

Tartalomjegyzék

Bevezetés 17

A könyvben követett elvek 18

A könyv felépítése 21

Szoftverek 23

I. A lineáris algebra forrásai 25

1 Vektorok 29

Vektorok a 2- és 3-dimenziós térben 29

Írányított szakasz, kötött és szabad vektor 29 • Vektor magadása egy irányított szakasszal 30 • Vektor megadása hossz és irány segítségével 31 • Vektorműveletek a 2- és 3-dimenziós térben 31
• A lineáris kombináció definíciója 33 • Lineáris függetlenség 35 • Speciális lineáris kombinációk 36

Távolság, szög, orientáció 39

Skaláris szorzás 39 • Hosszúság és szög 40
• Pithagorász-tétel 40 • Két fontos egyenlőtlenség 41
• Egységvektorral való szorzás és a merőleges vetítés 42
• Merőlegesség és orientáció 43 • Vektori szorzás 44
• Parallelepipedon térfogata, és előjeles térfogata 47 • Vegyes szorzat 47

Vektorok koordinátás alakban 50

Descartes-féle koordináta-rendszer 50 • Műveletek koordinátás alakban megadott vektorokkal 51 • A derékszögű koordináta-rendszer 53 • Az $\mathbb{R}^n$ halmaz 55 • $\mathbb{R}^n$ vektorainak összeadása és skalárral szorzása 55 • Lineáris kombináció, lineáris függetlenség, lineáris összefüggőség 57 • Skaláris szorzás $\mathbb{R}^n$ -ben 59 • Távolság és szög $\mathbb{R}^n$ -ben 60 • Korrelációs együttható 62 • Bitvektorok 63 • Kódvektorok, kódok 63
• Vektorműveletek $\mathbb{Z}_m^n$ -ben 64

2	<i>Lineáris egyenletrendszerek és megoldásuk</i>	69
	<i>Egyenes és sík egyenletei</i>	69
	Alakzatok és egyenletek 69 • Síkbeli egyenes egyenletei 71	
	• Síkbeli pont egyenletei 74 • A 3-dimenziós tér síkjainak egyenletei 75 • Térbeli egyenes egyenletei 77 • Térbeli pont egyenletei 80 • Egyenletek $\mathbb{R}^n$ -ben 81	
	<i>A lineáris egyenletrendszer és két modellje</i>	84
	Lineáris egyenlet és egyenletrendszer 84 • Ekvivalens lineáris egyenletrendszerek 86 • Mátrixok 87 • Egyenletrendszer mátrixa és bővített mátrixa 88 • Sormodell: hipersíkok metszete 89 • Oszlopmodell: vektor előállítása lineáris kombinációként 92	
	<i>Megoldás kiküszöböléssel</i>	95
	Elemi sorműveletek 95 • Lépcsős alak 95 • Gauss-módszer 96	
	• Redukált lépcsős alak 100 • Gauss–Jordan-módszer 101	
	• A redukált lépcsős alak egyértelműsége 103 • Szimultán egyenletrendszerek 104 • Kiküszöbölés $\mathbb{Z}_p$ -ben* 106	
	<i>Megoldás a gyakorlatban</i>	109
	A kiküszöbölés műveletigénye 109 • Numerikusan instabil egyenletrendszerek 109 • Részleges főelemkiválasztás 111	
	• Skálázás 113 • Iteratív módszerek 114 • Jacobi-iteráció 115	
	• Gauss–Seidel-iteráció 116 • Az iterációk konvergenciája 117	
3	<i>Megoldhatóság és a megoldások tere</i>	121
	<i>Homogén és inhomogén egyenletrendszerek megoldásai</i>	121
	Kötött változók száma, mátrix rangja 121 • Egyenletrendszer megoldhatóságának feltétele 123 • Homogén lineáris egyenletrendszer megoldásai 125 • Altér 126 • Kifeszített altér 128 • Az inhomogén lineáris egyenletrendszer megoldásai 130 • Lineáris függetlenség és összefüggőség 132	
	<i>Altér tulajdonságai és az egyenletrendszerek</i>	135
	Sor- és oszloptér 135 • Bázis 136 • Vektor egy bázisra vonatkozó koordinátás alakja 138 • Dimenzió és rang 140	
	• Elemi bázistranszformáció* 143	
	<i>A lineáris algebra alaptétele</i>	147
	A sortér és a nulltér merőlegessége 147 • Kiegészítő altér 148	
	• A lineáris egyenletrendszer megoldásainak jellemzése 151	
	<i>Megoldások</i>	155

II. Mátrixok algebrája és geometriája 161

4 *Mátrixműveletek definíciói* 165

Táblázatok 165

- Táblázatok összeadása 165 • Táblázat szorzása számmal 166
- Táblázatok szorzása 166 • Lineáris helyettesítés 167

Elemenkénti mátrixműveletek 170

- Alapfogalmak, jelölések 170 • Elemenkénti mátrixműveletek 172
- Mátrixok lineáris kombinációi 173

Mátrixszorzás 175

- Skaláris szorzat és diadikus szorzat mátrixszorzatos alakja 176
- Lineáris egyenletrendszer mátrixszorzatos alakja 177
- Lineáris helyettesítés mátrixszorzatos alakja 178 • Szorzás vektorral 179 • Szorzás standard egységvektorral 179 • A báziscsere mátrixszorzatos alakja 180 • Bázisfelbontás* 182
- Egységmátrix, elemi mátrixok 183 • Mátrixműveletek $\mathbb{Z}_m$ -ben* 185

Blokkmátrixok 185

- Műveletek blokkmátrixokkal 185 • Vektorokra partícionált mátrixok 187 • Lineáris egyenletrendszer megoldásának blokkmátrix alakja* 190

5 *Mátrixműveletek tulajdonságai* 195

Az alpműveletek algebrai tulajdonságai 195

- Az összeadás és a skalárral való szorzás tulajdonságai 195 • A szorzás tulajdonságai 196 • Mátrix hatványozása 198 • A transzponálás tulajdonságai 200

Mátrix inverze 201

- Az inverz 201 • Elemi mátrixok inverze 204 • Az inverz kiszámítása 205 • Az inverz tulajdonságai 207 • Az invertálhatóság és az egyenletrendszerek megoldhatósága 209
- Invertálhatóság, bázis, báziscsere 212

Műveletek speciális mátrixokkal 216

- Diagonális mátrixok 216 • Permutációs mátrixok és kigyók 216
- Háromszögmátrixok 218 • Szimmetrikus és ferde szimmetrikus mátrixok 219 • Mátrix és diád összegének inverze* 220 • Gyorsszorítás* 222

Az LU-felbontás 225

- Az LU-felbontás használata egyenletrendszer megoldására 226
- Mátrix invertálása LU-felbontással 227 • Az LU-felbontás kiszámítása 228 • PLU-felbontás 230 • Az LU-felbontás a gyakorlatban 233

Megoldások 235

6 *Determináns* 239

Parallelogramma előjeles területe 239 • Parallelepipedon előjeles térfogata 240

A determináns, mint sorvektorainak függvénye 241

A determináns definíciója 241 • A determináns értékének kiszámítása 243 • Mátrixműveletek és determináns 246
• Mikor 0 a determináns értéke 248

A determináns, mint elemeinek függvénye 254

Kígyók determinánsa 254 • Permutációs mátrix determinánsa* 256 • Előjeles aldetermináns 258 • Determináns kifejtése 261 • Cramer-szabály és a mátrix inverze 262
• Blokkmátrixok determinánsa* 266
• Vandermonde-determináns 267

Megoldások 273

7 *Mátrixleképezések és geometriájuk* 279

Mátrixleképezés, lineáris leképezés 279

A mátrixleképezés fogalma 279 • Műveletek mátrixleképezések között 280 • Mátrixleképezések tulajdonságai 281 • A mátrixleképezés hatásának szemléltetései 282 • Lineáris leképezés 285 • Lineáris leképezések alaptulajdonságai 288
• Lineáris leképezés mátrixa különböző bázisokban 289
• Hasonlóság 290 • Tartományok képe és mértékük változása 292 • Többváltozós függvények differenciálása* 293

2- és 3-dimenziós geometriai transzformációk mátrixa 301

Forgatás 301 • Merőleges vetítés 304 • Tükrözés 306
• Vetítés 306 • Eltolás 307

Merőleges vetítés és a legjobb közelítés 308

Merőleges vetítés $\mathbb{R}^n$ egy alterére 308 • Melyik mátrix merőleges vetítés mátrixa? 309 • Altértől való távolság 310
• Egyenletrendszer optimális megoldása 312 • Lineáris és polinomiális regresszió 313 • Vetítés 316

*Pszeudoinverz** 320

A pszeudoinverz fogalma 320 • A pszeudoinverz tulajdonságai 323 • A pszeudoinverz és a minimális abszolút értékű optimális megoldás 324

Ortonormált bázis, ortogonális mátrixok, ortogonalizáció 328

Ortogonalis és ortonormált bázis 328 • Ortogonalis mátrixok 330 • Ortogonalis mátrixok geometriája 332 • A 2- és 3-dimenziós tér ortogonalis transzformációi 333
• Givens-forgatás, Householder-tükrözés* 335
• Gram-Schmidt-ortogonalizáció* 337 • A QR-felbontás* 338
• Egyenletrendszer optimális megoldása QR-felbontással* 343

*Komplex és véges test feletti terek** 346

Komplex vektorok skaláris szorzata 346 • Önadjungált mátrixok 348 • Távolság és a merőleges vetítés komplex terekben 349 • Unitér mátrixok 349 • Fourier-mátrixok 349 • Diszkrét Fourier-transzformáció 352 • Periodikus összetevők szűrése 354 • Gyors Fourier-transzformáció 356 • Vektorok konvolúciója 359

Megoldások 359

III. Mátrixok sajátosságai 365

8 *Sajátérték, diagonalizálás* 369

Sajátérték, sajátvektor, sajátaltér 369

A sajátérték és a sajátvektor fogalma 369 • Karakterisztikus polinom 371 • A valós 2×2 -es mátrixok sajátaltereinek jellemzése 373 • Mátrix összes sajátértékének és sajátvektorának meghatározása 374 • A karakterisztikus egyenlet komplex gyökei 377 • A karakterisztikus egyenlet többszörös gyökei: az algebrai és a geometriai multiplicitás 378 • Sajátértékek és a mátrix hatványai 379 • Speciális mátrixok sajátértékei 381

Hasonlóság, diagonalizálhatóság 384

Lineáris transzformációk sajátértékei 384 • Hasonló mátrixok sajátértékei 385 • Mátrixok diagonalizálása 386 • Sajátértékek multiplicitása és a diagonalizálhatóság* 389 • Diagonalizálható mátrixok spektrálfelbontása* 392 • Mátrixok hatványai és egyéb függvényei 393 • Mátrixok ortogonális diagonalizálása, valós spektráltétel 394 • Spektráltétel önortogonális mátrixokra 397 • Mátrixok unitér diagonalizálása 397

Kvadratikus formák 397

Homogén másodfokú polinomok mátrixszorzatos alakja 397 • Főtengelytétel 399 • Kvadratikus formák és mátrixok definitisége 400 • Kúpszeletek osztályozása 401 • Definitiség és sajátértékek 401 • Szélsőérték 401 • Szélsőérték az egységgömbön 402

9 *Szinguláris érték* 403

Szinguláris érték, szinguláris vektor, SVD 403

Szinguláris érték és szinguláris vektorok 403 • Szinguláris felbontás 404 • A szinguláris értékek és a szinguláris felbontás meghatározása 407 • Szinguláris érték szerinti felbontás létezése 409 • Szinguláris felbontás geometriai

interpretációja 410 • Polárfelbontás 411 • Pszeudoinverz 413
 • Információtömörítés 414

Vektor- és mátrixnorma 416

Vektor abszolút értéke – az euklideszi norma 416 • A
 p -norma 416 • A norma általános fogalma 417
 • Vektornormák ekvivalenciája 419 • Vektornormák
 mátrixokon 420 • A mátrixnorma általános fogalma 422
 • Indukált norma 422 • Az 1-, 2- és ∞ -norma mátrixokra 424

Megoldások 427

10 *Jordan-féle normálalak* 429

Schur-felbontás 429 • Általánosított sajátvektorok és a
 Jordan-blokk 429 • Jordan normálalak 433 • A Jordan-alak
 egyértelműsége 435 • A Jordan-bázis konstrukciója 439
 • Mátrixfüggvények 444 • A Jordan normálalak használata a
 differenciálegyenletrendszerek megoldásában 445

11 *Nemnegatív mátrixok* 423

Mátrixok összehasonlítása 423

Pozitív mátrixok 424

Nemnegatív mátrixok 427

Irreducibilis mátrixok 431

Megoldások 434

A *Függelék* 437

Lebegőpontos számábrázolás 437

A lebegőpontos számábrázolás 437 • Műveletek lebegőpontos
 számokkal 439 • Algoritmusok műveletigénye: flop és flops 440

Komplex számok 442

Testek, gyűrűk 442

Prímelemű testek 445

Aritmetika véges halmazon 445

Polinomok 447

B *Lineáris algebra dióhéjban* 449

Irodalomjegyzék 451

Tárgymutató 453

Listák

Tételek, állítások, következmények

1.2. Parallelogramma-módszer	32	2.6. Síkbeli egyenes implicit vektoregyenlete	72
1.5. A vektorműveletek tulajdonságai	33	2.7. Síkbeli egyenes explicit egyenletrendszere	72
1.7. Vektorral párhuzamos vektorok	34	2.8. Síkbeli egyenes (implicit) egyenlete	72
1.8. Két vektorral egy síkba eső vektorok	34	2.10. Sík explicit vektoregyenlete	75
1.9. Térbeli vektorok	35	2.11. Sík implicit vektoregyenlete	75
1.11. Síkbeli vektor felbontása	36	2.12. Sík explicit egyenletrendszere	76
1.12. Térbeli vektor felbontása	36	2.13. Sík implicit egyenlete	76
1.13. Két ponton átmenő egyenes jellemzése	36	2.15. Térbeli egyenes explicit vektoregyenlete	77
1.14. Intervallum pontjainak jellemzése	37	2.16. Térbeli egyenes explicit egyenletrendszere	78
1.17. Mikor 0 a skaláris szorzat?	39	2.17. Térbeli egyenes implicit egyenletrendszere	78
1.18. A skaláris szorzás műveleti tulajdonságai	40	2.29. Ekvivalens átalakítások	86
1.19. Pithagorász-tétel	40	2.34. Sormodell	92
1.21. Cauchy–Bunyakovszkij–Schwarz-egyenlőtlenség	41	2.36. Oszlopmodell	93
1.22. Háromszög-egyenlőtlenség	41	2.42. Lépcsős alakra hozás	98
1.23. Egységvektorral való szorzás geometriai jelentése	42	2.50. A redukált lépcsős alak egyértelmű	103
1.24. Vektor felbontása merőleges összetevőkre	42	2.56. A kiküszöbölés műveletigénye	109
1.29. Mikor 0 a vektori szorzat?	46	2.65. Elégséges feltétel az iterációk konvergenciájára	118
1.30. Vektori szorzat abszolút értékének geometriai jelentése	46	3.1. Főelemek oszlopai	121
1.31. Vektori szorzás műveleti tulajdonságai	46	3.4. Köötött és szabad változók száma	122
1.35. Ekvivalencia reláció	49	3.6. A megoldhatóság mátrixrangos feltétele	123
1.38. Vektorműveletek koordinátás alakja	52	3.7. Homogén lineáris egyenletrendszer megoldhatósága	124
1.40. Skaláris szorzat ortonormált koordinátarendszerben	53	3.9. Megoldások lineáris kombinációja	125
1.41. Vektori szorzat ortonormált koordinátarendszerben	54	3.11. Alterek összege	127
1.44. Az összeadás és skalárral szorzás tulajdonságai	56	3.13. Megoldások altere	128
1.46. Lineáris függetlenség	57	3.16. A kifeszített altér altér	129
1.47. Lineáris összefüggőség	59	3.18. Homogén és inhomogén egyenletrendszer megoldásai	130
1.48. A skaláris szorzás tulajdonságai	59	3.20. Inhomogén egyenletrendszer megoldhatósága	131
1.51. Cauchy–Bunyakovszkij–Schwarz-egyenlőtlenség	61	3.22. Lineáris függetlenség eldöntése	132
1.52. Háromszög-egyenlőtlenség $\mathbb{R}^n$ -ben	61	3.24. Elemi sorműveletek hatása a sor- és oszlopvektorokra	135
1.53. Skaláris szorzat és abszolút érték $\mathbb{R}^n$ -ben	62	3.25. Mátrix lépcsős alakjának vektorai	136
2.5. Síkbeli egyenes explicit vektoregyenlete	71	3.29. Bázis ekvivalens definíciói	138
		3.31. Bázis-tétel	140
		3.34. Dimenzió = rang	141
		3.37. Dimenziótétel	142
		3.38. Elemi bázistranszformáció	144
		3.41. A sortér és a nulltér merőlegessége	148

3.42. Kiegészítő alterek tulajdonságai	149	6.25. Cramer-szabály	263
3.43. A merőleges kiegészítő alter tulajdonságai	150	6.27. Mátrix inverzének elemei	264
3.44. A lineáris algebra alaptétele	151	6.29. Determinánsok szorzata blokkmátrixban	266
3.45. A négy kitüntetett alter	151	6.30. 2×2 -es blokkmátrix determinánsa	267
3.46. Lineáris egyenletrendszer megoldásai	151	6.33. Vandermonde-determináns értéke	269
4.17. Mátrixszorzás és lineáris kombináció	179	7.2. Mátrixleképezések alapműveletei	280
4.18. Mátrix elemeinek, sor- és oszlopvektorainak elő- állítás	179	7.3. Inverz mátrixleképezések	281
4.22. Koordináták változása a bázis cseréjénél	182	7.4. A lineáris kombinációt megőrző leképezések	281
4.23. Bázisfelbontás	182	7.9. Síkbeli forgatás, tükrözés, vetítés	286
4.29. Elemi sorműveletek mátrixszorzással	185	7.10. Lineáris leképezés mátrixa	286
4.30. Műveletek blokkmátrixokkal	185	7.12. Lineáris leképezések alaptulajdonságai	288
4.34. A szorzat oszlopai és sorai	189	7.13. Lineáris leképezés mátrixai közti kapcsolat	290
4.36. A megoldás felírása blokkmátrixokkal	190	7.16. Hasonló mátrixok hatása	291
4.37. A nulltér bázisa	191	7.17. Hasonlóságra invariáns tulajdonságok	291
5.1. Összeadás és skalárral szorzás tulajdonságai	195	7.18. Tartomány mértékének változása lineáris transzformációban	293
5.4. Mátrixszorzás algebrai tulajdonságai	196	7.20. Jacobi-mátrix	294
5.5. Hatványozás azonosságai	199	7.23. Lányszabály	297
5.8. Transzponálás tulajdonságai	200	7.25. A forgatás mátrixa	301
5.13. Sorművelet inverzének mátrixa	204	7.28. Tengely körüli forgatás – Rodrigues-formula	302
5.14. Az inverz egyértelműsége	205	7.31. Egyenesre való merőleges vetítés mátrixa	304
5.15. Az inverz létezéséhez elég egy feltétel	205	7.32. Síkra való merőleges vetítés mátrixa	305
5.17. 2×2 -es mátrix inverze	207	7.34. Síkbeli tükrözés mátrixa	306
5.18. Az inverz alaptulajdonságai	207	7.35. Síkra való tükrözés mátrixa	306
5.20. Az invertálhatóság és az egyenletrendszerek	209	7.37. Altérre való vetítés mátrixa	308
5.24. Invertálhatóság és bázis	212	7.39. Merőleges vetítés mátrixai	309
5.25. Az áttérés mátrixának inverze	212	7.41. Legjobb közelítés tétele	311
5.28. Műveletek diagonális mátrixokkal	216	7.42. Vektor felbontása összetevőkre	311
5.32. Műveletek permutációs mátrixokkal	217	7.44. Egyenletrendszer optimális megoldása	312
5.35. Műveletek háromszögmátrixokkal	219	7.45. Lineáris regresszió	314
5.38. Műveletek (ferdén) szimmetrikus mátrixokkal	219	7.46. Linearizálható regressziós modellek	314
5.39. Felbontás szimmetrikus és ferdén szimmetrikus mátrix összegére	219	7.49. A projekció tulajdonságai	316
5.40. $A^T A$ és AA^T szimmetrikus	220	7.51. A vetítés ekvivalens definíciója	317
5.41. Sherman–Morrison-formula	220	7.52. Mikor merőleges egy vetítés?	318
5.49. Az LU-felbontás létezése és egyértelműsége	229	7.55. A pszeudoinverz létezése és egyértelműsége	321
6.2. Nullvektort tartalmazó determináns	243	7.57. Moore–Penrose-tétel	323
6.3. Elemi sorműveletek determináns	243	7.58. $A^+ A$ és AA^+ merőleges vetítés	324
6.4. Elemi mátrixok determinánsa	244	7.59. Optimális megoldás pszeudoinverzzel	324
6.5. Permutációs mátrix determinánsa	244	7.62. Ortogonális vektorok függetlensége	328
6.6. Háromszögmátrix determinánsa	244	7.63. Legjobb közelítés ONB esetén	329
6.8. Determinánsok szorzásszabálya	247	7.67. Szemiortogonális mátrixok ekvivalens definíciói	331
6.10. Transzponált determinánsa	247	7.68. Ortogonális mátrixok ekvivalens definíciói	331
6.12. Zérus értékű determináns	248	7.70. Ortogonális mátrixhoz tartozó mátrixleképezés	332
6.14. Egyenletrendszer megoldhatósága és a determi- náns	249	7.71. Ortogonális mátrixok tulajdonságai	333
6.15. Felbontás kigyók determinánsainak összegére	255	7.72.	333
6.16. Permutációs mátrix determinánsa	257	7.74. Egy vektor tükrözése egy másikba	336
6.18. Determinánsfüggvény létezése	257	7.76. Gram–Schmidt-ortogonalizáció	337
6.21. Determináns rendjének csökkentése	259	7.80. QR-felbontás létezése és egyértelműsége	340
6.23. Determinánsok kifejtési tétele	261	7.83. Legkisebb négyzetek QR-felbontással	343
		7.87. Az adjungált tulajdonságai	347
		7.88. A komplex skaláris szorzás tulajdonságai	347

7.91. Fourier-összeg helyettesítési értékei	350	. Zérusvektor	30
7.92. A Fourier-mátrixok tulajdonságai	351	. Vektor hossza	31
7.94. A DFT tulajdonságai	354	. Vektorok szöge	31
7.97. Gyors Fourier-transzformáció	357	1.1. Két vektor összege – háromszögmódszer	31
8.4. A sajátvektorok alterei	370	1.3. Vektorok különbsége	32
8.8. Háromszögmátrixok sajátértékei	373	1.4. Vektor szorzása skalárral	33
8.9. Determináns, nyom és a sajátértékek	373	1.6. Lineáris kombináció	33
8.11. 2×2 -es szimmetrikus mátrixok sajátalterei	374	1.10. Vektorok függetlensége	35
8.16. Mátrix invertálhatósága és a 0 sajátérték	379	1.15. Két vektor skaláris szorzata	39
8.17. Mátrix hatványainak sajátértékei és sajátvektorai	380	. Egységvektor	42
8.18. Mátrix hatványainak hatása	380	1.26. Vektori szorzat	45
8.19. Speciális mátrixok sajátértéke	381	1.33. Vegyes szorzat	48
8.20. Speciális komplex mátrixok sajátértéke	382	. Vektor koordinátás alakja 2D-ben	50
8.23. Sajátértékhez kapcsolódó invariánsok	385	. Vektor koordinátás alakja 3D-ben	50
8.25. Diagonalizálhatóság szükséges és elégséges fel- tétele	386		55
8.27. Különböző sajátértékek sajátvektorai	387	1.43. Vektorműveletek $\mathbb{R}^n$ -ben	55
8.28. Különböző sajátértékek és diagonalizálhatóság	388	1.49. Abszolút érték, szög, merőlegesség, távolság	60
8.30. Algebrai és geometriai multiplicitás kapcsolata	389	. Korrelációs együttható	62
8.31. Diagonalizálhatóság és a geometriai multiplicitás	390	1.54. Kód	64
8.33. Diagonalizálható mátrixok spektrálfelbontása	393	2.3. Alakzat (implicit) egyenletrendszer	70
8.36. Szimmetrikus mátrix sajátalterei	395	2.4. Alakzat (explicit) egyenletrendszer	71
8.37. Valós spektráltétel	395	2.21. Lineáris egyenlet	84
8.41. Főtengelytétel	399	2.25. Lineáris egyenletrendszer	85
8.45. Definitság meghatározása a sajátértékekből	401	2.26. Lineáris egyenletrendszer megoldása	86
9.6. A szinguláris értékek létezése és egyértelműsége	409	2.28. Ekvivalens egyenletrendszerek	86
9.7. Egységgömb képe	411	2.37. Elemi sorműveletek	95
9.8. Polárfelbontás	411	2.38. Lépcsős alak	95
9.10. A pszeudoinverz kiszámítása	413	2.45. Redukált lépcsős alak	100
9.16. Minden norma ekvivalens	420	. rref függvény	104
9.18. Frobenius-norma ekvivalens alakjai	421	2.51. Szimultán egyenletrendszerek	104
9.19.	421	2.64. Soronként domináns főátlójú mátrix	118
9.23. Indukált norma tulajdonságai	423	3.2. Mátrix rangja	122
9.24. 1-, 2- és ∞ -norma kiszámítása	424	3.10. Altér	126
10.6. Jordan normálalak	433	3.14. Nulltér	128
10.8. A Jordan-alak egyértelműsége	436	3.15. Kifeszített altér	128
10.13 Exponenciális függvény kiszámítása	444	3.19. Sortér, oszloptér	131
11.1. Perron-tétel: pozitív sajátérték és sajátvektor	424	3.26. Bázis	136
11.2. Perron-tétel: egyszeres és domináns sajátérték	425	3.32. Dimenzió	141
11.3. Perron–Frobenius-tétel – gyenge változat	427	3.35. Vektorrendszer rangja	142
11.4. Collatz–Wielandt-tétel	428	. Merőleges altér és merőleges kiegészítő altér	148
11.5. Nemnegatív mátrixok spektrálsugarának becs- lése	429	. Kiegészítő altér	148
11.8. Perron–Frobenius-tétel – erős változat	432	4.1. Lineáris helyettesítés	167
2.1. Mátrix rangja	449	4.4. Mátrixok egyenlősége	171
2.2. Invertálható négyzetes mátrixok	450	4.5. Adott típusú mátrixok tere	171
		4.6. Mátrixok összege, különbsége	172
		4.8. Zérusmátrix	172
		4.9. Mátrix szorzása skalárral	172
		4.11. Mátrixok szorzása	175
		4.13. Diadikus szorzat	176
		4.21. Áttérés mátrixa	181
		4.25. Egységmátrix	183

