda^{k-1}(\lambda \mathbf{x}) = \lambda^k \mathbf{x}.$$

Ha  $\mathbf{A}$  invertálható, akkor

$$\mathbf{A} \mathbf{x} = \lambda \mathbf{x}, \text{ amiből } \frac{1}{\lambda} \mathbf{x} = \mathbf{A}^{-1} \mathbf{x}, \text{ azaz } \lambda^{-1} \mathbf{x} = \mathbf{A}^{-1} \mathbf{x}.$$

Végül negatív kitevők esetén:

$$\mathbf{A}^k \mathbf{x} = \lambda^k \mathbf{x}, \text{ amiből } \lambda^{-k} \mathbf{x} = \mathbf{A}^{-k} \mathbf{x}. \quad \square$$

**8.18. TÉTEL (MÁTRIX HATVÁNYAINAK HATÁSA).** Tegyük fel, hogy  $\lambda_1, \lambda_2, \dots, \lambda_k$  sajátértékei az  $n \times n$ -es  $\mathbf{A}$  mátrixnak, és hogy  $\mathbf{x}_1, \dots, \mathbf{x}_k$  hozzájuk tartozó sajátvektorok. Ha egy  $n$ -dimenziós  $\mathbf{v}$  vektor előáll e sajátvektorok lineáris kombinációjaként, azaz

$$\mathbf{v} = c_1 \mathbf{x}_1 + c_2 \mathbf{x}_2 + \dots + c_k \mathbf{x}_k,$$

akkor bármely egész  $m$  esetén

$$\mathbf{A}^m \mathbf{v} = c_1 \lambda_1^m \mathbf{x}_1 + c_2 \lambda_2^m \mathbf{x}_2 + \dots + c_k \lambda_k^m \mathbf{x}_k.$$

**BIZONYÍTÁS.** A bizonyítás magától értetődő, hisz

$$\begin{aligned} \mathbf{A}^m \mathbf{v} &= \mathbf{A}^m (c_1 \mathbf{x}_1 + c_2 \mathbf{x}_2 + \dots + c_k \mathbf{x}_k) \\ &= c_1 \mathbf{A}^m \mathbf{x}_1 + c_2 \mathbf{A}^m \mathbf{x}_2 + \dots + c_k \mathbf{A}^m \mathbf{x}_k \\ &= c_1 \lambda_1^m \mathbf{x}_1 + c_2 \lambda_2^m \mathbf{x}_2 + \dots + c_k \lambda_k^m \mathbf{x}_k. \end{aligned} \quad \square$$

► Sajnos az nem igaz, hogy minden mátrixhoz találunk  $n$  független sajátvektort, amelyek lineáris kombinációjaként minden vektor felírható, így e tétel csak a sajátvektorok lineáris kombinációjaként előálló vektorokról szól! E tétel azt mutatja, fontos kérdés annak eldöntése, hogy egy mátrix sajátvektoraiból mikor alkotható bázis.

**Speciális mátrixok sajátértékei** A mátrixok egyes különleges tulajdonságai a sajátértékek bizonyos tulajdonságát is befolyásolják.

8.19. TÉTEL (SPECIÁLIS MÁTRIXOK SAJÁTÉRTÉKE). Legyen  $\mathbf{A}$  egy  $n$ -edrendű valós mátrix. Ekkor

- a) ha  $\mathbf{A}$  szimmetrikus, akkor minden sajátértéke valós,
- b) ha  $\mathbf{A}$  ferdén szimmetrikus, akkor minden sajátértéke imaginárius,
- c) ha  $\mathbf{A}$  ortogonális, akkor minden sajátértékének 1 az abszolút értéke,
- d)  $\mathbf{A}$  pontosan akkor nilpotens, ha minden sajátértéke 0, azaz karakterisztikus polinomja  $\lambda^n$  alakú.

**BIZONYÍTÁS.** a), b) Legyen  $(\lambda, \mathbf{x})$  egy  $\mathbf{A}$ -hoz tartozó saját pár. Az  $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$  egyenlőség mindkét oldalát balról szorozzuk be  $\mathbf{x}$  adjungáltjával (konjugáltjának transzponáltjával):

$$\mathbf{x}^H \mathbf{A} \mathbf{x} = \mathbf{x}^H \lambda \mathbf{x} = \lambda |\mathbf{x}|^2.$$

Vegyük mindkét oldal adjungáltját (konjugáltjának transzponáltját), kihasználva hogy mivel  $\mathbf{A}$  valós, ezért  $\mathbf{A}^H = \mathbf{A}^T$ :

$$\mathbf{x}^H \mathbf{A}^T \mathbf{x} = \bar{\lambda} |\mathbf{x}|^2.$$

Legyen  $\lambda = a + ib$ . Ha  $\mathbf{A}$  szimmetrikus, azaz  $\mathbf{A}^T = \mathbf{A}$ , akkor  $\lambda = \bar{\lambda}$ , azaz  $a + ib = a - ib$ . Így  $\lambda$  imaginárius része 0, tehát  $\lambda$  valós. Ha  $\mathbf{A}$  ferdén szimmetrikus, azaz  $\mathbf{A}^T = -\mathbf{A}$ , akkor  $a + ib = -a + ib$ , azaz  $\lambda$  valós része 0, így  $\lambda$  imaginárius.

c) Ha  $\mathbf{A}$  ortogonális, akkor bármely  $\mathbf{x}$  vektorra  $|\mathbf{A}\mathbf{x}| = |\mathbf{x}|$ . Így ha  $\mathbf{x}$  sajátvektor, akkor  $|\mathbf{x}| = |\mathbf{A}\mathbf{x}| = |\lambda\mathbf{x}| = |\lambda||\mathbf{x}|$ , amiből  $|\lambda| = 1$ .

d) Ha  $\mathbf{A}^k = \mathbf{O}$ , és  $\lambda$  sajátértéke  $\mathbf{A}$ -nak, akkor  $\lambda^k$  sajátértéke az  $\mathbf{A}^k = \mathbf{O}$  mátrixnak, annak viszont csak a 0 sajátértéke, így  $\mathbf{A}$ -nak is minden sajátértéke 0. Az állítás megfordítása a [Cayley–Hamilton-tételből](#) következik, melyet hamarosan bizonyítunk. Eszerint minden mátrix kielégíti karakterisztikus polinomját, így ha a karakterisztikus polinom  $\lambda^n = 0$ , akkor  $\mathbf{A}^n = \mathbf{O}$ , vagyis  $\mathbf{A}$  nilpotens.  $\square$

Az előző tétel bizonyításából minimális változtatással megkapható a következő tétel bizonyítása is:

8.20. TÉTEL (SPECIÁLIS KOMPLEX MÁTRIXOK SAJÁTÉRTÉKEI). Ha az  $n$ -edrendű komplex  $\mathbf{A}$  mátrix

- a) önadjungált, akkor minden sajátértéke valós,
- b) ferdén önadjungált, akkor minden sajátértéke imaginárius,
- c) unitér, akkor minden sajátértékének 1 az abszolút értéke.

*Feladatok*

játértéke van, akkor sajátalterei merőlegesek egymásra.

8.1. Igazoljuk a ?? tétel c) állítását, mely szerint ha egy  $2 \times 2$ -es szimmetrikus valós mátrixnak két különböző sa-

## Hasonlóság, diagonalizálhatóság

Egy lineáris transzformáció sajátértékei és karakterisztikus polinomja megőrződnek a különféle bázisokban fölírt mátrixaira is. Olyan bázist keresünk, melyben mátrixa a legegyszerűbb alakú.

**Lineáris transzformációk sajátértékei** A sajátérték, sajátvektor, sajátaltér fogalma természetes módon átvihető lineáris leképezésekre is.

**8.21. DEFINÍCIÓ (LINEÁRIS TRANSZFORMÁCIÓ SAJÁTÉRTÉKE, SAJÁTVEKTORA).** Azt mondjuk, hogy a  $\lambda$  szám az  $L$  lineáris transzformáció sajátértéke, ha létezik olyan nemnulla  $\mathbf{x}$  vektor, melyre  $L\mathbf{x} = \lambda\mathbf{x}$ . Az ilyen  $\mathbf{x}$  vektorokat az  $L$  lineáris transzformáció  $\lambda$  sajátértékéhez tartozó sajátvektorainak nevezzük.

Ha a lineáris transzformáció  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$  vagy  $\mathbb{R}^3 \rightarrow \mathbb{R}^3$  leképezés, mely valamilyen egyszerű geometriai transzformációt valósít meg, akkor néha a transzformáció mátrixának ismerete nélkül is könnyen meghatározhatjuk a sajátértékeket és sajátvektorokat.

**8.22. PÉLDA (LINEÁRIS TRANSZFORMÁCIÓ SAJÁTÉRTÉKE, SAJÁTALTERE).** Adjuk meg – pusztán geometriai szemléletünkre hagyatkozva – az alábbi lineáris leképezések sajátértékeit és a hozzájuk tartozó sajátaltereket.

- a sík vektorainak tükrözése egy egyenesre (vagy pontjainak tükrözése egy origón átmenő egyenesre);
- a sík vektorainak merőleges vetítése egy egyenesre (vagy pontjainak merőleges vetítése egy origón átmenő egyenesre);
- a tér vektorainak elforgatása egy egyenes körül a  $\pi$  egész számú többszörösétől különböző szöggel;
- a tér vektorainak merőleges vetítése egy síkra;
- a tér vektorainak tükrözése egy síkra.

**MEGOLDÁS.** Az előző fejezetben, így a 7.7. állításban bizonyítottakhoz hasonlóan látható, hogy mindegyik feladatbeli transzformáció lineáris.

a) Egy egyenesre való tükrözés esetén csak az egyenessel párhuzamos és rá merőleges vektorok mennek saját konstansszorosukba, mégpedig az egyenessel párhuzamos vektorok saját magukba, a rá merőlegesek a saját ellentettjükbe. Tehát e transzformációnak az 1 sajátértékhez tartozó sajátaltère az egyenessel párhuzamos vektorokból, a  $-1$ -hez tartozó sajátaltère a rá merőleges vektorokból áll. A pontokra vonatkozó állítás a pontokba mutató helyvektorokkal adódik.

b) A sík merőleges vetítése egy egyenesre – hasonlóan az előző esethez – helyben hagyja az egyenessel párhuzamos vektorokat, és a  $0$ -vektorba viszi a rá merőlegeseket. Tehát az 1 sajátértékhez tartozó

sajátaltér az egyenessel párhuzamos vektorokból, a 0-hoz tartozó sajátaltér a rá merőleges vektorokból áll.

c) A tér egyenes körüli elforgatása a forgástengellyel párhuzamos vektorokat önmagukba viszi, és ha a forgatás szöge különbözik  $\pi$  egész számú többszöröseitől, semelyik másik vektort sem viszi a saját skalárszorosába. Így az egyetlen sajátérték az 1, amelyhez tartozó sajátaltér a forgástengellyel párhuzamos vektorokból áll.

d) A tér vektorainak merőleges vetítése egy síkra helyben hagyja a sík összes vektorát, míg a síkra merőleges vektorokat a 0 vektorba viszi, tehát a két sajátérték 1 és 0. Az 1-hez tartozó sajátaltér a sík vektoraiból, a 0-hoz tartozó sajátaltér a rá merőleges vektorokból áll.

e) E feladatot megoldottuk a 8.1. példában. A két sajátérték 1 és  $-1$ , az 1 sajátértékhez tartozó sajátaltér a sík vektoraiból, a  $-1$ -hez tartozó sajátaltér a rá merőleges vektorokból áll.  $\square$

Egy lineáris leképezéshez bázisonként más-más mátrix tartozhat, de a sajátértékeik mégis ugyanazok, hisz egy vektor képe csak a leképezéstől függ, nem a választott bázistól.

*Hasonló mátrixok sajátértékei* A 7.15. és a 7.16. tételekből tudjuk, hogy egy lineáris leképezéshez különböző bázisokban tartozó mátrixok hasonlóak. Ezen felül tudjuk azt is, hogy fontos mátrixtulajdonságok invariánsak a hasonlóságra. E paragrafusban e tulajdonságok körét fogjuk bővíteni.

**8.23. TÉTEL (SAJÁTÉRTÉKHEZ KAPCSOLÓDÓ INVARIÁNSOK).** *Ha  $\mathbf{A} \sim \mathbf{B}$ , akkor  $\mathbf{A}$  és  $\mathbf{B}$  karakterisztikus polinomja azonos, így sajátértékei, azok algebrai, sőt geometriai multiplicitásai is megegyeznek.*

**BIZONYÍTÁS.** A bizonyítás során föltesszük, hogy valamely invertálható  $\mathbf{C}$  mátrixszal  $\mathbf{A} = \mathbf{C}^{-1}\mathbf{B}\mathbf{C}$ . Ekkor

$$\begin{aligned}\mathbf{A} - \lambda\mathbf{I} &= \mathbf{C}^{-1}\mathbf{B}\mathbf{C} - \lambda\mathbf{C}^{-1}\mathbf{I}\mathbf{C} \\ &= \mathbf{C}^{-1}(\mathbf{B}\mathbf{C} - \lambda\mathbf{I}\mathbf{C}) \\ &= \mathbf{C}^{-1}(\mathbf{B} - \lambda\mathbf{I})\mathbf{C},\end{aligned}$$

azaz  $\mathbf{A} - \lambda\mathbf{I}$  és  $\mathbf{B} - \lambda\mathbf{I}$  is hasonlóak. A 7.16. tétel szerint hasonló mátrixok determinánsa megegyezik, így  $\det(\mathbf{A} - \lambda\mathbf{I}) = \det(\mathbf{B} - \lambda\mathbf{I})$ , azaz megegyeznek  $\mathbf{A}$  és  $\mathbf{B}$  karakterisztikus polinomjai is. Ez maga után vonja, hogy megegyeznek sajátértékeik, és azok (algebrai) multiplicitásai. A geometriai multiplicitások egyenlőségéhez elég belátni, hogy  $\mathbf{A} - \lambda\mathbf{I}$  és  $\mathbf{B} - \lambda\mathbf{I}$  nullterének dimenziója megegyezik, azt viszont ugyancsak a 7.16. tételben igazoltuk.  $\square$

► Ismerjük a polinom együtthatói és gyökei közti összefüggéseket, így a sajátértékeknek a polinom együtthatóit adó függvényei is invari-

ánsak. Például a harmadfokú esetben:

$$\begin{aligned} & (\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda) \\ &= -\lambda^3 + (\lambda_1 + \lambda_2 + \lambda_3)\lambda^2 - (\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1)\lambda + \lambda_1\lambda_2\lambda_3. \end{aligned}$$

Tehát  $\lambda_1 + \lambda_2 + \lambda_3$ ,  $\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1$  és  $\lambda_1\lambda_2\lambda_3$  invariáns mennyiségek. Általában egy  $n$ -edrendű  $\mathbf{A}$  mátrixra invariánsak a sajátértékek következő függvényei:

$$\begin{aligned} e_1(\lambda_1, \lambda_2, \dots, \lambda_n) &= \sum_{1 \leq i \leq n} \lambda_i = \text{trace}(\mathbf{A}), \\ e_2(\lambda_1, \lambda_2, \dots, \lambda_n) &= \sum_{1 \leq i < j \leq n} \lambda_i \lambda_j, \\ e_3(\lambda_1, \lambda_2, \dots, \lambda_n) &= \sum_{1 \leq i < j < k \leq n} \lambda_i \lambda_j \lambda_k, \\ &\vdots \\ e_n(\lambda_1, \lambda_2, \dots, \lambda_n) &= \lambda_1 \lambda_2 \dots \lambda_n = \det(\mathbf{A}). \end{aligned}$$

Az itt felsorolt  $e_1, e_2, \dots, e_n$  függvényeket *elemi szimmetrikus polinomoknak* nevezzük. A sajátértékek utolsónak említett függvénye a determináns, melynek invariáns voltát megmutattuk a 7.16. tételben. Hamarosan igazoljuk, hogy minden mátrix hasonló egy felső háromszög-mátrixhoz, melynek főátlójában a sajátértékek vannak. Így hasonló mátrixok nyoma megegyezik, hisz a sajátértékek összegével egyenlő.

► Fontos következménye e tételnek, hogy van értelme *lineáris transzformáció karakterisztikus polinomjáról*, beszélni (legalábbis véges dimenziós esetben, pl.  $\mathbb{R}^n \rightarrow \mathbb{R}^n$  vagy  $\mathbb{C}^n \rightarrow \mathbb{C}^n$  transzformációk esetén).

**Mátrixok diagonalizálása és sajátfelbontása** Igen fontos kérdés, hogy egy adott lineáris leképezés sajátvektoraiból kiválasztható-e egy bázis. Ebben a bázisban ugyanis mátrixa – mint azt bizonyítani fogjuk – diagonális alakot ölt.

**8.24. DEFINÍCIÓ (DIAGONALIZÁLHATÓSÁG).** Az  $n \times n$ -es  $\mathbf{A}$  mátrix diagonalizálható, ha hasonló egy diagonális mátrixhoz, azaz ha létezik egy olyan diagonális  $\mathbf{\Lambda}$  és egy invertálható  $\mathbf{C}$  mátrix, hogy

$$\mathbf{\Lambda} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}. \quad (8.4)$$

**8.25. TÉTEL (DIAGONALIZÁLHATÓSÁG SZÜKSÉGES ÉS ELÉGSÉGES FELTÉTELE).** Az  $n \times n$ -es  $\mathbf{A}$  mátrix pontosan akkor diagonalizálható, azaz pontosan akkor létezik olyan  $\mathbf{C}$  mátrix, melyre  $\mathbf{C}^{-1}\mathbf{A}\mathbf{C}$  diagonális, ha  $\mathbf{A}$ -nak van  $n$  lineárisan független sajátvektora. Ekkor a diagonális mátrix az  $\mathbf{A}$  sajátértékeiből,  $\mathbf{C}$  a sajátvektoraiból áll.