Definíciók

. Irányított szakasz, kötött vektor	29
. Vektor	30

4.26. Elemi mátrixok	184	9.13. p -norma	417
5.9. Mátrix inverze	203	9.14. Norma	418
5.30. Permutációs mátrix, kígyó	217	9.15. Normák ekvivalenciája	420
5.34. Háromszögmátrix	218	9.17. Frobenius-norma	421
5.36. Szimmetrikus és ferdén szimmetrikus mátrixok	219	9.20. Mátrixnorma	422
5.45. LU-felbontás	226	9.21.	422
5.50. PLU-felbontás	231	9.22. Indukált norma	422
.	239	10.1. Általánosított sajátvektor	430
6.1. Determináns	241	10.3. Jordan-blokk	431
6.19. Előjeles aldetermináns	258	10.12. Mátrix exponenciális függvénye	444
6.32. Vandermonde-determináns	268	.	425
.	279	11.6. Reducibilis és irreducibilis mátrixok	431
7.7. Lineáris leképezés	285	1.1. Lebegőpontos számok	438
7.15. Hasonlóság	290	1.6. Test	442
. Lineáris leképezés rangja	292	1.10. $\mathbb{Z}_m$	446
. Lineáris leképezés determinánsa	292		
7.19. Differenciálhatóság	294	<i>Kidolgozott példák</i>	
. Altérre való merőleges vetület	308	1.16. Skaláris szorzat kiszámítása a definíció alapján	39
. Optimális megoldás	312	1.20. Skaláris szorzat kiszámítása	41
. Normálegyenlet-rendszer	312	1.25. Merőleges összetevőkre bontás	43
. Regressziós egyenes	314	1.27. Vektori szorzat meghatározása	45
7.48. Vetítés altérre	316	1.28. $\mathbf{i}, \mathbf{j}, \mathbf{k}$ vektori szorzata	45
7.53. A Moore–Penrose-féle pszeudoinverz	320	1.32. Parallelepipedon térfogata	47
. Ortogonális és ortonormált bázis	328	1.34. Vegyes szorzat	48
7.65. Ortogonális és szemiorotogonális mátrix	330	1.36. Vektorok koordinátái	50
. Givens-forgatás	335	1.37. Pontok koordinátái	51
. Householder-tükrözés	336	1.39. Skaláris szorzás koordinátarendszerben	52
7.78. QR-felbontás	338	1.42. Parallelogramma területe	54
7.85. Komplex mátrix adjungáltja	346	1.45.	57
7.86. Komplex vektorok skaláris szorzata	347	1.50. Vektorok szöge és távolsága	61
.	348	1.55. BCD-kód	64
. Komplex vektorok hossza, távolsága, szöge, merőlegessége	349	1.56. Lineáris kombináció $\mathbb{Z}_m^n$ -ben	64
7.90. Unitér mátrix	349	1.57. One time pad – a tökéletes titkosítás	65
.	349	1.58. Paritásbit	66
7.93. Diszkrét Fourier-transzformáció (DFT)	353	1.59. Ellenőrző összeg	66
.	356	2.1. Az $x + y = 1$ egyenlet	69
8.2. Sajátérték, sajátvektor	370	2.2. Az $x^2 + y^2 = 1$ egyenlet	69
8.5. Sajátaltér	370	2.9. Síkbeli egyenes egyenletei	74
.	371	2.14. Sík egyenletei	76
8.21. Lineáris transzformáció sajátértéke, sajátvektora	384	2.18. Térbeli egyenes egyenletrendszerei	79
8.24. Diagonalizálhatóság	386	2.19. Egyenes és sík explicit vektoregyenlete	81
8.35. Ortogonális diagonalizálhatóság	395	2.20. Hipersík egyenlete	81
8.39. Unitér diagonalizálhatóság	397	2.22. Lineáris egyenlet	84
.	397	2.23. Lineáris egyenlet azonos átalakítása	84
8.43. Kvadratikus formák és mátrixok definitsége	400	2.24. Lineáris egyenletrendszerek	85
9.1. Szinguláris érték	404	2.27. Egyenletrendszer egy megoldása	86
. Szinguláris felbontás	406	2.30. Mátrix használata a megoldáshoz	88
.	406	2.31. Sormodell két kétismeretlenes egyenlettel	89
. Polárfelbontás	411	2.32. Ha 0 lesz a bal oldal	90
9.12. Euklideszi norma	416	2.33. Sormodell három háromismeretlenes egyenlettel	90

2.35. A megoldás lépései az oszlopmodellben	93	4.27. Elemi mátrixok	184
2.39. Lépcsős alak	96	4.28. Mátrix balról szorzása elemi mátrixszal	184
2.40. Gauss-módszer, egy megoldás	96	4.31. Műveletek blokkmátrixokkal	186
2.41. Gauss-módszer, végtelen sok megoldás	97	4.32. 2×2 -es blokkmátrixok	187
2.43. Homogén lineáris egyenletrendszer megoldása	99	4.33. Szorzat előállítása diádok összegeként	188
2.44. Síkok metszésvonalának meghatározása	99	4.35. Nulltér bázisa	190
2.46. Redukált lépcsős alak	100	5.2. Egyszerűsítés mátrixszal	196
2.47. Redukált lépcsős alakra hozás	101	5.3. Nullosztó	196
2.48. Gauss–Jordan-módszer, egy megoldás	101	5.6. Mátrix hatványozása	199
2.49. Gauss–Jordan-módszer, végtelen sok megoldás	102	5.7. Polinom helyettesítési értéke	200
2.52. Szimultán egyenletrendszer megoldása	105	5.10. Mátrix inverze	203
2.53. Szimultán egyenletrendszer bővített mátrixa	105	5.11. Szinguláris mátrix	203
2.54. Egyenletrendszer $\mathbb{Z}_2$ fölött	106	5.12. $I - A$ inverze nilpotens A esetén	204
2.55. Egyenletrendszer $\mathbb{Z}_5$ fölött	107	5.16. Az inverz kiszámítása	206
2.57. Instabil egyenletrendszer	110	5.19. Inverz tulajdonságainak alkalmazása	208
2.58. Gauss-módszer lebegőpontos számokkal	111	5.21. Egyenletrendszer megoldása mátrixinvertálással	210
2.59. Részleges főelemkiválasztás	112	5.22. Mátrixegyenlet megoldása mátrixinvertálással	211
2.60. Sor szorzása	113	5.23. Mátrix elemi mátrixok szorzatára bontása	211
2.61. Jacobi-iteráció	115	5.26. Az áttérés mátrixának inverze	213
2.62. Gauss–Seidel-iteráció	116	5.27. Műveletek diagonális mátrixokkal	216
2.63. Divergens iteráció	117	5.29. Sorok permutációja mátrixszorzással	216
3.3. Mátrix rangjának kiszámítása	122	5.31. Kígyók	217
3.5. Kötött és szabad változók száma	122	5.33. Permutációs mátrix inverze	218
3.8. Egyenletrendszer megoldásainak száma	124	5.37. Szimmetrikus és ferdén szimmetrikus mátrixok	219
3.12. Altér	128	5.42. Inverz változása	221
3.17. Nulltér	129	5.43. Inverz változása számpéldán	221
3.21. Kifeszített altér vektorai	131	5.44. Gauss-kiküszöbölés mátrixszorzással	225
3.23. Vektorok lineáris függetlenségének eldöntése	132	5.46. Egyenletrendszer megoldása LU-felbontással	226
3.27. Altér bázisának meghatározása	137	5.47. Mátrix invertálása LU-felbontással	227
3.28. Vektor felírása a bázisvektorok lineáris kombinációjaként	137	5.51. PLU-felbontás előállítása	232
3.30. Vektor koordinátás alakja a $\mathcal{B}$ bázisban	139	6.7. Determináns kiszámítása háromszög alakra hozással	245
3.33. Mátrix transzponáltja	141	6.9. Determináns kiszámolása PLU-felbontásból	247
3.36. Dimenzió kiszámítása	142	6.11. Determináns kiszámítása elemi oszlopműveletekkel	248
3.39. Egyenletrendszer megoldása elemi bázistranszformációval	145	6.13. Zérus értékű determinánsok	249
3.40. Vektorokra merőleges altér	147	6.17. Inverziók száma és a determináns	257
3.47. Lineáris egyenletrendszer sortérbe eső megoldása	152	6.20. Előjeles aldetermináns	258
4.2. Lineáris helyettesítések kompozíciója	168	6.22. Determináns rendjének csökkentése	260
4.3. Mátrixok és elemeik	170	6.24. Kifejtési tétel	262
4.7. Mátrixok összege, különbsége	172	6.26. Cramer-szabály	263
4.10. Mátrixok lineáris kombinációja	173	6.28. Mátrix inverze	265
4.12. Mátrixok szorzása	175	6.31. Interpoláció másodfokú polinomokra	267
4.14. Skaláris és diadikus szorzat	176	7.1. Vektori szorzással definiált mátrixleképezés	280
4.15. Egyenletrendszer mátrixszorzatos alakja	177	7.5. Mátrixleképezés ábrázolása az egységgrács képével	283
4.16. Szimultán egyenletrendszer mátrixszorzatos alakja	178	7.6. Mátrixleképezés ábrázolása az egységkör képével	284
4.19. Áttérés standard bázisra	180	7.8. A deriválás és az integrálás lineáris leképezés	286
4.20. Báziscsere	180	7.11.	287
4.24. Bázisfelbontás	182	7.14. Lineáris leképezés mátrixa másik bázisban	290
		7.21. Jacobi-mátrix kiszámítása	295
		7.22. Függvényérték becslése Jacobi-mátrixszal	296

7.24. Láncszabály	297	8.14. Komplex sajátértékek és komplex elemű sajátvektorok	377
7.26. Forgatás egy tetszőleges pont körül	301	8.15. Sajátérték algebrai és geometriai multiplicitása	378
7.27. Koordinátatengely körüli forgatás a térben	302	8.22. Lineáris transzformáció sajátértéke, sajátaltere	384
7.29. Forgatás mátrixa	303	8.26. Mátrix diagonalizálása	387
7.30. A forgatás mátrixának inverze	304	8.29. Diagonalizálhatóság megállapítása	389
7.33. Síkra eső merőleges vetület kiszámítása	305	8.32. Lineáris transzformáció diagonalizálása	391
7.36. Vetítés síkra	306	8.34. Mátrixok nagy kitevős hatványai	394
7.38. Merőleges vetület kiszámítása	309	8.38.	396
7.40.	310	8.40. Másodfokú polinom mátrixszorzatos alakja	398
7.43.	311	8.42. Főtengely-transzformáció	399
7.47.	315	8.44. Definitiség meghatározása a sajátértékekből	401
7.50. Projekció mátrixa	317	9.2. Szinguláris értékek	404
7.54. Néhány pszeudoinverz	320	9.3. Szinguláris felbontások	406
7.56. A pszeudoinverz kiszámítása	322	9.4. Szinguláris értékek meghatározása	407
7.60. Egyenletrendszer optimális megoldása	325	9.5. Szinguláris felbontás	408
7.61. Egyenletrendszer optimális megoldása	326	9.9. Polárfelbontás kiszámítása	412
7.64. Egy pont síkra való merőleges vetülete	330	9.11. A pszeudoinverz kiszámítása SVD-ből	413
7.66. Ortogonális mátrixok	330	10.2. Jordan-lánc konstrukciója	430
7.69. Ortogonális mátrixok inverze	332	10.4. Jordan-lánchoz tartozó Jordan-blokk	431
7.73. Forgatás tengelye és szöge	334	10.5. Jordan-lánccok és Jordan-blokkok kapcsolata	432
7.75. Householder-tükrözés	336	10.7. Normáalakok	435
7.77. Gram–Schmidt-ortogonalizáció	338	10.9. Jordan-blokkok mérete	437
7.79. QR-felbontás kiszámítása	340	10.10 Jordan-blokkok mérete	438
7.81. QR-felbontás Givens-forgatásokkal	341	10.11 Jordan-bázis előállítás	441
7.82. QR-felbontás Householder-tükrözéssel	342	10.14 Mátrix exponenciális függvénye	444
7.84. Egyenletrendszer optimális megoldása	344	11.7.	431
7.89. Önadjungált mátrixok	349	1.2. Lebegőpontos számok értéke	438
7.95. DFT kiszámítása	354	1.3. Lebegőpontos számok halmaza	438
7.96. Magas frekvenciájú összetevők szűrése	355	1.4. Alapműveletek lebegőpontos számokkal	439
8.1. Jó bázis tükrözéshez	369	1.5. Flop és flops	440
8.3. Sajátérték, sajátvektor	370	1.7. Műveletek paritásokka	445
8.6. Sajátalter bázisának meghatározása	371	1.8. XOR és AND	445
8.7. Karakterisztikus polinom felírása	372	1.9. Számolás az órán	445
8.10. 2×2 -es mátrixok sajátvektorainak ábrázolása	373	1.11. Számolás $\mathbb{Z}_m$ -ben	446
8.12. Az összes sajátérték és sajátvektor meghatározása	375	1.12. Műveletábra	447
8.13. Magasabbfokú karakterisztikus egyenlet	376	1.13. Osztás, reciprok	447

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 vektorba 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 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 e három vektor koordinátás alakja a három standard egységvektor, így 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 λ 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 λ 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 λ egy sajátértéke, akkor a λ -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 λ 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 λ sajátértékéhez tartozó sajátvektorai és a nullvektor által alkotott alteret a λ 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

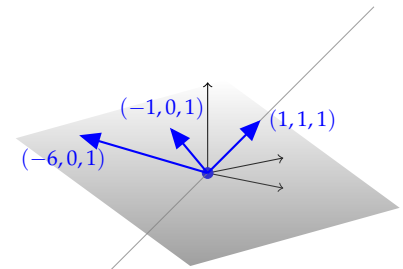
$$\mathbf{A} - 10\mathbf{I} = \begin{bmatrix} -7 & 6 & 1 \\ 1 & -2 & 1 \\ 1 & 6 & -7 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix},$$

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 6.12. tétel szerint pontosan akkor igaz, ha

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0. \quad (8.1)$$

Ez tehát azt jelenti, hogy λ 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.

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:

$$\begin{aligned} \det(\mathbf{A} - \lambda \mathbf{I}) &= \begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} = (a - \lambda)(d - \lambda) - bc \\ &= \lambda^2 - (a + d)\lambda + (ad - bc) \\ &= \lambda^2 - \text{trace}(\mathbf{A})\lambda + \det \mathbf{A}. \end{aligned} \quad (8.2)$$

Kimondható a következtetés: 2×2 -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

$$\begin{aligned} \det(\mathbf{C} - \lambda \mathbf{I}) &= \begin{vmatrix} a - \lambda & b & c \\ 1 & -\lambda & 0 \\ 0 & 1 & -\lambda \end{vmatrix} \\ &= (a - \lambda)\lambda^2 + c + b\lambda \\ &= -\lambda^3 + a\lambda^2 + b\lambda + c. \end{aligned}$$

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:

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.

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 $\mathbf{A} = [a_{ij}]$ mátrix karakterisztikus egyenlete

$$(a_{11} - \lambda)(a_{22} - \lambda) \dots (a_{nn} - \lambda) = 0$$

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*

$$\begin{aligned}\det(\mathbf{A}) &= \lambda_1 \lambda_2 \dots \lambda_n \\ \text{trace}(\mathbf{A}) &= \lambda_1 + \lambda_2 + \dots + \lambda_n\end{aligned}$$

Ezek az értékek megjelennek a karakterisztikus polinomban: a determináns a konstans tag, a nyom a $(-\lambda)^{n-1}$ együtthatója.

BIZONYÍTÁS. A karakterisztikus polinom gyöktényezői 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 2×2 -es mátrixok sajátaltéréinek jellemzése Olyan eredményekkel fogunk megismerkedni e paragrafusban, melyek általánosíthatók 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. 7.4 ábra).

8.10. PÉLDA (2×2 -ES MÁTRIXOK SAJÁTVEKTORAINAK ÁBRÁZOLÁSA). *Határozzuk meg a 7.5. és a 7.6. példákban is szereplő*

$$\mathbf{A} = \begin{bmatrix} 5 & 3 \\ 4 & 5 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 3 & 5 \\ 5 & 3 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} -5 & 3 \\ -3 & 5 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} -3 & 5 \\ -5 & 3 \end{bmatrix}.$$

mátrixok sajátértékeit és sajátvektorait. Szemléltessük ezeket az egységkör-ábrákban.

MEGOLDÁS. Egyszerű számolással meghatározható mind a négy mátrix karakterisztikus egyenlete, sajátértékei és sajátvektorai, bár a **D** mátrix esetén ezek komplex számokat is tartalmaznak. A karakterisztikus polinomot jelölje p , indexében a mátrix jelével. Ezután megadjuk a sajátértékeket, majd a sajátvektorokat:

$$p_A(\lambda) = \lambda^2 - \frac{5}{2}\lambda + 1, \quad \lambda_1 = 2, \quad \lambda_2 = \frac{1}{2}, \quad \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

$$p_B(\lambda) = \lambda^2 - \frac{3}{2}\lambda - 1, \quad \lambda_1 = 2, \quad \lambda_2 = -\frac{1}{2}, \quad \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

$$p_C(\lambda) = \lambda^2 - 1, \quad \lambda_1 = 1, \quad \lambda_2 = -1, \quad \mathbf{x}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

$$p_D(\lambda) = \lambda^2 + 1, \quad \lambda_1 = i, \quad \lambda_2 = -i, \quad \mathbf{x}_1 = \begin{bmatrix} \frac{3}{5} - \frac{4}{5}i \\ 1 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} \frac{3}{5} + \frac{4}{5}i \\ 1 \end{bmatrix}.$$

Némi nehézséget a **D** mátrix okoz, ezért az ahhoz tartozó számításokat részletezzük:

$$|\mathbf{D} - \lambda \mathbf{I}| = \begin{vmatrix} -\frac{3}{4} - \lambda & \frac{5}{4} \\ -\frac{5}{4} & -\frac{3}{4} - \lambda \end{vmatrix} = \lambda^2 + 1$$

$$\lambda_1 = i: \begin{bmatrix} -\frac{3}{4} - i & \frac{5}{4} \\ -\frac{5}{4} & -\frac{3}{4} - i \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -\frac{3}{4} + \frac{4}{5}i \\ 0 & 0 \end{bmatrix} \text{ amiből } \mathbf{x}_1 = \begin{bmatrix} \frac{3}{5} - \frac{4}{5}i \\ 1 \end{bmatrix},$$

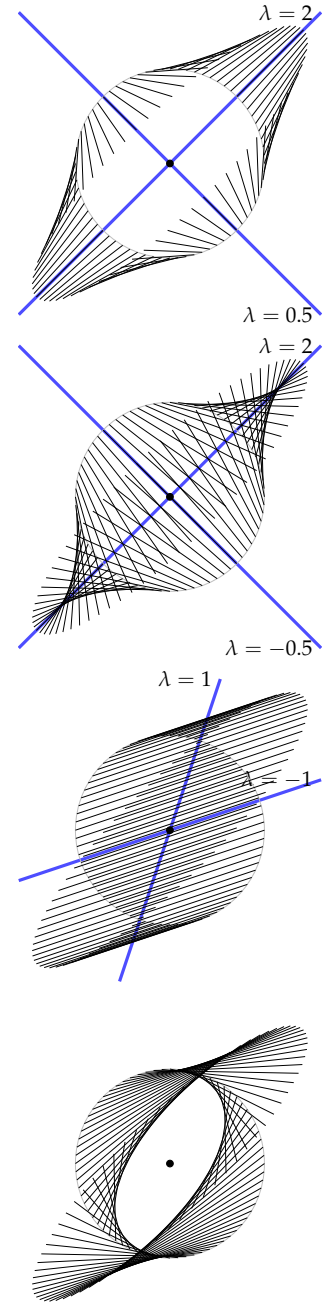
$$\lambda_2 = -i: \begin{bmatrix} -\frac{3}{4} + i & \frac{5}{4} \\ -\frac{5}{4} & -\frac{3}{4} + i \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -\frac{3}{4} - \frac{4}{5}i \\ 0 & 0 \end{bmatrix} \text{ amiből } \mathbf{x}_2 = \begin{bmatrix} \frac{3}{5} + \frac{4}{5}i \\ 1 \end{bmatrix}.$$

A négy mátrixhoz tartozó egységkörábra a 8.2 ábrán látható. $\square$

8.11. TÉTEL (2×2 -ES SZIMMETRIKUS MÁTRIXOK SAJÁTALTEREI). Legyen **A** egy 2×2 -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 2×2 -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 (8.2) 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 csak 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. $\square$



8.2. ábra: A négy leképezés sajátirányai.

Mátrix összes sajátértékének és sajátvektorának meghatározása Az előző paragrafusokban leírtak alapján egy mátrix sajátértékeinek és sajátvektorainak meghatározása két lépésben elvégezhető:

1. megoldjuk a $\det(\mathbf{A} - \lambda \mathbf{I}) = 0$ karakterisztikus egyenletet, ennek gyökei a sajátértékek,
2. minden λ sajátértékhez meghatározzuk az $\mathbf{A} - \lambda \mathbf{I}$ nullterének egy bázisát, az általa kifeszített altér nemzérus vektorai a λ -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 egyenlet 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 számítógépet, vagy például a függelékben megtalálható Rolle-féle gyöktételt. 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. □

A karakterisztikus egyenlet komplex gyökei Ha valós elemű mátrixot vizsgálunk, megeshet, hogy a karakterisztikus egyenletnek vannak komplex 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.

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 a $\frac{1}{2} + \frac{\sqrt{3}}{2}i$ 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 a $\frac{1}{2} + \frac{\sqrt{3}}{2}i$ sajátértékhez tartozó sajátaltér egy bázisa az $(i, 1)$ vektorból áll.

A $\frac{1}{2} - \frac{\sqrt{3}}{2}i$ 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 a $\frac{1}{2} - \frac{\sqrt{3}}{2}i$ sajátértékhez tartozó sajátalteret a $(-i, 1)$ sajátvektor feszíti ki. $\square$

A karakterisztikus egyenlet többszörös gyökei: az algebrai és a geometriai multiplicitás Ha λ a karakterisztikus egyenlet k -szoros gyöke, vagy más szóval λ multiplicitása vagy *algebrai multiplicitása* k , akkor a λ -hoz tartozó sajátaltér d dimenziójára $1 \leq d \leq k$. Ezt az állítást később bebizonyítjuk. A sajátaltér dimenzióját szokták a λ *geometriai multiplicitásának* nevezni. A 8.12. és a 8.13. példák olyan eseteket mutattak, amikor a sajátértékek algebrai és geometriai multiplicitása azonos, azaz minden sajátaltér épp annyi dimenziós, amennyi a gyök (algebrai) multiplicitása. A következő feladat azt mutatja, hogy a sajátaltér 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 $\mathbf{A}$ sajátaltère 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. Meghatározzuk sajátaltèreiket. $\lambda = 1$ 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

$$\mathbf{B} - \lambda \mathbf{I} = \mathbf{B} - 2\mathbf{I} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

ahonnan 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. □

Sajátértékek és a mátrix hatványai A mátrixok függvényeinek számolása szoros kapcsolatban van a sajátértékekkel. E témában első lépés a mátrixokhatványok sajátértékeinek és sajátvektorának 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. □

8.17. TÉTEL (MÁTRIX HATVÁNYAINAK SAJÁTÉRTÉKEI ÉS SAJÁTVEKTORAI). Ha λ az $\mathbf{A}$ mátrix egy sajátértéke és $\mathbf{x}$ egy hozzá tartozó sajátvektor, akkor bármely egész n esetén λ^n sajátértéke az $\mathbf{A}^n$ mátrixnak és $\mathbf{x}$ egy hozzá tartozó sajátvektor, amennyiben λ^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 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 $\mathbf{x}_1, \dots, \mathbf{x}_k$ a 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 λ^n alakú.

BIZONYÍTÁS. a), b) Írjuk fel a sajátpár elemeit komplex algebrai alakban, azaz legyen $\lambda = a + bi$ a valós $\mathbf{A}$ mátrix sajátértéke, és legyen $\mathbf{x} = \mathbf{y} + \mathbf{z}i$ a hozzá tartozó sajátvektor. Tehát $\mathbf{Ax} = \mathbf{A}(\mathbf{y} + \mathbf{z}i) = (a + bi)(\mathbf{y} + \mathbf{z}i)$. Ha ennek mindkét oldalát balról beszorozzuk az $\mathbf{x}$ konjugáltjának transzponáltjával, azaz az $\bar{\mathbf{x}}' = (\mathbf{y} - \mathbf{z}i)'$ vektorral, akkor kapjuk, hogy

$$\bar{\mathbf{x}}'\mathbf{Ax} = (\mathbf{y} - \mathbf{z}i)'(a + bi)(\mathbf{y} + \mathbf{z}i) = (a + bi)(|\mathbf{y}|^2 + |\mathbf{z}|^2).$$

Itt kihasználtuk, hogy $\mathbf{y}'\mathbf{y} = |\mathbf{y}|^2$ és $\mathbf{z}'\mathbf{z} = |\mathbf{z}|^2$. Véve mindkét oldal konjugáltját és kihasználva hogy $\mathbf{A}$ valós, tehát $\bar{\mathbf{A}} = \mathbf{A}$, kapjuk, hogy

$$\mathbf{x}'\mathbf{A}\bar{\mathbf{x}} = (a - bi)(|\mathbf{y}|^2 + |\mathbf{z}|^2).$$

Ezután vegyük mindkét oldal transzponáltját. A jobb oldal nem változik, hisz ott egy szám, azaz egy 1×1 -es mátrix áll:

$$\bar{\mathbf{x}}'\mathbf{A}'\mathbf{x} = (a - bi)(|\mathbf{y}|^2 + |\mathbf{z}|^2).$$

Ezt összevetve a (119) egyenlőség két oldalán álló kifejezéssel, és kihasználva, hogy $\mathbf{x} \neq \mathbf{0}$ miatt $|\mathbf{y}|^2 + |\mathbf{z}|^2 \neq 0$, kapjuk, hogy

$$a + bi = \begin{cases} a - bi, & \text{ha } \mathbf{A} \text{ szimmetrikus,} \\ -a + bi, & \text{ha } \mathbf{A} \text{ ferdén szimmetrikus.} \end{cases}$$

Eszerint $b = 0$, ha $\mathbf{A}$ szimmetrikus, és $a = 0$, ha $\mathbf{A}$ ferdén szimmetrikus, ami bizonyítja az első két állítást.

c) Ha $\mathbf{A}$ ortogonális, akkor bármely $\mathbf{x}$ vektorra $|\mathbf{Ax}| = |\mathbf{x}|$. Így a sajátvektorra is, azaz $|\mathbf{x}| = |\mathbf{Ax}| = |\lambda\mathbf{x}| = |\lambda||\mathbf{x}|$, amiből $|\lambda| = 1$.

d) Ha $\mathbf{A}^k = \mathbf{O}$, és λ sajátértéke $\mathbf{A}$ -nak, akkor λ^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 megfordítását hamarosan igazoljuk. $\square$

Az előző tétel bizonyításából minimális változtatással megkapható a következő tétel bizonyítása is, melyet feladatként tűzünk ki:

8.20. TÉTEL (SPECIÁLIS KOMPLEX MÁTRIXOK SAJÁTÉRTÉKE). *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 2×2 -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 λ 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 λ 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 π egész számú többszörösétől különböző szöggel;
- a tér vektorainak merőleges vetítése egy síkra;
- a tér vektorainak tükrözése egy síkra.

MEGOLDÁS. Az előző fejezetben, így a ?? példában 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 sajátaltér az egyenessel párhuzamos vektorokból, a -1 -hez tartozó sajátaltér 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 π egész számú többszörösétől különböző szöggel való elforgatása egy egyenes körül a forgástengellyel párhuzamos vektorokat önmagukba viszi, és semelyik másikat 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 $\mathbf{0}$ 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 ha egy leképezés egy vektort a λ -szorosába viszi, akkor azt λ -szorosába viszi minden bázisban.

Hasonló mátrixok sajátértékei A 7.16. és a 7.17. tételekből tudjuk, hogy egy lineáris leképezéshez különböző bázisokban tartozó mátrixok hasonlóak, másrészt 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 és 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 7.17. 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 7.17. tételben igazoltuk. $\square$

► E tétel egyik következménye, hogy hasonló mátrixok determinánsa megegyezik, hisz ezek a sajátértékek szorzatával egyenlők. Ezt már

bizonyítottuk a 7.17. tételben. Hasonlóan igaz, hogy hasonló mátrixok nyoma is megegyezik, hisz ezek a sajátértékek összegével egyenlők.

► Fontos következménye tehát e tételnek, hogy van értelme *lineáris transzformáció karakterisztikus polinomjáról, sajátértékeiről, sajátaltéréről* beszélni. Legalábbis véges dimenziós esetben, pl. $\mathbb{R}^n \rightarrow \mathbb{R}^n$ vagy $\mathbb{C}^n \rightarrow \mathbb{C}^n$ transzformációk esetén ezt látjuk.

Mátrixok diagonalizálása Korábban láttuk, hogy egy mátrix hatványainak hatását milyen könnyű a sajátvektorok lineáris kombinációin kiszámítani. Ez akkor lenne igazán hatékony eszköz, ha sajátvektorokból egy bázist tudnánk választani. Ebben a bázisban ugyanis a mátrix – 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

$$\mathbf{\Lambda} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}. \quad (8.3)$$

8.25. TÉTEL (DIAGONALIZÁLHATÓSÁG SZÜKSÉGES ÉS ELÉGSÉGES FELTÉTELE). Az $n \times n$ -es $\mathbf{A}$ mátrix pontosan akkor diagonalizálható, ha 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 $\mathbf{C} = [\mathbf{c}_1 \ \mathbf{c}_2 \ \dots \ \mathbf{c}_n]$ és $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, akkor

$$[\mathbf{c}_1 \ \mathbf{c}_2 \ \dots \ \mathbf{c}_n] \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} = \mathbf{A}[\mathbf{c}_1 \ \mathbf{c}_2 \ \dots \ \mathbf{c}_n], \quad (8.4)$$

és itt a bal oldali mátrix i -edik oszlopa $\lambda_i \mathbf{c}_i$, a jobb oldali mátrixé $\mathbf{A}\mathbf{c}_i$, amik megegyeznek, azaz $\mathbf{A}\mathbf{c}_i = \lambda_i \mathbf{c}_i$, tehát $\mathbf{c}_i$ a λ_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, a sajátvektorokból pedig egy $\mathbf{C}$ mátrixot úgy, hogy a $\mathbf{\Lambda}$ i -edik oszlopába kerülő λ_i sajátértékhez tartozó sajátvektor a $\mathbf{C}$ mátrix i -edik oszlopába kerüljön. Mivel $\lambda_i \mathbf{c}_i = \mathbf{A}\mathbf{c}_i$, ezért fönnáll a (8.4) összefüggés, azaz $\mathbf{\Lambda}$ hasonló $\mathbf{A}$ -hoz. $\square$

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 sajátértékeit 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:

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad \square$$

A diagonalizálhatóság fontos voltára tekintettel érdemes további feltételeket gyűjteni, melyek könnyen ellenőrizhetők. Több elégséges feltétel származtatható a következő tételből:

8.27. 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 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

$$\mathbf{x}_i = c_1 \mathbf{x}_1 + \dots + c_{i-1} \mathbf{x}_{i-1}, \quad (8.5)$$

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:

$$\lambda_i \mathbf{x}_i = c_1 \lambda_1 \mathbf{x}_1 + \dots + c_{i-1} \lambda_{i-1} \mathbf{x}_{i-1}. \quad (8.6)$$

Ezután a (8.5) egyenlet mindkét oldalát λ_i -vel szorozva kapjuk, hogy

$$\lambda_i \mathbf{x}_i = c_1 \lambda_i \mathbf{x}_1 + \dots + c_{i-1} \lambda_i \mathbf{x}_{i-1}. \quad (8.7)$$

Végül a (8.7) egyenletből a (8.6) egyenletet kivonva kapjuk, hogy

$$0 = c_1(\lambda_i - \lambda_1)\mathbf{x}_1 + \dots + c_{i-1}(\lambda_i - \lambda_{i-1})\mathbf{x}_{i-1},$$

Mivel az $\mathbf{x}_1, \dots, \mathbf{x}_{i-1}$ vektorok már lineárisan függetlenek, és a $\lambda_1, \dots, \lambda_i$ értékek különbözőek, így $c_1 = \dots = c_{i-1} = 0$. 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 bárhogy választunk közülük 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 vektorrá, 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.

8.28. KÖVETKEZMÉNY (KÜLÖNBÖZŐ SAJÁTÉRTÉKEK ÉS DIAGONALIZÁLHATÓSÁG). Ha az n -edrendű $\mathbf{A}$ mátrixnak n darab különböző sajátértéke van, akkor diagonalizálható.

BIZONYÍTÁS. A 8.25. és a 8.27. tételek szerint n különböző sajátértékhez n független sajátvektor tartozik, ami épp azt jelenti, hogy a mátrix diagonalizálható. $\square$

Végezetül összefoglalásul 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 komplex gyökei, a komplex sajátértékekhez ugyanis nem található valós sajátvektor, így nem találhatunk n független sajátvektort, hisz a valós sajátértékek száma kisebb n -nél. 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, sajátvektoraiból ortonormált bázis is kiválasztható. Ezt hamarosan belátjuk (8.36. tétel).

8.29. PÉLDA (DIAGONALIZÁLHATÓSÁG MEGÁLLAPÍTÁSA). *Döntsük el, hogy az alábbi mátrixok közül melyik diagonalizálható valós mátrixként!*

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 1 & -4 \\ 4 & 1 \end{bmatrix}, \mathbf{C} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}, \mathbf{D} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

MEGOLDÁS. Az $\mathbf{A}$ mátrix nilpotens, mert $\mathbf{A}^2 = \mathbf{O}$, így nem diagonalizálható. A $\mathbf{B}$ mátrixnak vannak komplex 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 csak a figyelmet ellenőrző kérdés, e mátrix diagonális, tehát diagonalizálva van, azaz biztosan 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.

8.30. TÉTEL (ALGEBRAI ÉS GEOMETRIAI MULTIPLICITÁS KAPCSOLATA). *Egy mátrix valamely sajátértékének geometriai multiplicitása nem lehet nagyobb az algebrai multiplicitásánál.*

BIZONYÍTÁS. Az $\mathbf{A}$ mátrix egy μ sajátértékének geometriai multiplicitását jelölje g . Ez azt jelenti, hogy $\mathbf{A} - \mu\mathbf{I}$ nullterének van g vektorból álló bázisa. Legyen ez $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_g$. Egészítsük ki e bázist az egész tér bázisává a $\mathbf{x}_{g+1}, \dots, \mathbf{x}_n$ 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, azaz $\mathbf{C} = [\mathbf{X} | \mathbf{Y}]$. Mivel $\mathbf{X}$ oszlopai a μ -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 fölhasználva 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 6.29. 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 μ legalább g -szeres algebrai multiplicitású gyöke. $\square$

8.31. 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 az azonos sajátértékhez tartozó vektorok által kifeszített sajátaltér dimenziója megegyezik e 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 komplex 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 8.27. 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 sajátalterei dimenziójának összege n .

8.32. PÉLDA (LINEÁRIS TRANSZFORMÁCIÓ DIAGONALIZÁLÁSA). Az alábbi lineáris leképezésekhez keressünk – pusztán geometriai szemléletünkre hagyatkozva – olyan bázist, melyben a mátrixa diagonális. Használjuk fel a 8.22. példa eredményeit.

- a) a sík vektorainak tükrözése egy egyenesre (vagy pontjainak tükrözése egy origón átmenő egyenesre);
- b) a sík vektorainak merőleges vetítése egy egyenesre (vagy pontjainak merőleges vetítése egy origón átmenő egyenesre);
- c) a tér vektorainak elforgatása egy egyenes körül a π egész számú többszörösétől különböző szöggel;
- d) a tér vektorainak merőleges vetítése egy síkra;
- e) a tér vektorainak tükrözése egy síkra.

MEGOLDÁS. A 8.22. példában meghatároztuk e leképezések sajátaltereit. 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 sajátaltere 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 választva 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 választva 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:

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 \mathbf{I} & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & \lambda_2 \mathbf{I} & \dots & \mathbf{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & \mathbf{O} & \dots & \lambda_k \mathbf{I} \end{bmatrix}, \quad \mathbf{C} = [\mathbf{X}_1 \quad \mathbf{X}_2 \quad \dots \quad \mathbf{X}_k], \quad \mathbf{C}^{-1} = \begin{bmatrix} \mathbf{Y}'_1 \\ \mathbf{Y}'_2 \\ \vdots \\ \mathbf{Y}'_k \end{bmatrix}.$$

Írjuk föl e mátrixokkal az $\mathbf{A} = \mathbf{C}\mathbf{\Lambda}\mathbf{C}^{-1}$ felbontást, majd fejtsük ki a blokkműveleteket:

$$\begin{aligned} \mathbf{A} = \mathbf{C}\mathbf{\Lambda}\mathbf{C}^{-1} &= [\mathbf{X}_1 \quad \mathbf{X}_2 \quad \dots \quad \mathbf{X}_k] \begin{bmatrix} \lambda_1 \mathbf{I} & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & \lambda_2 \mathbf{I} & \dots & \mathbf{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & \mathbf{O} & \dots & \lambda_k \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{Y}'_1 \\ \mathbf{Y}'_2 \\ \vdots \\ \mathbf{Y}'_k \end{bmatrix} \\ &= \lambda_1 \mathbf{X}_1 \mathbf{Y}'_1 + \lambda_2 \mathbf{X}_2 \mathbf{Y}'_2 + \dots + \lambda_k \mathbf{X}_k \mathbf{Y}'_k \\ &= \lambda_1 \mathbf{P}_1 + \lambda_2 \mathbf{P}_2 + \dots + \lambda_k \mathbf{P}_k, \end{aligned}$$

ahol $\mathbf{P}\mathbf{P}_i = \mathbf{X}_i \mathbf{Y}'_i$ a λ_i sajátértékhez tartozó mátrix, melyről a következőket fogjuk megmutatni:

8.33. ÁLLÍTÁS (DIAGONALIZÁLHATÓ MÁTRIXOK SPEKTRÁLFELBONTÁSA). Minden $\{\lambda_1, \lambda_2, \dots, \lambda_k\}$ spektrumú diagonalizálható $\mathbf{A}$ mátrix felírható

$$\mathbf{A} = \lambda_1 \mathbf{P}_1 + \lambda_2 \mathbf{P}_2 + \dots + \lambda_k \mathbf{P}_k$$

alakba, 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 megmutatható, hogy a fenti három feltétel szükséges és elégséges feltétele annak, hogy $\mathbf{A}$ diagonalizálható legyen. E tételt nevezik a diagonalizálható mátrixok spektráltételének.

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 $\mathbf{P} \mathbf{P}_i = \mathbf{X}_i \mathbf{Y}_i'$ egyenlőségeket is kapjuk, hogy $\mathbf{P}_1 + \mathbf{P}_2 + \dots + \mathbf{P}_k = \mathbf{I}$.

b) A $\mathbf{A} \mathbf{C}^{-1} \mathbf{C} = \mathbf{I}$ egyenlőség blokkmátrixokalakja viszont az $\mathbf{Y}_i' \mathbf{X}_i = \mathbf{I}$, és az $\mathbf{Y}_i' \mathbf{X}_j = \mathbf{O}$ ($i \neq j$) egyenlőségre vezet.

c) Az előzőekből adódik, hogy $\mathbf{P}_i^2 = \mathbf{P}_i$, ugyanis $\mathbf{P}_i^2 = \mathbf{X}_i (\mathbf{Y}_i' \mathbf{X}_i) \mathbf{Y}_i' = \mathbf{X}_i \mathbf{Y}_i' = \mathbf{P}_i$. 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 $\mathcal{O}(\mathbf{X} \mathbf{Y}) \subseteq \mathcal{O}(\mathbf{X})$.

$$\mathcal{O}(\mathbf{P}_i) = \mathcal{O}(\mathbf{X}_i \mathbf{Y}_i') \subseteq \mathcal{O}(\mathbf{X}_i) = \mathcal{O}(\mathbf{X}_i \mathbf{Y}_i' \mathbf{X}_i) = \mathcal{O}(\mathbf{P}_i \mathbf{X}_i) \subseteq \mathcal{O}(\mathbf{P}_i).$$

Tehát mindenütt egyenlőség áll fenn, és $\mathcal{O}(\mathbf{P}_i) = \mathcal{O}(\mathbf{X}_i) = \mathcal{N}(\mathbf{A} - \lambda_i \mathbf{I})$. 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 $\mathcal{N}(\mathbf{P}_i) = \mathcal{O}(\mathbf{A} - \lambda_i \mathbf{I})$. $\square$

Mátrixok hatványai és egyéb függvényei Ha egy folyamat egy $\mathbf{x}_k$ állapotát a következővel egy lineáris $\mathbf{x}_{k+1} = \mathbf{A} \mathbf{x}_k$ kapcsolat fűzi össze, akkor az $\mathbf{x}_k = \mathbf{A}^k \mathbf{x}_0$ összefüggés miatt a folyamatot az $\mathbf{A}$ mátrix hatványai jellemzik. Kérdés lehet például a mátrixhatványok aszimptotikus viselkedése, vagy a nagy kitevőjű hatványok gyors kiszámításának módja.

Diagonális mátrix hatványai igen könnyen számolhatók, csak a főátló elemeit kell hatványozni. Másrészt $(\mathbf{C}^{-1}\mathbf{M}\mathbf{C})^k = \mathbf{C}^{-1}\mathbf{M}^k\mathbf{C}$, ezért a diagonalizálható mátrixok igen könnyen hatványozhatók.

8.34. PÉLDA (MÁTRIXOK NAGY KITEVŐS HATVÁNYAI). Tekintsük az alábbi két „majdnem egyenlő” mátrixot:

$$\mathbf{A} = \begin{bmatrix} -0.3 & 1.8 \\ -0.6 & 1.8 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -0.3 & 1.8 \\ -0.5 & 1.8 \end{bmatrix}$$

Vizsgáljuk meg hatványaik határértékét, ha a kitevő tart a végtelenhez!

MEGOLDÁS. Mindkét mátrixot diagonalizáljuk:

$$\mathbf{\Lambda}_1 = \mathbf{C}^{-1}\mathbf{A}\mathbf{C} = \begin{bmatrix} 0.6 & 0.0 \\ 0.0 & 0.9 \end{bmatrix}, \quad \text{ahol } \mathbf{C} = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{3}{\sqrt{13}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{13}} \end{bmatrix}$$

valamint

$$\mathbf{\Lambda}_2 = \mathbf{D}^{-1}\mathbf{B}\mathbf{D} = \begin{bmatrix} 1.2 & 0.0 \\ 0.0 & 0.3 \end{bmatrix}, \quad \text{ahol } \mathbf{D} = \begin{bmatrix} 1 & 1 \\ \frac{5}{6} & \frac{1}{3} \end{bmatrix}, \quad \mathbf{D}^{-1} = \begin{bmatrix} -\frac{2}{3} & 2 \\ \frac{5}{3} & -2 \end{bmatrix}.$$

Így a k -adik hatvány könnyen számolható:

$$\mathbf{A}^k = \mathbf{C} \begin{bmatrix} 0.6 & 0.0 \\ 0.0 & 0.9 \end{bmatrix}^k \mathbf{C}^{-1} = \mathbf{C} \begin{bmatrix} 0.6^k & 0.0 \\ 0.0 & 0.9^k \end{bmatrix} \mathbf{C}^{-1}$$

Mivel mindkét sajátérték abszolút értéke kisebb 1-nél, ezért $\mathbf{\Lambda}_1^k \rightarrow \mathbf{O}$ és így $\mathbf{A}^k \rightarrow \mathbf{O}$, ha $k \rightarrow \infty$. A $\mathbf{B}$ mátrix esetén

$$\mathbf{B}^k = \mathbf{D} \begin{bmatrix} 1.2 & 0.0 \\ 0.0 & 0.3 \end{bmatrix}^k \mathbf{D}^{-1} = \mathbf{D} \begin{bmatrix} 1.2^k & 0.0 \\ 0.0 & 0.3^k \end{bmatrix} \mathbf{D}^{-1},$$

ami arra vezet, hogy $\mathbf{\Lambda}_2^k \rightarrow \begin{bmatrix} \infty & 0 \\ 0 & 0 \end{bmatrix}$ és a $\mathbf{D}$ és a $\mathbf{D}^{-1}$ elemeinek előjelét is figyelembe véve így $\mathbf{A}^k \rightarrow \begin{bmatrix} -\infty & \infty \\ \infty & \infty \end{bmatrix}$, ha $k \rightarrow \infty$. $\square$

Mátrixok ortogonális diagonalizálása, valós spektráltétel Diagonalizálni egy mátrixot azzal ekvivalens, hogy a hozzá tartozó mátrixleképezéshez találni egy olyan bázist, melyben mátrixa diagonális. A gyakorlati alkalmazásokban különösen szerencsés, ha e bázis még ortonormált is! Valós mátrixok közt pontosan a szimmetrikusak azok, amelyek így diagonalizálhatók.

A 8.25. tétel szerint a diagonalizálhatóság szükséges és elégséges feltétele, hogy létezzen a mátrix rendjével egyező számú független sajátvektora. Ha e vektorok ortonormált rendszert alkotnak, akkor a belőlük alkotott mátrix ortogonális mátrix.

8.35. DEFINÍCIÓ (ORTOGONÁLIS DIAGONALIZÁLHATÓSÁG). Az $\mathbf{A}$ mátrix ortogonálisan diagonalizálható, ha találunk egy ortogonális $\mathbf{Q}$ és egy diagonális $\mathbf{\Lambda}$ mátrixot, hogy $\mathbf{Q}'\mathbf{A}\mathbf{Q} = \mathbf{\Lambda}$.

Először megmutatjuk, hogy szimmetrikus mátrix esetén a sajátalterek nem csak függetlenek egymástól, de merőlegesek is egymásra.

8.36. 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 $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ és $\mathbf{A}\mathbf{y} = \mu\mathbf{y}$. Ebből adódik, hogy

$$\lambda(\mathbf{x}'\mathbf{y}) = (\lambda\mathbf{x})'\mathbf{y} = (\mathbf{A}\mathbf{x})'\mathbf{y} = \mathbf{x}'\mathbf{A}'\mathbf{y} = \mathbf{x}'\mathbf{A}\mathbf{y} = \mathbf{x}'\mu\mathbf{y} = \mu(\mathbf{x}'\mathbf{y}).$$

Eszerint $(\lambda - \mu)(\mathbf{x}'\mathbf{y}) = 0$, de $\lambda - \mu \neq 0$, ezért $\mathbf{x}'\mathbf{y} = \mathbf{x} \cdot \mathbf{y} = 0$, azaz a két vektor merőleges egymásra. $\square$

8.37. TÉTEL (VALÓS SPEKTRÁLTÉTEL). A valós $\mathbf{A}$ mátrix pontosan akkor diagonalizálható ortogonálisan, ha szimmetrikus.

BIZONYÍTÁS. ($\Rightarrow$) Tegyük fel, hogy $\mathbf{A}$ ortogonálisan diagonalizálható, azaz létezik olyan ortogonális $\mathbf{Q}$ és diagonális $\mathbf{\Lambda}$ mátrix, hogy $\mathbf{Q}'\mathbf{A}\mathbf{Q} = \mathbf{\Lambda}$. Ezt balról a $\mathbf{Q}$, jobbról a $\mathbf{Q}'$ mátrixszal szorozva kapjuk, hogy $\mathbf{Q}\mathbf{Q}'\mathbf{A}\mathbf{Q}\mathbf{Q}' = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}'$. Mivel $\mathbf{Q}$ ortogonális, ezért $\mathbf{Q}\mathbf{Q}' = \mathbf{I}$, így $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}'$. Ekkor

$$\mathbf{A}' = (\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}')' = (\mathbf{Q}')'\mathbf{\Lambda}'\mathbf{Q}' = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}' = \mathbf{A}.$$

Tehát $\mathbf{A}$ szimmetrikus.

($\Leftarrow$) Tegyük fel, hogy $\mathbf{A}$ szimmetrikus, azaz $\mathbf{A}' = \mathbf{A}$. Ekkor minden sajátértéke valós, sajátértékei algebrai multiplicitásának összege tehát n . Először megmutatjuk, hogy minden sajátérték geometriai multiplicitása megegyezik algebrai multiplicitásával, tehát a mátrix diagonalizálható. Legyen μ egy tetszőleges sajátérték, geometriai multiplicitása legyen g . Ekkor sajátalteréből kiválaszthatunk egy $\mathbf{b}_1, \dots, \mathbf{b}_g$ bázist, amelyet kiegészítünk a teljes tér $\mathbf{b}_1, \dots, \mathbf{b}_g, \dots, \mathbf{b}_n$ bázisává. Az e vektorokból képzett $\mathbf{B}$ mátrixra tehát igaz az

$$\mathbf{A}\mathbf{B} = \mathbf{B} \begin{bmatrix} \mu\mathbf{I}_g & \mathbf{X} \\ \mathbf{O} & \mathbf{Y} \end{bmatrix}$$

összefüggés. Mivel $\mathbf{B}$ oszlopai lineárisan függetlenek, ezért $\mathbf{B}$ invertálható, és így $\mathbf{B}^{-1}\mathbf{A}\mathbf{B} = \begin{bmatrix} \mu\mathbf{I}_g & \mathbf{X} \\ \mathbf{O} & \mathbf{Y} \end{bmatrix}$. Eszerint $\mathbf{A} \sim \begin{bmatrix} \mu\mathbf{I}_g & \mathbf{X} \\ \mathbf{O} & \mathbf{Y} \end{bmatrix}$, tehát a két

mátrixnak megegyeznek a sajátértékei. Mivel $\mathbf{Y}$ -nak már nem lehet μ a sajátértéke, és a karakterisztikus polinom

$$\det(\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} (\mu - \lambda)\mathbf{I}_g & \mathbf{X} \\ \mathbf{O} & \mathbf{Y} - \lambda \mathbf{I} \end{vmatrix} = (\mu - \lambda)^g \det(\mathbf{Y} - \lambda \mathbf{I}),$$

ezért μ algebrai multiplicitása is g , amivel bizonyítottuk, hogy $\mathbf{A}$ diagonalizálható. Már csak annak bizonyítása maradt, hogy a sajátvektorból ortonormált bázis is kiválasztható, ez viszont abból következik, hogy szimmetrikus mátrix sajátalterei merőlegesek egymásra, így ha minden sajátaltérből kiválasztunk egy ortonormált bázist, akkor azok egyesítése az egész tér ortonormált bázisa lesz (ld. még erről a 8.27. tételt és az azt követő második megjegyzést). $\square$

8.38. 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 ortonormáltat. 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

$$\begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{\sqrt{2}}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}.$$

□

Spektráltétel önortogonális mátrixokra

Mátrixok unitér diagonalizálása A valós mátrixok ortogonális diagonalizálhatóság mintájára a komplex mátrixok unitér diagonalizálhatósága is alapvető kérdés.