**BIZONYÍTÁS.** Ha  $\mathbf{A}$  hasonló egy diagonális mátrixhoz, azaz van olyan  $\mathbf{C}$  mátrix, hogy  $\mathbf{\Lambda} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$  diagonális, akkor  $\mathbf{C}$ -vel balról szorozva a  $\mathbf{C}\mathbf{\Lambda} = \mathbf{A}\mathbf{C}$  egyenlőséget kapjuk. Ha  $\mathbf{C} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_n]$  és  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ , akkor

$$[\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_n] \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} = \mathbf{A}[\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_n]. \quad (8.5)$$

Itt a bal oldali mátrix  $i$ -edik oszlopa  $\lambda_i \mathbf{x}_i$ , a jobb oldali mátrixé  $\mathbf{A}\mathbf{x}_i$ . Ezek megegyeznek, azaz  $\mathbf{A}\mathbf{x}_i = \lambda_i \mathbf{x}_i$ , tehát  $\mathbf{x}_i$  a  $\lambda_i$  sajátértékhez tartozó sajátvektor. Mivel  $\mathbf{C}$  invertálható, ezért oszlopvektorai függetlenek, ami bizonyítja az állításunk egyik felét. Tegyük most fel, hogy van  $\mathbf{A}$ -nak  $n$  független sajátvektora. Képezzünk a sajátértékekből egy  $\mathbf{\Lambda}$  diagonális mátrixot, úgy hogy a  $\mathbf{C}$  mátrix  $i$ -edik oszlopába kerülő  $\mathbf{x}_i$  vektorhoz tartozó  $\lambda_i$  sajátérték a  $\mathbf{\Lambda}$  mátrix  $i$ -edik oszlopába kerüljön. Mivel  $\lambda_i \mathbf{x}_i = \mathbf{A}\mathbf{x}_i$ , ezért fennáll a (8.5) összefüggés, azaz  $\mathbf{A}$  hasonló  $\mathbf{A}$ -hoz.  $\square$

► A  $\mathbf{\Lambda} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$  átírható

$$\mathbf{A} = \mathbf{C}\mathbf{\Lambda}\mathbf{C}^{-1} \quad (8.6)$$

alakba, amit az  $\mathbf{A}$  mátrix *sajátfelbontásának* nevezünk.

8.26. PÉLDA (MÁTRIX DIAGONALIZÁLÁSA). *Diagonalizálható-e a 8.12. példabeli*

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

*mátrix?*

**MEGOLDÁS.** Az  $\mathbf{A}$  mátrix sajátértégeit és sajátvektorait meghatároztuk a 8.12. példában. Mivel  $\lambda_1 = 0$ ,  $\lambda_2 = \lambda_3 = 2$ , a hozzájuk tartozó sajátvektorok  $(1, 0, 0)$ ,  $(1/2, 1, 0)$  és  $(1/2, 0, 1)$ , és ezek a vektorok lineárisan függetlenek, ezért  $\mathbf{A}$  hasonló a  $\mathbf{\Lambda}$  diagonális mátrixhoz, ahol

$$\mathbf{\Lambda} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad \text{és} \quad \mathbf{C} = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Ez könnyen igazolható a  $\mathbf{C}\mathbf{\Lambda} = \mathbf{A}\mathbf{C}$  összefüggés ellenőrzésével:

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad \square$$



*Bal sajátvektorok és a sajátfelbontás diadikus alakja\** Az  $\mathbf{Ax} = \lambda\mathbf{x}$  egyenlet helyett vizsgálhatjuk az  $\mathbf{y}^T\mathbf{A} = \lambda\mathbf{y}^T$  egyenletet is.

Az  $\mathbf{y}^T\mathbf{A} = \lambda\mathbf{y}^T$  egyenlet  $\mathbf{y}^T \neq \mathbf{0}^T$  feltételnek megfelelő sorvektora- it az  $\mathbf{A}$  mátrix *bal sajátvektorainak* nevezzük. E megfogalmazásban a sajátvektorokat szokás jobb sajátvektoroknak is nevezni.

A fenti egyenletet transzponálva kapjuk, hogy

$$\mathbf{A}^T\mathbf{y} = \lambda\mathbf{y},$$

vagyis a bal sajátvektorok a transzponált sajátvektorainak transzponáltjai. Mivel  $\det(\mathbf{A} - \lambda\mathbf{I}) = \det((\mathbf{A} - \lambda\mathbf{I})^T) = \det(\mathbf{A}^T - \lambda\mathbf{I})$ , azaz  $\mathbf{A}$  és  $\mathbf{A}^T$  karakterisztikus polinomja azonos, ezért a bal és jobb sajátvektorokhoz tartozó sajátértékek azonosak. A bal és jobb sajátvektorok azonban általában nem egyeznek meg.

Ha  $\mathbf{A}$  diagonalizálható, azaz  $\mathbf{A} = \mathbf{C}^{-1}\mathbf{AC}$ , akkor a 8.25. tételhez hasonlóan a  $\mathbf{AC}^{-1} = \mathbf{C}^{-1}\mathbf{A}$  mátrixegyenletből azt kapjuk, hogy  $\mathbf{C}^{-1}$  sorvektorai  $\mathbf{A}$  bal sajátvektorai. Jelölje ugyanis a  $\mathbf{C}^{-1}$  mátrix  $i$ -edik sorvektorát  $\mathbf{y}_i^T$ , ekkor

$$\begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{y}_1^T \\ \mathbf{y}_2^T \\ \vdots \\ \mathbf{y}_n^T \end{bmatrix} = \begin{bmatrix} \mathbf{y}_1^T \\ \mathbf{y}_2^T \\ \vdots \\ \mathbf{y}_n^T \end{bmatrix} \mathbf{A} \quad (8.7)$$

tehát  $\lambda_i\mathbf{y}_i^T = \mathbf{y}_i^T\mathbf{A}$  ( $i = 1, 2, \dots, n$ ), vagyis  $\mathbf{y}_i^T$  valóban bal sajátvektor. Ekkor a sajátfelbontás azonnal diádok összegévé bontható:

$$\begin{aligned} \mathbf{A} &= \mathbf{CAC}^{-1} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_n] \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{y}_1^T \\ \mathbf{y}_2^T \\ \vdots \\ \mathbf{y}_n^T \end{bmatrix} \\ &= \lambda_1\mathbf{x}_1\mathbf{y}_1^T + \lambda_2\mathbf{x}_2\mathbf{y}_2^T + \dots + \lambda_n\mathbf{x}_n\mathbf{y}_n^T \end{aligned} \quad (8.8)$$

Ezt nevezzük a *sajátfelbontás diadikus alakjának*.

**8.27. PÉLDA** (SAJÁTFELBONTÁS DIADIKUS ALAKJA ÉS A BAL SAJÁTVEKTOROK). Határozzuk meg a 8.26. és a 8.12. példában szereplő

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

mátrix bal sajátvektorait, sajátfelbontását és annak diadikus alakját.

**MEGOLDÁS.** A bal sajátvektorok megegyeznek a transzponált sajátvektorainak transzponáltjaival. Így a bal sajátvektorok a szokásos technikával kiszámolhatók. Ha azonban a sajátfelbontást is kiszámoljuk,

a bal sajátvektorok kiolvashatók a  $\mathbf{C}^{-1}$  mátrixból is. A 8.26. példában meghatároztuk a  $\mathbf{C}$  mátrixot, annak inverzét kiszámolva kapjuk a sajátfelbontást:

$$\mathbf{A} = \mathbf{C}\mathbf{\Lambda}\mathbf{C}^{-1} = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Ebből a diádokká való átirás a következő:

$$\begin{aligned} \mathbf{A} &= 0 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} + 2 \begin{bmatrix} \frac{1}{2} \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} + 2 \begin{bmatrix} \frac{1}{2} \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}. \end{aligned} \quad \square$$

*Diagonalizálható mátrixok polinomjai és a Cayley–Hamilton-tétel\** Belátjuk, hogy ha egy mátrixot behelyettesítünk saját karakterisztikus polinomjába, nullmátrixot kapunk.

Ha  $p(x) = c_n x^n + c_{n-1} x^{n-1} + \dots + c_1 x + c_0$  egy tetszőleges polinom, akkor értelmezhető e polinomnak egy tetszőleges négyzetes mátrixban fölvetett értéke, vagyis értelmezhető négyzetes mátrix polinomja a következő képlettel:

$$p(\mathbf{A}) = c_n \mathbf{A}^n + c_{n-1} \mathbf{A}^{n-1} + \dots + c_1 \mathbf{A} + c_0 \mathbf{I}.$$

Ha az  $\mathbf{A}$  mátrix diagonalizálható, és sajátfelbontása  $\mathbf{A} = \mathbf{C}\mathbf{\Lambda}\mathbf{C}^{-1}$ , akkor  $\mathbf{A}^2 = \mathbf{C}\mathbf{\Lambda}\mathbf{C}^{-1}\mathbf{C}\mathbf{\Lambda}\mathbf{C}^{-1} = \mathbf{C}\mathbf{\Lambda}^2\mathbf{C}^{-1}$ . Hasonlóképp tetszőleges nemnegatív  $k$  egészre

$$\mathbf{A}^k = \mathbf{C}\mathbf{\Lambda}^k\mathbf{C}^{-1}.$$

Eszerint bármely  $p(x)$  polinomra  $p(\mathbf{A}) = \mathbf{C}p(\mathbf{\Lambda})\mathbf{C}^{-1}$ . Másrészt bármely diagonális mátrix polinomja a diagonális elemek polinomjával számolható, azaz

$$p(\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)) = \text{diag}(p(\lambda_1), p(\lambda_2), \dots, p(\lambda_n)).$$

Tehát bizonyítottuk a következő állítást:

**8.28. ÁLLÍTÁS (DIAGONALIZÁLHATÓ MÁTRIX POLINOMJA).** Legyen  $\mathbf{A} = \mathbf{C}\mathbf{\Lambda}\mathbf{C}^{-1}$ , ahol  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ , és  $p(x)$  egy tetszőleges polinom. Ekkor

$$p(\mathbf{A}) = \mathbf{C} \begin{bmatrix} p(\lambda_1) & 0 & \dots & 0 \\ 0 & p(\lambda_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & p(\lambda_n) \end{bmatrix} \mathbf{C}^{-1}.$$

Jelölje most  $\chi_A$  az  $A$  mátrix karakterisztikus polinomját! Tehát ha  $\lambda$  az  $A$  sajátértéke, akkor  $\chi_A(\lambda) = 0$ . Innen azonnal adódik, hogy  $\chi_A(A) = O$ , ugyanis

$$\begin{aligned}\chi_A(A) &= C \chi_A(\Lambda) C^{-1} \\ &= C \operatorname{diag}(\chi_A(\lambda_1), \chi_A(\lambda_2), \dots, \chi_A(\lambda_n)) C^{-1} \\ &= O.\end{aligned}$$

Ez az állítás nem csak diagonalizálható mátrixokra igaz, hanem általánosan is. Erre azonnal mutatunk egy egyszerű, de rossz bizonyítást! Mivel a karakterisztikus polinom  $\chi_A(\lambda) = \det(A - \lambda I)$ , gondolhatnánk, hogy akkor  $\chi_A(A) = \det(A - AI) = \det(O) = 0$ . Ez azonban hibás érvelés:  $\lambda$  helyébe egy mátrixban mátrixokat helyettesítettünk, ráadásul a jobb oldalon skalár, a bal oldalon mátrix áll!

**8.29. TÉTEL (CAYLEY–HAMILTON-TÉTEL).** *Ha  $A$  egy tetszőleges négyzetes mátrix, melynek karakterisztikus polinomja  $\chi_A$ , akkor  $\chi_A(A) = O$ .*

**BIZONYÍTÁS.** Legyen  $B = A - \lambda I$ . Az  $A$  karakterisztikus polinomja így

$$\det B = \chi_A(\lambda) = (-1)^n \lambda^n + p_{n-1} \lambda^{n-1} + \dots + p_1 \lambda + p_0. \quad (8.9)$$

$B$  bármely eleméhez tartozó előjeles aldetermináns  $\lambda$  egy legfőljebb  $n-1$ -edfokú polinomja, így léteznek olyan konstans elemű  $C_0, C_1, \dots, C_{n-1}$  mátrixok, hogy

$$\operatorname{adj} B = \lambda^{n-1} C_{n-1} + \dots + \lambda C_1 + C_0. \quad (8.10)$$

A mátrix inverzéről szóló 6.28. tételbeli (6.3) képlet szerint  $\det(B)I = B \operatorname{adj}(B)$ . Ennek jobb oldalán elvégezzük a (8.10) szerinti behelyettesítést:

$$\begin{aligned}B \operatorname{adj} B &= (A - \lambda I) \left( \sum_{k=0}^{n-1} \lambda^k C_k \right) \\ &= AC_0 + \left( \sum_{k=1}^{n-1} \lambda^k (AC_k - C_{k-1}) \right) - \lambda^n C_{n-1}.\end{aligned}$$

Felírjuk a  $\det(B)I = B \operatorname{adj}(B)$  egyenlőség bal és jobb oldalán álló együtthatók egyenlőségét, és mindegyiket beszorozzuk  $A$  megfelelő hatványával az alábbiak szerint:

$$\begin{array}{ll}(-1)^n I = -C_{n-1} & \cdot A^n \\ p_{n-1} I = AC_{n-1} - C_{n-2} & \cdot A^{n-1} \\ \vdots & \vdots \\ p_2 I = AC_2 - C_1 & \cdot A^2 \\ p_1 I = AC_1 - C_0 & \cdot A \\ p_0 I = AC_0 & \cdot A^0.\end{array}$$

A beszorzás után kapott egyenlőségeket összeadva kapjuk, hogy

$$(-1)^n \mathbf{A}^n + p_{n-1} \mathbf{A}^{n-1} + \dots + p_1 \mathbf{A} + p_0 \mathbf{I} = \mathbf{O},$$

a jobb oldalon ugyanis teleszkópszerűen minden tag kiesik. Ezzel tehát bizonyítottuk, hogy  $\chi_{\mathbf{A}}(\mathbf{A}) = \mathbf{O}$ .  $\square$

**Különböző sajátértékek sajátalterei** A diagonalizálhatóság fontos voltára tekintettel érdemes további feltételeket gyűjteni, melyek könnyen ellenőrizhetők.

Több elégséges feltétel származtatható a következő tételből:

**8.30. TÉTEL (KÜLÖNBÖZŐ SAJÁTÉRTÉKEK SAJÁTVEKTORAI).** Ha  $\lambda_1, \lambda_2, \dots, \lambda_k$  különböző sajátértékei az  $n \times n$ -es  $\mathbf{A}$  mátrixnak, akkor a hozzájuk tartozó  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$  sajátvektorok lineárisan függetlenek.

**BIZONYÍTÁS.** Indirekt módon bizonyítunk. Tegyük fel, hogy e vektorok lineárisan összefüggők. Ekkor van a vektorok közt olyan, amely csak a kisebb indexűek lineáris függvénye. Legyen ezek közül a legkisebb indexű  $\mathbf{x}_i$ , azaz

$$\mathbf{x}_i = c_1 \mathbf{x}_1 + \dots + c_{i-1} \mathbf{x}_{i-1}, \quad (8.11)$$

de az  $i$ -nél kisebb indexű vektorok már lineárisan függetlenek. Szorozzuk meg az egyenlőség mindkét oldalát balról az  $\mathbf{A}$  mátrixszal:

$$\mathbf{A}\mathbf{x}_i = \mathbf{A}(c_1 \mathbf{x}_1 + \dots + c_{i-1} \mathbf{x}_{i-1}) = c_1 \mathbf{A}\mathbf{x}_1 + \dots + c_{i-1} \mathbf{A}\mathbf{x}_{i-1},$$

majd használjuk ki, hogy e vektorok sajátvektorok:

$$\lambda_i \mathbf{x}_i = c_1 \lambda_1 \mathbf{x}_1 + \dots + c_{i-1} \lambda_{i-1} \mathbf{x}_{i-1}. \quad (8.12)$$

Ezután a (8.11) egyenlet mindkét oldalát  $\lambda_i$ -vel szorozva kapjuk, hogy

$$\lambda_i \mathbf{x}_i = c_1 \lambda_i \mathbf{x}_1 + \dots + c_{i-1} \lambda_i \mathbf{x}_{i-1}. \quad (8.13)$$

Végül a (8.13) egyenletből a (8.12) egyenletet kivonva kapjuk, hogy

$$\mathbf{0} = c_1 (\lambda_i - \lambda_1) \mathbf{x}_1 + \dots + c_{i-1} (\lambda_i - \lambda_{i-1}) \mathbf{x}_{i-1},$$

Mivel az  $\mathbf{x}_1, \dots, \mathbf{x}_{i-1}$  vektorok már lineárisan függetlenek, és a  $\lambda_1, \dots, \lambda_i$  értékek különbözőek, ezért  $c_1 = \dots = c_{i-1} = 0$ . Eszerint

$$\mathbf{x}_i = 0\mathbf{x}_1 + \dots + 0\mathbf{x}_{i-1} = \mathbf{0},$$

ami ellentmondás, hisz  $\mathbf{x}_i$  sajátvektor, tehát nem lehet a  $\mathbf{0}$ . Ez bizonyítja az indirekt feltevés helytelen voltát, azaz igazolja állításunkat.  $\square$

► Szokás úgy fogalmazni, hogy a különböző sajátértékekhez tartozó sajátalterek lineárisan függetlenek, hisz bárhogyan választunk mindegyikükből egy-egy nemzérus vektort, azok lineárisan függetlenek lesznek.

- Másik fontos következménye e tételnek, hogy ha különböző sajátértékekhez tartozó sajátalterek mindegyikéből lineárisan független vektorokat választunk, akkor még ezek egyesítése is lineárisan független lesz. Ha ugyanis lineárisan összefüggnének, akkor az egy altérbe eső vektorok lineáris kombinációit összevonva egyetlen vektorra, minden sajátértékhez egy-egy sajátvektort kapnánk, melyek összefüggők lennének, ami a fenti tételnek ellentmond.
- Speciálisan az is igaz, hogy különböző sajátértékekhez tartozó sajátalterek mindegyikéből egy bázist választva, azok egyesítése is lineárisan független vektorrendszert ad.