8.39. DEFINÍCIÓ (UNITÉR DIAGONALIZÁLHATÓSÁG). Az $\mathbf{A}$ mátrix unitéren diagonalizálható, ha találunk egy unitér $\mathbf{U}$ és egy diagonális $\mathbf{\Lambda}$ mátrixot, hogy $\mathbf{U}^* \mathbf{A} \mathbf{U} = \mathbf{\Lambda}$.

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. Mivel $xy = yx$, ezért ezt a polinomot ekvivalens módon több más alakba is át lehet írni, például: $2x^2 + 3xy + yx - y^2$,

$2x^2 - 5xy + 9yx - y^2, \dots, 2x^2 + 2xy + 2yx - y^2$. Az utolsó formula egy mátrixszorzatos fölrás lehetőségét sejteti:

$$2x^2 + 2xy + 2yx - y^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Valóban, 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:

$$\begin{aligned} & ax^2 + 2bxy + 2cxz + dy^2 + 2eyz + fz^2 \\ &= ax^2 + bxy + cxz + byx + dy^2 + eyz + czx + ezy + fz^2 \\ &= \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \end{aligned}$$

8.40. 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 együttthatójú 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

$$\mathbf{x}'\mathbf{A}\mathbf{x}$$

alakra hozható, ahol $\mathbf{A}$ szimmetrikus mátrix. Ezt nevezzük kvadratikusan alaknak, vagy kvadratikusan formának. Tipikus alkalmazásokban a mátrix valós elemű. Ennek megfelelően e szakaszban a továbbiakban kvadratikusan formán mindig valós kvadratikusan formát értünk, tehát kvadratikusan formának nevezzük azt az

$$\mathbb{R}^n \rightarrow \mathbb{R}; \mathbf{x} \mapsto \mathbf{x}'\mathbf{A}\mathbf{x}$$

függvényt, ahol $\mathbf{A}$ szimmetrikus mátrix.

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 $\mathbf{Q}'\mathbf{A}\mathbf{Q} = \mathbf{\Lambda}$. Tudjuk, hogy az $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ mátrixleképezés mátrixa a $\mathbf{Q}$ oszlopvektorai alkotta $\mathcal{Q}$ 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:

$$\mathbf{x}'\mathbf{A}\mathbf{x} = (\mathbf{Q}\mathbf{y})'\mathbf{A}(\mathbf{Q}\mathbf{y}) = \mathbf{y}'\mathbf{Q}'\mathbf{A}\mathbf{Q}\mathbf{y} = \mathbf{y}'\mathbf{\Lambda}\mathbf{y}.$$

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:

8.41. TÉTEL (FŐTENGELYTÉTEL). Ha $\mathbf{A}$ egy n -edrendű valós szimmetrikus mátrix, és a $\mathbf{Q}$ mátrix ortogonálisan diagonalizálja, azaz $\mathbf{Q}'\mathbf{A}\mathbf{Q} = \mathbf{\Lambda}$ diagonális, akkor az $\mathbf{x} = \mathbf{Q}\mathbf{y}$ helyettesítés az $\mathbf{x}'\mathbf{A}\mathbf{x}$ kvadratikus formát a $\mathbf{y}'\mathbf{\Lambda}\mathbf{y}$ kvadratikus formába transzformálja, mely kifejtve csak négyzetes tagokat tartalmaz, azaz

$$\mathbf{x}'\mathbf{A}\mathbf{x} = \mathbf{y}'\mathbf{\Lambda}\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2, \quad (8.8)$$

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 $\mathbf{x}'\mathbf{A}\mathbf{x} = c$ egyenletű felületnek a $\mathcal{Q}$ bázis vektorai mind szimmetriatengelyei, melyeket főtengelyeknek is nevezünk.

► 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 formán nem változtat, hisz abban csak a sajátértékek szerepelnek.

► A főtengelytétel alkalmazását egy kvadratikus formán *főtengely-transzformációnak* nevezzük.

8.42. PÉLDA (FŐTENGELY-TRANSZFORMÁCIÓ). Végezzük el a főtengely-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

$$\begin{vmatrix} 2/9 & -5/3\sqrt{5} & 0 \\ -2/9 & -4/3\sqrt{5} & 1/\sqrt{5} \\ 1/9 & 2/3\sqrt{5} & 2/\sqrt{5} \end{vmatrix} = -1,$$

tehát egy megfelelő ortonormált bázis: $(\frac{2}{9}, -\frac{2}{9}, \frac{1}{9})$, $(\frac{5}{3\sqrt{5}}, \frac{4}{3\sqrt{5}}, -\frac{2}{3\sqrt{5}})$, $(0, \frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}})$. $\square$

Kvadratikus formák és mátrixok definitisége A főtengetyétel könnyen áttekinthetővé teszi a kvadratikus forma által fölvetett értékek lehetséges előjelét. Ez lehetővé teszi a kvadratikus formák egy fontos osztályozását.

8.43. DEFINÍCIÓ (KVADRATIKUS FORMÁK ÉS MÁTRIXOK DEFINITSÉGE).

Azt mondjuk, hogy az $f(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}$ 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.

Azt mondjuk, hogy a szimmetrikus $\mathbf{A}$ mátrix pozitív/negatív definit/szemidefinit, illetve indefinit, ha a hozzá tartozó kvadratikus forma az.

► Ha $\mathbf{A}$ negatív definit, akkor $-\mathbf{A}$ pozitív definit. Hasonló állítás igaz a szemidefinitségre is.

► Ha $\mathbf{A} = [a]$, azaz $\mathbf{A}$ 1×1 -es, akkor $\mathbf{A}$ pontosan akkor pozitív definit, ha $a > 0$.

►

► 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 főtengetyítétel szerint minden kvadrátikus forma egyenlő változók négyzeteinek a sajátértékekkel vett lineáris kombinációjával, ezért a definitség típusa pusztán csak a sajátértékek előjeleinek ismeretében eldönthető.

8.44. PÉLDA (DEFINITISÉG 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 kvadrátikus forma minden értéke pozitív, ha a változók nem mindegyike 0. Tehát e kvadrátikus forma pozitív definit. Hasonlóképp a $\mathbf{B}$ sajátértékei -1 , -1 és 2 , a főtengety-transzformáció után kapott alak $-\xi^2 - \eta^2 + 4\zeta^2$. E forma negatív értéket vesz fel például az $(1, 0, 0)$ helyen és pozitívat 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. $\square$

8.45. TÉTEL (DEFINITISÉG MEGHATÁROZÁSA A SAJÁTÉRTÉKEKBŐL). A valós szimmetrikus $\mathbf{A}$ mátrix, illetve az $\mathbf{x}'\mathbf{A}\mathbf{x}$ kvadrátikus forma pontosan akkor

- pozitív definit, ha $\mathbf{A}$ minden sajátértéke pozitív;
- pozitív szemidefinit, ha $\mathbf{A}$ minden sajátértéke nemnegatív;
- negatív definit, ha $\mathbf{A}$ minden sajátértéke negatív;
- negatív szemidefinit, ha $\mathbf{A}$ minden sajátértéke nempozitív;
- indefinit, ha $\mathbf{A}$ -nak van pozitív és negatív sajátértéke is.

Kúpszeletek osztályozása

Definitség és sajátértékek

Szélsőérték

Szélsőérték az egységgömbön

9

Szinguláris érték

A szimmetrikus mátrix ortogonális diagonalizálását fogjuk általánosítani tetszőleges mátrixra, egy helyett két ortonormált bázis megkeresésével. Az egyik bázis a mátrixleképezés értelmezési tartományának, a másik az értékkészletnek lesz bázisa. Ebben az általánosításban a sajátértékek szerepét a szinguláris értékek veszik át, a spektrálfelbontását a szinguláris érték szerinti felbontás (SVD). Az alkalmazások közül kiemelkednek az információtömörítéssel kapcsolatosak, de az egyenletrendszerek megoldásához használt leghatékonyabbak közé tartozó algoritmusok is ide sorolhatók.

Szinguláris érték, szinguláris vektor, SVD

Azt tudjuk, hogy egy mátrixleképezés kölcsönösen egyértelmű a sortér és az oszloptér között. Olyan ortonormált bázisát keressük a sortérnek, és az oszloptérnek, melyek közt a mátrixleképezés természetes kapcsolatot létesít.

Szinguláris érték és szinguláris vektorok Az, hogy az $n \times n$ -es $\mathbf{A}$ mátrix ortogonálisan diagonalizálható, azt jelenti, hogy létezik egy olyan $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ ortonormált bázis, és léteznek olyan λ_i ($i = 1, 2, \dots, n$) számok, hogy $\mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_i$ minden i indexre. Ha a mátrix téglalapalakú, például $m \times n$ -es, azaz az értelmezési tartomány és az értékkészlet különböző tér is lehet, akkor mindkettőben választanunk kell egy bázist. Legyen $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ az értelmezési tartomány ONB-a, az értékkészlet pedig $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\}$. Az analógia $\mathbf{A}\mathbf{v}_i = \sigma_i\mathbf{u}_i$ alakú feltételek kikötését kívánja. Az $m > n$ és $m < n$ esetek nehézségeit leküzdendő szorítkozzunk a sortérre, azt ugyanis tudjuk, hogy az $A : \mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ leképezés kölcsönösen egyértelmű a sortér és az oszloptér között. Így először csak e két altérben keressünk megfelelő bázist. Közös dimenziójuk a ranggal egyenlő, jelölje ezt r .

9.1. DEFINÍCIÓ (SZINGULÁRIS ÉRTÉK). Azt mondjuk, hogy a pozitív $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ számok az r -rangú valós vagy komplex $\mathbf{A}$ mátrix szinguláris értékei, ha van olyan $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ ortonormált bázisa a sortérnek, és $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ ortonormált bázisa az oszloptérnek, hogy

$$\mathbf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i, \quad i = 1, \dots, r.$$

A $\mathbf{v}_i$ vektorokat bal, az $\mathbf{u}_i$ vektorokat jobb szinguláris vektoroknak nevezzük.

► E definíció fontos következménye, hogy mivel $|\mathbf{u}_i| = 1$ ($i = 1, 2, \dots, r$), ezért $|\mathbf{A}\mathbf{v}_i| = \sigma_i$.

► A szinguláris értékekre adható olyan definíció is, melyben a két bázis létezése nem szerepel, csak az $\mathbf{A}$ mátrix – ezt később megemlíjtük –, de didaktikai okból nem ezt használjuk.

► A szinguláris értékeket úgy is definiálhatjuk, hogy ha $k = \min(m, n)$, akkor a szinguláris értékek száma k , és csak annyit kötünk ki róluk, hogy nemnegatívak. Látni fogjuk, hogy ha $k > r$, akkor e definíció mellett $\sigma_{r+1} = \sigma_{r+2} = \dots = \sigma_k = 0$.

9.2. PÉLDA (SZINGULÁRIS ÉRTÉKEK). Igazoljuk, hogy a

$$\begin{aligned} \{\mathbf{v}_1, \mathbf{v}_2\} &= \{(4/5, -3/5), (3/5, 4/5)\}, \\ \{\mathbf{u}_1, \mathbf{u}_2\} &= \{(-5/13, 12/13), (12/13, 5/13)\} \end{aligned}$$

ONB-ok választása mellett az

$$\mathbf{A} = \begin{bmatrix} -4/13 & 6 \\ 111/13 & -4 \end{bmatrix}$$

mátrix szinguláris értékei $\sigma_1 = 10$ és $\sigma_2 = 5$.

MEGOLDÁS. Ez nyilvánvalóan igaz, hisz

$$\begin{bmatrix} -4/13 & 6 \\ 111/13 & -4 \end{bmatrix} \begin{bmatrix} 4/5 \\ -3/5 \end{bmatrix} = 10 \begin{bmatrix} -5/13 \\ 12/13 \end{bmatrix}, \quad \begin{bmatrix} -4/13 & 6 \\ 111/13 & -4 \end{bmatrix} \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix} = 5 \begin{bmatrix} 12/13 \\ 5/13 \end{bmatrix},$$

azaz fennállnak az $\mathbf{A}\mathbf{v}_1 = \sigma_1 \mathbf{u}_1$ és az $\mathbf{A}\mathbf{v}_2 = \sigma_2 \mathbf{u}_2$ összefüggések. $\square$

Egyelőre nem tudjuk sem azt, hogy egy tetszőleges valós vagy komplex mátrixnak vannak-e szinguláris értékei, és ha igen, egyértelműek-e, azaz például függnek-e a bázis megválasztásától.

Szinguláris felbontás Képezzük a szinguláris értékekből a diagonális

$$\Sigma_1 = \text{diag}(\sigma_1, \dots, \sigma_r) = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_r \end{bmatrix}$$

A szinguláris érték fogalmát Erhard Schmidt vezette be 1907-ben, de ő még *sajátértéknek* nevezte. Mai nevét 1937-ben kapta, mert különösen akkor tűnt hasznos eszköznek – például az egyenletrendszerek megoldásában –, amikor az együtthatómátrix szinguláris.

mátrixot, valamint a bázisvektorokból az $\mathbf{U}_1 = \{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ és a $\mathbf{V}_1 = \{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ mátrixokat. Ekkor az $\mathbf{A}\mathbf{v}_i = \sigma\mathbf{u}_i$ egyenlőségek az

$$\mathbf{A}\mathbf{V}_1 = \mathbf{U}_1\mathbf{\Sigma}_1, \quad (9.1)$$

azaz az

$$\mathbf{A} \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_r \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_r \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_r \end{bmatrix} \quad (9.2)$$

alakot öltik. Egészítsük ki a $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ vektorrendszert a teljes n -dimenziós tér ONB-ává, és jelöljük a belőlük képzett $n \times n$ -es ortogonális mátrixot $\mathbf{V}$ -vel, $n > r$ esetén az új vektorokból képzett mátrixot $\mathbf{V}_2$ -vel, azaz $\mathbf{V}_2 = \begin{bmatrix} \mathbf{v}_{r+1} & \dots & \mathbf{v}_n \end{bmatrix}$. Mivel $\mathbf{V}_2$ oszlopvektorai merőlegesek a sortérre, ezért a nulltérben vannak, tehát $r < i \leq n$ esetén $\mathbf{A}\mathbf{v}_i = \mathbf{0}$.

Hasonlóképp az előzőkhöz egészítsük ki az $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ vektorrendszert a teljes m -dimenziós tér ONB-ává, és jelöljük e vektorokból képzett $m \times m$ -es mátrixot $\mathbf{U}$ -val, és $m > r$ esetén legyen $\mathbf{U}_2 = \begin{bmatrix} \mathbf{u}_{r+1} & \dots & \mathbf{u}_m \end{bmatrix}$. Végül a $\mathbf{\Sigma}_1 = \text{diag}(\sigma_1, \dots, \sigma_r)$ mátrixot egészítsük ki egy $m \times n$ -es mátrixszá nullblokkok hozzávételével, jelölje e mátrixot $\mathbf{\Sigma}$, tehát $\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$. Ekkor a (9.2) egyenlőség a következőképp módosítható:

$$\begin{aligned} \mathbf{A}\mathbf{V} &= \begin{bmatrix} \mathbf{A}\mathbf{v}_1 & \dots & \mathbf{A}\mathbf{v}_r & | & \mathbf{A}\mathbf{v}_{r+1} & \dots & \mathbf{A}\mathbf{v}_n \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1\mathbf{u}_1 & \dots & \sigma_r\mathbf{u}_r & | & \mathbf{0} & \dots & \mathbf{0} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{u}_1 & \dots & \mathbf{u}_r & | & \mathbf{u}_{r+1} & \dots & \mathbf{u}_m \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \dots & 0 & | & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 & | & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & | & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \sigma_r & | & 0 & \dots & 0 \\ \hline 0 & 0 & \dots & 0 & | & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & | & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 & | & 0 & \dots & 0 \end{bmatrix} \\ &= \mathbf{U}\mathbf{\Sigma} \end{aligned}$$

A mátrixok méreteit is kiírva $\mathbf{A}_{m \times n} \mathbf{V}_{n \times n} = \mathbf{U}_{m \times m} \mathbf{\Sigma}_{m \times n}$, blokkmátrix alakba átírva

$$\mathbf{A} \begin{bmatrix} \mathbf{V}_1 & | & \mathbf{V}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{U}_1 & | & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_1 & | & \mathbf{0} \\ \hline \mathbf{0} & | & \mathbf{0} \end{bmatrix}$$

Ha $r = n$, illetve $r = m$, akkor $\mathbf{V}_2$, illetve $\mathbf{U}_2$ üresek, azaz 0 számú oszlopból állnak, ami értelemszerűen változtat e képletben. Mivel a

négyzetes V mátrix oszlopvektorai ONB-t alkotnak, ezért valós esetben V ortogonális, így $V^{-1} = V'$, komplex esetben unitér, és $V^{-1} = V^*$. Ezt fölhasználva, az $AV = U\Sigma$ egyenlőségből kapjuk, hogy $A = U\Sigma V'$ ($A = U\Sigma V^*$). Ezt nevezzük a valós (komplex) A mátrix *szinguláris érték szerinti felbontásának*, vagy röviden *szinguláris felbontásának*.

Tekintsük e felbontás blokkmátrixalakját. A műveleteket blokkonként elvégezve, a Σ -ban lévő nullmátrixok miatt a következőt kapjuk:

$$A = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \left[\begin{array}{c|c} \Sigma_1 & O \\ \hline O & O \end{array} \right] \begin{bmatrix} V'_1 \\ V'_2 \end{bmatrix} = U_1 \Sigma_1 V'_1.$$

Az $A = U_1 \Sigma_1 V'_1$ (komplex esetben $A = U_1 \Sigma_1 V_1^*$) felbontást *redukált szinguláris felbontásnak* nevezzük. Ha ezt a szorzatot az $U_1 \Sigma_1$ oszlopvektorokra, és a V'_1 sorvektorokra blokkosított alakján végezzük, az A mátrix egy diadikus felbontását kapjuk, melyet *szinguláris érték szerinti diadikus felbontásnak* vagy a *szinguláris felbontás diadikus alakjának* nevezünk:

$$A = \sigma_1 u_1 v'_1 + \sigma_2 u_2 v'_2 + \cdots + \sigma_r u_r v'_r.$$

(E felbontások mindegyikének komplex alakjában transzponált helyett adjungált áll.)

9.3. PÉLDA (SZINGULÁRIS FELBONTÁSOK). *Igazoljuk, hogy az*

$$\begin{bmatrix} 0 & -2 & 2 \\ -2 & 3 & -2 \\ 4 & -2 & 0 \end{bmatrix} = \begin{bmatrix} 1/3 & -2/3 & -2/3 \\ -2/3 & 1/3 & -2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix} \begin{bmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2/3 & -2/3 & 1/3 \\ 2/3 & 1/3 & -2/3 \\ 1/3 & 2/3 & 2/3 \end{bmatrix}$$

felbontás szinguláris felbontás. Ennek alapján írjuk fel az egyenlőség bal oldalán álló mátrix redukált szinguláris felbontását, és a szinguláris felbontás diadikus alakját!

MEGOLDÁS. Először ellenőriznünk kell, hogy valóban egy $A = U\Sigma V'$ alakú egyenlőségről van szó, azaz hogy

- az egyenlőség fennáll,
- az U és V mátrixok ortogonálisak, Σ diagonális.

Könnyen ellenőrizhető, hogy igen. Ebből már következik, hogy az U első két oszlopát u_1 -gyel, illetve u_2 -vel, a V' első két sorát v'_1 -vel, illetve v'_2 -vel jelölve fennállnak az $Av_1 = 6u_1$ és az $Av_2 = 3u_2$ egyenlőségek. Mivel Σ -ról látható, hogy $r(A) = 2$, ezért U első 2 oszlopát, V' első 2 sorát, Σ bal felső 2×2 -es blokkját megtartva a redukált alakot kapjuk:

$$\begin{bmatrix} 0 & -2 & 2 \\ -2 & 3 & -2 \\ 4 & -2 & 0 \end{bmatrix} = \begin{bmatrix} 1/3 & -2/3 \\ -2/3 & 1/3 \\ 2/3 & 2/3 \end{bmatrix} \begin{bmatrix} 6 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2/3 & -2/3 & 1/3 \\ 2/3 & 1/3 & -2/3 \end{bmatrix}$$

Végül ennek alapján a szinguláris érték szerinti diadikus felbontás:

$$\begin{aligned} \begin{bmatrix} 0 & -2 & 2 \\ -2 & 3 & -2 \\ 4 & -2 & 0 \end{bmatrix} &= 6 \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{2}{3} \end{bmatrix} \begin{bmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix} + 3 \begin{bmatrix} -\frac{2}{3} \\ \frac{1}{3} \\ \frac{2}{3} \end{bmatrix} \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \end{bmatrix} \\ &= \begin{bmatrix} 4/3 & -4/3 & 2/3 \\ -8/3 & 8/3 & -4/3 \\ 8/3 & -8/3 & 4/3 \end{bmatrix} + \begin{bmatrix} -4/3 & -2/3 & 4/3 \\ 2/3 & 1/3 & -2/3 \\ 4/3 & 2/3 & -4/3 \end{bmatrix} \end{aligned}$$

Egy 2-rangú mátrix két 1-rangú összegére való bontása többféleképp is lehetséges. E felbontás tulajdonságaira hamarosan visszatérünk. $\square$

A szinguláris értékek és a szinguláris felbontás meghatározása Kérdés, hogyan számolható ki a szinguláris felbontás? Egy egyszerű ötlettel visszavezethető egy sajátértékszámítási feladatra. Tekintsük az $\mathbf{A}'\mathbf{A}$ (komplex esetben az $\mathbf{A}^*\mathbf{A}$) mátrixot! Mivel $\mathbf{U}'\mathbf{U} = \mathbf{I}$ és $\mathbf{\Sigma}'\mathbf{\Sigma} = \mathbf{\Sigma}^2$, ezért

$$\mathbf{A}'\mathbf{A} = (\mathbf{U}\mathbf{\Sigma}\mathbf{V}')'\mathbf{U}\mathbf{\Sigma}\mathbf{V}' = \mathbf{V}\mathbf{\Sigma}'\mathbf{U}'\mathbf{U}\mathbf{\Sigma}\mathbf{V}' = \mathbf{V}\mathbf{\Sigma}^2\mathbf{V}'.$$

Ez épp $\mathbf{A}'\mathbf{A}$ spektrálfelbontása, hisz $\mathbf{V}$ ortogonális (komplex esetben unitér), és $\mathbf{\Sigma}^2$ diagonális. Mivel $\mathbf{\Sigma}'\mathbf{\Sigma} = \mathbf{\Sigma}^2 = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2, 0, \dots, 0)$, ezért $\mathbf{A}'\mathbf{A}$ sajátértékei épp a szinguláris értékek négyzetei.

9.4. PÉLDA (SZINGULÁRIS ÉRTÉKEK MEGHATÁROZÁSA). Számítsuk ki az

$$\mathbf{A} = \begin{bmatrix} 0 & -2 & 2 \\ -2 & 3 & -2 \\ 4 & -2 & 0 \end{bmatrix}$$

mátrix szinguláris értékeit, és írjuk fel szinguláris felbontását!

MEGOLDÁS. A szinguláris értékek megegyeznek $\mathbf{A}'\mathbf{A}$ nemnulla sajátértékeinek gyökeivel.

$$\mathbf{A}'\mathbf{A} = \begin{bmatrix} 20 & -14 & 4 \\ -14 & 17 & -10 \\ 4 & -10 & 8 \end{bmatrix}$$

Ennek karakterisztikus polinomja $x^3 - 45x^2 + 324x$, melynek gyökei 36, 9 és 0. Tehát a szinguláris értékek 6 és 3. Az $\mathbf{A}'\mathbf{A}$ mátrix egységnyi sajátvektorai:

$$\begin{aligned} \lambda_1 &= 36 & \mathbf{v}_1 &= (2/3, -2/3, 1/3) \\ \lambda_2 &= 9 & \mathbf{v}_2 &= (2/3, 1/3, -2/3) \\ \lambda_3 &= 0 & \mathbf{v}_3 &= (1/3, 2/3, 2/3). \end{aligned}$$

Mivel $\mathbf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i$, ezért $\mathbf{u}_i = \mathbf{A}\mathbf{v}_i / \sigma_i$. Így kiszámolható $\mathbf{u}_1$ és $\mathbf{u}_2$ is:

$$\mathbf{u}_1 = \frac{\mathbf{A}\mathbf{v}_1}{\sigma_1} = \frac{(2, -4, 4)}{6} = \left(\frac{1}{3}, -\frac{2}{3}, \frac{2}{3}\right)$$

$$\mathbf{u}_2 = \frac{\mathbf{A}\mathbf{v}_2}{\sigma_2} = \frac{(-2, 1, 2)}{3} = \left(-\frac{2}{3}, \frac{1}{3}, \frac{2}{3}\right)$$

A harmadik vektor ilyen módon nem számolható ki, mivel $\mathbf{A}\mathbf{v}_3 = \mathbf{0}$, így az $\{\mathbf{u}_1, \mathbf{u}_2\}$ rendszert nekünk kell kiegészítenünk $\mathbb{R}^3$ bázisává. Ezt több módon is megtehetjük. Mivel $\{\mathbf{u}_1, \mathbf{u}_2\}$ az oszloptér bázisa, ezért a merőleges kiegészítő altér – vagyis $\mathbf{A}'$ nullterének – bázisát keressük. E példában legegyszerűbb megoldás vektori szorzással számolni: $\mathbf{u}_3 = \mathbf{u}_1 \times \mathbf{u}_2 = (-2/3, -2/3, -1/3)$. A szinguláris és a redukált szinguláris felbontás tehát:

$$\mathbf{A} = \frac{1}{3} \begin{bmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ 2 & 2 & -1 \end{bmatrix} \begin{bmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \frac{1}{3} \begin{bmatrix} 2 & -2 & 1 \\ 2 & 1 & -2 \\ 1 & 2 & 2 \end{bmatrix},$$

$$\mathbf{A} = \frac{1}{3} \begin{bmatrix} 1 & -2 \\ -2 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 6 & 0 \\ 0 & 3 \end{bmatrix} \frac{1}{3} \begin{bmatrix} 2 & -2 & 1 \\ 2 & 1 & -2 \end{bmatrix}$$

A felírásban az $\mathbf{U}$ és $\mathbf{V}$ mátrixból is kiemeltünk $\frac{1}{3}$ -ot, de ez is a mátrixhoz tartozik – egyébként nem lenne ortogonális. $\square$

Az $\mathbf{U}$ mátrix meghatározására egy további módszer is adódik. Az $\mathbf{A}'\mathbf{A}$ helyett vizsgáljuk meg az $\mathbf{A}\mathbf{A}'$ mátrixot.

$$\mathbf{A}\mathbf{A}' = \mathbf{U}\Sigma\mathbf{V}'(\mathbf{U}\Sigma\mathbf{V}')' = \mathbf{U}\Sigma\mathbf{V}'\mathbf{V}\Sigma'\mathbf{U}' = \mathbf{U}\Sigma^2\mathbf{U}'.$$

Eszerint a szinguláris értékek az $\mathbf{A}\mathbf{A}'$ mátrixból is meghatározhatók. Ennek sajátvektorai az $\mathbf{U}$ mátrix első r oszlopát adják. Mivel az $\mathbf{A}'\mathbf{A}$ mátrix σ_i^2 értékhez tartozó sajátvektora $\mathbf{v}_i$, ezért $\mathbf{A}'\mathbf{A}\mathbf{v}_i = \sigma_i^2 \mathbf{v}_i$, másrészt $\mathbf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i$, így e két összefüggést összevetve kapjuk, hogy $\mathbf{A}'\mathbf{A}\mathbf{v}_i = \mathbf{A}'(\sigma_i \mathbf{u}_i) = \sigma_i^2 \mathbf{v}_i$, azaz

$$\mathbf{A}'\mathbf{u}_i = \sigma_i \mathbf{v}_i, \text{ azaz } \mathbf{v}_i = \frac{\mathbf{A}'\mathbf{u}_i}{\sigma_i}.$$

Érdeemes lehet az $\mathbf{A}\mathbf{A}'$ pozitív sajátértékekhez tartozó sajátvektorait keresni, ha $m < n$, mert ekkor csak m -dimenziós vektorokkal kell számolni (ld. a 9.2. feladatot).

9.5. PÉLDA (SZINGULÁRIS FELBONTÁS). Számítsuk ki az

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}$$

mátrix szinguláris érték szerinti felbontását!

MEGOLDÁS. $\mathbf{A}'\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, melynek karakterisztikus polinomja $\lambda^2 - 4\lambda + 3 = (\lambda - 3)(\lambda - 1)$. Az $\mathbf{A}'\mathbf{A}$ sajátértékei 3 és 1, tehát $\mathbf{A}$ szinguláris értékei $\sqrt{3}$ és 1. A hozzájuk tartozó egységnyi hosszú sajátvektorok $\mathbf{v}_1 = (1/\sqrt{2}, 1/\sqrt{2})$, $\mathbf{v}_2 = (-1/\sqrt{2}, 1/\sqrt{2})$. Így

$$\mathbf{V} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}, \quad \mathbf{\Sigma} = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Az $\mathbf{u}_i = \mathbf{A}\mathbf{v}_i/\sigma_i$ összefüggés alapján $\mathbf{u}_1 = \frac{1}{\sqrt{6}}(1, 2, 1)$, $\mathbf{u}_2 = \frac{1}{\sqrt{2}}(1, 0, -1)$. Az előző példához hasonlóan $\mathbf{u}_3$ kiszámítható az $\mathbf{u}_3 = \mathbf{u}_1 \times \mathbf{u}_2$ képlettel is, de most inkább számoljunk úgy, hogy keressük $\mathbf{A}'$ nullterének bázisát. A nulltér meghatározásához meg kell oldani a $\mathbf{A}' = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ együtthatómátrixú homogén lineáris egyenletrendszert. Innen is az adódik, hogy $\mathbf{u}_3 = \frac{1}{\sqrt{3}}(1, -1, 1)$. Itt választhatjuk e vektor ellentettjét is, mert $\mathbf{A}\mathbf{u}_3 = \mathbf{0}$, vagyis az előjelnek itt nincs szerepe. Így

$$\mathbf{U} = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ 2/\sqrt{6} & 0 & 1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & -1/\sqrt{3} \end{bmatrix}.$$

Tehát a szinguláris felbontás

$$\begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ 2/\sqrt{6} & 0 & 1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & -1/\sqrt{3} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}.$$

□

Szinguláris érték szerinti felbontás létezése Kérdés még, hogy bármely valós vagy komplex mátrixnak léteznek-e szinguláris értékei, azok egyértelműek-e, és ha igen, a szinguláris felbontás egyértelmű-e.

9.6. TÉTEL (A SZINGULÁRIS ÉRTÉKEK LÉTEZÉSE ÉS EGYÉRTELMŰSÉGE). Minden r -rangú valós vagy komplex $\mathbf{A}$ mátrixnak létezik r szinguláris értéke. Ezek valós esetben megegyeznek az $\mathbf{A}'\mathbf{A}$, illetve az $\mathbf{A}\mathbf{A}'$ (komplex esetben az $\mathbf{A}^*\mathbf{A}$, illetve az $\mathbf{A}\mathbf{A}^*$) pozitív sajátértékeinek négyzetgyökével. A szinguláris értékek monoton csökkenő sorozata egyértelmű.

BIZONYÍTÁS. A bizonyítást valós esetre írjuk le, komplexre lényegében azonos. Az $\mathbf{A}'\mathbf{A}$ mátrix szimmetrikus ($\mathbf{A}^*\mathbf{A}$ önadjungált), mert $(\mathbf{A}'\mathbf{A})' = \mathbf{A}'(\mathbf{A}')' = \mathbf{A}'\mathbf{A}$. Ennek következtében minden sajátértéke valós, és ortogonálisan diagonalizálható. Másrészt minden sajátértéke nemnegatív, ugyanis

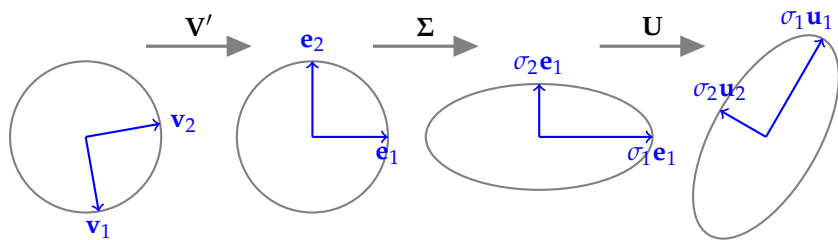
$$|\mathbf{A}\mathbf{x}|^2 = (\mathbf{A}\mathbf{x})'(\mathbf{A}\mathbf{x}) = \mathbf{x}'\mathbf{A}'\mathbf{A}\mathbf{x} = \mathbf{x}'(\mathbf{A}'\mathbf{A}\mathbf{x}) = \mathbf{x}'(\lambda\mathbf{x}) = \lambda|\mathbf{x}|^2.$$

(Másként fogalmazva $\mathbf{A}'\mathbf{A}$ pozitív szemidefinit.) A 0-tól különböző sajátértékek száma megegyezik $\mathbf{A}'\mathbf{A}$ rangjával, hisz az megegyezik diagonális alakja nemnulla elemeinek számával. Másrészt ???? szerint

$r(\mathbf{A}'\mathbf{A}) = r(\mathbf{A}) = r$. Tehát, ha nagyság szerinti sorba rendezzük a sajátértékeket ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$), akkor $\lambda_i > 0$, ha $1 \leq i \leq r$, és $\lambda_i = 0$, ha $r < i \leq n$. Eszerint $\sigma_i = \sqrt{\lambda_i} > 0$, ha $1 \leq i \leq r$. Végül, mivel $\mathbf{A}'\mathbf{A}$ sajátértékei egyértelműek, ezért $\mathbf{A}$ szinguláris értékei is. Azt, hogy $\mathbf{A}'\mathbf{A}$ és $\mathbf{A}\mathbf{A}'$ 0-tól különböző sajátértékei megegyeznek, korábban beláttuk. $\square$

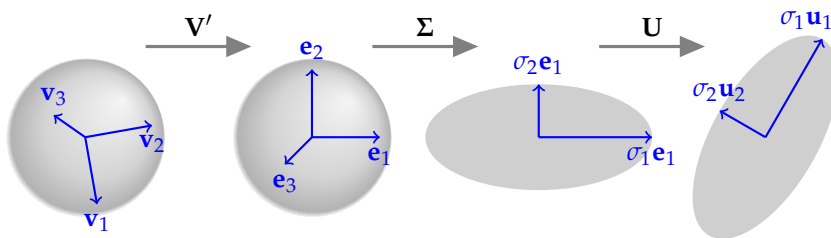
Szinguláris felbontás geometriai interpretációja A szinguláris felbontás segítségével jól szemléltethető, hogy egy lineáris leképezés hatására mi a képe egy egységgömbnek.

Először szemléltessük egy 2×2 -es, valós, 2-rangú mátrix szinguláris felbontását, tényezőinek hatását ábrázolva. Mivel a felbontás $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}'$, először $\mathbf{V}'$ hat a sík vektoraira. $\mathbf{V}'$ ortogonális, tehát vagy egy forgatás, vagy egy tükrözés. Mivel $\mathbf{V}$ oszlopai épp a $\mathbf{v}_i$ vektorok, ezért $\mathbf{V}'\mathbf{v}_i = \mathbf{e}_i$. Ezután a $\mathbf{\Sigma}$ a két tengely irányában nyújt/összenyom: $\mathbf{\Sigma}\mathbf{e}_i = \sigma_i\mathbf{e}_i$. Végül $\mathbf{U}$ ismét egy forgatás vagy tükrözés: $\mathbf{U}\sigma\mathbf{e}_i = \sigma\mathbf{U}\mathbf{e}_i = \sigma\mathbf{u}_i$.



9.1. ábra: Az egységkör képe. Legyen $\mathbf{A}$ egy 2×2 -es, valós, 2-rangú mátrix. A $\mathbf{v}_i \mapsto \mathbf{A}\mathbf{v}_i = \sigma_i\mathbf{u}_i$ leképezés hatása az egységkörtől lépésként jól szemléltethető: $\mathbf{V}'\mathbf{v}_i = \mathbf{e}_i$, $\mathbf{\Sigma}\mathbf{e}_i = \sigma_i\mathbf{e}_i$, $\mathbf{U}\sigma\mathbf{e}_i = \sigma\mathbf{U}\mathbf{e}_i = \sigma\mathbf{u}_i$, azaz $\mathbf{V}'$ a $\{\mathbf{v}_i\}$ bázist a standardba viszi ortogonális leképezéssel, ott $\mathbf{\Sigma}$ tengelyirányban nyújtja/összenyomja, végül az ortogonális $\mathbf{U}$ hat rá.

Ezután szemléltessük egy 2×3 -as, valós, 2-rangú mátrix szinguláris felbontását. Először $\mathbf{V}'$ hat a tér vektoraira. $\mathbf{V}'$ ortogonális, és a $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ ortonormált bázist a standard bázisba viszi: $\mathbf{V}'\mathbf{v}_i = \mathbf{e}_i$ ($i = 1, 2, 3$). Ezután a $\mathbf{\Sigma}$ a két első tengely irányában nyújt/összenyom: $\mathbf{\Sigma}\mathbf{e}_i = \sigma_i\mathbf{e}_i$ ($i = 1, 2$), azonban a harmadik tengely irányával párhuzamosan vetíti: $\mathbf{\Sigma}\mathbf{e}_3 = \mathbf{0}$. A kép itt nem egy ellipszison, hanem a teljes általa határolt tartomány. Végül az ortogonális $\mathbf{U}$ ezt elforgatja vagy tükrözi egy egyenesre.



9.2. ábra: Az egységgömb képe. Legyen $\mathbf{A}$ egy 2×3 -as, valós, 2-rangú mátrix. A $\mathbf{v}_i \mapsto \mathbf{A}\mathbf{v}_i = \sigma_i\mathbf{u}_i$ leképezés hatása az egységgömb felületén: $\mathbf{V}'\mathbf{v}_i = \mathbf{e}_i$ ($i = 1, 2, 3$), $\mathbf{\Sigma}\mathbf{e}_i = \sigma_i\mathbf{e}_i$ ($i = 1, 2$), $\mathbf{U}\sigma\mathbf{e}_i = \sigma\mathbf{U}\mathbf{e}_i = \sigma\mathbf{u}_i$, azaz $\mathbf{V}'$ a $\{\mathbf{v}_i\}$ bázist a standardba viszi, ott $\mathbf{\Sigma}$ az első két tengelyirányban nyújtja/összenyomja, de a harmadik tengelyirányban vetíti, így a gömbfelület képe egy ellipszistartomány, amire végül $\mathbf{U}$ hat.

9.7. TÉTEL (EGYSÉGGÖMB KÉPE). Legyen $\mathbf{A}$ egy r -rangú, $m \times n$ -es valós mátrix. Az $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ leképezés $\mathbb{R}^n$ egységgömbjének felületét, azaz az $\mathbf{e}'\mathbf{e} = 1$ egyenletet kielégítő pontokat az $\mathbb{R}^m$ egy r -dimenziós altere

- a) egy ellipszoidjának felületére képzi, ha $r = n$, és
b) egy ellipszoidja által határolt tartományára képzi, ha $r < n$.

BIZONYÍTÁS. Tekintsük $\mathbf{A}$ szinguláris felbontásának diadikus alakját:

$$\mathbf{A} = \sigma_1 \mathbf{u}_1 \mathbf{v}_1' + \sigma_2 \mathbf{u}_2 \mathbf{v}_2' + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r'.$$

Ha $\mathbf{e} \in \mathbb{R}^n$ egy egységvektor, akkor $\mathbf{V}'\mathbf{e}$ is egységvektor, azaz $(\mathbf{v}_1'\mathbf{e})^2 + (\mathbf{v}_2'\mathbf{e})^2 + \cdots + (\mathbf{v}_n'\mathbf{e})^2 = 1$, hisz $\mathbf{V}$ ortogonális mátrix. Így a fenti diadikus alakot használva

$$\begin{aligned} \mathbf{A}\mathbf{e} &= \sigma_1 \mathbf{u}_1 \mathbf{v}_1' \mathbf{e} + \sigma_2 \mathbf{u}_2 \mathbf{v}_2' \mathbf{e} + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r' \mathbf{e} \\ &= (\sigma_1 \mathbf{v}_1' \mathbf{e}) \mathbf{u}_1 + (\sigma_2 \mathbf{v}_2' \mathbf{e}) \mathbf{u}_2 + \cdots + (\sigma_r \mathbf{v}_r' \mathbf{e}) \mathbf{u}_r \\ &= x_1 \mathbf{u}_1 + x_2 \mathbf{u}_2 + \cdots + x_r \mathbf{u}_r, \end{aligned}$$

ahol $x_i = \sigma_i \mathbf{v}_i' \mathbf{e}$ ($i = 1, 2, \dots, r$). Az $x_i = 0$, ha $i = r+1, \dots, m$ és $\mathbf{x} = (x_1, x_2, \dots, x_m)$ értékadással kapjuk, hogy $\mathbf{A}\mathbf{e} = \mathbf{U}\mathbf{x}$. Így $\mathbf{U}$ ortogonalitása miatt $|\mathbf{A}\mathbf{e}| = |\mathbf{U}\mathbf{x}| = |\mathbf{x}|$. Ennek alapján fölírható az az egyenlet, melyet $\mathbf{A}\mathbf{e}$ pontjai kielégítenek, mivel

$$\begin{aligned} \left(\frac{x_1}{\sigma_1}\right) + \left(\frac{x_2}{\sigma_2}\right) + \cdots + \left(\frac{x_r}{\sigma_r}\right) &= (\mathbf{v}_1'\mathbf{e})^2 + (\mathbf{v}_2'\mathbf{e})^2 + \cdots + (\mathbf{v}_r'\mathbf{e})^2 \\ &\leq (\mathbf{v}_1'\mathbf{e})^2 + (\mathbf{v}_2'\mathbf{e})^2 + \cdots + (\mathbf{v}_n'\mathbf{e})^2 = 1. \end{aligned}$$

Eszerint az egyenlet

$$\begin{aligned} \left(\frac{x_1}{\sigma_1}\right) + \left(\frac{x_2}{\sigma_2}\right) + \cdots + \left(\frac{x_r}{\sigma_r}\right) &= 1, \quad \text{ha } r = n, \\ \left(\frac{x_1}{\sigma_1}\right) + \left(\frac{x_2}{\sigma_2}\right) + \cdots + \left(\frac{x_r}{\sigma_r}\right) &\leq 1, \quad \text{ha } r < n, \end{aligned} \quad \square$$

Polárfelbontás A komplex számok exponenciális alakja – azaz az $re^{i\varphi}$ alak – egy nemnegatív nyújtási tényező (r) és egy egységnyi abszolút értékű komplex szám ($e^{i\varphi}$, ami a komplex síkon φ -vel való forgatás) szorzata. A komplex síkon e szám polárkoordinátás alakja (r, φ) . Az analóg mátrixfelbontás több mérnöki alkalmazásban, pl. az anyagtranszformációk leírásánál használható.

Polárfelbontáson egy négyzetes mátrixnak egy pozitív szemidefinit és egy ortogonális mátrix szorzatára való felbontását értjük.

9.8. TÉTEL (POLÁRFELBONTÁS). Bármely komplex (valós) négyzetes $\mathbf{A}$ mátrix előáll

$$\mathbf{A} = \mathbf{P}\mathbf{Q}$$

alakban, ahol $\mathbf{P}$ pozitív szemidefinit önadjungált (szimmetrikus) mátrix, $\mathbf{Q}$ pedig unitér (ortogonális). Ha $\mathbf{A}$ invertálható, akkor $\mathbf{P}$ pozitív definit, és a felbontás egyértelmű.

BIZONYÍTÁS. A felbontás az $\mathbf{A}$ szinguláris felbontásából megkapható:

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^* = \mathbf{U}\Sigma\mathbf{U}^*\mathbf{U}\mathbf{V}^* = (\mathbf{U}\Sigma\mathbf{U}^*)(\mathbf{U}\mathbf{V}^*),$$

ahonnan $\mathbf{P} = \mathbf{U}\Sigma\mathbf{U}^*$, $\mathbf{Q} = \mathbf{U}\mathbf{V}^*$. A $\mathbf{P}$ pozitív szemidefinit, hisz hasonló a pozitív szemidefinit Σ mátrixhoz. Amennyiben $\mathbf{A}$ invertálható, akkor Σ pozitív definit. $\mathbf{P}$ önadjungált, hisz $(\mathbf{U}\Sigma\mathbf{U}^*)^* = \mathbf{U}\Sigma^*\mathbf{U}^* = \mathbf{U}\Sigma\mathbf{U}^*$. $\mathbf{Q}$ unitér (ortogonális), hisz két unitér (ortogonális) mátrix szorzata. A $\mathbf{P}$ egyértelmű nem csak a szemidefinit esetben, ugyanis

$$\mathbf{A}\mathbf{A}^* = \mathbf{P}\mathbf{Q}\mathbf{Q}^*\mathbf{P}^* = \mathbf{P}\mathbf{P}^* = \mathbf{P}^2, \quad \square$$

azaz $\mathbf{P} = \sqrt{\mathbf{A}\mathbf{A}^*}$, és pozitív szemidefinit önadjungált mátrix négyzetgyöke egyértelmű. Ha $\mathbf{P}$ pozitív definit, akkor invertálható, így $\mathbf{Q} = \mathbf{P}^{-1}\mathbf{A}$ is egyértelmű. $\square$

► A polárfelbontás nem csak analóg a komplex számok exponenciális alakjával, de determinánsa is épp ezt az alakot adja: ha $\det \mathbf{P} = r$, $\det \mathbf{Q} = e^{i\varphi}$ (hisz $\mathbf{Q}$ unitér, így determinánsának abszolút értéke 1), akkor $\det \mathbf{A} = re^{i\varphi}$.

► Hasonló állítás mondható fordított sorrenddel is, ráadásul azonos unitér (ortogonális) mátrixszal, azaz létezik olyan pozitív szemidefinit önadjungált $\hat{\mathbf{P}}$ mátrix, hogy $\mathbf{A} = \mathbf{Q}\hat{\mathbf{P}}$.

► Valós térben a polárfelbontás geometriai jelentése az, hogy minden mátrixleképezés két olyan leképezés egymás utánjaként áll elő, amelyekből az egyik forgatja vagy forgatva tükrözi a teret ($\mathbf{Q}$), a másik pedig egy ortonormált bázis tengelyei mentén nyújtja/összenyomja a teret minden tengelyirányban egy-egy nemnegatív tényező szerint.

9.9. PÉLDA (POLÁRFELBONTÁS KISZÁMÍTÁSA). Számítsuk ki a 9.4. példában is szereplő

$$\mathbf{A} = \begin{bmatrix} 0 & -2 & 2 \\ -2 & 3 & -2 \\ 4 & -2 & 0 \end{bmatrix}$$

mátrix polárfelbontását!

MEGOLDÁS. A 9.4. példában megadtuk a valós $\mathbf{A}$ mátrix szinguláris felbontását. Így a $\mathbf{P} = \mathbf{U}\Sigma\mathbf{U}'$, $\mathbf{Q} = \mathbf{U}\mathbf{V}'$ képletekbe való helyettesítés megadja a választ:

$$\mathbf{A} = \begin{bmatrix} 2 & -2 & 0 \\ -2 & 3 & -2 \\ 0 & -2 & 4 \end{bmatrix} \begin{bmatrix} -4/9 & -8/9 & 1/9 \\ -4/9 & 1/9 & -8/9 \\ 7/9 & -4/9 & -4/9 \end{bmatrix}. \quad \square$$

Pszseudoinverz A szinguláris felbontás egy új lehetőséget ad a pszeudoinverz kiszámítására, sőt akár a komplexek fölött való definiálására is.

A 7.54. e) pontjának azonnali következménye, hogy ha Σ az $\mathbf{A}$ mátrix diagonális alakja a szinguláris felbontásában, akkor Σ^+ főátlójának i -edik eleme $1/\sigma_i$ ($i = 1, 2, \dots, r$), minden más elem 0.

9.10. TÉTEL (A PSZEUDOINVERZ KISZÁMÍTÁSA). Legyen $\mathbf{A}$ egy valós mátrix és legyen a redukált szinguláris felbontása $\mathbf{A} = \mathbf{U}_1 \Sigma_1 \mathbf{V}_1'$, a szinguláris felbontása $\mathbf{A} = \mathbf{U} \Sigma \mathbf{V}'$. Ekkor

$$\mathbf{A}^+ = \mathbf{V}_1 \Sigma_1^{-1} \mathbf{U}_1' = \mathbf{V} \Sigma^+ \mathbf{U}'.$$

BIZONYÍTÁS. Az $\mathbf{A} = \mathbf{U}_1 (\Sigma_1 \mathbf{V}_1')$ felbontásban $\mathbf{U}_1$ teljes oszloprangú, $\Sigma_1 \mathbf{V}_1'$ teljes sorrangú, így alkalmazható $\mathbf{A}$ -ra a (7.15) képlet. Eszerint

$$\begin{aligned} \mathbf{A}^+ &= (\Sigma_1 \mathbf{V}_1')' (\Sigma_1 \mathbf{V}_1' (\Sigma_1 \mathbf{V}_1')')^{-1} (\mathbf{U}_1' \mathbf{U}_1)^{-1} \mathbf{U}_1' = \mathbf{V}_1 \Sigma_1^{-2} \mathbf{U}_1' \\ &= \mathbf{V}_1 \Sigma_1^{-1} \mathbf{U}_1'. \end{aligned}$$

Ebből következik a másik egyenlőség is, mivel

$$\mathbf{V} \Sigma^+ \mathbf{U}' = \begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_2 \end{bmatrix} \begin{bmatrix} \Sigma_1^{-1} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix} \begin{bmatrix} \mathbf{U}_1' \\ \mathbf{U}_2' \end{bmatrix} = \mathbf{V}_1 \Sigma_1^{-1} \mathbf{U}_1'. \quad \square$$

► A pszeudoinverzet valós mátrixokra definiáltuk, e tétel képleteivel komplexekre is természetes módon kiterjeszthető. Ha $\mathbf{A} = \mathbf{U} \Sigma \mathbf{V}^* = \mathbf{U}_1 \Sigma_1 \mathbf{V}_1^*$ az $\mathbf{A}$ szinguláris és redukált szinguláris felbontása, akkor legyen $\mathbf{A}^+ = \mathbf{V} \Sigma^+ \mathbf{U}^* = \mathbf{V}_1 \Sigma_1^{-1} \mathbf{U}_1^*$.

9.11. PÉLDA (A PSZEUDOINVERZ KISZÁMÍTÁSA SVD-BŐL). Számítsuk ki a 9.5. példában is szereplő

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}$$

mátrix pszeudoinverzét!

MEGOLDÁS. Ha ismerjük egy mátrix szinguláris felbontását, akkor annak felhasználásával csak képletbehelyettesítés dolga a pszeudoinverz kiszámítása. Persze, akkor már a redukált alakot is ismerjük, így felesleges a teljes szinguláris alakot használni, a redukált alak használatának kevesebb a számításigénye. A 9.5. példában meghatároztuk az $\mathbf{A}$ mátrix szinguláris alakját. Abból a redukált alak

$$\mathbf{A} = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{2} \\ 2/\sqrt{6} & 0 \\ 1/\sqrt{6} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix},$$

amiből a pszeudoinverz

$$\begin{aligned}\mathbf{A}^+ &= \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{3} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{bmatrix} \\ &= \begin{bmatrix} -1/3 & 1/3 & 2/3 \\ 2/3 & 1/3 & -1/3 \end{bmatrix} \quad \square\end{aligned}$$

Információtömörítés

Feladatok

9.1. SZINGULÁRIS FELBONTÁSOK Az $\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$ mátrixnak egyetlen szinguláris értéke van, $\sigma_1 = 2$. Igazoljuk, hogy az

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} [2] \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix}$$

felbontások az $\mathbf{A}$ mátrix szinguláris és redukált szinguláris felbontásai. (Segítségül a szinguláris felbontásban a blokk-struktúrát is jelöltük.)