**8.31. KÖVETKEZMÉNY (KÜLÖNBÖZŐ SAJÁTÉRTÉKEK ÉS A DIAGONALIZÁLHATÓSÁG).** *Ha az  $n$ -edrendű  $A$  mátrixnak  $n$  darab különböző sajátértéke van, akkor diagonalizálható.*

**BIZONYÍTÁS.** A 8.30. tétel szerint  $n$  különböző sajátértékhez  $n$  független sajátvektor tartozik, ami a 8.25. tétel szerint épp azt jelenti, hogy a mátrix diagonalizálható.  $\square$

Végezetül felsorolunk néhány mátrixosztályt, melyekbe tartozó mátrixok mindegyike egyformán viselkedik a diagonalizálhatóságra nézve:

- a) Egy valós  $n$ -edrendű mátrix nem diagonalizálható a valós mátrixok körében, ha karakterisztikus egyenletének vannak nem valós gyökei, mert a diagonalizált alakban volnának nem valós számok. Például a

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

mátrix a valósok fölött nem diagonalizálható, de a komplexek fölött igen (ld. a ?? feladatot).

- b) Nem diagonalizálhatók a nilpotens mátrixok, például a

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

mátrix (?? feladat).

- c) Diagonalizálható minden szimmetrikus mátrix, sőt, sajátvektorai-ból ortonormált bázis is kiválasztható. Ezt hamarosan belátjuk (9.2. tétel).

**8.32. PÉLDA (DIAGONALIZÁLHATÓSÁG MEGÁLLAPÍTÁSA).** *Döntsük el,*

hogy az alábbi mátrixok közül melyik diagonalizálható valós mátrixként!

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 1 & -4 \\ 4 & 1 \end{bmatrix}, \mathbf{C} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}, \mathbf{D} = \begin{bmatrix} 6 & 9 \\ 9 & 6 \end{bmatrix}.$$

**MEGOLDÁS.** Az  $\mathbf{A}$  mátrix nilpotens, mert  $\mathbf{A}^2 = \mathbf{O}$ , így nem diagonalizálható. A  $\mathbf{B}$  mátrixnak vannak nem valós sajátértékei, így a valósok fölött nem diagonalizálható, de a komplexek fölött igen, mert két különböző sajátértéke van. A  $\mathbf{C}$  mátrixnak különbözőek a sajátértékei, és mind valósak (1, 4, 6), így diagonalizálható. A  $\mathbf{D}$  mátrix szimmetrikus, tehát diagonalizálható.  $\square$

**Sajátértékek multiplicitása és a diagonalizálhatóság\*** A sajátértékek algebrai és geometriai multiplicitása, valamint a diagonalizálhatóság közt egyszerű, de fontos összefüggés van. Nevezetesen a geometriai multiplicitás sosem nagyobb az algebrainál, másrészt a diagonalizálhatóság ekvivalens azzal, hogy a geometriai és algebrai multiplicitások minden sajátérték esetén megegyeznek.

**8.33. TÉTEL (ALGEBRAI ÉS GEOMETRIAI MULTIPLICITÁS KAPCSOLATA).** Egy mátrix valamely sajátértékének geometriai multiplicitása nem lehet nagyobb az algebrai multiplicitásánál.

**BIZONYÍTÁS.** Az  $\mathbf{A}$  mátrix egy  $\mu$  sajátértékének geometriai multiplicitását jelölje  $g$ . Ez azt jelenti, hogy  $\mathbf{A} - \mu\mathbf{I}$  nullterének  $g$  a dimenziója. Legyen egy bázisa  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_g\}$ . Egészítsük ki e bázist az egész tér bázisává az  $\mathbf{x}_{g+1}, \dots, \mathbf{x}_n$  vektorokkal. E független vektorokból képzett  $\mathbf{C} = [\mathbf{x}_1 \dots \mathbf{x}_g | \mathbf{x}_{g+1} \dots \mathbf{x}_n]$  mátrix invertálható. Írjuk  $\mathbf{C}$ -t blokkmátrix alakba:  $\mathbf{X}$  legyen az első  $g$  oszlopból álló blokk,  $\mathbf{Y}$  a maradék, azaz  $\mathbf{C} = [\mathbf{X} | \mathbf{Y}]$ . Mivel  $\mathbf{X}$  oszlopai a  $\mu$ -höz tartozó sajátvektorok, ezért  $\mathbf{AX} = \mu\mathbf{X}$ . A  $\mathbf{C}^{-1}$  inverzet első  $g$  sora után bontsuk blokkokra:

$$\mathbf{C}^{-1} = \begin{bmatrix} \mathbf{Z} \\ \mathbf{W} \end{bmatrix}.$$

Írjuk fel az  $\mathbf{I} = \mathbf{C}^{-1}\mathbf{C}$  összefüggést blokkmátrix alakban:

$$\begin{bmatrix} \mathbf{I}_g & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_{n-g} \end{bmatrix} = \begin{bmatrix} \mathbf{Z} \\ \mathbf{W} \end{bmatrix} \begin{bmatrix} \mathbf{X} & \mathbf{Y} \end{bmatrix} = \begin{bmatrix} \mathbf{ZX} & \mathbf{ZY} \\ \mathbf{WX} & \mathbf{WY} \end{bmatrix}.$$

Innen leolvasható, hogy  $\mathbf{WX} = \mathbf{O}$ ,  $\mathbf{ZY} = \mathbf{O}$ ,  $\mathbf{ZX} = \mathbf{I}_g$ ,  $\mathbf{WY} = \mathbf{I}_{n-g}$ . Ezeket felhasználva kapjuk, hogy

$$\mathbf{C}^{-1}\mathbf{AC} = \begin{bmatrix} \mathbf{Z} \\ \mathbf{W} \end{bmatrix} \mathbf{A} \begin{bmatrix} \mathbf{X} & \mathbf{Y} \end{bmatrix} = \begin{bmatrix} \mathbf{ZAX} & \mathbf{ZAY} \\ \mathbf{WAX} & \mathbf{WAY} \end{bmatrix} = \begin{bmatrix} \mu\mathbf{I}_g & \mathbf{ZAY} \\ \mathbf{O} & \mathbf{WAY} \end{bmatrix},$$

ugyanis  $\mathbf{ZAX} = \mathbf{Z}\mu\mathbf{X} = \mu\mathbf{ZX} = \mu\mathbf{I}_g$ , és  $\mathbf{WAX} = \mu\mathbf{WX} = \mathbf{O}$ . Az így kapott mátrix karakterisztikus polinomja

$$\begin{vmatrix} \mu\mathbf{I}_g - \lambda\mathbf{I}_g & \mathbf{ZAY} \\ \mathbf{O} & \mathbf{WAY} - \lambda\mathbf{I}_{n-g} \end{vmatrix},$$

ami a 6.30. tétel szerint  $(\mu - \lambda)^g \det(\mathbf{WAY} - \lambda\mathbf{I}_{n-g})$ . Ez pedig azt jelenti, hogy  $\mathbf{C}^{-1}\mathbf{AC}$  és ezzel együtt  $\mathbf{A}$  karakterisztikus polinomjának  $\mu$  legalább  $g$ -szeres algebrai multiplicitású gyöke.  $\square$

**8.34. TÉTEL (DIAGONALIZÁLHATÓSÁG ÉS A GEOMETRIAI MULTIPLICITÁS).** Egy  $n$ -edrendű négyzetes mátrix pontosan akkor diagonalizálható, ha a sajátértékeihez tartozó geometriai multiplicitások összege  $n$ .

**BIZONYÍTÁS.**  $(\Rightarrow)$  Ha a mátrix diagonalizálható, akkor egy adott sajátértékhez tartozó sajátaltér dimenziója megegyezik a sajátérték geometriai multiplicitásával. A geometriai multiplicitások összege tehát épp  $n$ , hisz egyetlen sajátvektor sem lehet két sajátaltérben.

$(\Leftarrow)$  A geometriai multiplicitás nem nagyobb az algebrainál, az algebraiak összege pedig legfeljebb  $n$  (komplex esetben pontosan  $n$ , valós esetben lehet  $n$ -nél kisebb is, ha a karakterisztikus polinomnak vannak nem valós gyökei). Így ha a geometriai multiplicitások összege  $n$ , akkor minden sajátaltérből kiválasztva egy bázist, és véve ezek egyesítését, egy  $n$  sajátvektorból álló független vektorrendszert kapunk (ld. a 8.30. tétel utáni megjegyzéseket). Így tehát a mátrix diagonalizálható.  $\square$

► E tétel elegáns megfogalmazása így hangzik: egy  $\mathbb{T}$  test fölötti  $n$ -edrendű mátrix pontosan akkor diagonalizálható, ha a  $\mathbb{T}^n$  tér előáll sajátaltérekének direkt összegeként.

► A tétel lineáris leképezésekre is kimondható: az  $A : \mathbb{T}^n \rightarrow \mathbb{T}^n$  lineáris leképezés pontosan akkor diagonalizálható, ha sajátaltérei dimenziójának összege  $n$ .

**8.35. PÉLDA (LINEÁRIS TRANSZFORMÁCIÓ DIAGONALIZÁLÁSA).** Az alábbi lineáris leképezésekhez keressünk – pusztán geometriai szemléletünkre hagyatkozva – olyan bázist, melyben a mátrixuk diagonális. Használjuk fel a 8.22. példa eredményeit.

- a sík vektorainak tükrözése egy egyenesre (vagy pontjainak tükrözése egy origón átmenő egyenesre);
- a sík vektorainak merőleges vetítése egy egyenesre (vagy pontjainak merőleges vetítése egy origón átmenő egyenesre);
- a tér vektorainak elforgatása egy egyenes körül a  $\pi$  egész számú többszörösétől különböző szöggel;
- a tér vektorainak merőleges vetítése egy síkra;
- a tér vektorainak tükrözése egy síkra.

**MEGOLDÁS.** A 8.22. példában meghatároztuk e leképezések sajátaltérét. Ezeket használjuk a következőkben.

a) Az egyenes – melyre tükrözünk – egyik irányvektora legyen  $\mathbf{a}$ , egy rá merőleges nemnulla vektor legyen  $\mathbf{b}$ . Ekkor  $T\mathbf{a} = \mathbf{a}$  és  $T\mathbf{b} = -\mathbf{b}$ , ahol  $T$  a tükröző lineáris leképezés. Ennek az  $\{\mathbf{a}, \mathbf{b}\}$  bázisban a mátrixa

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

b) Az egyenes – melyre vetítünk – egyik irányvektora legyen  $\mathbf{a}$ , egy rá merőleges nemnulla vektor legyen  $\mathbf{b}$ . Ekkor  $P\mathbf{a} = \mathbf{a}$  és  $P\mathbf{b} = \mathbf{0}$ , ahol  $P$  a vetítő lineáris leképezés. Ennek az  $\{\mathbf{a}, \mathbf{b}\}$  bázisban a mátrixa

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

c) E leképezésnek nincs valós diagonális mátrixa, mert csak egyetlen valós sajátaltér van, és az csak 1-dimenziós: ez a tengely irányvektora által kifeszített altér. A forgástengelyre merőleges sík ugyan nem sajátaltér, de a forgatás önmagába viszi (az ilyet nevezik *invariáns altérnek*), így ennek bázisával egy „diagonálshoz közeli” alakot kaphatunk. Ha a forgás tengelyének egy irányvektora  $\mathbf{a}$ , a rá merőleges sík egy ortonormált bázisa  $\{\mathbf{b}, \mathbf{c}\}$ , ahol a  $\mathbf{b}$  vektor  $\pi/2$  radiánnal való elforgatottja épp  $\mathbf{c}$ , akkor az  $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$  bázisban a forgató  $F$  leképezés mátrixa

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix},$$

ugyanis  $F\mathbf{a} = \mathbf{a}$ ,  $F\mathbf{b} = \cos \alpha \mathbf{b} + \sin \alpha \mathbf{c}$ ,  $F\mathbf{c} = -\sin \alpha \mathbf{b} + \cos \alpha \mathbf{c}$ .

d) A sík, melyre vetítünk az 1 sajátértékhez tartozik. Ha ebben választunk egy  $\{\mathbf{a}, \mathbf{b}\}$  bázist, és  $\mathbf{c}$  egy a síkra merőleges nemzérus vektor, akkor  $T\mathbf{a} = \mathbf{a}$ ,  $T\mathbf{b} = \mathbf{b}$ ,  $T\mathbf{c} = \mathbf{0}$ , így  $T$  mátrixa

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

e) A sík, melyre tükrözünk az 1 sajátértékhez tartozik. Ha ebben választunk egy  $\{\mathbf{a}, \mathbf{b}\}$  bázist, és  $\mathbf{c}$  egy a síkra merőleges nemzérus vektor, akkor  $T\mathbf{a} = \mathbf{a}$ ,  $T\mathbf{b} = \mathbf{b}$ ,  $T\mathbf{c} = -\mathbf{c}$ , így  $T$  mátrixa

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

□



**Diagonalizálható mátrixok spektrálfelbontása\*** A diagonalizálható  $\mathbf{A}$  mátrix  $\mathbf{A} = \mathbf{C}\mathbf{\Lambda}\mathbf{C}^{-1}$  alakja egy hasznos felbontását, az ún. spektrálfelbontását adja a mátrixnak.

Legyen  $\mathbf{A}$  spektruma  $\{\lambda_1, \lambda_2, \dots, \lambda_k\}$ . Bontsuk fel a  $\mathbf{\Lambda}$  mátrixot blokkdiagonális alakra úgy, hogy az azonos sajátértékek egy blokkba kerüljenek, majd ennek megfelelően bontsuk fel a  $\mathbf{C}$  és  $\mathbf{C}^{-1}$  mátrixot is blokkokra a következők szerint:

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 \mathbf{I} & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & \lambda_2 \mathbf{I} & \dots & \mathbf{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & \mathbf{O} & \dots & \lambda_k \mathbf{I} \end{bmatrix}, \quad \mathbf{C} = [\mathbf{X}_1 \quad \mathbf{X}_2 \quad \dots \quad \mathbf{X}_k], \quad \mathbf{C}^{-1} = \begin{bmatrix} \mathbf{Y}_1^T \\ \mathbf{Y}_2^T \\ \vdots \\ \mathbf{Y}_k^T \end{bmatrix}.$$

Írjuk föl e mátrixokkal az  $\mathbf{A} = \mathbf{C}\mathbf{\Lambda}\mathbf{C}^{-1}$  felbontást, majd fejtjük ki a blokkműveleteket:

$$\begin{aligned} \mathbf{A} = \mathbf{C}\mathbf{\Lambda}\mathbf{C}^{-1} &= [\mathbf{X}_1 \quad \mathbf{X}_2 \quad \dots \quad \mathbf{X}_k] \begin{bmatrix} \lambda_1 \mathbf{I} & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & \lambda_2 \mathbf{I} & \dots & \mathbf{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & \mathbf{O} & \dots & \lambda_k \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{Y}_1^T \\ \mathbf{Y}_2^T \\ \vdots \\ \mathbf{Y}_k^T \end{bmatrix} \\ &= \lambda_1 \mathbf{X}_1 \mathbf{Y}_1^T + \lambda_2 \mathbf{X}_2 \mathbf{Y}_2^T + \dots + \lambda_k \mathbf{X}_k \mathbf{Y}_k^T \\ &= \lambda_1 \mathbf{P}_1 + \lambda_2 \mathbf{P}_2 + \dots + \lambda_k \mathbf{P}_k, \end{aligned}$$

ahol  $\mathbf{P}_i = \mathbf{X}_i \mathbf{Y}_i^T$  a  $\lambda_i$  sajátértékhez tartozó mátrix, melyről a következőket fogjuk megmutatni:

8.36. ÁLLÍTÁS (DIAGONALIZÁLHATÓ MÁTRIXOK SPEKTRÁLFELBONTÁSA). Minden  $\{\lambda_1, \lambda_2, \dots, \lambda_k\}$  spektrumú diagonalizálható  $\mathbf{A}$  mátrix felírható

$$\mathbf{A} = \lambda_1 \mathbf{P}_1 + \lambda_2 \mathbf{P}_2 + \dots + \lambda_k \mathbf{P}_k$$

alakban, ahol

- a)  $\mathbf{P}_1 + \mathbf{P}_2 + \dots + \mathbf{P}_k = \mathbf{I}$ ,
- b)  $\mathbf{P}_i \mathbf{P}_j = \mathbf{O}$ , ha  $i \neq j$ ,
- c)  $\mathbf{P}_i$  az  $\mathcal{N}(\mathbf{A} - \lambda_i \mathbf{I})$  sajátaltérre való  $\mathcal{O}(\mathbf{A} - \lambda_i \mathbf{I})$  altér menti vetítés.

► Valójában több is igaz, nevezetesen megmutatható, hogy a fenti három feltétel szükséges és elégséges feltétele annak, hogy  $\mathbf{A}$  diagonalizálható legyen. E tételt a diagonalizálható mátrixok spektráltételének is nevezik.