9.2. Számítsuk ki a

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

mátrix szinguláris érték szerinti felbontását!

Vektor- és mátrixnorma

A vektorokhoz hasonlóan a mátrixok bizonyos tulajdonságainak – például sorozataik konvergenciájának – vizsgálatában is hasznosak az olyan mennyiségek, melyek a köztük lévő különbségeket a távolságra emlékeztető módon mérik. Ehhez az abszolút érték fogalmának általánosításán keresztül vezet út. A mátrixnormák intim kapcsolatban vannak a szinguláris értékekkel.

Vektornorma

Vektor abszolút értéke – az euklideszi norma A 2- és 3-dimenziós vektorok abszolút értékét a ?? fejezetben definiáltuk, majd az n -dimenziós terekre is kiterjesztettük. Ennek segítségével két vektor távolságát is definiálni tudtuk. A következőkben olyan – az alkalmazásokban is fontos – függvényeket definiálunk, amelyek az abszolút érték „origótól való távolság” tulajdonságát általánosítják. E függvényeket *normának* nevezzük. Mindenekelőtt ilyen nevet adunk a vektor abszolút értékének is, és egyúttal valós vektorokról komplexekre is kiterjesztjük a definíciót.

9.12. DEFINÍCIÓ (EUKLIDESZI NORMA). Az $\mathbf{x}$ vektor euklideszi normája vagy más néven abszolút értéke

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{\mathbf{x}'\mathbf{x}}, \quad \text{ha } \mathbf{x} \in \mathbb{R}^n, \quad (9.3)$$

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2} = \sqrt{\mathbf{x}^*\mathbf{x}}, \quad \text{ha } \mathbf{x} \in \mathbb{C}^n. \quad (9.4)$$

Például az $\mathbf{x} = (1 + i, 1 - 2i, 3)$ vektor euklideszi normája

$$\|\mathbf{x}\|_2 = \sqrt{(1+i)(1-i) + (1-2i)(1+2i) + 3^2} = \sqrt{2+5+9} = 4.$$

- Vektor euklideszi normájára az $|\cdot|$, $\|\cdot\|$ és a $\|\cdot\|_2$ jelölések egyaránt használatosak.
- A komplex vektorokra adott definíciónak a valós speciális esete, így akár ez az egy is megfelelne. Sőt, így a definíció $\mathbb{C}$ tetszőleges részteste fölötti vektorokra is érvényes.
- Ha $\mathbf{x}$ egy tetszőleges nemzérus vektor, akkor $\mathbf{x} / \|\mathbf{x}\|_2$ egységvektor. Egy vektorból az azonos irányú egységvektor ilyen módon való képzését *normálásnak* nevezzük, és azt mondjuk, hogy az $\mathbf{x}$ vektort *normáljuk*.

A p -norma Ha egy négyzethálós utcaszerkezetű város egy kereszteződésében állunk (legyen ez az origó), akkor egy x háznyival keletre és y

háznnyival északra fekvő ponthoz vezető legrövidebb út hossza gyalog vagy taxival $x + y$. Általában e „mérték” szerint az origóból az (x, y) koordinátájú kereszteződéshez vezető legrövidebb út hossza $|x| + |y|$. Mivel itt csak a koordinátarendszer rácsvonalain haladhatunk, szokás e normát *rácsnormának* nevezni.

Egy másik normához jutunk a következő számítógépes képméretező feladattal. Ki van jelölve egy kép közepe. Egy (x, y) képpont tőle való távolsága legyen az a legkisebb c szám, hogy e pont a $(-c, -c)$ és (c, c) pontok által meghatározott négyzetbe még épp beférjen. Világos, hogy $c = \max\{|x|, |y|\}$. E normát *maximum normának* is nevezik.

Az euklideszi norma, a rácsnorma és a maximum norma is származtatható a következő általánosabb normából:

9.13. DEFINÍCIÓ (p -NORMA). A $p \geq 1$ valósra az $\mathbf{x} \in \mathbb{C}^n$ vektor p -normája $\|\mathbf{x}\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$, míg ennek határértéke a ∞ -norma, azaz $\|\mathbf{x}\|_\infty = \lim_{p \rightarrow \infty} \|\mathbf{x}\|_p$.

Például $\|(3, 4, 5)\|_3 = \sqrt[3]{27 + 64 + 125} = 6$, $\|(1 + i, 0)\|_1 = 1 + \sqrt{2}$.

- Világos, hogy a 2-norma megegyezik az euklideszi-normával, az 1-norma a rácsnormával.
- A maximum norma megegyezik a ∞ -normával, azaz

$$\|\mathbf{x}\|_\infty = \lim_{p \rightarrow \infty} \|\mathbf{x}\|_p = \lim_{p \rightarrow \infty} \left(\sum_{i=1}^n |x_i|^p \right)^{1/p} = \max_i |x_i|.$$

Ennek bizonyításához jelöljük a legnagyobb abszolút értékű koordinátát $x_{\max}$ -szal. Ekkor minden x_i koordinátára $|x_i|/|x_{\max}| \leq 1$, és így

$$1 \leq \sum_{i=1}^n |x_i/x_{\max}|^p \leq n.$$

Mindegyik kifejezést $1/p$ -edik hatványra emelve, majd $|x_{\max}|$ -szal beszorozva kapjuk, hogy

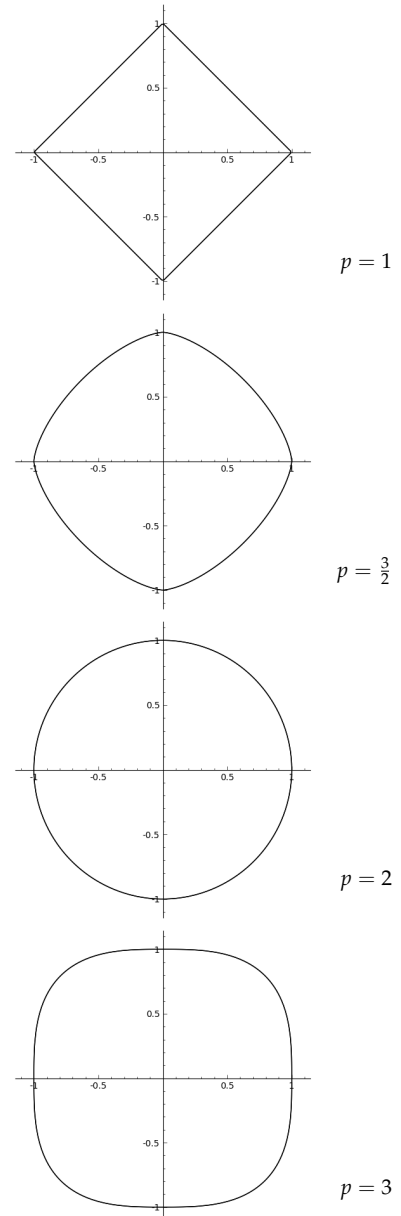
$$|x_{\max}| \leq |x_{\max}| \left(\sum_{i=1}^n \left| \frac{x_i}{x_{\max}} \right|^p \right)^{1/p} \leq |x_{\max}| n^{1/p},$$

és $n^{1/p} \rightarrow 1$, ha $p \rightarrow \infty$, ami bizonyítja az állítást.

- Érdekes megtekinteni az origótól valamely normában egységnyi távolságra lévő pontok halmazát, vagyis az egységgömböt. A 9.3 ábra az 1-, $\frac{3}{2}$ -, 2-, 3- és ∞ -normához tartozó egységköröket (2-dimenziós egységgömböket) mutatja.

A norma általános fogalma Az előzőekben az abszolút értékhez – vagyis az origótól való távolsághoz – kerestünk hasonló függvényt. Kérdés azonban, hogy milyen tulajdonságok fontosak számunkra, melyeket

Az angolszász tankönyvek a rácsnormára gyakran a *Manhattan norm* vagy a *taxicab norm* kifejezést használják.



meg akarunk őrizni. Az abszolút értékkel definiált távolság használatakor a következő tulajdonságok tűnnek fontosnak:

- (a) $|\mathbf{x}| \geq 0$, azaz vektor abszolút értéke *nem negatív*.
 - (b) $|\mathbf{x}| = 0$ pontosan akkor áll fenn, ha $\mathbf{x} = \mathbf{0}$. Ennek fontos tartalma, hogy a $d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|$ képlettel definiált távolságfüggvény *szeparálja a pontokat*, azaz két különböző pont távolsága sosem 0.
 - (c) $|c\mathbf{x}| = |c||\mathbf{x}|$, ami a lineáris leképezéseknél megismert homogenitásra emlékeztető tulajdonság: szokás *pozitív homogenitásnak* nevezni.
 - (d) $|\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|$, amit *háromszögegyenlőtlenség* néven ismerünk.
- E tulajdonságok a következő definícióhoz vezetnek:

9.14. DEFINÍCIÓ (NORMA). Egy $f : \mathbb{R}^n \rightarrow \mathbb{R}$, vagy $f : \mathbb{C}^n \rightarrow \mathbb{R}$ függvényt *normának* nevezzük, ha fennállnak a következők:

1. $f(\mathbf{x}) \geq 0$ minden $\mathbf{x}$ vektorra, és $f(\mathbf{x}) = 0$ pontosan akkor áll fenn, ha $\mathbf{x} = \mathbf{0}$,
2. $f(c\mathbf{x}) = |c|f(\mathbf{x})$ minden $\mathbf{x}$ vektorra,
3. $f(\mathbf{x} + \mathbf{y}) \leq f(\mathbf{x}) + f(\mathbf{y})$.

Az $f(\mathbf{x})$ értéket $\mathbf{x}$ normájának nevezzük.

► A normát általában az abszolút értékre emlékeztető $\|\cdot\|$ zárójellel jelöljük, azaz $\mathbf{x}$ normáját $\|\mathbf{x}\|$ jelöli. E jelöléssel tehát a norma egy $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$, vagy $\|\cdot\| : \mathbb{C}^n \rightarrow \mathbb{R}$ függvény.

► $\|\mathbf{x}\| = \|\mathbf{-x}\|$ bármely $\|\cdot\|$ normára igaz, hisz $\|\mathbf{-x}\| = |-1| \|\mathbf{x}\| = \|\mathbf{x}\|$.

► Igen hasznos a háromszögegyenlőtlenség különbségre fölírt következő alakja:

$$\|\mathbf{z} - \mathbf{x}\| \geq \|\mathbf{z}\| - \|\mathbf{x}\| \quad (9.5)$$

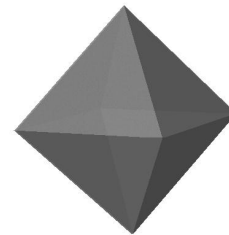
Ez a következőképp igazolható: legyen $\mathbf{z} = \mathbf{x} + \mathbf{y}$, ekkor a háromszögegyenlőtlenségből kapjuk, hogy $\|\mathbf{z} - \mathbf{x}\| \geq \|\mathbf{z}\| - \|\mathbf{x}\|$, de $\mathbf{x}$ és $\mathbf{z}$ szerepét fölcserélve $\|\mathbf{x} - \mathbf{z}\| \geq \|\mathbf{x}\| - \|\mathbf{z}\|$ is igaz, így $\|\mathbf{z} - \mathbf{x}\| = \|\mathbf{x} - \mathbf{z}\|$ igazolja az egyenlőtlenséget.

► Axiomatikus felépítésben kevesebb is megkövetelhető a norma definíciójában, nevezetesen az első pont két feltétele helyett elég egy is, azaz elég a következő:

- 1' ha $f(\mathbf{x}) = 0$, akkor $\mathbf{x} = \mathbf{0}$,
- 2' $f(c\mathbf{x}) = |c|f(\mathbf{x})$ minden $\mathbf{x}$ vektorra,
- 3' $f(\mathbf{x} + \mathbf{y}) \leq f(\mathbf{x}) + f(\mathbf{y})$.

A definíció utolsó két tulajdonságából adódik, hogy bármely $\mathbf{x}$ vektorra $f(\mathbf{x}) \geq 0$, és $f(\mathbf{0}) = 0$, így 1.–3. ekvivalens 1'–3'-vel (ld. a 9.4. feladat).

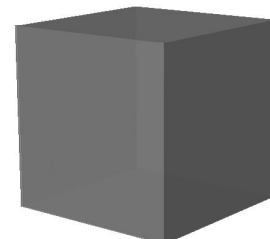
► A p -norma minden $1 \leq p \leq \infty$ esetben norma. Ennek bizonyítása meglehetősen technikai jellegű, ezért csak a feladatok közt közöljük (ld. 9.13.). A bizonyítás két nevezetes egyenlőtlenségre – a Hölder- és



$p = 1$



$p = 2$



$p = \infty$

a Minkowski-egyenlőtlenségre – épül, melyeket ugyancsak feladatként tűzünk ki (ld. 9.11., 9.12.). Valójában a *Minkowski-egyenlőtlenség* maga a háromszögegyenlőtlenség:

$$\|\mathbf{x} + \mathbf{y}\|_p \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p. \quad (9.6)$$

A Hölder-egyenlőtlenség a CBS-egyenlőtlenség általánosítása:

$$|\mathbf{x}^* \mathbf{y}| \leq \|\mathbf{x}\|_p \|\mathbf{y}\|_q, \text{ ahol } \frac{1}{p} + \frac{1}{q} = 1. \quad (9.7)$$

► A legfontosabb esetekben, vagyis a $p = 1$, $p = 2$ és $p = \infty$ esetben annak bizonyítása, hogy a p -norma norma, eddigi ismereteinket felhasználva egyszerű, ezért annak meggondolását minden olvasónak ajánljuk (ld. 9.5., 9.6. feladatok).

► Normából további normák származtathatóak. Ha $\mathbf{x} \mapsto \|\mathbf{x}\|$ egy norma, és A egy egy-egy értelmű lineáris leképezés, akkor az $\mathbf{x} \mapsto \|A\mathbf{x}\|$ leképezés is norma (9.9. feladat), továbbá norma az

$$\mathbf{x} \mapsto \sup_{\mathbf{y} \neq \mathbf{0}} \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{y}\|}$$

függvény is (ld. 9.10.).

► Azonnal látszik, hogy

$$\max_i \{|x_i|\} \leq \sqrt{|x_1|^2 + \dots + |x_n|^2} \leq |x_1| + \dots + |x_n|$$

azaz

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1. \quad (9.8)$$

Másrészt az is könnyen igazolható (ld. 9.7., hogy

$$\|\mathbf{x}\|_1 \leq \sqrt{n} \|\mathbf{x}\|_2, \|\mathbf{x}\|_2 \leq \sqrt{n} \|\mathbf{x}\|_\infty \text{ és } \|\mathbf{x}\|_1 \leq n \|\mathbf{x}\|_\infty. \quad (9.9)$$

Ezek az egyenlőtlenségek vezetnek a normák ekvivalenciájának fogalmához, ami a következő paragrafus témája.

► Minden norma folytonos függvény. Ez például a (9.5) egyenlőtlenségnek és a (9.9) első becslésének következménye (ld. 9.8. feladat).

Vektornormák ekvivalenciája A (9.8) és a (9.9) egyenlőtlenségek szerint

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_1 \text{ és } \|\mathbf{x}\|_1 \leq n \|\mathbf{x}\|_\infty,$$

vagyis mindkét normának felső korlátját adja a másik egy megfelelő konstansszorosa. Ez azt jelenti, hogy például a konvergenciakérdések eldöntésében e két norma egyformán viselkedik, vagyis egy vektorozat pontosan akkor konvergens az egyik szerint, ha a másik szerint is.

9.15. **DEFINÍCIÓ** (NORMÁK EKVIVALENCIÁJA). Azt mondjuk, hogy az $\|\cdot\|_a$ és $\|\cdot\|_b$ normák ekvivalensek, ha van olyan c és d pozitív valós, hogy $\|\cdot\|_a \leq c \|\cdot\|_b$ és $\|\cdot\|_b \leq d \|\cdot\|_a$.

A (9.8) és a (9.9) egyenlőtlenségek azt mutatják, hogy az 1-, 2- és ∞ -normák mind ekvivalensek.

9.16. **TÉTEL** (MINDEN NORMA EKVIVALENS). Legyen $\mathbb{K} = \mathbb{C}$ vagy $\mathbb{R}$. A $\mathbb{K}^n$ téren értelmezett bármely két norma ekvivalens.

BIZONYÍTÁS. Megmutatjuk, hogy tetszőleges $\mathbb{K}^n$ -en értelmezett $\|\cdot\|$ norma ekvivalens az 1-normával. Ebből azonnal következik, hogy bármely két norma ekvivalens egymással.

A háromszögegyenlőtlenséget alkalmazva az $\mathbf{x} = x_1 \mathbf{e}_1 + \dots + x_n \mathbf{e}_n$ felbontásra kapjuk, hogy

$$\|\mathbf{x}\| \leq \sum_{i=1}^n |x_i| \|\mathbf{e}_i\| \leq c \sum_{i=1}^n |x_i| = c \|\mathbf{x}\|_1,$$

ahol $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ a standrad bázis, és $c = \max_i \|\mathbf{e}_i\|$. Ezzel bizonyítottuk, hogy $\|\mathbf{x}\| \leq c \|\mathbf{x}\|_1$.

Az $\|\mathbf{x}\|_1 \leq d \|\mathbf{x}\|$ egyenlőtlenség bizonyításához meg kell mutatnunk, hogy $\|\mathbf{x}\|_1 / \|\mathbf{x}\|$ felülről korlátos a nemnulla vektorok halmazán. Indirekt módon tegyük fel, hogy van olyan $\{\mathbf{x}_k\}$ sorozat, hogy $\|\mathbf{x}_k\|_1 / \|\mathbf{x}_k\| \rightarrow \infty$, ha $k \rightarrow \infty$. Ekkor $\|\mathbf{x}_k\| / \|\mathbf{x}_k\|_1 \rightarrow 0$, azaz az $\mathbf{y}_k = \mathbf{x}_k / \|\mathbf{x}_k\|_1$ olyan sorozat, hogy $\|\mathbf{y}_k\| \rightarrow 0$, és $\|\mathbf{y}_k\|_1 = 1$. Mivel az 1-normájú egységgömb korlátos, ezért feltételezhető, hogy az $\mathbf{y}_k$ sorozat konvergens (egyébként vegyük egy konvergens részsorozatot), melynek $\mathbf{y}$ -nal jelölt határértéke is az egységgömbön van, azaz $\|\mathbf{y}\|_1 = 1$, így $\mathbf{y} \neq \mathbf{0}$. Másrészt $\|\cdot\|$ folytonos, így $\|\mathbf{y}_k\| \rightarrow \|\mathbf{y}\|$, tehát $\|\mathbf{y}\| = 0$, ami ellentmondás. $\square$

► Fölvetődik a kérdés, hogy ha az összes norma ekvivalens, akkor mi értelme bevezetni normák ekvivalenciájának fogalmát! A válasz az, hogy az ekvivalenciát csak véges dimenziós terekre bizonyítottuk, és valóban, végtelen dimenziós terekben nem teljesül. Az viszont fontos következmény, hogy véges dimenziós terekben vektorok konvergenciakérdéseinek eldöntéséhez mindig olyan normát választhatunk, ami a legkényelmesebben használható, hisz az eredmény a normaválasztástól független.

Mátrixnorma

Vektornormák mátrixokon Egy $m \times n$ -es mátrix tekinthető egy mn -dimenziós vektornak is, így a vektorokra definiált normák mátrixokra is alkalmazhatók. Ezek között legfontosabb a 2-norma mátrixokra való kiterjesztése, mely több ekvivalens alakban is felírható.

9.17. DEFINÍCIÓ (FROBENIUS-NORMA). Az $\mathbf{A} \in \mathbb{C}^{m \times n}$ mátrix Frobenius-normája

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2} = \sqrt{\sum_{i=1}^m \|\mathbf{A}_{i*}\|_2^2} = \sqrt{\sum_{j=1}^n \|\mathbf{A}_{*j}\|_2^2}.$$

Itt azért nem a 2-norma elnevezést használjuk, mert azt más normára tartogatjuk. A Frobenius-norma további módokon is számolható:

9.18. TÉTEL (FROBENIUS-NORMA EKVIVALENS ALAKJAI).

$$\|\mathbf{A}\|_F = \sqrt{\text{trace}(\mathbf{A}^* \mathbf{A})} = \sqrt{\sum_{i=1}^{\min(m,n)} \sigma_i^2}. \quad (9.10)$$

BIZONYÍTÁS. Az első alak azonnal következik a nyom definíciójából, hisz $\mathbf{A}^* \mathbf{A}$ átlójának i -edik eleme éppen $\|\mathbf{A}_{*i}\|_2^2$ -tel egyezik meg. Mivel pedig egy mátrix nyoma megegyezik sajátértékeinek összegével, és $\mathbf{A}^* \mathbf{A}$ sajátértékei pedig a szinguláris értékek négyzetei, ezért $\text{trace}(\mathbf{A}^* \mathbf{A}) = \sum_{i=1}^{\min(m,n)} \sigma_i^2$, ami bizonyítja az állítást. $\square$

Egy vektornorma akkor nyújthat igazán hasznos információt a mátrixokról is, ha valamilyen módon kapcsolatban van a mátrix olyan sajátosságaival, ami mn -dimenziós vektorként nehezen leírható. A mátrixszal való szorzás például ilyen. A vektorok 2-normája és a mátrixok Frobenius-normája közt például a következő összefüggés áll fenn:

9.19. ÁLLÍTÁS. Bármely $\mathbf{x} \in \mathbb{C}^n$ vektorra és $\mathbf{A} \in \mathbb{C}^{m \times n}$ mátrixra

$$\|\mathbf{A}\mathbf{x}\|_2 \leq \|\mathbf{A}\|_F \|\mathbf{x}\|_2. \quad (9.11)$$

BIZONYÍTÁS. Igazolására a Cauchy–Bunyakovszkij–Schwarz-egyenlőtlenséget alkalmazzuk:

$$\|\mathbf{A}\mathbf{x}\|_2^2 = \sum_{i=1}^m |\mathbf{A}_{i*} \mathbf{x}|^2 \leq \sum_{i=1}^m \|\mathbf{A}_{i*}\|_2^2 \|\mathbf{x}\|_2^2 = \|\mathbf{A}\|_F^2 \|\mathbf{x}\|_2^2. \quad \square$$

► E tulajdonság általában nem igaz minden mátrixokra alkalmazott vektornormára. Ilyen a maximum norma is, melyet a következőképp vihetünk át mátrixokra:

$$\|\mathbf{A}\|_{\max} = \max_{i,j} \{|a_{ij}|\}.$$

Például az $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ mátrix maximum normája 2. E mátrix különböző vektorokkal vett szorzatának normája és a normák szorzata közt

mindhárom reláció fennállhat. Például az

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

szorzatokban a normákra a $2 \cdot 1 > 1$, $2 \cdot 1 = 2$, $2 \cdot 1 < 3$ relációk teljesülnek.

► A (9.11) tulajdonság mátrixok szorzatára is átvihető, azaz

$$\|\mathbf{AB}\|_F \leq \|\mathbf{A}\|_F \|\mathbf{B}\|_F$$

igaz bármely $\mathbf{A} \in \mathbb{C}^{m \times n}$ és $\mathbf{B} \in \mathbb{C}^{n \times k}$ mátrixokra.

A mátrixnorma általános fogalma A vektornormák alkalmazhatók mátrixokra is. Sok könyv azonban – és így teszünk mi is – egy normát csak akkor tekint mátrixnormának, ha a vektornorma axiómái mellett egy mátrixszorzásra vonatkozóan is eleget tesz.

9.20. DEFINÍCIÓ (MÁTRIXNORMA). Legyen $\mathbb{K} = \mathbb{R}$ vagy $\mathbb{C}$. Egy K fölötti mátrixokon értelmezett valós értékű $\|\cdot\|$ függvény mátrixnorma, ha tetszőleges azonos méretű $\mathbf{A}$ és $\mathbf{B}$ mátrixra és összeszorozható $\mathbf{A}$ és $\mathbf{C}$ mátrixra

1. $\|\mathbf{A}\| \geq 0$, és $\|\mathbf{A}\| = 0$ pontosan akkor áll fenn, ha $\mathbf{A} = \mathbf{O}$,
2. $\|c\mathbf{A}\| = |c| \|\mathbf{A}\|$,
3. $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$,
4. $\|\mathbf{AC}\| \leq \|\mathbf{A}\| \|\mathbf{C}\|$.

A korábbiak szerint tehát a Frobenius-norma mátrixnorma, míg a maximum normát nem tekintjük mátrixnormának.

9.21. DEFINÍCIÓ. Azt mondjuk, hogy a $\|\cdot\|_M$ mátrixnorma valamint a $\|\cdot\|_a$ és a $\|\cdot\|_b$ vektornormák illeszkednek vagy konzisztensek, ha tetszőleges $\mathbf{A}$ mátrixra és megfelelő dimenziójú $\mathbf{x}$ vektorra

$$\|\mathbf{Ax}\|_b \leq \|\mathbf{A}\|_M \|\mathbf{x}\|_a.$$

Például a (9.11) szerint a Frobenius-norma illeszkedik a 2-normához.

Indukált norma E paragrafusban vektornormákból kiindulva újabb mátrixnormákhoz jutunk.

9.22. DEFINÍCIÓ (INDUKÁLT NORMA). Legyen $\|\cdot\|$ egy tetszőleges vektornorma. Ekkor az

$$\|\mathbf{A}\| = \max_{\|\mathbf{x}\|=1} \|\mathbf{Ax}\| \quad (9.12)$$

egyenlőséggel definiált függvényt a vektornorma által indukált mátrixnormának nevezzük.

► Az indukált mátrixnormára a vektornorma jelölését szokás használni, így például a mátrix p -norma definíciója

$$\|\mathbf{A}\|_p = \max_{\|\mathbf{x}\|_p=1} \|\mathbf{Ax}\|_p.$$

► Ha mátrix helyett lineáris leképezésre értelmezzük a fenti definíciót, *operátornormáról* beszélünk.

► A normák ekvivalenciájából következik, hogy bármely normában az egységgömb korlátos és zárt. Így a rajta értelmezett folytonos $\mathbf{x} \mapsto \mathbf{Ax}$ függvénynek van maximuma és minimuma, tehát a definíció értelmes.