**BIZONYÍTÁS.** A fent konstruált felbontásról megmutatjuk, hogy eleget tesz a feltételeknek.

a) A  $\mathbf{C}\mathbf{C}^{-1} = \mathbf{I}$  egyenlőséget blokkmátrixokként kezelve és fölhasználva a  $\mathbf{P}_i = \mathbf{X}_i \mathbf{Y}_i^T$  egyenlőségeket kapjuk, hogy  $\mathbf{P}_1 + \mathbf{P}_2 + \dots + \mathbf{P}_k = \mathbf{I}$ .

b) A  $\mathbf{C}^{-1}\mathbf{C} = \mathbf{I}$  egyenlőség blokkmátrixalakja viszont az  $\mathbf{Y}_i^T \mathbf{X}_i = \mathbf{I}$ , és

az  $\mathbf{Y}_i^T \mathbf{X}_j = \mathbf{O}$  ( $i \neq j$ ) egyenlőségre vezet, ahonnan  $\mathbf{P}_i \mathbf{P}_j = \mathbf{X}_i \mathbf{Y}_i^T \mathbf{X}_j \mathbf{Y}_j^T = \mathbf{O}$ .

c) Az előzőekből adódik, hogy  $\mathbf{P}_i^2 = \mathbf{P}_i$ , ugyanis  $\mathbf{P}_i^2 = \mathbf{X}_i (\mathbf{Y}_i^T \mathbf{X}_i) \mathbf{Y}_i^T = \mathbf{X}_i \mathbf{Y}_i^T = \mathbf{P}_i$ . Eszerint tehát  $\mathbf{P}_i$  vetítés. Meg kell még mutatnunk, hogy  $\mathcal{O}(\mathbf{P}_i) = \mathcal{N}(\mathbf{A} - \lambda_i \mathbf{I})$ . Ehhez fölhasználjuk, hogy bármely két  $\mathbf{X}, \mathbf{Y}$  mátrixra  $\mathcal{O}(\mathbf{XY}) \subseteq \mathcal{O}(\mathbf{X})$ .

$$\mathcal{O}(\mathbf{P}_i) = \mathcal{O}(\mathbf{X}_i \mathbf{Y}_i^T) \subseteq \mathcal{O}(\mathbf{X}_i) = \mathcal{O}(\mathbf{X}_i \mathbf{Y}_i^T \mathbf{X}_i) = \mathcal{O}(\mathbf{P}_i \mathbf{X}_i) \subseteq \mathcal{O}(\mathbf{P}_i).$$

Tehát mindenütt egyenlőség áll fenn, és  $\mathcal{O}(\mathbf{P}_i) = \mathcal{O}(\mathbf{X}_i) = \mathcal{N}(\mathbf{A} - \lambda_i \mathbf{I})$ , hiszen  $\mathcal{O}(\mathbf{X}_i)$  a  $\lambda_i$ -hez tartozó sajátaltér. Végül megmutatjuk, hogy a vetítés nulltere  $\mathcal{N}(\mathbf{P}_i) = \mathcal{O}(\mathbf{A} - \lambda_i \mathbf{I})$ . Kihasználva a fönt bizonyítottakat kapjuk, hogy

$$\mathbf{P}_i (\mathbf{A} - \lambda_i \mathbf{I}) = \mathbf{P}_i \left( \sum_{j=1}^k \lambda_j \mathbf{P}_j - \lambda_i \sum_{j=1}^k \mathbf{P}_j \right) = \sum_{j=1}^k (\lambda_j - \lambda_i) \mathbf{P}_i \mathbf{P}_j = \mathbf{O}.$$

Eszerint  $\mathcal{O}(\mathbf{A} - \lambda_i \mathbf{I}) \subseteq \mathcal{N}(\mathbf{P}_i)$ . Másrészt  $\mathcal{N}(\mathbf{A} - \lambda_i \mathbf{I}) = \mathcal{O}(\mathbf{P}_i)$ , így a dimenziótétel miatt  $\dim \mathcal{O}(\mathbf{A} - \lambda_i \mathbf{I}) = \dim \mathcal{N}(\mathbf{P}_i)$ , ami bizonyítja, hogy  $\mathcal{N}(\mathbf{P}_i) = \mathcal{O}(\mathbf{A} - \lambda_i \mathbf{I})$ .  $\square$

8.37. PÉLDA (SPEKTRÁLFELBONTÁS). Határozzuk meg a 8.12., a 8.26. és a 8.27. példákban szereplő

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

mátrix spektrálfelbontását.

**MEGOLDÁS.** E felbontáshoz felhasználhatjuk a 8.27. példában már kiszámolt sajátfelbontás diadikus alakját összevonva az azonos sajátértékhez tartozó diádokat, de kiemelve a sajátértéket:

$$\begin{aligned} \mathbf{A} &= 0 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} + 2 \begin{bmatrix} \frac{1}{2} \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} + 2 \begin{bmatrix} \frac{1}{2} \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \\ &= 0 \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + 2 \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Tehát

$$\mathbf{P}_1 = \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{P}_2 = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

$\mathbf{P}_1 + \mathbf{P}_2 = \mathbf{I}$ ,  $\mathbf{P}_1 \mathbf{P}_2 = \mathbf{O}$ , e két mátrix valóban vetítő mátrix, hisz  $\mathbf{P}_1^2 = \mathbf{P}_1$ ,  $\mathbf{P}_2^2 = \mathbf{P}_2$ , és láthatóan a sajátaltérre vetítenek (ld. 8.12.).  $\square$

**Sajátalterek direkt összege** Eddigi eredményeinket úgy foglalhatjuk össze, hogy egy  $n$ -edrendű valós négyzetes mátrix pontosan akkor diagonalizálható, ha  $\mathbb{R}^n$  minden nem zérus vektora egyértelműen felbomlik sajátvektorok összegére.

Az  $\mathbb{R}^n$  teret többször is felbontottuk két altér direkt összegére. Most általánosítjuk a direkt összeg fogalmát.

**8.38. DEFINÍCIÓ.** Legyenek  $\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_k$  a  $\mathcal{V}$  vektortér alterei. Azt mondjuk, hogy a  $\mathcal{V}$  tér a  $\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_k$  alterek direkt összege – jelölésben  $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \dots \oplus \mathcal{V}_k$  –, ha az alterek páronkénti metszete csak a nullvektorból áll, és az alterek összege az egész tér, azaz ha

- a)  $\mathcal{V}_i \cap \mathcal{V}_j = \{\mathbf{0}\}$ , ha  $i \neq j$  és
- b)  $\mathcal{V} = \mathcal{V}_1 + \mathcal{V}_2 + \dots + \mathcal{V}_k$ .

A két altér direkt összegére kimondott és bizonyított 7.36. tétel természetes módon átvihető több altér összegére is. Az ott adott bizonyítás minimális változtatással itt is működik:

**8.39. TÉTEL (A DIREKT ÖSSZEG TULAJDONSÁGAI).** Legyenek  $\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_k$  az  $n$ -dimenziós  $\mathcal{V}$  vektortér alterei. Az alábbi állítások ekvivalensek:

- a)  $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \dots \oplus \mathcal{V}_k$
- b)  $\mathcal{V}$  minden vektora egyértelműen előáll egy  $\mathcal{V}_1$ -es, egy  $\mathcal{V}_2$ -es és egy  $\mathcal{V}_k$ -beli vektor összegeként,
- c)  $\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_k$  alterek egy-egy bázisának egyesítése a  $\mathcal{V}$  bázisát adja,
- d)  $\mathcal{V}_i \cap \mathcal{V}_j = \{\mathbf{0}\}$ , ha  $i \neq j$  és  $\dim \mathcal{V}_1 + \dim \mathcal{V}_2 + \dots + \dim \mathcal{V}_k = n$ .

► Az  $\mathbb{R}^3$  tér standard bázisvektorai által generált 1-dimenziós alterek legyenek  $\mathcal{V}_1 = \text{span } \mathbf{i}$ ,  $\mathcal{V}_2 = \text{span } \mathbf{j}$  és  $\mathcal{V}_3 = \text{span } \mathbf{k}$ . Ekkor  $\mathbb{R}^3 = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \mathcal{V}_3$ .

► A 3-dimenziós tér egy síkjára való tükrözés sajátalterei a sík, melyre tükrözünk (ez a  $\lambda = 1$  sajátértékhez tartozó altér) és a síkra merőleges egyenes (mely a  $\lambda = -1$  sajátértékhez tartozik).  $\mathbb{R}^3$  e két altér direkt összege, mivel  $\mathbb{R}^3$  minden vektora egyértelműen felbomlik egy síkba eső és egy rá merőleges vektor összegére.

A fentiek alapján **diagonalizálható mátrixok spektrálfelbontásáról** szóló tétel egy gyönyörű megfogalmazáshoz vezet:

**8.40. TÉTEL (DIAGONALIZÁLHATÓ MÁTRIXOK SAJÁTALTEREI).** Az  $\mathbf{A} \in \mathbb{R}^{n \times n}$  mátrix pontosan akkor diagonalizálható, ha  $\mathbb{R}^n$  előáll az  $\mathbf{A}$  sajátaltéréinek direkt összegeként.

**BIZONYÍTÁS.** Legyen  $\mathbf{x} \in \mathbb{R}^n$  egy tetszőleges vektor. Ha  $\mathbf{A}$  diagonalizálható, akkor a 8.36. tétel szerint  $\mathbf{A} = \lambda_1 \mathbf{P}_1 + \lambda_2 \mathbf{P}_2 + \dots + \lambda_k \mathbf{P}_k$  az  $\mathbf{A}$  mátrix spektrálfelbontása, így

$$\mathbf{A}\mathbf{x} = \lambda_1 \mathbf{P}_1 \mathbf{x} + \lambda_2 \mathbf{P}_2 \mathbf{x} + \dots + \lambda_k \mathbf{P}_k \mathbf{x} = \mathbf{v}_1 + \mathbf{v}_2 + \dots + \mathbf{v}_k,$$

ahol a  $\mathbf{v}_i = \lambda_i \mathbf{P}_i \mathbf{x}$  vektor valóban a  $\lambda_i$ -hez tartozó sajátaltérben van, tehát a tér sajátaltéreinak összege. Mivel a sajátaltéreknek csak a nullvektor a közös elemük, ezért ez az összeg egyúttal direkt összeg.

Ha  $\mathbb{R}^n$  előáll  $\mathbf{A}$  sajátaltéreinak direkt összegeként, akkor a sajátaltérek dimenzióinak összege  $n$ , azaz a geometriai multiplicitások összege  $n$ , így az  $\mathbf{A}$  mátrix diagonalizálható.  $\square$

► A tétel nem csak mátrixokra, de lineáris leképezésekre is kimondható: egy  $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$  lineáris leképezés pontosan akkor diagonalizálható, ha  $\mathbb{R}^n$  felbomlik  $L$  sajátaltéreinak direkt összegére.

► A 3-dimenziós térben szemléltető példák a síkra való tükrözés, a síkra való vetítés, az egyenesre való tükrözés és az egyenesre való vetítés. Mindhárom példában  $\mathbb{R}^3$  egy 2- és egy 1-dimenziós altér direkt összegére bomlik.

► A 3-dimenziós tér egy egyenesre körüli  $\alpha$  szögű ( $\alpha \neq k\pi$ ) elforgatás nem diagonalizálható, bár látjuk, hogy a forgatás síkja és tengelye két olyan altér, melyek direkt összege az egész tér.

### A sajátérték kiszámítása

---

*A sajátértékek karakterisztikus polinomból való kiszámítása igen számításigényes módszer. A sajátértékek becslésére, közelítésére használt – többnyire iteratív technikákat használó – algoritmusok sokkal hatékonyabbak. Ráadásul az ezekben használt elvek nem csak a numerikus számítások szempontjából fontosak.*

---

**Gersgorin-körök** Egy mátrix elemeiből nagyon egyszerű számításokkal becslések adhatók a sajátértékekre. A komplex számsík Gersgorin-körei például tartalmazzák az összes sajátértéket.

Az  $n \times n$ -es valós vagy komplex  $\mathbf{A}$  mátrix Gersgorin-körein az  $a_{ii}$  közepű, és  $r_i^{\text{sor}}$  sugarú  $G_i^{\text{sor}}$ , illetve  $r_i^{\text{osz}}$  sugarú  $G_i^{\text{osz}}$  köröket értjük ( $i = 1, 2, \dots, n$ ), ahol

$$r_i^{\text{sor}} = \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|, \quad r_i^{\text{osz}} = \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ji}|. \quad (8.14)$$

Más szavakkal pl. a  $G_i^{\text{sor}}$  kör középpontja  $a_{ii}$ , sugara az  $\mathbf{A}$  mátrix  $i$ -edik sorában lévő, nem a főátlón lévő elemek abszolút értékeinek összege, míg  $G_i^{\text{osz}}$  sugara az  $i$ -edik oszlop nem a főátlón lévő elemei abszolút értékének összege.

E körök Szemjon Aranovics Gersgorin orosz matematikus nevét viselik, mely a ciril betűkről való átírás miatt többféle alakban is előfordul, például Gershgorin, Gerschgorin, Geršgorin.

8.41. PÉLDA (GERSGORIN-KÖRÖK). Rajzoljuk föl a

$$\mathbf{A} = \begin{bmatrix} 4 & -4 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 8 \end{bmatrix}$$

mátrix Gersgorin köreit!

A megoldás a 8.3 ábrán látható.

8.42. TÉTEL (GERSGORIN-KÖRÖK TULAJDONSÁGAI). Legyen  $\mathbf{A}$  valós vagy komplex  $n \times n$ -es mátrix. Ekkor igazak a következők:

- Minden sajátérték benne van a  $G_i^{\text{sor}}$  körök egyesítésében.
- Minden sajátérték benne van a  $G_i^{\text{osz}}$  körök egyesítésében.
- Minden sajátérték benne van az előző két halmaz metszetében.
- Ha a  $G_i^{\text{sor}}$  körök egy  $k$ -elemű részhalmaza diszjunkt a maradék  $n - k$  kör mindegyikétől, akkor uniójuk multiplicitással számolva pontosan  $k$  sajátértéket tartalmaz.

**BIZONYÍTÁS.** Jelölje  $\sigma(\mathbf{A})$  az  $\mathbf{A}$  mátrix spektrumát!

a) Az állítás szerint  $\sigma(\mathbf{A}) \subseteq \bigcup_i G_i^{\text{sor}}$ . A  $\lambda$  sajátértékhez tartozó egyik sajátvektort osszuk el legnagyobb abszolút értékű koordinátájával. E vektort jelölje  $\mathbf{x}$ , legyen  $x_i = 1$ , tehát  $|x_j| \leq 1$  ( $j = 1, 2, \dots, n$ ). Mivel  $\lambda = \lambda x_i = [\lambda \mathbf{x}]_i = [\mathbf{A}\mathbf{x}]_i = \sum_j a_{ij}x_j$ , így  $\lambda - a_{ii} = \sum_{j \neq i} a_{ij}x_j$ , tehát

$$|\lambda - a_{ii}| = \left| \sum_{j \neq i} a_{ij}x_j \right| \leq \sum_{j \neq i} |a_{ij}| |x_j| \leq \sum_{j \neq i} |a_{ij}| = r_i^{\text{sor}}.$$

b) Mivel  $\mathbf{A}$  és  $\mathbf{A}^T$  sajátértékei megegyeznek, ezért sorok helyett oszlopokkal ugyanezek a számítások megismételhetők.

c) A sajátértékek az előző két pontban megadott halmazok mindegyikében, így metszetükben is benne vannak.

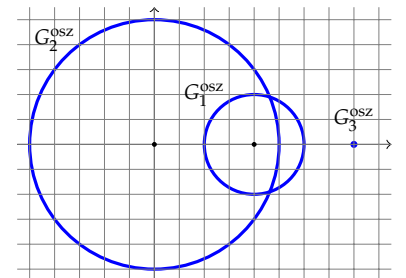
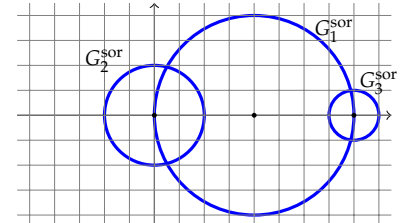
d) Legyen  $\mathbf{B}(r)$  az a mátrix, melynek elemeire  $b_{ii} = a_{ii}$  és  $b_{ij} = ra_{ij}$ , ha  $i \neq j$ . Mátrixműveletekkel kifejezve

$$\mathbf{B}(r) = r\mathbf{A} + (1-r)\text{diag}(a_{11}, \dots, a_{nn}).$$

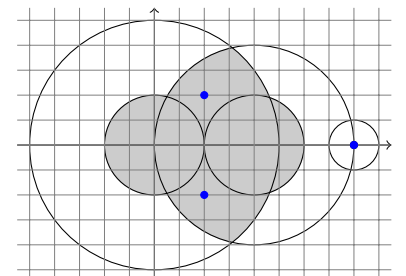
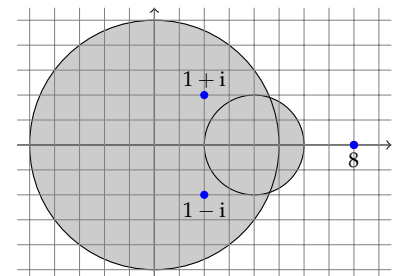
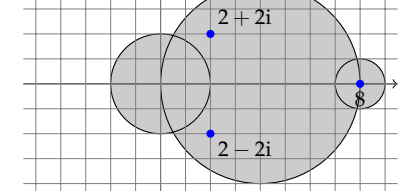
Eszerint  $\mathbf{B}(0) = \text{diag}(a_{11}, \dots, a_{nn})$ ,  $\mathbf{B}(1) = \mathbf{A}$ . Változzék  $r$  folyamatosan 0-tól 1-ig. Eközben  $\mathbf{B}(r)$  Gersgorin-körei 0-sugarúból indulva nőnek az  $\mathbf{A}$  Gersgorin-köreiig, középpontjaik közben nem változnak. Mivel a sajátértékek a mátrix elemeinek folytonos függvényei, ezért az egymással fedésbe kerülő, de a többitől diszjunkt maradó  $k$  számú kör által lefedett sajátértékek száma  $k$  marad.  $\square$

► Az nem igaz, hogy mindegyik Gersgorin-körben van legalább egy sajátérték. Például az  $\mathbf{A} = \begin{bmatrix} 2 & -2 \\ 1 & 0 \end{bmatrix}$  mátrix sajátértékei  $1+i$  és  $1-i$ . Ezek nem esnek bele sem a 0-közepű, 1-sugarú sem a 2 középpő 1-sugarú Gersgorin-körbe (ld. 8.4 a), b) ábra).

| kör                | közép-<br>pont | sugár                               |
|--------------------|----------------|-------------------------------------|
| $G_1^{\text{sor}}$ | $a_{11} = 4$   | $r_1^{\text{sor}} =  -4  = 4$       |
| $G_2^{\text{sor}}$ | $a_{22} = 0$   | $r_2^{\text{sor}} = 2$              |
| $G_3^{\text{sor}}$ | $a_{33} = 8$   | $r_3^{\text{sor}} = 1$              |
| $G_1^{\text{osz}}$ | $a_{11} = 4$   | $r_1^{\text{osz}} = 2$              |
| $G_2^{\text{osz}}$ | $a_{22} = 0$   | $r_2^{\text{osz}} =  -4  +  1  = 5$ |
| $G_3^{\text{osz}}$ | $a_{33} = 8$   | $r_3^{\text{osz}} = 0$              |



8.3. ábra: Az  $\mathbf{A}$ -mátrix Gersgorin körei



8.4. ábra: Az  $\mathbf{A}$  mátrix sajátértékeit tartalmazza a) a sorösszeg szerinti halmaz, b) az oszlopösszeg szerinti halmaz, c) a két halmaz metsze

► Az előző megjegyzésnek megfelelően nem lehet úgy keresni a sajátértékeket, hogy minden főátlóbeli elemhez a sor- és oszlopösszegek közül a kisebbiket választjuk. Helyesen a tétel *a)* és *b)* pontjaiban konstruált halmazok metszetét kell venni, ami például esetünkben a 8.4 *c)* ábrán látható.

8.43. PÉLDA (GERSGORIN-KÖRÖK HASZNÁLATA). *Mutassuk meg a Gersgorin-körök segítségével, hogy a*

$$\mathbf{B} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 5 & 1 \\ 2 & 1 & 9 \end{bmatrix}$$

*mátrixnak minden sajátértéke valós.*

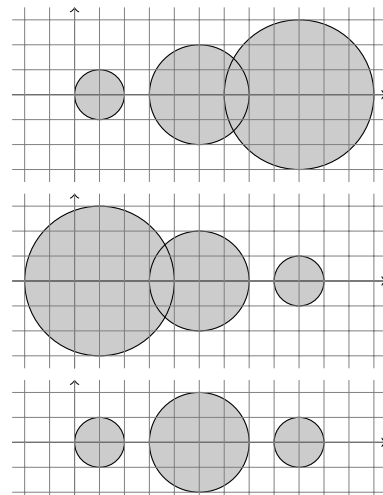
**MEGOLDÁS.** Mivel valós elemű mátrix komplex gyökei párosával egymás konjugáltjai, ezért ha egy Gersgorin-kör diszjunkt a többitől, abban csak egy valós sajátérték lehet. A  $\mathbf{B}$  mátrix sorok szerinti Gersgorin-köröi azt mutatják, hogy az 1-közepű 1-sugarú körben egy valós sajátérték van. Mivel a másik két kör metszi egymást, ezért a másik két sajátérték még lehetne nem valós. Az oszlopok szerinti 9-közepű 1-sugarú Gersgorin-kör viszont egy újabb valós sajátértéket garantál, így a harmadik is szükségképpen az. Ezt mutatja a sorok és oszlopok szerinti halmazok metszete is. A sajátértékek a  $[0, 2]$ ,  $[3, 7]$  és  $[8, 10]$  intervallumokba esnek (ld. 8.5 ábra). □

8.44. ÁLLÍTÁS (DOMINÁNS FŐÁTLÓJÚ MÁTRIX INVERTÁLHATÓSÁGA). *Bármely soronként domináns főátlójú valós vagy komplex mátrix invertálható. Hasonló igaz az oszloponként domináns főátlójú mátrixokra is.*

**BIZONYÍTÁS.** A 2.58. definíció szerint az  $\mathbf{A}$  mátrix soronként domináns főátlójú, ha bármely főátlóbeli elemére  $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$ . Ez épp azt jelenti, hogy  $\mathbf{A}$  minden Gersgorin-körének középpontja messzebb van az origótól, mint amekkora a sugara, azaz a 0 szám egyik Gersgorin-körben sincs benne, tehát a 0 nem sajátérték, így  $\mathbf{A}$  invertálható. □

**Hatványmódszer** A sajátértékek a karakterisztikus polinom gyökei, a négynél magasabb fokú polinomokra viszont nem létezik megoldóképlet, így elvileg sem létezhet olyan módszer, mely a sajátértékeket mindig pontosan ki tudja számolni. A sajátértékek teszőlegesen pontos közelítésére viszont hatékony algoritmusok léteznek. Ezek legalapvetőbbike a hatványmódszer.

Négyzetes mátrix egy sajátértékét *szigorúan dominánsnak* nevezzük, ha egyszeres multiplicitású, és abszolút értékben nagyobb az összes többinél. A hozzá tartozó sajátvektort és sajátalteret *szigorúan domináns sajátvektornak*, ill. *sajátalternek*, a belőlük alkotott párt *szigorúan*



8.5. ábra: A  $\mathbf{B}$  mátrix sorösszeg és oszlopösszeg szerinti Gersgorin köröi és ezek metszete. Utóbbi mutatja, hogy  $\mathbf{B}$ -nek három valós sajátértéke van.

domináns saját párnak nevezzük.

A *hatványmódszer* megadja négyzetes mátrix szigorúan domináns saját páriját.

Tegyük fel, hogy az  $\mathbf{A} \in \mathbb{R}^{n \times n}$  mátrixnak  $\lambda_1$  szigorúan domináns sajátértéke. Összes sajátértékét indexeljük úgy, hogy fennálljon az

$$|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_m|$$

összefüggés. Legyen  $\mathbf{v}_i$  egy  $\lambda_i$ -hez tartozó sajátvektor. Világos, hogy  $\lambda_1$  valós, egyébként  $\bar{\lambda}_1$  egy tőle különböző, de azonos abszolút értékű sajátérték lenne.

Legyen  $\mathbf{x} \in \mathbb{R}^n$  egy olyan vektor, mely előáll a sajátvektorok lineáris kombinációjaként. Ha  $\mathbf{A}$  diagonalizálható, akkor minden vektor ilyen. Legyen tehát  $\mathbf{x} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_m \mathbf{v}_m$ . Ekkor tetszőleges  $k$  nemnegatív egészre

$$\begin{aligned} \mathbf{A}^k \mathbf{x} &= c_1 \mathbf{A}^k \mathbf{v}_1 + c_2 \mathbf{A}^k \mathbf{v}_2 + \dots + c_m \mathbf{A}^k \mathbf{v}_m \\ &= c_1 \lambda_1^k \mathbf{v}_1 + c_2 \lambda_2^k \mathbf{v}_2 + \dots + c_m \lambda_m^k \mathbf{v}_m. \end{aligned}$$

Ekkor  $\lambda_1^k$ -val való osztás után  $k \rightarrow \infty$  esetén

$$\frac{1}{\lambda_1^k} \mathbf{A}^k \mathbf{x} = c_1 \mathbf{v}_1 + c_2 \left( \frac{\lambda_2}{\lambda_1} \right)^k \mathbf{v}_2 + \dots + c_m \left( \frac{\lambda_m}{\lambda_1} \right)^k \mathbf{v}_m \rightarrow c_1 \mathbf{v}_1,$$

ugyanis  $(\lambda_i/\lambda_1)^k \rightarrow 0$ , ha  $i > 1$ . Eszerint ha  $c_1 \neq 0$ , akkor  $\mathbf{A}^k \mathbf{x}$  iránya tart a domináns sajátvektor irányához. Ezt egy példán szemléltetjük.

8.45. PÉLDA. Tekintsük az

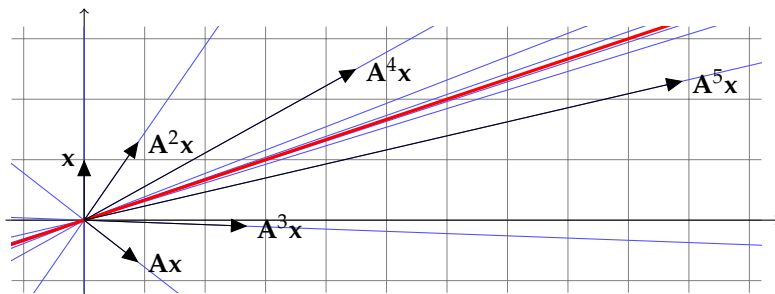
$$\mathbf{A} = \begin{bmatrix} 1.7 & 0.9 \\ 0.9 & -0.7 \end{bmatrix}$$

mátrixot! Határozzuk meg domináns sajátértékét és sajátalterét! Legyen  $\mathbf{x} = (0, 1)$ . Számítsuk ki az  $\mathbf{A}^k \mathbf{x}$  vektorokat néhány  $k$  értékre szemléltetve irányuknak a domináns sajátvektor irányához való tartását!

**MEGOLDÁS.** Az  $\mathbf{A}$  karakterisztikus polinomja  $\det(\mathbf{A} - x\mathbf{I}) = x^2 - x - 2$ , ennek gyökei  $\lambda_1 = 2$ ,  $\lambda_2 = -1$ , így a szigorúan domináns sajátérték  $\lambda_1 = 2$ . A hozzá tartozó sajátalteret a  $(3, 1)$  vektor feszíti ki. Egy programmal kiszámoltuk az  $\mathbf{A}^k \mathbf{x}$  vektorokat a  $k = 0, 1, \dots, 8$  értékekre:

| $k$                       | 0                                      | 1                                           | 2                                          | 3                                           | 4                                          | 5                                          | 6                                           | 7                                            |
|---------------------------|----------------------------------------|---------------------------------------------|--------------------------------------------|---------------------------------------------|--------------------------------------------|--------------------------------------------|---------------------------------------------|----------------------------------------------|
| $\mathbf{A}^k \mathbf{x}$ | $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ | $\begin{bmatrix} 0.9 \\ -0.7 \end{bmatrix}$ | $\begin{bmatrix} 0.9 \\ 1.3 \end{bmatrix}$ | $\begin{bmatrix} 2.7 \\ -0.1 \end{bmatrix}$ | $\begin{bmatrix} 4.5 \\ 2.5 \end{bmatrix}$ | $\begin{bmatrix} 9.9 \\ 2.3 \end{bmatrix}$ | $\begin{bmatrix} 18.9 \\ 7.3 \end{bmatrix}$ | $\begin{bmatrix} 38.7 \\ 11.9 \end{bmatrix}$ |

A vektorok hossza láthatóan végtelenhez konvergál, de az általuk kifeszített altereknek a domináns sajátalterhez való tartása így is jól leolvasható a 8.6 ábráról.

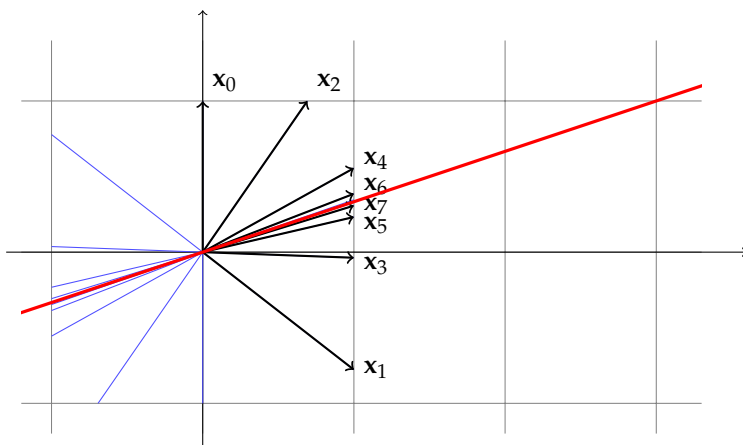


8.6. ábra: Az  $A^k x$  vektorok ( $k = 0, 1, \dots, 5$ ) jelölt irányai, valamint ezek határértéke: az  $A$  mátrix  $\lambda = 2$  sajátértékhez tartozó sajátaltére (pirossal színezve).

Ha a domináns sajátérték abszolút értéke kisebb lenne 1-nél, akkor az  $A^k x$  vektorsorozat a nullvektorhoz konvergálna. Ezért érdemes e sorozatot normálni, például osztani a hosszával. Még egyszerűbb elosztani a vektort a legnagyobb abszolút értékű koordinátájával. (Az így kapott vektorsorozat nem biztos, hogy konvergens lesz. Lehet, hogy pl. egy-egy koordinátahelyen alternáló divergens sorozatot kapunk. Ez elkerülhető, ha valamelyik nem nullához konvergáló koordináta előjelét egy esetleges  $-1$ -gyel szorzással mindig pozitívvá tesszük.) Jelölje  $x_k$  az  $A^k x$  vektor legnagyobb koordinátájával való osztása után kapott vektort, és jelölje  $i$  ennek a koordinátának az indexét. Tehát  $[x_k]_i = 1$ . Az alábbi táblázat az  $x_k$  és az  $[Ax_k]_i$  értékeket mutatja (utóbbi azt adja meg, hogy az 1 értékű koordináta hányszorosára változik az  $A$ -val való szorzás után):

| $k$        | 0                                      | 1                                               | 2                                              | 3                                               | 4                                              | 5                                              | 6                                              | 7                                              |
|------------|----------------------------------------|-------------------------------------------------|------------------------------------------------|-------------------------------------------------|------------------------------------------------|------------------------------------------------|------------------------------------------------|------------------------------------------------|
| $x_k$      | $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ | $\begin{bmatrix} 1.000 \\ -0.778 \end{bmatrix}$ | $\begin{bmatrix} 0.692 \\ 1.000 \end{bmatrix}$ | $\begin{bmatrix} 1.000 \\ -0.037 \end{bmatrix}$ | $\begin{bmatrix} 1.000 \\ 0.556 \end{bmatrix}$ | $\begin{bmatrix} 1.000 \\ 0.232 \end{bmatrix}$ | $\begin{bmatrix} 1.000 \\ 0.386 \end{bmatrix}$ | $\begin{bmatrix} 1.000 \\ 0.307 \end{bmatrix}$ |
| $[Ax_k]_i$ |                                        | -0.700                                          | 1.000                                          | -0.769                                          | 1.667                                          | 2.200                                          | 1.909                                          | 2.047                                          |

Az  $x_k$  vektorokat szemlélteti a 8.7 ábra. Sorozatuk határértéke az  $(1, 1/3)$  vektor, ami a domináns sajátvektor. Általánosan is igaz,



8.7. ábra: Az  $x_k$  vektorok és a sajátaltér.



hogy az így kapott  $\mathbf{x}_k$  sorozat elemei a szigorúan domináns sajátvektor becslését adják. Egyúttal a domináns sajátérték becslésére is kiadódik. Könnyen bizonyítható ugyanis, hogy ha  $\mathbf{x}_k$   $i$ -edik koordinátája 1, akkor  $\mathbf{A}\mathbf{x}_k$   $i$ -edik koordinátája a  $\lambda_1$  egy becslését adja, pontosabban e sorozat határértéke  $\lambda_1$ . E konkrét példában a fenti táblázat alsó sora épp ezt a sorozatot tartalmazza, mely a  $\lambda_1 = 2$  értékhez konvergál.  $\square$

**8.46. TÉTEL (HATVÁNYMÓDSZER).** *Ha  $\lambda_1$  az  $\mathbf{A} \in \mathbb{R}^{n \times n}$  mátrix szigorúan domináns sajátértéke, akkor létezik olyan  $\mathbf{x}_0$  vektor, hogy az  $\mathbf{A}^k \mathbf{x}_0$  vektorok által kifeszített alterek sorozata a domináns sajátaltérhez konvergál.*

E tételt fentebb bizonyítottuk abban a speciális esetben, amikor  $\mathbf{x}_0$  a sajátvektorok lineáris kombinációja. Valójában elég csak azt kikötni, hogy az  $\mathbf{x}_0$  vektornak a domináns sajátvektor irányába eső összetevője ne a zérusvektor legyen.

*Feladatok**Gersgorin-körök*

8.2. Mutassuk meg, hogy az

$$A = \begin{bmatrix} 2 & 0 & 5 & 0 \\ 0 & 9 & 0 & 1 \\ -1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 6 \end{bmatrix}$$

mátrixnak legalább két sajátértéke valós.

8.3. utassuk meg, hogy az alábbi mátrixok minden sajátér-

téke valós! a)  $\begin{bmatrix} 9 & 3 & 2 \\ 0 & 4 & 2 \\ 1 & 0 & 1 \end{bmatrix}$  b)  $\begin{bmatrix} 9 & 0 & 3 & 0 \\ 0 & 4 & 0 & 1 \\ -3 & 0 & -1 & 0 \\ 2 & 1 & 2 & 0 \end{bmatrix}$

---

### Megoldások

8.2. A 9-közepű 1-sugarú Gersgorin-körben csak 1 gyök lehet, így az valós. Mivel összesen 4 gyöke van és a komplexek párosan fordulnak elő, ezért kell még valós gyöknek lennie.

8.3. a) A sorokhoz tartozó Gersgorin-körök közül az 1-közepű, 1-sugarú kör diszjunkt a többtől, így ebben a körben egy van egy valós sajátérték, az oszlopokhoz tartozók közül a 9-közepű, 1-sugarú diszjunkt a többtől, így ebben a körben is van egy valós sajátérték, és ha két sajátérték valós, akkor a harmadik is. Az a)-beli mátrixban a sorokra, a b)-beliben az oszlopokra alkalmazva



## 9

# Diagonalizálás ortonormált bázisban

Számtalan műszaki és tudományos probléma valós szimmetrikus mátrixok vizsgálatára vezet. Ezek szerencsés esetek, mert ilyenkor sajátvektorokból álló ortonormált bázis is található, és ez sok számítást egyszerűvé és numerikusan is stabilabbá tesz. A szimmetrikus mátrixok használatát a kvadratikus alakok jeírásán demonstráljuk.

## Ortogonalis és unitér diagonalizálás

---

Egy mátrixot diagonalizálni azzal ekvivalens, hogy a hozzá tartozó mátrixlekepezéshez egy olyan bázist találni, melyben mátrixa diagonális. Különösen szerencsés, ha e bázis még ortonormált is.