► Az előző megjegyzést is figyelembe véve könnyen igazolható, hogy a definíció a következő ekvivalens alakokba is átírható:

$$\|\mathbf{A}\| = \sup_{\|\mathbf{x}\| \neq 0} \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|} = \max_{\|\mathbf{x}\| \neq 0} \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|}. \quad (9.13)$$

Ez abból következik, hogy az $\mathbf{y} = \mathbf{x} / \|\mathbf{x}\|$ jelöléssel

$$\|\mathbf{A}\| = \max_{\|\mathbf{y}\|=1} \|\mathbf{Ay}\| = \max_{\mathbf{x} \neq 0} \left\| \mathbf{A} \left(\frac{\mathbf{x}}{\|\mathbf{x}\|} \right) \right\| = \max_{\mathbf{x} \neq 0} \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|}$$

► Azt még igazolnunk kell, hogy a mátrixnorma elnevezés e függvényre valóban jogos.

9.23. TÉTEL (INDUKÁLT NORMA TULAJDONSÁGAI). Legyen $\|\cdot\|$ egy tetszőleges vektornorma, ekkor az indukált, a (9.12) képlettel definiált mátrixfüggvény

- a) mátrixnorma, azaz fennáll a 9.20. definíció mind a négy feltétele,
- b) illeszkedik az indukáló vektornormához, azaz

$$\|\mathbf{Ax}\| \leq \|\mathbf{A}\| \|\mathbf{x}\|.$$

BIZONYÍTÁS. Először az illeszkedést igazoljuk. Ha $\mathbf{x} = \mathbf{0}$, akkor az egyenlőtlenség teljesül, hisz mindkét oldalán $\mathbf{0}$ áll. Ha $\mathbf{x} \neq \mathbf{0}$, akkor a (9.13) szerint

$$\|\mathbf{A}\| = \max_{\mathbf{z} \neq \mathbf{0}} \frac{\|\mathbf{Az}\|}{\|\mathbf{z}\|} \geq \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|},$$

azaz $\|\mathbf{Ax}\| \leq \|\mathbf{A}\| \|\mathbf{x}\|$.

A mátrixnormát definiáló 4 feltétel közül az első három nyilvánvalóan teljesül. A negyedik igazolásához legyen $\mathbf{y}$ egy olyan vektor, amelyben az $\mathbf{x} \mapsto \|\mathbf{ABx}\|$ felveszi a maximumát az egységgömbön, azaz amelyre $\|\mathbf{y}\| = 1$, és

$$\|\mathbf{ABy}\| = \max_{\|\mathbf{x}\|=1} \|\mathbf{ABx}\| = \|\mathbf{AB}\|.$$

Ekkor az illeszkedés kétszeri alkalmazásával

$$\|\mathbf{AB}\| = \|\mathbf{ABy}\| \leq \|\mathbf{A}\| \|\mathbf{By}\| \leq \|\mathbf{A}\| \|\mathbf{B}\| \|\mathbf{y}\| = \|\mathbf{A}\| \|\mathbf{B}\|. \quad \square$$

Az 1-, 2- és ∞ -norma mátrixokra A fõnt definiált p -normák közül mátrixokra is az 1-, 2- és a ∞ -norma a legfontosabb. Kiszámításukra a definíciónál egyszerűbb módszer is adódik.

9.24. TÉTEL (1-, 2- ÉS ∞ -NORMA KISZÁMÍTÁSA). Legyen $\mathbf{A} \in \mathbb{C}^{m \times n}$, ekkor

$$\|\mathbf{A}\|_1 = \max_j \sum_{i=1}^n |a_{ij}| = \text{legnagyobb abszolút oszlopösszeg}, \quad (9.14)$$

$$\|\mathbf{A}\|_\infty = \max_i \sum_{j=1}^m |a_{ij}| = \text{legnagyobb abszolút sorösszeg}, \quad (9.15)$$

$$\|\mathbf{A}\|_2 = \|\mathbf{A}^*\|_2 = \max_{\|\mathbf{x}\|_2=1} \max_{\|\mathbf{y}\|_2=1} |\mathbf{y}^* \mathbf{A} \mathbf{x}| = \sigma_1, \quad (9.16)$$

ahol σ_1 az $\mathbf{A}$ legnagyobb szinguláris értéke, azaz $\mathbf{A}^* \mathbf{A}$ legnagyobb sajátértékének gyöke. Ha az $\mathbf{A} \in \mathbb{C}^{n \times n}$ mátrix invertálható, akkor

$$\|\mathbf{A}^{-1}\|_2 = \max_{\|\mathbf{x}\|_2=1} \frac{1}{\|\mathbf{A} \mathbf{x}\|_2} = \frac{1}{\min_{\|\mathbf{x}\|_2=1} \|\mathbf{A} \mathbf{x}\|_2} = \frac{1}{\sigma_n}, \quad (9.17)$$

ahol σ_n az $\mathbf{A}$ legkisebb szinguláris értéke.

BIZONYÍTÁS. $p = 1$: Bármely $\mathbf{x}$ vektorra $\|\mathbf{x}\|_1 = 1$ esetén a skalárookra vonatkozó háromszögegyenlõtlenség miatt

$$\begin{aligned} \|\mathbf{A} \mathbf{x}\|_1 &= \sum_{i=1}^m \left| \sum_{j=1}^n a_{ij} x_j \right| \leq \sum_{i=1}^m \sum_{j=1}^n |a_{ij}| |x_j| = \sum_{j=1}^n |x_j| \sum_{i=1}^m |a_{ij}| \\ &\leq \left(\sum_{j=1}^n |x_j| \right) \max_j \sum_{i=1}^m |a_{ij}| = \max_j \sum_{i=1}^m |a_{ij}| \end{aligned}$$

Ez a maximum el is érhető, mert ha a k -adik oszlopban a legnagyobb az abszolút értékek összege, akkor $\|\mathbf{A} \mathbf{e}_k\|_1 = \max_j \sum_{i=1}^m |a_{ij}|$.

$p = \infty$: Bármely $\mathbf{x}$ vektorra $\|\mathbf{x}\|_\infty = 1$ esetén

$$\|\mathbf{A} \mathbf{x}\|_\infty = \max_i \left| \sum_{j=1}^n a_{ij} x_j \right| \leq \max_i \sum_{j=1}^n |a_{ij}| |x_j| \leq \max_i \sum_{j=1}^n |a_{ij}|.$$

Ez a maximum el is érhető, mert ha a k -adik sorban a legnagyobb az abszolút értékek összege, akkor az

$$\mathbf{x} = \begin{bmatrix} \overline{a_{kj}} \\ |a_{kj}| \end{bmatrix}$$

vektorra $\|\mathbf{x}\|_\infty = 1$ és $\|\mathbf{A} \mathbf{x}\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|$.

$p = 2$: A CBS-egyenlõtlenség szerint $|\mathbf{y}^* \mathbf{A} \mathbf{x}| \leq \|\mathbf{y}\|_2 \|\mathbf{A} \mathbf{x}\|_2$, így

$$\max_{\|\mathbf{x}\|_2=1} \max_{\|\mathbf{y}\|_2=1} |\mathbf{y}^* \mathbf{A} \mathbf{x}| \leq \max_{\|\mathbf{x}\|_2=1} \|\mathbf{A} \mathbf{x}\|_2 = \|\mathbf{A}\|_2.$$

Így csak azt kell megmutatni, hogy van olyan $\mathbf{x}_0$ és $\mathbf{y}_0$ egységvektor, melyre az előbbi egyenlőtlenségben egyenlőség áll. Legyen $\mathbf{x}_0$ az a vektor, melyben $\|\mathbf{Ax}\|_2$ a maximumot adja, és $\mathbf{y}$ ennek normált képe, azaz

$$\|\mathbf{Ax}_0\|_2 = \max_{\|\mathbf{x}\|_2=1} \|\mathbf{Ax}\|_2 = \|\mathbf{A}\|_2, \quad \mathbf{y}_0 = \frac{\mathbf{Ax}_0}{\|\mathbf{Ax}_0\|_2} = \frac{\mathbf{Ax}_0}{\|\mathbf{A}\|_2}.$$

Ekkor

$$\mathbf{y}_0^* \mathbf{Ax}_0 = \frac{\mathbf{x}_0^* \mathbf{A}^* \mathbf{Ax}_0}{\|\mathbf{A}\|_2} = \frac{\|\mathbf{Ax}_0\|_2^2}{\|\mathbf{A}\|_2} = \frac{\|\mathbf{A}\|_2^2}{\|\mathbf{A}\|_2} = \|\mathbf{A}\|_2.$$

Az $\|\mathbf{A}\|_2 = \sigma_1$ igazolásához a következő maximumot keressük:

$$\|\mathbf{A}\|_2^2 = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{Ax}\|_2^2}{\|\mathbf{x}\|_2^2} = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^* \mathbf{A}^* \mathbf{Ax}}{\mathbf{x}^* \mathbf{x}}.$$

Mivel $\mathbf{A}^* \mathbf{A}$ önadjungált, ezért létezik sajátvektoraiból álló ortonormált bázisa. Vektorai legyenek $\mathbf{u}_1, \dots, \mathbf{u}_n$, a hozzájuk tartozó sajátértékek $\lambda_1, \dots, \lambda_n$, melyek közül λ_1 legyen a legnagyobb. Ekkor egyrészt

$$\lambda_1 = \frac{\mathbf{u}_1^* \mathbf{A}^* \mathbf{Au}_1}{\mathbf{u}_1^* \mathbf{u}_1},$$

másrészt bármely $\mathbf{x} = \sum_j c_j \mathbf{u}_j \neq \mathbf{0}$ vektorra

$$\lambda_1 - \frac{\mathbf{x}^* \mathbf{A}^* \mathbf{Ax}}{\mathbf{x}^* \mathbf{x}} = \lambda_1 - \frac{\sum_{j=1}^n \lambda_j c_j^2}{\sum_{j=1}^n c_j^2} = \frac{\sum_{j=1}^n (\lambda_1 - \lambda_j) c_j^2}{\sum_{j=1}^n c_j^2} \geq 0,$$

tehát $\|\mathbf{A}\|_2^2 = \lambda_1$, azaz $\|\mathbf{A}\|_2 = \sqrt{\lambda_1} = \sigma_1$.

$p = 2$ esetén $\mathbf{A}^{-1}$ normája is egyszerűen számolható:

$$\begin{aligned} \frac{1}{\min_{\|\mathbf{x}\|_2=1} \|\mathbf{Ax}\|_2} &= \max_{\|\mathbf{x}\|_2=1} \frac{1}{\|\mathbf{Ax}\|_2} = \max_{\mathbf{y} \neq \mathbf{0}} \frac{1}{\left\| \mathbf{A} \frac{\mathbf{A}^{-1} \mathbf{y}}{\|\mathbf{A}^{-1} \mathbf{y}\|_2} \right\|_2} = \max_{\mathbf{y} \neq \mathbf{0}} \frac{\|\mathbf{A}^{-1} \mathbf{y}\|_2}{\|\mathbf{y}\|_2} \\ &= \max_{\mathbf{y} \neq \mathbf{0}} \left\| \mathbf{A}^{-1} \left(\frac{\mathbf{y}}{\|\mathbf{y}\|_2} \right) \right\|_2 = \max_{\|\mathbf{x}\|_2=1} \|\mathbf{A}^{-1} \mathbf{x}\|_2 = \|\mathbf{A}^{-1}\|_2. \end{aligned}$$

Mivel $\mathbf{A}^{-1}$ szinguláris értékei az $\mathbf{A}$ szinguláris értékeinek reciprokai, ezért $\mathbf{A}^{-1}$ legnagyobb szinguláris értéke az $\mathbf{A}$ legkisebb szinguláris értékének reciproka. $\square$

► Az 1-, a ∞ - és a 2-normára szokásos másik elnevezés: *oszlopnorma*, *sornorma* és *spektrálnorma*.

Feladatok

9.3. Számítsuk ki az alábbi vektorok megadott normáit!

1. $\mathbf{x} = (\sqrt{3} - i, 6i, 3)$, $\mathbf{y} = (0.1, -0.2, -0.2)$, $p = 1, 2, \infty$;
2. $(1, 2, 2)$, $(2, 3, 6)$, $(1, 4, 8)$, $(4, 4, 7)$, $p = 2$;
3. $(i, 2, \sqrt{2} - \sqrt{2}i, -4i)$, $p = 1, 2, \infty$;
4. $(3, 4, 5)$, $(11, 12, 13, 14)$, $p = 3$;
5. $\|(95800, 217519, 414560)\|_4$, $\|(27, 84, 110, 133)\|_5$.

9.4. Mutassuk meg, hogy a 9.14. definíció 1.–3. pontja és az utána következő megjegyzés 1'–3' pontja ekvivalensek.

9.5. Mutassuk meg, hogy az 1-norma norma.

9.6. Mutassuk meg, hogy az ∞ -norma norma.

9.7. Mutassuk meg, hogy $\|\mathbf{x}\|_p \leq c \|\mathbf{x}\|_q$, ahol c a következő táblázatból kiolvasható, ahol p értékei a sorok, q értékei az oszlopok fejlécében vannak.

	1	2	∞
1		$\sqrt{n}$	n
2	1		$\sqrt{n}$
∞	1	1	

9.8. Mutassuk meg, hogy minden norma folytonos függvény.

9.9. Mutassuk meg, hogy ha $\|\cdot\|$ egy norma, és A egy egy-egy értelmű lineáris leképezés, akkor az $\mathbf{x} \mapsto \|A\mathbf{x}\|$ leképezés is norma.

9.10. Mutassuk meg, hogy ha $\|\cdot\|$ egy norma, akkor az

$$\mathbf{x} \mapsto \sup_{\mathbf{y} \neq 0} \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{y}\|}$$

függvény is az. E normát duálnormának is szokás nevezni.

9.11. **HÖLDER-EGYENLŐTLENSÉG** Igazoljuk, hogy bármely $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ vektor és $p, q \geq 1$ valósok esetén

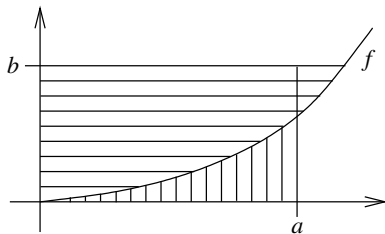
$$\sum_{i=1}^n |x_i y_i| \leq \|\mathbf{x}\|_p \|\mathbf{y}\|_q, \text{ ahol } \frac{1}{p} + \frac{1}{q} = 1. \quad (9.18)$$

A következő lépéseket javasoljuk:

1. Igazoljuk, hogy $a, b > 0$, $p, q \geq 1$ és $\frac{1}{p} + \frac{1}{q} = 1$ esetén

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Ennek igazolására határozzuk meg az $f : x \mapsto x^{p-1}$ függvényhez tartozó két alábbi satírozott tartomány területét!



2. Az előző egyenlőtlenségben végezzük el az

$$a = \frac{|x_i|}{\|\mathbf{x}\|_p}, \quad b = \frac{|y_i|}{\|\mathbf{y}\|_q}$$

helyettesítéseket, majd ezzel igazoljuk a Hölder-egyenlőtlenség (9.18) alakját.

3. Végül bizonyítsuk a Hölder-egyenlőtlenség (9.7) alakját is.

9.12. **MINKOWSKI-EGYENLŐTLENSÉG** Igazoljuk, hogy bármely $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ vektor és $p \geq 1$ valós esetén

$$\|\mathbf{x} + \mathbf{y}\|_p \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p. \quad (9.19)$$

A következő lépéseket javasoljuk:

1. Igazoljuk, majd alkalmazzuk az x_i, y_i számokra az

$$|a + b|^p = |a + b| |a + b|^{p/q} \leq |a| |a + b|^{p/q} + |b| |a + b|^{p/q}$$

egyenlőtlenséget, ahol $1/p + 1/q = 1$.

2. Alkalmazzuk a Hölder-egyenlőtlenséget a $\sum_{i=1}^n |x_i| |x_i| + |y_i|^{p/q}$ kifejezésre.

9.13. Mutassuk meg, hogy a p -norma norma.

9.14. Számítsuk ki az alábbi mátrixok Frobenius-, 1-, 2- és ∞ -normáját!

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 3 & 0 \\ 4 & 0 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 2 & -2 \\ 1 & 2 \end{bmatrix}.$$

9.15. Számítsuk ki az alábbi mátrixok Frobenius-, 1-, 2- és ∞ -normáját!

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 4 & 0 & 0 \\ 0 & 0 & 8 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 2 & 2 & -1 \\ -1 & 2 & 2 \\ 2 & -1 & 2 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & 2 \end{bmatrix}.$$

9.16. Konstruáljunk olyan $\mathbf{A}$, $\mathbf{B}$ és $\mathbf{C}$ mátrixokat, hogy maximum normájukra $\|\mathbf{AB}\|_{\max} < \|\mathbf{A}\|_{\max} \|\mathbf{B}\|_{\max}$, $\|\mathbf{AC}\|_{\max} = \|\mathbf{A}\|_{\max} \|\mathbf{C}\|_{\max}$ és $\|\mathbf{BC}\|_{\max} > \|\mathbf{B}\|_{\max} \|\mathbf{C}\|_{\max}$ legyen.

9.17. Igazoljuk, hogy minden indukált $\|\cdot\|$ mátrixnormára $\|\mathbf{I}\| = 1$, ugyanakkor $\|\mathbf{I}\|_F = \sqrt{n}$.

9.18. Igazoljuk, hogy tetszőleges mátrixnormára

$$\rho(\mathbf{A}) \leq \|\mathbf{A}\|,$$

ahol $\rho(\mathbf{A})$ az $\mathbf{A}$ spektrálsugara.

9.19. Bizonyítsuk be, hogy ha $\mathbf{A}$ normális ($\mathbf{A}^* \mathbf{A} = \mathbf{A} \mathbf{A}^*$), akkor $\|\mathbf{A}\|_2 = \rho(\mathbf{A})$.

Megoldások

9.1. Ellenőriznünk kell, hogy

- az egyenlőség fennáll,
- az $\mathbf{U}$ és $\mathbf{V}$ ortogonális mátrixok, $\mathbf{\Sigma}$ diagonális,
- az $\mathbf{u}_1 = (1/\sqrt{2}, 1/\sqrt{2})$ és $\mathbf{v}_1 = (1/\sqrt{2}, 1/\sqrt{2}, 0)$ vektorokra $\mathbf{A}\mathbf{v}_1 = 2\mathbf{u}_1$.

Ezek mind nyilvánvalóak, vagyis az első felbontás szinguláris. Mivel $r(\mathbf{A}) = 1$, ezért $\mathbf{U}$ első oszlopát és $\mathbf{V}'$ első sorát, valamint $\mathbf{\Sigma}$ bal felső elemét meghagyva valóban a második alakot kapjuk.

9.2. A $\mathbf{B}'\mathbf{B}$ mátrix karakterisztikus polinomja $-\lambda^3 + 4\lambda^2 - 3\lambda = -(\lambda - 3)(\lambda - 1)\lambda$, így sajátértékei 3, 1 és 0. A szinguláris értékek $\sqrt{3}$ és 1. A sajátvektorok számításához 3-dimenziós vektorokkal kell számolnunk. Talán jobban járunk, ha inkább a $\mathbf{B}\mathbf{B}'$ mátrixszal próbálkozunk. Mivel $\mathbf{B}\mathbf{B}' = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, ezért a karakterisztikus polinom $\lambda^2 - 4\lambda + 3 = (\lambda - 3)(\lambda - 1)$. A szinguláris értékek tehát $\sqrt{3}$ és 1, összhangban az előbbi számítással. A hozzájuk tartozó egységnyi hosszú sajátvektorok most nem a $\mathbf{V}_1$, hanem $\mathbf{U}_1$ oszlopait adják: $\mathbf{u}_1 = (1/\sqrt{2}, 1/\sqrt{2})$, $\mathbf{u}_2 = (-1/\sqrt{2}, 1/\sqrt{2})$. A $\mathbf{v}_i = \mathbf{A}'\mathbf{u}_i/\sigma_i$ képletet használva a $\mathbf{V}$ mátrix is meghatározható. Tehát

$$\mathbf{B} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ 2/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} \\ -1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix}$$

a $\mathbf{B}$ mátrix szinguláris felbontása.

9.3.

- $\|\mathbf{x}\|_1 = 11$, $\|\mathbf{x}\|_2 = 7$, $\|\mathbf{x}\|_\infty = 6$, $\|\mathbf{y}\|_1 = 0.5$, $\|\mathbf{y}\|_2 = 0.3$, $\|\mathbf{x}\|_\infty = 0.2$.
- Ezek az úgynevezett Pitagorászi számnégyesekből képzett vektorok, amelyekben a koordináták négyzetösszege négyzetszám, így a 2-normájuk egész. A normák 3, 7, 9, 9.
- 9, 5, 4;
- 6, 20;
- e két példa a $p = 4$ és $p = 5$ értékre a legkisebb

olyan $p - 1$ -dimenziós pozitív egész vektor, melynek p -normája egész: $\|(95800, 217519, 414560)\|_4 = 422481$, $\|(27, 84, 110, 133)\|_5 = 144$. Euler még azt sejtette, hogy ilyen nincs.

$$\begin{aligned} 9.14. \quad & \|\mathbf{A}\|_F = 5, \|\mathbf{A}\|_1 = 6, \|\mathbf{A}\|_2 = 5, \|\mathbf{A}\|_\infty = 6. \\ & \|\mathbf{B}\|_F = 5, \|\mathbf{B}\|_1 = 7, \|\mathbf{B}\|_2 = 5, \|\mathbf{B}\|_\infty = 7. \\ & \|\mathbf{C}\|_F = \sqrt{13}, \|\mathbf{C}\|_1 = 4, \|\mathbf{C}\|_2 = 3, \|\mathbf{C}\|_\infty = 4. \end{aligned}$$

$$\begin{aligned} 9.15. \quad & \|\mathbf{A}\|_F = 9, \|\mathbf{A}\|_1 = 8, \|\mathbf{A}\|_2 = 8, \|\mathbf{A}\|_\infty = 8. \\ & \|\mathbf{B}\|_F = 3\sqrt{3}, \|\mathbf{B}\|_1 = 5, \|\mathbf{B}\|_2 = 3, \|\mathbf{B}\|_\infty = 5. \\ & \|\mathbf{C}\|_F = 3\sqrt{3}, \|\mathbf{C}\|_1 = 5, \|\mathbf{C}\|_2 = 5, \|\mathbf{C}\|_\infty = 5. \end{aligned}$$

9.16. Tekintsük például az alábbi három mátrixot:

$$\mathbf{A} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 1 \\ 2 & 2 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

Ezek mindegyikében 2 az elemek maximuma, így bármely két mátrix maximum normájának szorzata 4. Szorzataik:

$$\mathbf{AB} = \begin{bmatrix} 0 & 2 \\ 2 & 2 \end{bmatrix}, \quad \mathbf{AC} = \begin{bmatrix} 2 & 4 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{BC} = \begin{bmatrix} 0 & 1 \\ 2 & 6 \end{bmatrix}.$$

Ezek elemeinek maximuma rendre 2, 4, 6.

9.18. Ha λ egy tetszőleges sajátértéke $\mathbf{A}$ -nak, és $\mathbf{x}$ a hozzá tartozó egyik sajátvektor, azaz $\mathbf{Ax} = \lambda\mathbf{x}$, akkor

$$\mathbf{A}\mathbf{x}\mathbf{x}^* = \lambda\mathbf{x}\mathbf{x}^* \rightsquigarrow \|\mathbf{A}\| \|\mathbf{x}\mathbf{x}^*\| \geq \|\mathbf{A}\mathbf{x}\mathbf{x}^*\| = |\lambda| \|\mathbf{x}\mathbf{x}^*\|,$$

és mivel $\mathbf{x} \neq \mathbf{0}$, így $\mathbf{x}\mathbf{x}^* \neq \mathbf{0}$, azaz $\|\mathbf{x}\mathbf{x}^*\| \neq 0$, vagyis leosztva vele $|\lambda| \leq \|\mathbf{A}\|$ adódik. Ez minden sajátértékre, így a spektrálsugárra is igaz.

9.19. Ha $\mathbf{A}$ normális, akkor unitéren hasonló egy diagonális $\mathbf{D}$ mátrixhoz, azaz $\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{Q}^*$ valamely unitér $\mathbf{Q}$ mátrixszal. Ekkor $\mathbf{A}^*\mathbf{A} \sim \mathbf{D}^*\mathbf{D}$ is fönnáll, ugyanis $\mathbf{A}^*\mathbf{A} = (\mathbf{Q}\mathbf{D}\mathbf{Q}^*)^*(\mathbf{Q}\mathbf{D}\mathbf{Q}^*) = \mathbf{Q}\mathbf{D}^*\mathbf{D}\mathbf{Q}$, tehát $\mathbf{A}^*\mathbf{A}$ és $\mathbf{D}^*\mathbf{D}$ sajátértékei megegyeznek. Másrészt $\mathbf{D}^*\mathbf{D}$ minden sajátértéke $|\lambda|^2$ alakú, ahol λ az $\mathbf{A}$ valamely sajátértéke. Összegezve: mivel $\|\mathbf{A}\|_2 = \sigma_1$, azaz az $\mathbf{A}^*\mathbf{A}$ legnagyobb sajátértékének gyöke, ami viszont megegyezik $\mathbf{A}$ legnagyobb sajátértékével, azaz a $\rho(\mathbf{A})$ spektrálsugárral.

Jordan-féle normálalak

Mind bizonyos problémák megértésében, mind a gyakorlati alkalmazásokban fontos lehet, hogy egy lineáris leképezés mátrixát milyen bázisban írjuk fel. Láttuk milyen sokat segít, ha a bázis sajátvektorokból áll, sőt, ha még ortonormált is. Ekkor a mátrix diagonális alakot ölt. Kérdés azonban, milyen mátrixok diagonalizálhatók, és amelyek nem, azokkal mit tudunk kezdeni.