---

**Valós mátrixok ortogonalis diagonalizálása, valós spektráltétel** Megmutatjuk, hogy a valós mátrixok közt pontosan a szimmetrikusak azok, amelyekhez található olyan ortonormált bázis, melyben az diagonális alakot ölt.

A 8.25. tétel szerint a diagonalizálhatóság szükséges és elégséges feltétele, hogy létezzék a mátrix rendjével egyező számú független sajátvektora. Ha e vektorok ortonormált rendszert alkotnak, akkor a belőlük alkotott mátrix ortogonalis mátrix.

**9.1. DEFINÍCIÓ (ORTOGONÁLIS DIAGONALIZÁLHATÓSÁG).** Az  $\mathbf{A}$  mátrix ortogonalisan diagonalizálható, ha találunk egy ortogonalis  $\mathbf{Q}$  és egy diagonális  $\mathbf{\Lambda}$  mátrixot, hogy  $\mathbf{Q}^T \mathbf{A} \mathbf{Q} = \mathbf{\Lambda}$ .

- A definícióbeli egyenlőséggel ekvivalens alak:  $\mathbf{A} = \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^T$ .
- Az nyilvánvaló, hogy ha egy mátrix ortogonalisan diagonalizálható, akkor szimmetrikus, ha ugyanis  $\mathbf{A} = \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^T$ , akkor

$$\mathbf{A}^T = (\mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^T)^T = (\mathbf{Q}^T)^T \mathbf{\Lambda}^T \mathbf{Q}^T = \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^T = \mathbf{A}.$$

E fejezet fő célja az állítás megfordításának igazolása. Ennek érdekében először megmutatjuk, hogy szimmetrikus mátrix sajátalterei nem csak függetlenek egymástól, de merőlegesek is egymásra.

**9.2. TÉTEL (SZIMMETRIKUS MÁTRIX SAJÁTALTEREI).** *Szimmetrikus mátrix bármely két különböző sajátaltére merőleges egymásra.*

**BIZONYÍTÁS.** Két különböző sajátaltér két különböző sajátértékhez tartozik. Megmutatjuk, hogy az egyik altér bármelyik vektora merőleges a másik altér bármely vektorára. Legyen tehát  $(\lambda, \mathbf{x})$  és  $(\mu, \mathbf{y})$  két saját pár, ahol  $\lambda \neq \mu$  két különböző sajátértéke  $\mathbf{A}$ -nak. Így  $\mathbf{Ax} = \lambda\mathbf{x}$  és  $\mathbf{Ay} = \mu\mathbf{y}$ . Ebből adódik, hogy

$$\lambda(\mathbf{x}^T \mathbf{y}) = (\lambda\mathbf{x})^T \mathbf{y} = (\mathbf{Ax})^T \mathbf{y} = \mathbf{x}^T \mathbf{A}^T \mathbf{y} = \mathbf{x}^T \mathbf{Ay} = \mathbf{x}^T \mu\mathbf{y} = \mu(\mathbf{x}^T \mathbf{y}).$$

Eszerint  $(\lambda - \mu)(\mathbf{x}^T \mathbf{y}) = 0$ , de  $\lambda - \mu \neq 0$ , ezért  $\mathbf{x}^T \mathbf{y} = \mathbf{x} \cdot \mathbf{y} = 0$ , azaz a két vektor merőleges egymásra.  $\square$

**9.3. TÉTEL (VALÓS SPEKTRÁLTÉTEL).** *A valós  $\mathbf{A}$  mátrix pontosan akkor diagonalizálható ortogonálisan, ha szimmetrikus.*

**BIZONYÍTÁS.** Az állítás egyik felét megmutattuk a 9.7. definíció után. A másik felét az  $\mathbf{A}$  mátrix rendjére vonatkozó teljes indukcióval bizonyítjuk.  $n = 1$  esetén nincs mit bizonyítani, ekkor  $\mathbf{A}$  szimmetrikus és diagonális alakú. Tegyük fel, hogy minden legfőljebb  $n - 1$ -edrendű mátrixra igaz az állítás, azaz hogy ha szimmetrikus, akkor ortogonálisan hasonló egy diagonális mátrixhoz. Mivel  $\mathbf{A}$  szimmetrikus, ezért **minden sajátértéke valós**. Legyen ezek egyike  $\lambda$ , a hozzá tartozó egyik egységnyi hosszú sajátvektor  $\mathbf{u}_1$ . Egészítsük ki  $\mathbf{u}_1$ -et a teljes tér egy  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$  ONB-ává. Ekkor a  $\mathbf{Q}_0 = [\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_n]$  mátrix ortogonális, és

$$\begin{aligned} \mathbf{Q}_0^T \mathbf{A} \mathbf{Q}_0 &= \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix} \mathbf{A} [\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_n] = \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix} [\mathbf{A}\mathbf{u}_1 \ \mathbf{A}\mathbf{u}_2 \ \dots \ \mathbf{A}\mathbf{u}_n] \\ &= \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix} [\lambda\mathbf{u}_1 \ \mathbf{A}\mathbf{u}_2 \ \dots \ \mathbf{A}\mathbf{u}_n] = \begin{bmatrix} \lambda & * \\ \mathbf{0} & \mathbf{A}_1 \end{bmatrix} = \mathbf{B}, \end{aligned} \quad (9.1)$$

ugyanis  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$  ONB, így  $\mathbf{u}_1^T(\lambda\mathbf{u}_1) = \lambda(\mathbf{u}_1^T \mathbf{u}_1) = \lambda$ , és  $i > 1$  esetén  $\mathbf{u}_i^T(\lambda\mathbf{u}_1) = \lambda(\mathbf{u}_i^T \mathbf{u}_1) = 0$ . A  $\mathbf{B}$  blokkmátrixban tehát  $\mathbf{A}_1$  egy  $(n - 1) \times (n - 1)$ -es valós mátrix. Másrészt

$$\mathbf{B}^T = (\mathbf{Q}_0^T \mathbf{A} \mathbf{Q}_0)^T = \mathbf{Q}_0^T \mathbf{A}^T \mathbf{Q}_0 = \mathbf{Q}_0^T \mathbf{A} \mathbf{Q}_0 = \mathbf{B},$$

azaz  $\mathbf{B} = \begin{bmatrix} \lambda & * \\ 0 & \mathbf{A}_1 \end{bmatrix}$  szimmetrikus, tehát  $\mathbf{B} = \begin{bmatrix} \lambda & 0^T \\ 0 & \mathbf{A}_1 \end{bmatrix}$ , és  $\mathbf{A}_1$  is szimmetrikus!

Mivel  $\mathbf{B}$  hasonló  $\mathbf{A}$ -hoz, ezért sajátértékeik megegyeznek, tehát  $\mathbf{A}_1$  minden sajátértéke  $\mathbf{A}$ -nak is sajátértéke. A teljes indukció miatt viszont  $\mathbf{A}_1$ -hez létezik olyan  $\mathbf{Q}_1$  ortogonális és  $\Lambda_1$  diagonális mátrix, hogy  $\mathbf{A}_1 = \mathbf{Q}_1 \Lambda_1 \mathbf{Q}_1^T$ . Ekkor viszont a  $\mathbf{Q} = \mathbf{Q}_0 \begin{bmatrix} 1 & 0^T \\ 0 & \mathbf{Q}_1 \end{bmatrix}$  mátrix ortogonális, hisz két ortogonális mátrix szorzata, és  $\mathbf{A}$ -t hasonlóvá teszi egy diagonális mátrixhoz:

$$\begin{aligned} \mathbf{Q}^T \mathbf{A} \mathbf{Q} &= \left( \mathbf{Q}_0 \begin{bmatrix} 1 & 0^T \\ 0 & \mathbf{Q}_1 \end{bmatrix} \right)^T \mathbf{A} \left( \mathbf{Q}_0 \begin{bmatrix} 1 & 0^T \\ 0 & \mathbf{Q}_1 \end{bmatrix} \right) = \begin{bmatrix} 1 & 0^T \\ 0 & \mathbf{Q}_1 \end{bmatrix}^T \mathbf{Q}_0^T \mathbf{A} \mathbf{Q}_0 \begin{bmatrix} 1 & 0^T \\ 0 & \mathbf{Q}_1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0^T \\ 0 & \mathbf{Q}_1 \end{bmatrix}^T \begin{bmatrix} \lambda & 0^T \\ 0 & \mathbf{A}_1 \end{bmatrix} \begin{bmatrix} 1 & 0^T \\ 0 & \mathbf{Q}_1 \end{bmatrix} \\ &= \begin{bmatrix} \lambda & 0^T \\ 0 & \mathbf{Q}_1^T \mathbf{A}_1 \mathbf{Q}_1 \end{bmatrix} = \begin{bmatrix} \lambda & 0^T \\ 0 & \Lambda_1 \end{bmatrix}. \end{aligned}$$

Ezzel bizonyítottuk, hogy a szimmetrikus  $\mathbf{A}$  mátrix is ortogonálisan diagonalizálható.  $\square$

► A tétel bizonyítása egyúttal ötletet ad a diagonalizálás megvalósításához is. Határozzuk meg az  $\mathbf{A}$  mátrix egy  $\lambda$  sajátértékét, és állítsuk elő a

$$\begin{bmatrix} \lambda & 0^T \\ 0 & \mathbf{A}_1 \end{bmatrix}, \quad \text{és az} \quad \begin{bmatrix} 1 & 0^T \\ 0 & \mathbf{Q}_1 \end{bmatrix}$$

mátrixokat. A következő lépésben ismételjük meg e lépést az  $\mathbf{A}_1$  mátrixszal, de az ott előálló  $(n-2) \times (n-2)$ -es  $\mathbf{Q}_2$  mátrixból az

$$\begin{bmatrix} 1 & 0 & 0^T \\ 0 & 1 & 0^T \\ 0 & 0 & \mathbf{Q}_2 \end{bmatrix}$$

ortogonális mátrixot képezzük, stb. Végül az így kapott mátrixok szorzata lesz  $\mathbf{Q}$ , azaz

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{Q}_1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0^T \\ 0 & 1 & 0^T \\ 0 & 0 & \mathbf{Q}_2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0^T \\ 0 & 1 & 0 & 0^T \\ 0 & 0 & 1 & 0^T \\ 0 & 0 & 0 & \mathbf{Q}_3 \end{bmatrix} \cdots$$

9.4. PÉLDA. Diagonalizáljuk az alábbi mátrixot ortogonálisan!

$$\begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}.$$

Határozzuk meg az áttérés mátrixát a standardról arra a bázisra, melyben a mátrix diagonális alakot vesz fel!

**MEGOLDÁS.** A karakterisztikus polinom:

$$\begin{vmatrix} 3-\lambda & 1 & 1 \\ 1 & 3-\lambda & 1 \\ 1 & 1 & 3-\lambda \end{vmatrix} = -\lambda^3 + 9\lambda^2 - 24\lambda + 20,$$

melynek gyökei 2, 2 és 5. Tehát a diagonális alak  $\text{diag}(2, 2, 5)$ . Az áttérés mátrixához szükségünk lesz a sajátvektorokra.  $\lambda = 2$  esetén:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

az  $x + y + z = 0$  egyenletrendszer megoldása pedig  $(x, y, z) = (-1, 1, 0)s + (-1, 0, 1)t$ . Ennek az altérnek egy bázisa tehát a  $(-1, 1, 0)$  és a  $(-1, 0, 1)$  vektorokból áll. A  $\lambda = 5$  esetén:

$$\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix},$$

amely egyenletrendszer megoldása  $(x, y, z) = (1, 1, 1)t$ . A két különböző sajátaltérből való vektorok merőlegesek egymásra, de a  $\lambda = 2$ -höz tartozó sajátaltér két sajátvektora nem alkot ortogonális rendszert, ezért az általuk kifeszített térben új bázist keresünk, egy ortonormál-tat. Legyen az  $\mathbf{a} = (-1, 1, 0)/\sqrt{2}$  vektor az egyik, ekkor

$$(-1, 0, 1) - \left( (-1, 0, 1) \cdot \frac{(-1, 1, 0)}{\sqrt{2}} \right) \frac{(-1, 1, 0)}{\sqrt{2}} = \left( -\frac{1}{2}, -\frac{1}{2}, 1 \right).$$

Ezt normálva kapjuk a  $\mathbf{b} = (-\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{\sqrt{2}}{\sqrt{3}})$  vektort, végül a  $\lambda = 5$ -höz tartozó normált vektor  $\mathbf{c} = \frac{1}{\sqrt{3}}(1, 1, 1)$ . A standard bázisra való áttérés mátrixa tehát az  $[\mathbf{a}|\mathbf{b}|\mathbf{c}]$  mátrix, azaz

$$\begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}.$$

Ennek inverze lesz a standard bázisról való áttérés mátrixa, mely – ortogonális mátrixról lévén szó – a transzponáltja, azaz

$$\begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{\sqrt{2}}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}.$$

□



**Schur-felbontás\*** Olyan ortonormált bázist találni, melyben egy mátrix egyszerűbb alakú, akkor is fontos, ha az az alak nem a diagonális. Ilyen például a felsőháromszög-mátrix-alak.

A **valós spektráltétel** bizonyításának csekély változtatása egy másik hasznos tételre vezet. Az ott konstruált **B** mátrixban ugyanis elimináltuk az **A** első oszlopának főátló alatti elemeit. Ezt a lépést ismételve elérhető, hogy a mátrixot felső háromszög-alakra hozzuk.

#### 9.5. TÉTEL (SCHUR-FELBONTÁS).

a) Minden valós négyzetes **A** mátrix, melynek összes sajátértéke valós, ortogonálisan hasonló egy **T** felső háromszögmátrixhoz, azaz van olyan **Q** ortogonális mátrix, hogy  $\mathbf{A} = \mathbf{QTQ}^T$ .

b) Minden komplex négyzetes **A** mátrix unitéren hasonló egy **T** felső háromszögmátrixhoz, azaz van olyan **U** unitér mátrix, hogy  $\mathbf{A} = \mathbf{UTU}^H$ .

► Fontos megjegyezni, hogy a tétel valós és komplex mátrixokra vonatkozó része közt csak annyi a különbség, hogy valós mátrixoknál megköveteltük a sajátértékek valós voltát, míg komplexeknél nem tettünk semmi kikötést – sejtetően azért, mert komplex mátrix sajátértékei mindig komplexek.

**BIZONYÍTÁS.** A bizonyítást csak a valós esetre írjuk le, a komplex eset tárgyalása lényegében azonos.

Követve a 9.3. tételt, teljes indukcióval bizonyítunk.  $n = 1$  esetén az állítás nyilván igaz. A feltételek szerint **A** minden sajátértéke valós, legyen ezek egyike  $\lambda$ , a hozzá tartozó egységnyi sajátvektorok egyike  $\mathbf{u}_1$ . Innen a 9.3. tétel bizonyítását megismételjük egészen a (9.1) mátrix előállításáig, azaz kapjuk, hogy az  $\mathbf{u}_1$  vektort kiegészítve a teljes tér egy  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$  ONB-ává, az ortogonális  $\mathbf{Q}_0 = [\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_n]$  mátrixszal

$$\mathbf{Q}_0^T \mathbf{A} \mathbf{Q}_0 = \begin{bmatrix} \lambda & \mathbf{v}^T \\ \mathbf{0} & \mathbf{A}_1 \end{bmatrix} = \mathbf{B}.$$

Mivel **B** hasonló **A**-hoz, ezért sajátértékeik megegyeznek, tehát  $\mathbf{A}_1$  minden sajátértéke **A**-nak is sajátértéke. A teljes indukció miatt viszont  $\mathbf{A}_1$ -hez létezik olyan  $\mathbf{Q}_1$  ortogonális és  $\mathbf{T}_1$  felső háromszög mátrix, hogy  $\mathbf{A}_1 = \mathbf{Q}_1 \mathbf{T}_1 \mathbf{Q}_1^T$ . Ekkor viszont a  $\mathbf{Q} = \mathbf{Q}_0 \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{Q}_1 \end{bmatrix}$  mátrix ortogonális, hisz két ortogonális mátrix szorzata, és **A**-t hasonlóvá teszi egy

felsőháromszög-mátrixhoz:

$$\begin{aligned} \mathbf{Q}^T \mathbf{A} \mathbf{Q} &= \left( \mathbf{Q}_0 \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{Q}_1 \end{bmatrix} \right)^T \mathbf{A} \left( \mathbf{Q}_0 \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{Q}_1 \end{bmatrix} \right) = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{Q}_1 \end{bmatrix}^T \mathbf{Q}_0^T \mathbf{A} \mathbf{Q}_0 \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{Q}_1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{Q}_1^T \end{bmatrix} \begin{bmatrix} \lambda & \mathbf{v}^T \\ \mathbf{0} & \mathbf{A}_1 \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{Q}_1 \end{bmatrix} \\ &= \begin{bmatrix} \lambda & \mathbf{v}^T \mathbf{Q}_1 \\ \mathbf{0} & \mathbf{Q}_1^T \mathbf{A}_1 \mathbf{Q}_1 \end{bmatrix} = \begin{bmatrix} \lambda & \mathbf{v}^T \mathbf{Q}_1 \\ \mathbf{0} & \mathbf{T}_1 \end{bmatrix}. \end{aligned}$$

Ez utóbbi mátrix pedig felsőháromszög-mátrix.  $\square$

9.6. PÉLDA (SCHUR-FELBONTÁS). Hozzuk ortogonális hasonlósági transzformációval felső háromszög alakra az

$$\mathbf{A} = \begin{bmatrix} 7 & 6 & -3 \\ 3 & 17 & -6 \\ -12 & 14 & 4 \end{bmatrix}$$

mátrixot!

**MEGOLDÁS.** A karakterisztikus polinom  $-x^3 + 28x^2 - 245x + 686 = (7-x)^2(14-x)$ . A 7 kétszeres sajátérték, a sajátaltér 1-dimenziós, sajátvektor  $\mathbf{x}_1 = (2, 3, 6)$ , a 14-hez tartozó sajátvektor  $\mathbf{x}_2 = (9, 17, 13)$ , diagonalizálni nem lehet, mivel a geometriai multiplicitások összege (2) kisebb az algebraiak összegénél (3).

Az első sajátvektorhoz választunk egy ortonormált bázist, abból képezzük a  $\mathbf{Q}_0$  és a  $\mathbf{Q}_0^T \mathbf{A} \mathbf{Q}_0$  mátrixot:

$$\mathbf{Q}_0 = [\mathbf{x}_1 | \mathbf{u}_2 | \mathbf{u}_3] = \frac{1}{7} \begin{bmatrix} 2 & -6 & 3 \\ 3 & -2 & -6 \\ 6 & 3 & 2 \end{bmatrix}, \quad \mathbf{Q}_0^T \mathbf{A} \mathbf{Q}_0 = \begin{bmatrix} 7 & 0 & -21 \\ 0 & 14 & 0 \\ 0 & 7 & 7 \end{bmatrix}$$

tehát

$$\mathbf{A}_1 = \begin{bmatrix} 14 & 0 \\ 7 & 7 \end{bmatrix}.$$

A 7 sajátértékhez tartozó sajátvektor  $(0, 1)$ , rá merőleges a  $(1, 0)$ . Így

$$\mathbf{Q}_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{Q}_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Innen

$$\begin{aligned} \mathbf{Q} &= \mathbf{Q}_0 \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_1 \end{bmatrix} = \frac{1}{7} \begin{bmatrix} 2 & 3 & -6 \\ 3 & -6 & -2 \\ 6 & 2 & 3 \end{bmatrix}, \text{ és ebből} \\ \mathbf{Q}^T \mathbf{A} \mathbf{Q} &= \begin{bmatrix} 7 & -21 & 0 \\ 0 & 7 & 7 \\ 0 & 0 & 14 \end{bmatrix} \end{aligned}$$

E példa mátrixához található racionális elemű ortogonális mátrix, mi a számítások könnyebb követhetősége érdekében ezt adtuk meg. Ilyen mátrix keresése nem része a példának, mivel a gyakorlatban nemigen találkozni ilyen speciális esettel.

$\square$

**Mátrixok unitér diagonalizálása\*** A valós mátrixok ortogonális diagonalizálhatóságához hasonlóan a komplex mátrixok unitér diagonalizálhatósága is alapvető kérdés. Itt azonban a normális mátrixok játsszák a főszerepet.

9.7. DEFINÍCIÓ (UNITÉR DIAGONALIZÁLHATÓSÁG). Az  $\mathbf{A}$  mátrix unitéren diagonalizálható, ha találunk egy  $\mathbf{U}$  unitér és egy  $\mathbf{\Lambda}$  diagonális mátrixot, melyre  $\mathbf{U}^H \mathbf{A} \mathbf{U} = \mathbf{\Lambda}$  (illetve  $\mathbf{A} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H$ ).

► A szimmetrikus és önadjungált mátrixok közti analógiai alapján azt várjuk, hogy az önadjungált mátrixok lesznek az unitéren diagonalizálhatók. (A [valós spektráltétel](#) bizonyításának első része, azaz a  $\Rightarrow$  irány bizonyítása nem vihető át valós szimmetrikus mátrixokról komplex önadjungált mátrixokra, mivel  $\mathbf{\Lambda}^H = \mathbf{\Lambda}$  csak valós mátrixokra igaz.) A valóságban mátrixok egy jóval tágabb köre fog az unitéren diagonalizálhatók közé tartozni: ezek lesznek a normális mátrixok.

9.8. DEFINÍCIÓ (NORMÁLIS MÁTRIX). Azt mondjuk, hogy az  $\mathbf{A} \in \mathbb{C}^{n \times n}$  mátrix normális, ha  $\mathbf{A}^H \mathbf{A} = \mathbf{A} \mathbf{A}^H$ , azaz ha felcserélhető saját adjungáltjával.

► Könnyen ellenőrizhető, hogy normális az összes komplex önadjungált, ferdén önadjungált és unitér, valamint az összes valós szimmetrikus, ferdén szimmetrikus és ortogonális mátrix.  
► Vannak olyan normális mátrixok is, melyek nem tartoznak a fenti listában felsoroltak közé. Például az

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

mátrix normális, mert

$$\mathbf{A}^H \mathbf{A} = \mathbf{A} \mathbf{A}^H = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}.$$

► A normális mátrixok sok szép tulajdonsággal rendelkeznek, melyekre a továbbiakban visszatérünk, a legfontosabbat a következő tétel mondja ki.

9.9. TÉTEL (UNITÉR DIAGONALIZÁLHATÓSÁG). Az  $\mathbf{A} \in \mathbb{C}^{n \times n}$  mátrix pontosan akkor unitéren diagonalizálható, ha normális.

**BIZONYÍTÁS.** ( $\Rightarrow$ ) Tegyük fel, hogy  $\mathbf{A} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H$ , azaz  $\mathbf{A}$  unitéren diagonalizálható. Mivel bármely komplex  $z$  számra  $\bar{z}z = z\bar{z}$ , ezért minden

komplex diagonális mátrix normális, így  $\Lambda^H \Lambda = \Lambda \Lambda^H$ . Eszerint

$$\begin{aligned} \mathbf{A}^H \mathbf{A} &= (\mathbf{U} \Lambda \mathbf{U}^H)^H (\mathbf{U} \Lambda \mathbf{U}^H) = \mathbf{U} \Lambda^H \mathbf{U}^H \mathbf{U} \Lambda \mathbf{U}^H = \mathbf{U} \Lambda^H \Lambda \mathbf{U}^H \\ &= \mathbf{U} \Lambda \Lambda^H \mathbf{U}^H = \mathbf{U} \Lambda \mathbf{U}^H \mathbf{U} \Lambda^H \mathbf{U}^H = (\mathbf{U} \Lambda \mathbf{U}^H) (\mathbf{U} \Lambda^H \mathbf{U}^H)^H = \mathbf{A} \mathbf{A}^H. \end{aligned}$$

( $\Leftarrow$ ) A [Schur-felbontás](#) szerint minden komplex négyzetes  $\mathbf{A}$  mátrix előáll

$$\mathbf{A} = \mathbf{U} \mathbf{T} \mathbf{U}^H$$

alakban, ahol  $\mathbf{U}$  unitér,  $\mathbf{T}$  felsőháromszög-mátrix. Tegyük fel, hogy  $\mathbf{A}$  normális. Ekkor  $\mathbf{T}$  is normális, ugyanis a fenti levezetéshez hasonlóan

$$\begin{aligned} \mathbf{T}^H \mathbf{T} &= (\mathbf{U}^H \mathbf{A} \mathbf{U})^H (\mathbf{U}^H \mathbf{A} \mathbf{U}) = \mathbf{U}^H \mathbf{A}^H \mathbf{U} \mathbf{U}^H \mathbf{A} \mathbf{U} = \mathbf{U}^H \mathbf{A}^H \mathbf{A} \mathbf{U} \\ &= \mathbf{U}^H \mathbf{A} \mathbf{A}^H \mathbf{U} = \mathbf{U}^H \mathbf{A} \mathbf{U} \mathbf{U}^H \mathbf{A}^H \mathbf{U} = (\mathbf{U}^H \mathbf{A} \mathbf{U}) (\mathbf{U}^H \mathbf{A} \mathbf{U})^H = \mathbf{T} \mathbf{T}^H. \end{aligned}$$

A  $\mathbf{T}$  mátrix alakja

$$\begin{bmatrix} t_{11} & t_{12} & \dots & t_{1n} \\ 0 & t_{22} & \dots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & t_{nn} \end{bmatrix},$$

ezért  $[\mathbf{T}^H \mathbf{T}]_{11} = |t_{11}|^2$ ,  $[\mathbf{T} \mathbf{T}^H]_{11} = |t_{11}|^2 + |t_{12}|^2 + \dots + |t_{1n}|^2$ , amiből  $t_{12} = \dots = t_{1n} = 0$  adódik. Hasonlóan fölírva a  $[\mathbf{T}^H \mathbf{T}]_{22}$  és a  $[\mathbf{T} \mathbf{T}^H]_{22}$  elemeket kapjuk, hogy  $t_{23} = \dots = t_{2n} = 0$ , stb. Tehát  $\mathbf{T}$  diagonális.  $\square$

*Feladatok*

## Kvadratikus formák

---

*A csupa másodfokú tagot tartalmazó többváltozós polinomok mátrixok sajátértékeinek és sajátvektorainak ismeretében egyszerűbb alakra hozhatók, így könnyebben vizsgálhatók. E témának számtalan lineáris algebrán kívüli matematikai és matematikán kívüli alkalmazása is van.*

---

**Homogén másodfokú polinomok mátrixszorzatos alakja** Egy polinom egy tagja *másodfokú*, ha abban az ismeretlenek fokszámainak összege 2. Például az  $x$ ,  $y$  és  $z$  változókban másodfokú tagok az alábbiak:  $3x^2$ ,  $axy$ ,  $2b^3xz$ ,  $-\pi^2z^2$ . Az olyan többváltozós polinomot, melyben csak másodfokú tagok vannak, többváltozós *homogén másodfokú polinomnak* nevezzük. Például  $2x^2 + 4xy - y^2$  egy 2-változós homogén másodfokú polinom. A  $4xy = 2xy + 2yx$  felbontással e polinom egy szimmetrikus mátrixszal való mátrixszorzatos alakba írható:

$$2x^2 + 2xy + 2yx - y^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Általában is igaz, hogy

$$ax^2 + 2bxy + cy^2 = ax^2 + bxy + byx + cy^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Hasonlóképp a háromváltozós homogén másodfokú polinomok is mátrixszorzatos alakba írhatók egy szimmetrikus mátrixszal:

$$\begin{aligned} & ax^2 + 2bxy + 2cxz + dy^2 + 2eyz + fz^2 \\ &= ax^2 + bxy + cxz + byx + dy^2 + eyz + czx + ezy + fz^2 \\ &= \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}. \end{aligned}$$

**9.10. PÉLDA (MÁSODFOKÚ POLINOM MÁTRIXSZORZATOS ALAKJA).** Írjuk fel az  $x_1^2 + 2x_2^2 + 2x_3^2 - 5x_1x_2 - 3x_2x_1 + 5x_1x_3 - x_3x_1$  kifejezést mátrixszorzatos alakban!

**MEGOLDÁS.** A vegyes tagokat először összevonva, majd két egyenlő

együtthatójú részre bontva kapjuk, hogy

$$\begin{aligned}
 & x_1^2 + 2x_2^2 + 2x_3^2 - 8x_1x_2 + 4x_1x_3 \\
 &= x_1^2 + 2x_2^2 + 2x_3^2 - 4x_1x_2 - 4x_2x_1 + 2x_1x_3 + 2x_3x_1 \\
 &= x_1^2 - 4x_1x_2 + 2x_1x_3 - 4x_2x_1 + 2x_2^2 + 0x_2x_3 + 2x_3x_1 + 0x_3x_2 + 2x_3^2 \\
 &= \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 1 & -4 & 2 \\ -4 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}. \quad \square
 \end{aligned}$$

A fentieket követve az  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  vektor koordinátáitól függő homogén másodfokú polinomok mindegyike

$$\mathbf{x}^T \mathbf{A} \mathbf{x}$$

alakra hozható, ahol  $\mathbf{A}$  szimmetrikus mátrix.

**9.11. DEFINÍCIÓ (KVADRATIKUS FORMA).** Valós kvadratikus formának nevezzük azt az

$$\mathbb{R}^n \rightarrow \mathbb{R}; \mathbf{x} \mapsto \mathbf{x}^T \mathbf{A} \mathbf{x}$$

függvényt, ahol  $\mathbf{A}$  szimmetrikus mátrix. Analóg módon komplex kvadratikus forma az a

$$\mathbb{C}^n \rightarrow \mathbb{C}; \mathbf{x} \mapsto \mathbf{x}^H \mathbf{A} \mathbf{x}$$

függvény, ahol  $\mathbf{A}$  önadjungált.

E szakaszban a továbbiakban kvadratikus formán – ha mást nem mondunk – valós kvadratikus formát értünk.

**Főtengelytétel** Egy kvadratikus formához tartozó szimmetrikus mátrix diagonalizálásával a kvadratikus forma is egyszerű alakra hozható.

A **spektráltétel** szerint minden valós szimmetrikus mátrix ortogonálisan diagonalizálható, azaz létezik egy olyan ortogonális  $\mathbf{Q}$  mátrix, és egy diagonális  $\mathbf{\Lambda}$  mátrix, melyre  $\mathbf{Q}^T \mathbf{A} \mathbf{Q} = \mathbf{\Lambda}$ . Tudjuk, hogy az  $\mathbf{x} \mapsto \mathbf{A} \mathbf{x}$  mátrixleképezés mátrixa a  $\mathbf{Q}$  oszlopvektorai által alkotott  $\mathcal{Q}$  ortonormált bázisban  $\mathbf{\Lambda}$ . Ha egy tetszőleges  $\mathbf{x}$  vektor alakja e bázisban  $\mathbf{y}$ , akkor  $\mathbf{x} = \mathbf{Q} \mathbf{y}$ . E helyettesítést elvégezve ugyanennek a függvénynek a  $\mathcal{Q}$  bázisban fölírt alakját kapjuk:

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = (\mathbf{Q} \mathbf{y})^T \mathbf{A} (\mathbf{Q} \mathbf{y}) = \mathbf{y}^T \mathbf{Q}^T \mathbf{A} \mathbf{Q} \mathbf{y} = \mathbf{y}^T \mathbf{\Lambda} \mathbf{y}.$$

Eszerint a kvadratikus forma e bázisban nagyon egyszerűvé válik, csak négyzetes tagokat tartalmaz:  $\lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2$ , ahol  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ . Ezzel bizonyítottuk az alábbi tételt:

**9.12. TÉTEL (FŐTENGELYTÉTEL).** Legyen  $\mathbf{A}$  egy  $n$ -edrendű valós szimmetrikus mátrix, melyet a  $\mathbf{Q}$  mátrix ortogonálisan diagonalizál, azaz  $\mathbf{Q}^T \mathbf{A} \mathbf{Q} = \mathbf{\Lambda}$  diagonális. Ekkor az  $\mathbf{x} = \mathbf{Q} \mathbf{y}$  helyettesítés az  $\mathbf{x}^T \mathbf{A} \mathbf{x}$  kvadratikus formát

az  $\mathbf{y}^T \mathbf{A} \mathbf{y}$  kvadratikus formába transzformálja, mely kifejtve csak négyzetes tagokat tartalmaz, azaz

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{y}^T \mathbf{A} \mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2, \quad (9.2)$$

ahol  $\lambda_1, \lambda_2, \dots, \lambda_n$  az  $\mathbf{A}$  mátrix sajátértékei.

► A tétel nevét később fogjuk részletesen megmagyarázni, most csak annyit, hogy az  $\mathbf{x}^T \mathbf{A} \mathbf{x} = c$  egyenletű felületnek a  $\mathcal{Q}$  bázis vektorai mind szimmetriatengelyei, melyeket főtengelyeknek is nevezzük.

► Mivel  $\mathbf{Q}$  ortogonális mátrix, ezért  $\det \mathbf{Q} = 1$  vagy  $\det \mathbf{Q} = -1$ . Gyakorlati (például bizonyos 3-dimenziós) alkalmazásokban fontos lehet, hogy a  $\mathcal{Q}$  bázis is jobbsodrású legyen, azaz hogy  $\det \mathbf{Q} = 1$  legyen. Így a standard bázis beleforgatható az új bázisba. Ez elérhető, ha  $\det \mathbf{Q} = -1$  esetén  $\mathbf{Q}$  bármelyik oszlopát  $-1$ -szeresére változtatjuk. Ez a kvadratikus formát nem befolyásolja, hisz abban csak a sajátértékek szerepelnek.

► A főtengetytétel alkalmazását egy kvadratikus formán *főtengely-transzformációnak* nevezzük.

9.13. PÉLDA (FŐTENGETY-TRANSZFORMÁCIÓ). Végezzük el a főtengety-transzformációt az

$$f(x, y, z) = x^2 + 2y^2 + 2z^2 - 8xy + 4xz$$

kvadratikus formán. Keressünk olyan jobbsodrású ortonormált bázist, melyben épp ez a kvadratikus forma alakja. Mi az áttérés mátrixa?

**MEGOLDÁS.** A kvadratikus forma mátrixszorzat-alakja

$$\begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 1 & -4 & 2 \\ -4 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Mátrixának karakterisztikus polinomja  $-\lambda^3 + 5\lambda^2 + 12\lambda - 36$ , ennek gyökei, azaz a sajátértékek 6,  $-3$ , 2, a hozzájuk tartozó sajátvektorok rendre  $(2, -2, 1)$ ,  $(-5, -4, 2)$ ,  $(0, 1, 2)$ . Így a keresett kvadratikus forma

$$\begin{bmatrix} \xi & \eta & \zeta \end{bmatrix} \begin{bmatrix} 6 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} \xi \\ \eta \\ \zeta \end{bmatrix} = 6\xi^2 - 3\eta^2 + 2\zeta^2.$$

A sajátvektorokat normálva megkapjuk az ortonormált bázist, melynek vektoraiból képzett determináns

$$\begin{vmatrix} \frac{2}{3} & -\frac{5}{3\sqrt{5}} & 0 \\ -\frac{2}{3} & -\frac{4}{3\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{3} & \frac{2}{3\sqrt{5}} & \frac{2}{\sqrt{5}} \end{vmatrix} = -1,$$



tehát egy megfelelő ortonormált bázis:  $(\frac{2}{3}, -\frac{2}{3}, \frac{1}{3}), (\frac{5}{3\sqrt{5}}, \frac{4}{3\sqrt{5}}, -\frac{2}{3\sqrt{5}}), (0, \frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}})$ .  $\square$

**Kvadratikus formák és mátrixok definitisége** A főtengetlytétel könnyen áttekinthetővé teszi a kvadratikus forma által fölvetett értékek lehetséges előjelét. Ez lehetővé teszi a kvadratikus formák egy fontos osztályozását.

**9.14. DEFINÍCIÓ (KVADRATIKUS FORMÁK ÉS MÁTRIXOK DEFINITSÉGE).**

Azt mondjuk, hogy az  $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$  kvadratikus forma

a) pozitív definit, ha  $f(\mathbf{x}) > 0$ ,

b) pozitív szemidefinit, ha  $f(\mathbf{x}) \geq 0$ ,

c) negatív definit, ha  $f(\mathbf{x}) < 0$ ,

d) negatív szemidefinit, ha  $f(\mathbf{x}) \leq 0$

bármely  $\mathbf{x} \neq \mathbf{0}$  vektor esetén, és azt mondjuk, hogy  $f$

e) indefinit, ha pozitív és negatív értékeket is fölvesz.

A szimmetrikus  $\mathbf{A}$  mátrixot pozitív/negatív definitnek/szemidefinitnek, illetve indefinitnek nevezzük, ha a hozzá tartozó kvadratikus forma az.

► Ha  $\mathbf{A}$  negatív definit, akkor  $-\mathbf{A}$  pozitív definit. Hasonló állítás igaz a szemidefinitiségre is.

► Ha  $\mathbf{A} = [a]$ , azaz  $\mathbf{A}$   $1 \times 1$ -es, akkor  $\mathbf{A}$  pontosan akkor pozitív definit, ha  $a > 0$ .

► Az identikus mátrix pozitív definit, ugyanis  $\mathbf{x}^T \mathbf{I} \mathbf{x} = \mathbf{x}^T \mathbf{x} = |\mathbf{x}|^2$ , ami pozitív, ha  $\mathbf{x} \neq \mathbf{0}$ .

► Tetszőleges  $\mathbf{A}$  valós mátrix esetén  $\mathbf{A}^T \mathbf{A}$  pozitív szemidefinit, ugyanis  $\mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} = (\mathbf{A} \mathbf{x})^T (\mathbf{A} \mathbf{x}) = |\mathbf{A} \mathbf{x}|^2 \geq 0$ .

► Világos, hogy ha egy kvadratikus formában csak négyzetes tagok szerepelnek, akkor azonnal leolvasható definitiségének típusa. Például az  $f(x, y) = x^2 + 2y^2$ ,  $g(x, y) = x^2 - 2y^2$ ,  $h(x, y) = -x^2 - 2y^2$ ,  $k(x, y, z) = x^2 + 2y^2$  formákról látható, hogy  $f$  pozitív definit, hisz az  $(x, y) \neq (0, 0)$  esetén értéke mindig pozitív,  $g$  indefinit,  $h$  negatív definit, és  $k$  pozitív szemidefinit, hisz értéke  $(x, y, z) \neq (0, 0, 0)$  esetén is lehet 0 (ha  $x = y = 0$ , de  $z \neq 0$ ). Miután a főtengetlytétel szerint minden kvadratikus forma egyenlő a változók négyzeteinek a sajátértékekkel vett lineáris kombinációjával, ezért a definitség típusa pusztán csak a sajátértékek előjeleinek ismeretében eldönthető.

**9.15. PÉLDA (DEFINITSÉG MEGHATÁROZÁSA A SAJÁTÉRTÉKEKBŐL).** Határozzuk meg az

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}$$

mátrixok definitségének típusát!

**MEGOLDÁS.** Az  $\mathbf{A}$  mátrix sajátértékei 1, 1 és 4. Így a főtengety-transzformáció után kapott

$$\begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \xi & \eta & \zeta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} \xi \\ \eta \\ \zeta \end{bmatrix} = \xi^2 + \eta^2 + 4\zeta^2$$

alakból látható, hogy e kvadratikusan minden értéke pozitív, ha a változók nem mindegyike 0. Tehát e kvadratikusan pozitív definit. Hasonlóképp a  $\mathbf{B}$  sajátértékei  $-1$ ,  $-1$  és  $2$ , a főtengety-transzformáció után kapott alak  $-\xi^2 - \eta^2 + 2\zeta^2$ . E forma negatív értéket vesz fel például az  $(1, 0, 0)$  helyen, és pozitív a  $(0, 0, 1)$  helyen, tehát indefinit. Végül  $\mathbf{C}$  sajátértékei  $-3$ ,  $-3$  és  $0$ , így a főtengety-transzformáció után kapott alak  $-3\xi^2 - 3\eta^2 + 0\zeta^2 = -3\xi^2 - 3\eta^2$ . Ennek értéke a  $(0, 0, 1)$  helyen 0, és pozitív értéket nem vesz fel, tehát negatív szemidefinit.  $\square$

**9.16. TÉTEL (DEFINITISÉG MEGHATÁROZÁSA A SAJÁTÉRTÉKEKBŐL).** A valós szimmetrikus  $\mathbf{A}$  mátrix, illetve az  $\mathbf{x}^\top \mathbf{A} \mathbf{x}$  kvadratikusan forma pontosan akkor

- a) pozitív definit, ha  $\mathbf{A}$  minden sajátértéke pozitív;
- b) pozitív szemidefinit, ha  $\mathbf{A}$  minden sajátértéke nemnegatív;
- c) negatív definit, ha  $\mathbf{A}$  minden sajátértéke negatív;
- d) negatív szemidefinit, ha  $\mathbf{A}$  minden sajátértéke nempozitív;
- e) indefinit, ha  $\mathbf{A}$ -nak van pozitív és negatív sajátértéke is.

**Definitesség és főminorok** Mátrix definitisége gyakran könnyen eldönthető főminorainak vagy vezető főminorainak értékéből.

Válasszuk ki egy négyzetes mátrix néhány sorát, és ugyanannyiadik sorszámu oszlopát, a többi sort és oszlopot hagyjuk el. Az így kapott négyzetes részmátrix determinánsát a mátrix *főminorának* nevezzük. Ha az első  $k$  sort és az első  $k$  oszlopot választjuk ki, *vezető főminoráról* beszélünk, pontosabban a  $k$ -adrendű vagy  $k$ -adik vezető főminoráról.

Ha egy mátrix diagonális alakú és pozitív definit, azaz minden sajátértéke pozitív, akkor minden főminora is pozitív, ha pozitív szemidefinit, akkor minden főminora nemnegatív. Ha e diagonális mátrix minden sajátértéke negatív, akkor vezető főminorai felváltva  $- + - + - + \dots$  előjelűek. E megfigyelések átvihetők nem diagonális alakú mátrixokra is.

**9.17. TÉTEL (DEFINIT MÁTRIXOK FŐMINORAI ÉS VEZETŐ FŐMINORAI).** A valós szimmetrikus  $\mathbf{A}$  mátrix, illetve az  $\mathbf{x}^\top \mathbf{A} \mathbf{x}$  kvadratikusan forma pontosan akkor

- a) pozitív definit, ha  $\mathbf{A}$  minden főminora pozitív;
- b) pozitív szemidefinit, ha  $\mathbf{A}$  minden főminora nemnegatív;
- c) pozitív definit, ha  $\mathbf{A}$  minden vezető főminora pozitív;

d) negatív definit, ha  $\mathbf{A}$  minden páratlan rendű vezető főminora negatív, páros rendű vezető főminora pozitív.

BIZONYÍTÁS. ... □

► Az állítás vezető főminorokra vonatkozó pontjai nem vihetők át minden további feltétel nélkül szemidefinit mátrixokra. Az ugyan igaz, hogy ha egy mátrix pozitív szemidefinit, akkor vezető főminorainak sorozata egy darabig pozitív, majd onnan 0. Ennek megfordítása viszont már nem igaz.

**Szélsőérték** A többváltozós függvények szélsőértékének első és második parciális deriváltjaira vonatkozó feltételei azonnal érthetővé válnak a definitség fogalmának segítségével.

A függvényanalízisből ismert Taylor tételének többváltozós alakja szerint egy legalább kétszer differenciálható  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  függvény a következő alakban írható fel:

$$f(\mathbf{x}) = f(\mathbf{a}) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{a})(x_i - a_i) + \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a})(x_i - a_i)(x_j - a_j) + \sum_{i,j} \varepsilon_{ij}(\mathbf{x})(x_i - a_i)(x_j - a_j),$$

ahol  $\mathbf{a} \in \mathbb{R}^n$  az  $f$  értelmezési tartományának egy belső pontja.

---

**Megoldások**



## *Irodalomjegyzék*

Wolf Holzmann. Uniqueness of reduced row echelon form. <http://www.cs.uleth.ca/~holzmann/notes/reduceduniq.pdf>, 2002.



# Tárgymutató

$\pi$ -transzponált 166

adjungált 257, 343

affin altér 128

alakzat egyenletrendszere 64

alapvektorok 41

algebrai multiplicitása 376

alsó háromszögmátrix 208

általános megoldás 91

altér 124

    affin 128

    invariáns 392

    kiegészítő 302

    komplementer 302

    merőleges kiegészítő 141

alterek merőlegessége 140

altér eltoltja 128

alulhatározott 79

ASCII-kód 55

áttérés mátrixa 174

balrendszer 34

bal sajátvektor 385

bázis 41, 133

    standard 134

báziscsere 173

bázisfelbontás 175

bázisoszlop 88

bázisvektorok 41

BCD-kód 55

bináris reláció 39

bitvektor 55

blokkmátrix 162

bővített mátrix 81

csoport 23

deriváltleképezés 284

determináns 233

DFT 350

diád 170

diadikus szorzat 170

diagonalizálhatóság 383

differentiálhatóság 283

dimenzió 137

direkt összeg 304

diszkrét Fourier-összeg 347

diszkrét Fourier-transzformáció 350

egyenletrendszer

    numerikusan instabil 104

egységmátrix 176

egységvektor 32

egyűthetőségmátrix 81

Einstein-konvenció 189

ekvivalenciareláció 39

ekvivalens

    átalakítások 79

    lineáris egyenletrendszerek 79

elemi mátrix 176

elemi sorműveletek 88

ellenőrző összeg 58

előjeles aldetermináns 250

előjelestérfoogat 37

előjeles terület 231

euklideszi norma 21

explicit 65

fejléc (táblázat) 155

felső háromszögmátrix 208

ferdén szimmetrikus 209

FFT 354

főátló 80

főelem 88

főminor 418

főoszlop 88

forogatónyomaték 34

Fourier-mátrix 348

Fourier-összeg 347

Gauss–Jordan-módszer 94

Gauss–Seidel-iteráció 110

Gauss-módszer 89

geometriai multiplicitás 376

Givens-forgatás 331

gradiens 286

Gram-mátrix 147

gyors Fourier-transzformáció 354

hajlásszög 51

Hamming-kód 100

háromszögmódszer 21

hasonló mátrixok 280

hatványmódszer 399

Hermite, Charles 343

Hermite mátrix 346

hiper-kockamátrix 165

hipermátrix 165

    ferdén szimmetrikus 166

    kontrakciós szorzat 185

    külső szorzat 167

    szimmetrikus 166

hipersík 75

homogén lineáris egyenletrendszer

    inhomogénhez tartozó 91

Householder-módszer 338

Householder-tükrözés 332

idempotens 307

implicit 64

inkonzisztens 79

invariáns altér 392

invertálható 194

invertálható művelet 193

inverz

    elemé 193

inverzió 238  
 irányított szakasz 19  
 irányított szög 34  
 irányvektor 65  
 ISO 31-11 20

Jacobi-determináns 289  
 Jacobi-iteráció 110  
 Jacobi-mátrix 285  
 jobbrendszer 34  
 jól kondicionált 104

karakterisztikus egyenlet 369  
 karakterisztikus polinom 369  
 képtér 271  
 kernel 271  
 kiegészítő altér 302  
 kifeszített altér 126  
 kígyó 207, 247  
 kitüntetett altér 141  
 klasszikus adjungált 257  
 kód  
   hossza 56  
   kódszó 56  
 kódvektor 56  
 kollineáris vektor 21  
 komplanáris 22  
 komplementer altér 302  
 kompozíció  
   lineáris helyettesítéseké 158  
 konjugált 281  
 konstans tag 77  
 kontrakciós szorzat 185  
 konzisztens 79  
 koordináta 41  
 koordinátarendszer 41  
 korrelációs együttható 54  
 kötött változó 90  
 kötött vektor 19  
 Kronecker-szorzat 164  
 külső szorzat 167  
   Segre-féle 169  
 kvadratikus forma 415  
 legjobb közelítés 307  
 legkisebb négyzetek elve 309  
 lépcsős alak 88  
 levéldiagram 124  
 lineáris  
   egyenlet 77  
   egyenletrendszer 78

kombináció 23  
 lineárisan független 25, 49  
 lineárisan összefüggő 25  
 lineáris egyenletrendszer  
   konzisztens 79  
 lineáris egyenletrendszer  
   alulhatározott 79  
   túlhatározott 79  
 lineáris egyenletrendszerek  
   homogén 78  
   megoldása 79  
 lineáris egyenletrendszerek  
   ekvivalens 79  
   inhomogén 78  
 lineáris helyettesítés  
   mátrixa 172  
 lineáris helyettesítés 157  
 lineáris leképezés 274  
   képtere 271  
   magtere 271  
 lineáris transzformáció  
   karakterisztikus polinomja 383  
   sajátértéke 381  
   sajátértékei 383  
 lineáris transzformáció 274  
 LU-felbontás 215  
 magtér 271  
 másodfokú tag 414  
 mátrix 80, 158  
   áttérés mátrixa 174  
   diagonális 159  
   elemi 176  
   ellentettje 160  
   ferdén szimmetrikus 209  
   négyzetes 159  
   normális 411  
   önadjungált 346  
   ortogonális 326  
   rangja 120  
   ritka 81  
   soronként domináns főátlójú 112  
   sűrű 81  
   szemiortogonális 326  
   szimmetrikus 209  
   szinguláris 194  
   teljes oszloprangú 305  
 mátrixleképezés 271  
 mátrixok tere 159  
 mátrixszorzat  
   diádok összegére bontása 179

megoldás  
   általános 91  
   partikuláris 91  
   triviális 122  
 megoldásvektor 79  
 megoldható 79  
 merőleges összetevő 305  
 merőleges vetület  
   altérre eső 305  
 Moore–Penrose-féle pszeudoinverz 316  
 multilineáris mátrixszorzat 169  
 negatív (szemi)definit 417  
 négy kitüntetett altér 141  
 nilpotens 194  
 norma  
   euklideszi 21  
 normálás 324  
 normálegyenlet 309  
 normálegyenlet-rendszer 309  
 normális mátrix 411  
 nullosztó 188  
 nulltér 126  
 nullvektor 20  
 numerikusan instabil 104  
 numerikusan stabil 104  
 önadjungált 346  
 optimális megoldás 309  
 origó 20  
 ortogonális 44  
 ortogonális bázis (OB) 324  
 ortogonális diagonalizálás 405  
 ortogonális mátrix 326  
 ortonormált bázis 44  
 ortonormált bázis (ONB) 324  
 oszlopmátrix 81  
 oszloptér 129  
 oszlopvektor 42, 81  
 osztályozás 39  
 paralelepipedon 27  
   előjeles térfogata 37  
 parallelogramma 27  
   előjeles területe 231  
 párhuzamos vektor 21  
 paritásbit 58  
 partícionálás 39  
 partikuláris megoldás 91  
 permutációmátrix 207



- permutáló mátrix 207
- pivotelem 88
- PLU-felbontás 221
- polinom
  - elemi szimmetrikus 383
  - homogén másodfokú 414
- pozitív (szemi)definit 417
- precedencia-elv 190
- projekció 313
- pszeudoinverz 316
- QR-felbontás 335
  - redukált 335
  - teljes 335
- rang 120, 138
  - lineáris leképezése 281
- redukált lépcsős alak 93
- reflexív 40
- regressziós egyenes 311
- reláció 39
- részleges főelemkiválasztás 106
- részleges pivotálás 106
- ritka mátrix 81
- rosszul kondicionált 104
- rref függvény 97
- sajátaltér 368
- sajátérték 368
  - lineáris transzformációé 381
- sajátfelbontás 384
  - diadikus alakja 385
- sajátpár 368
- sajátvektor 368
  - bal 385
- sakktáblaszabály 250
- Sarrus-szabály 249
- skalár 19
- skaláris szorzat 29
- skalázás 107
- sorlépcsős alak 88
- sormátrix 81
- soronként domináns főátló 112
- sortér 129
- sorvektor 81
- standard bázis 134
- sudoku 182
- szabad változó 90
- szabad vektor 20
- személyi szám 55
- szemiortogonális mátrix 326
- szigorúan domináns saját pár 398
- szimmetrikus mátrix 209
- szimmetrikus reláció 40
- szimultán egyenletrendszer 97
- szinguláris 194
- szög 51
- táblázat 155
- távolság 30
  - altértől 307
- tenzorszorzat 164
- torzor 23
- transzponált 137
  - Hermite-féle 343
- transzverzális 207
- transzverzális (kígyó) 247
- tranzitív reláció 40
- triviális megoldás 122
- túlhatározott 79
- unitér 346
- unitér diagonalizálás 411
- Vandermonde-determináns 254
- Vandermonde-mátrix 254
- vec függvény 164
- vegyes szorzat 38
- vektor 20
  - abszolút értéke 21, 346
  - azonos irányú 21
  - egyirányú 21
  - ellenkező irányú 21
  - hossza 21, 30, 346
  - jelölése 20, 81
  - kollineáris 21
  - koordinátáinak elválasztása 81
  - koordinátás alakja 41
  - mátrix alakja 81
  - normálása 324
  - összeg 21
  - párhuzamos 21
- vektoregyenlet 64
- vektori szorzat 35
- vektorok
  - merőlegessége 51, 346
  - szöge 30, 51, 346
  - távolsága 51, 346
- vetítés 313
- vetület 312
- vezető főminor 418
- zérustér 125
- zérusvektor 20