Schur-felbontás

Általánosított sajátvektorok és a Jordan-blokk Tekintsük a 8.15.. példában szereplő

$$\mathbf{A} = \begin{bmatrix} 4 & 1 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{bmatrix}$$

mátrix hatását a standard bázis vektorain:

$$\mathbf{A}\mathbf{e}_1 = 4\mathbf{e}_1$$

$$\mathbf{A}\mathbf{e}_2 = \mathbf{e}_1 + 4\mathbf{e}_2$$

$$\mathbf{A}\mathbf{e}_3 = \mathbf{e}_2 + 4\mathbf{e}_3$$

Átrendezés után kapjuk, hogy:

$$(\mathbf{A} - 4\mathbf{I})\mathbf{e}_1 = \mathbf{0}$$

$$(\mathbf{A} - 4\mathbf{I})\mathbf{e}_2 = \mathbf{e}_1$$

$$(\mathbf{A} - 4\mathbf{I})\mathbf{e}_3 = \mathbf{e}_2$$

Az $\mathbf{A} - 4\mathbf{I}$ mátrix fenti hatását a következő diagrammal fogjuk szemléltetni:

$$\mathbf{0} \xleftarrow{\mathbf{A}-4\mathbf{I}} \mathbf{e}_1 \xleftarrow{\mathbf{A}-4\mathbf{I}} \mathbf{e}_2 \xleftarrow{\mathbf{A}-4\mathbf{I}} \mathbf{e}_3$$

Eszerint $\mathbf{e}_1$ sajátvektor, és mint a 8.15.. példában láttuk, $\mathbf{A}$ -nak más sajátvektora nincs. Viszont az előző egyenletekből a következő össze-

függés adódik:

$$\begin{aligned}(\mathbf{A} - 4\mathbf{I})^2 \mathbf{e}_2 &= \mathbf{0} \\ (\mathbf{A} - 4\mathbf{I})^3 \mathbf{e}_3 &= \mathbf{0}.\end{aligned}$$

Az ilyen vektorokra külön elnevezést fogunk alkalmazni:

10.1. DEFINÍCIÓ (ÁLTALÁNOSÍTOTT SAJÁTVEKTOR). Az $\mathbf{x} \neq \mathbf{0}$ vektort a négyzetes $\mathbf{A}$ mátrix λ sajátértékéhez tartozó általánosított sajátvektornak nevezzük, ha valamilyen k természetes számra $(\mathbf{A} - \lambda\mathbf{I})^k \mathbf{x} = \mathbf{0}$. $k = 1$ esetén $\mathbf{x}$ sajátvektor. Az általánosított sajátvektorokból álló $\mathbf{x}_i$ ($i = 1, 2, \dots, k$) sorozatot Jordan-láncnak nevezzük, ha $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x}_i = \mathbf{x}_{i-1}$ és $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x}_1 = \mathbf{0}$.

Az előzőekben vizsgált konkrét példában tehát a térnek sajátvektorokból álló bázisa nincs, hisz a sajátaltér csak 1-dimenziós, de általánosított sajátvektorokból álló bázisa van.

10.2. PÉLDA (JORDAN-LÁNC KONSTRUKCIÓJA). Keressünk az alábbi mátrixok minden sajátaltérének minden bázisvektorához egy abban végződő Jordan-láncot:

$$\mathbf{A} = \begin{bmatrix} 6 & -1 & -3 \\ -1 & 5 & 2 \\ 2 & -1 & 1 \end{bmatrix} \text{ és a } \mathbf{B} = \begin{bmatrix} 3 & 0 & 1 \\ 0 & 4 & 0 \\ -1 & 0 & 5 \end{bmatrix}$$

MEGOLDÁS. A karakterisztikus polinom mindkét mátrix esetén $-\lambda^3 + 12\lambda^2 - 48\lambda + 64 = (4 - \lambda)^3$, ezért $\lambda = 4$ háromszoros algebrai multiplicitású sajátérték.

Az $\mathbf{A}$ mátrix esetén a sajátaltér 1-dimenziós, melyet az $\mathbf{x}_1 = (1, -1, 1)$ vektor feszít ki.

Mivel $(\mathbf{A} - 4\mathbf{I})^3 = \mathbf{0}$, ezért legföljebb három hosszú sorozatra számíthatunk: olyan $\mathbf{x}_2$ és $\mathbf{x}_3$ vektort keresünk, melyre $(\mathbf{A} - 4\mathbf{I})\mathbf{x}_2 = \mathbf{x}_1$ és $(\mathbf{A} - 4\mathbf{I})\mathbf{x}_3 = \mathbf{x}_2$. Ez két egyenletrendszer megoldását jelenti. Elég a megoldások közül egyet kiválasztani:

$$\mathbf{A} - 4\mathbf{I} = \begin{bmatrix} 2 & -1 & -3 \\ -1 & 1 & 2 \\ 2 & -1 & -3 \end{bmatrix}, \text{ egy-egy megoldás: } \mathbf{x}_2 = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} -1 \\ -2 \\ 0 \end{bmatrix}.$$

Ezek a következő láncot adják:

$$\mathbf{0} \xleftarrow{\mathbf{A}-4\mathbf{I}} \mathbf{x}_1 = (1, -1, 1) \xleftarrow{\mathbf{A}-4\mathbf{I}} \mathbf{x}_2 = (0, -1, 0) \xleftarrow{\mathbf{A}-4\mathbf{I}} \mathbf{x}_3 = (-1, -2, 0)$$

A $\mathbf{B}$ mátrix esetén a sajátaltér 2-dimenziós, melyet az $\mathbf{x}_1 = (1, 0, 1)$ és az $\mathbf{y}_1(0, 1, 0)$ vektorok feszítenek ki.

Mivel $(\mathbf{B} - 4\mathbf{I})^2 = \mathbf{0}$, ezért legföljebb kettő hosszú láncra számíthatunk: olyan $\mathbf{x}_2$ és $\mathbf{y}_2$ vektort keresünk, melyre $(\mathbf{B} - 4\mathbf{I})\mathbf{x}_2 = \mathbf{x}_1$ és

$(\mathbf{B} - 4\mathbf{I})\mathbf{y}_2 = \mathbf{y}_1$. Ezek közül az első megoldható, és megoldásainak egyike $(-1, 0, 0)$, míg a második nem oldható meg. Ez két láncra vezet:

$$\begin{aligned} \mathbf{0} &\xleftarrow{\mathbf{B}-4\mathbf{I}} \mathbf{x}_1 = (1, 0, 1) \xleftarrow{\mathbf{B}-4\mathbf{I}} \mathbf{x}_2 = (-1, 0, 0) \\ \mathbf{0} &\xleftarrow{\mathbf{B}-4\mathbf{I}} \mathbf{y}_1 = (0, 1, 0) \end{aligned}$$

Az egyik lánc kettő hosszú, a másik egy hosszú. $\square$

A bevezetőben mutatott mátrix-típus fontos szerepet játszik a továbbiakban:

10.3. DEFINÍCIÓ (JORDAN-BLOKK). *Azt a négyzetes mátrixot, melynek főátlójában azonos λ értékek, fölötte 1-esek, egyebütt 0-k állnak, azaz melynek alakja*

$$\mathbf{J}_\lambda = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{bmatrix} \quad (10.1)$$

Jordan-blokknak nevezzük.

A fentiek mintájára könnyen látható, hogy az ilyen mátrixoknak a standard bázis minden vektora általánosított sajátvektora, tehát ki mondható a következő megállapítás: minden Jordan-bloknak van általánosított sajátvektorokból álló bázisa.

10.4. PÉLDA (JORDAN-LÁNCHÓZ TARTOZÓ JORDAN-BLOKK). *Írjuk fel a 10.2.. példában szereplő $\mathbf{A}$ mátrixot az ott konstruált Jordan-lánc vektoraiból álló bázisban!*

MEGOLDÁS. Mivel $\mathbf{A}\mathbf{x}_3 = 4\mathbf{x}_3 + \mathbf{x}_2$, $\mathbf{A}\mathbf{x}_2 = 4\mathbf{x}_2 + \mathbf{x}_1$ és $\mathbf{A}\mathbf{x}_1 = 4\mathbf{x}_1$, ezért az $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ bázisban a leképezés mátrixa

$$\mathbf{J} = \begin{bmatrix} 4 & 1 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{bmatrix}$$

Ezt az eredményt az áttérés $\mathbf{C} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \mathbf{x}_3]$ mátrixával való közvetlen számolással is megkaphatjuk, ha elvégezzük az alábbi mátrixszorzásokat:

$$\mathbf{J} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C} = \begin{bmatrix} -2 & -3 & 0 \\ -1 & -1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 6 & -1 & -3 \\ -1 & 5 & 2 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -3 & 0 \\ -1 & 2 & 0 \\ 1 & -3 & 1 \end{bmatrix} \quad \square$$

10.5. PÉLDA (JORDAN-LÁNCOK ÉS JORDAN-BLOKKOK KAPCSOLATA). Tudjuk, hogy az $\mathbf{A}$ mátrixnak két különböző sajátértéke van, $\lambda_1 = 2$ és $\lambda_2 = 4$, valamint hogy a $\mathbf{C}$ mátrix oszlopvektorai az $\mathbf{A}$ egy Jordan-bázisát alkotják, ahol

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 1 & -1 & 0 & 1 & -1 \\ 0 & 2 & 1 & -1 & 0 & 1 & -1 \\ 2 & -2 & 2 & 2 & 0 & -1 & 1 \\ 1 & -1 & 0 & 3 & 1 & 1 & -1 \\ -1 & 1 & 0 & -1 & 5 & 1 & -1 \\ -1 & 1 & 0 & -1 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}, \mathbf{C} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

Rajzoljuk fel a Jordan-láncok diagrammját, és határozzuk meg a $\mathbf{J} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$ mátrixot a szorzások elvégzése nélkül!

MEGOLDÁS. A $\mathbf{C}$ oszlopai Jordan-bázist alkotnak, azaz minden oszlopvektor egy Jordan-lánc eleme. Mivel az $\mathbf{A}$ mátrixnak csak két különböző sajátértéke van, ez azt jelenti, hogy minden oszlopvektort vagy az $\mathbf{A} - 2\mathbf{I}$ vagy az $\mathbf{A} - 4\mathbf{I}$ mátrix vagy a zérusvektorba, vagy egy másik oszlopvektorba visz (előbbi esetben az oszlopvektor sajátvektor, utóbbi esetben csak általánosított sajátvektor). E hatást az $(\mathbf{A} - 2\mathbf{I})\mathbf{C}$ és az $(\mathbf{A} - 4\mathbf{I})\mathbf{C}$ szorzatok kiszámításával könnyen megkaphatjuk:

$$\mathbf{A} - 2\mathbf{I} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 3 & 1 & 0 \\ 0 & 0 & 0 & 2 & 3 & 3 & 0 \\ 0 & 0 & 0 & 0 & 2 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 & 2 & 2 \end{bmatrix} \quad \mathbf{A} - 4\mathbf{I} = \begin{bmatrix} -2 & -1 & 0 & 0 & 0 & 0 & 0 \\ -2 & -1 & -2 & 0 & 0 & 0 & 0 \\ 0 & -2 & -2 & 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Az első szorzatból látszik, hogy $(\mathbf{A} - 2\mathbf{I})[\mathbf{c}_1 \ \mathbf{c}_2 \ \mathbf{c}_3] = [\mathbf{0} \ \mathbf{c}_1 \ \mathbf{0}]$, míg a másodikból, hogy $(\mathbf{A} - 4\mathbf{I})[\mathbf{c}_4 \ \mathbf{c}_5 \ \mathbf{c}_6 \ \mathbf{c}_7] = [\mathbf{0} \ \mathbf{c}_4 \ \mathbf{c}_5 \ \mathbf{0}]$ (a második szorzatban már nem is kellett volna a $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3$ vektorokkal szorozni látva az előző szorzás eredményét). Ebből fölrajzolható a diagram:

$$\begin{array}{ccccccc} 0 & \xleftarrow{\mathbf{A}-2\mathbf{I}} & \mathbf{c}_1 & \xleftarrow{\mathbf{A}-2\mathbf{I}} & \mathbf{c}_2 & & \\ 0 & \xleftarrow{\mathbf{A}-2\mathbf{I}} & \mathbf{c}_3 & & & & \\ 0 & \xleftarrow{\mathbf{A}-4\mathbf{I}} & \mathbf{c}_4 & \xleftarrow{\mathbf{A}-4\mathbf{I}} & \mathbf{c}_5 & \xleftarrow{\mathbf{A}-4\mathbf{I}} & \mathbf{c}_6 \\ 0 & \xleftarrow{\mathbf{A}-4\mathbf{I}} & \mathbf{c}_7 & & & & \end{array}$$

A diagramból kiolvasható az $\mathbf{A}$ hatása a $\mathbf{c}_i$ vektorokra: $\mathbf{A}\mathbf{c}_1 = 2\mathbf{c}_1$, $\mathbf{A}\mathbf{c}_2 = 2\mathbf{c}_2 + \mathbf{c}_1$, $\mathbf{A}\mathbf{c}_3 = 2\mathbf{c}_3$, $\mathbf{A}\mathbf{c}_4 = 4\mathbf{c}_4$, $\mathbf{A}\mathbf{c}_5 = 4\mathbf{c}_5 + \mathbf{c}_4$, $\mathbf{A}\mathbf{c}_6 = 4\mathbf{c}_6 + \mathbf{c}_5$,

$\mathbf{A}\mathbf{c}_7 = 4\mathbf{c}_7$. Ebből felírható a leképezés mátrixa:

$$\mathbf{J} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C} = \left[\begin{array}{ccc|ccc|c} 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 4 \end{array} \right]$$

□

Jordan normálalak Nem hozható minden mátrix diagonális alakra, de egy ahhoz közeli alakra – ahogy azt a 10.5. példa mutatja – igen. Ebben csak közvetlenül a főátló fölött lehetnek nemnulla elemek, és azok is csak 1-esek. Ráadásul az új bázis ismeretében ez az alak jól leírja a mátrixleképezés hatását is.

10.6. TÉTEL (JORDAN NORMÁLALAK). Bármely $\mathbb{C}^{n \times n}$ -beli mátrix hasonló egy Jordan-blokkokból álló blokkdiagonális mátrixhoz. Részletesen megfogalmazva: minden $\mathbf{A} \in \mathbb{C}^{n \times n}$ mátrixhoz létezik olyan $\mathbf{C}$ mátrix, hogy a $\mathbf{J} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$ mátrix alakja

$$\mathbf{J} = \left[\begin{array}{cccc} \mathbf{J}_1 & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & \mathbf{J}_2 & \dots & \mathbf{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & \mathbf{O} & \dots & \mathbf{J}_k \end{array} \right] \quad (10.2)$$

ahol k az $\mathbf{A}$ független sajátvektorainak maximális száma, és $\mathbf{J}_i$ az i -edik sajátvektorhoz tartozó Jordan-blokk.

- A tételbeli (10.2) mátrixot az $\mathbf{A}$ mátrix Jordan-féle normálalakjának nevezzük.
- A különböző Jordan-blokkok különböző sajátvektorokhoz tartoznak, de mivel több sajátvektor is tartozhat ugyanahhoz a sajátértékhez, ezért egy sajátérték több Jordan-blokkban is szerepelhet.

BIZONYÍTÁS. Megmutatjuk, hogy ha találunk k független sajátvektort, akkor találunk k Jordan-láncot is, melyek vektorai a tér bázisát adják, és ebben a bázisban a leképezés mátrixa a tételbeli Jordan-alakot ölti.

A mátrix méretére vonatkozó teljes indukcióval bizonyítunk. $n = 1$ esetén az állítás nyilván igaz. Tegyük fel, hogy igaz az állítás minden n -nél kisebb méretű mátrixra.

Legyen λ az $n \times n$ -es $\mathbf{A}$ mátrix egy sajátértéke, és legyen $\mathbf{x}$ egy hozzá tartozó sajátvektor. Az $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ megoldásainak terét, vagyis a λ -hoz tartozó sajátalteret jelölje $\mathcal{N}_\lambda$. Ennek dimenzióját, vagyis $(\mathbf{A} - \lambda\mathbf{I})$ nullitását jelölje r .

Mivel $r > 0$, ezért a dimenziótétel miatt $(\mathbf{A} - \lambda \mathbf{I})$ oszlopterének dimenziója $n - r < n$. Jelölje az oszlopteret $\mathcal{O}_\lambda$. Az $\mathcal{O}_\lambda$ invariáns altere $\mathbf{A}$ -nak, azaz $\mathbf{A}(\mathcal{O}_\lambda) \subseteq \mathcal{O}_\lambda$, ugyanis $\mathcal{O}_\lambda$ elemei $(\mathbf{A} - \lambda \mathbf{I})\mathbf{v}$ alakúak, ahol $\mathbf{v}$ tetszőleges, és $\mathbf{A}(\mathbf{A} - \lambda \mathbf{I})\mathbf{v} = (\mathbf{A}^2 - \lambda \mathbf{A})\mathbf{v} = (\mathbf{A} - \lambda \mathbf{I})(\mathbf{A}\mathbf{v})$ ugyancsak eleme $\mathcal{O}_\lambda$ -nak. Az $\mathbf{A}$ mátrixleképezést szűkítsük le $\mathcal{O}_\lambda$ -ra, jelölje ennek mátrixát $\hat{\mathbf{A}}$.

Az $\hat{\mathbf{A}}$ mátrix $n - r$ méretű, ezért az indukciós feltevés szerint van Jordan-láncokból álló bázisa. A következő diagram e láncokat szemlélteti, és azt, hogy $\hat{\mathbf{A}} - \lambda \mathbf{I}$ hogy hat rajtuk. Mivel azonban e halmazon ez megegyezik $\mathbf{A} - \lambda \mathbf{I}$ hatásával, a nyilak fölé is ezt írjuk:

$$\begin{array}{ccccccc} \mathbf{0} & \xleftarrow{\mathbf{A} - \lambda_1 \mathbf{I}} & \mathbf{x}_1^1 & \xleftarrow{\mathbf{A} - \lambda_1 \mathbf{I}} & \dots & \xleftarrow{\mathbf{A} - \lambda_1 \mathbf{I}} & \mathbf{x}_{s_1}^1 \\ \mathbf{0} & \xleftarrow{\mathbf{A} - \lambda_2 \mathbf{I}} & \mathbf{x}_1^2 & \xleftarrow{\mathbf{A} - \lambda_2 \mathbf{I}} & \dots & \xleftarrow{\mathbf{A} - \lambda_2 \mathbf{I}} & \mathbf{x}_{s_2}^2 \\ \vdots & & \vdots & & \vdots & & \vdots \\ \mathbf{0} & \xleftarrow{\mathbf{A} - \lambda_p \mathbf{I}} & \mathbf{x}_1^p & \xleftarrow{\mathbf{A} - \lambda_p \mathbf{I}} & \dots & \xleftarrow{\mathbf{A} - \lambda_p \mathbf{I}} & \mathbf{x}_{s_p}^p \end{array}$$

Itt az $\mathbf{x}_k^j$ vektor felső indexe a lánc sorszámát jelöli.

Az $\mathbf{A} - \lambda \mathbf{I}$ oszloptere és magtere metszetének dimenzióját jelölje q , azaz legyen $q = \dim(\mathcal{O}_\lambda \cap \mathcal{N}_\lambda)$. Mivel $\mathcal{N}_\lambda$ elemei az $\mathbf{A}$ sajátvektorai, azok $\mathcal{O}_\lambda$ -ba eső báziselemei pedig az $\mathbf{x}_1^j$ vektorok, ezért q közülük a λ sajátvektora, azaz q sajátérték megegyezik λ -val, pl. az első q , azaz $\lambda_k = \lambda$ ($k = 1, 2, \dots, q$). A láncok végén lévő $\mathbf{x}_{s_k}^k$ vektorok elemei $\mathcal{O}_\lambda$ -nak, tehát mindegyikhez van olyan $\mathbf{y}_k$ vektor, hogy $(\mathbf{A} - \lambda \mathbf{I})\mathbf{y}_k = \mathbf{x}_{s_k}^k$ ($k = 1, 2, \dots, q$).

Az $\mathcal{N}_\lambda$ alter r -dimenziós, annak q -dimenziós alteréhez találunk olyan $\mathcal{Z}$ alteret, mely $(r - q)$ -dimenziós, és minden vektora sajátvektor. Így a Jordan-láncok így alakulnak:

$$\begin{array}{ccccccc} \mathbf{0} & \xleftarrow{\mathbf{A} - \lambda \mathbf{I}} & \mathbf{x}_1^1 & \xleftarrow{\mathbf{A} - \lambda \mathbf{I}} & \dots & \xleftarrow{\mathbf{A} - \lambda \mathbf{I}} & \mathbf{x}_{s_1}^1 & \xleftarrow{\mathbf{A} - \lambda \mathbf{I}} & \mathbf{y}_1 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ \mathbf{0} & \xleftarrow{\mathbf{A} - \lambda \mathbf{I}} & \mathbf{x}_1^q & \xleftarrow{\mathbf{A} - \lambda \mathbf{I}} & \dots & \xleftarrow{\mathbf{A} - \lambda \mathbf{I}} & \mathbf{x}_{s_q}^q & \xleftarrow{\mathbf{A} - \lambda \mathbf{I}} & \mathbf{y}_q \\ \mathbf{0} & \xleftarrow{\mathbf{A} - \lambda_{q+1} \mathbf{I}} & \mathbf{x}_1^{q+1} & \xleftarrow{\mathbf{A} - \lambda_{q+1} \mathbf{I}} & \dots & \xleftarrow{\mathbf{A} - \lambda_{q+1} \mathbf{I}} & \mathbf{x}_{s_{q+1}}^{q+1} & & \\ \vdots & & \vdots & & \vdots & & \vdots & & \\ \mathbf{0} & \xleftarrow{\mathbf{A} - \lambda_p \mathbf{I}} & \mathbf{x}_1^p & \xleftarrow{\mathbf{A} - \lambda_p \mathbf{I}} & \dots & \xleftarrow{\mathbf{A} - \lambda_p \mathbf{I}} & \mathbf{x}_{s_p}^p & & \\ \mathbf{0} & \xleftarrow{\mathbf{A} - \lambda \mathbf{I}} & \mathbf{z}_1 & & & & & & \\ \vdots & & \vdots & & & & & & \\ \mathbf{0} & \xleftarrow{\mathbf{A} - \lambda \mathbf{I}} & \mathbf{z}_{r-q} & & & & & & \end{array} \quad (10.3)$$

Az $\mathbf{x}$ -vektorok száma $n - r$, az $\mathbf{y}$ -vektorok száma q , a $\mathbf{z}$ -vektorok száma $r - q$, ezek összege pedig $(n - r) + q + (r - q) = n$, tehát van elég vektor egy bázishoz. Már csak a függetlenségüket kell belátni. A konstrukció olyan volt, hogy az $\mathbf{x}$ - és $\mathbf{z}$ -vektorok mind függetlenek

egymástól, csak az $\mathbf{y}$ -vektorok tőlük való függetlenségét kell bizonyítani. Tegyük fel, hogy

$$\sum_{j,k} \zeta_{j,k} \mathbf{x}_k^j + \sum_t \eta_t \mathbf{y}_t + \sum_r \zeta_r \mathbf{z}_r = \mathbf{0},$$

ahol nem minden η_t nulla. Szorozzuk meg az egyenlőséget (balról) az $\mathbf{A} - \lambda \mathbf{I}$ mátrixszal. Mivel az első q lánc végén lévő $\mathbf{x}_{s_t}^t$ -vektorok kisebb indexűekbe mennek, másrészt $(\mathbf{A} - \lambda \mathbf{I})\mathbf{y}_t = \mathbf{x}_{s_t}^t$, ezért olyan összefüggéshez jutunk, amelyben már csak $\mathbf{x}$ -vektorok lesznek, de nem minden együttható nulla, hisz van nemnulla η együttható, és ez ellentmondás. $\square$

10.7. PÉLDA (NORMÁLALAKOK). Soroljuk fel az összes lehetséges Jordan normálalakját annak a mátrixnak, melyről csak annyit tudunk, hogy $(1 - \lambda)^4$ a karakterisztikus polinomja. Ne tekintsünk különbözőnek két normálalkot, ha azok csak a Jordan-blokkok sorrendjében különböznek egymástól!

MEGOLDÁS. A karakterisztikus polinom negyedfokú, így a mátrix 4×4 -es. Mivel minden sajátérték 1, ezért a Jordan-alak főátlójában csupa 1-es szerepel. A lehetséges öt alak elemi leszámplálással megkapható:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Nem tekintjük különbözőnek a blokkok cseréjével egymásból megkapható alakokat. Például az

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

mátrix a negyedik alakból a két blokk cseréjével megkapható. $\square$

A Jordan-alak egyértelműsége A $\mathbb{C}^{n \times n}$ -beli mátrixok a Jordan normálalakjuk szerint osztályokba sorolhatók. Egy osztályba azok a mátrixok kerülnek, melyek normálalakja azonos. Egy ilyen osztályozás azonban csak akkor létezik, ha minden mátrixnak csak egyetlen normálalakja van, azaz a normálalak egyértelmű. Ezt biztosítja a következő tétel.

10.8. TÉTEL (A JORDAN-ALAK EGYÉRTELMŰSÉGE). Egy mátrix Jordan normálalakja a Jordan-blokkok sorrendjétől eltekintve egyértelmű.

BIZONYÍTÁS. A felbontás egyértelműségének bizonyításához elég belátni, hogy bármely két hasonló mátrix Jordan-alakjának meghatározó adatai a hasonlóságra nézve invariánsak.

A Jordan-blokkok, és így a Jordan-láncok száma megegyezik a független sajátvektorok maximális számával – ez invariáns.

Az egyszerűség kedvéért először tegyük fel, hogy $\mathbf{A}$ minden sajátértéke azonos, jelölje λ . A továbbiak könnyebb megértésére lássunk egy konkrét példát. Legyen az $\mathbf{A}$ mátrix karakterisztikus polinomja $(\lambda - x)^{13}$, és tegyük fel, hogy Jordan-bázisa a következőképp néz ki:

$$\begin{array}{ccccccc}
 0 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_1^1 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_2^1 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_3^1 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_4^1 \\
 0 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_1^2 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_2^2 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_3^2 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_4^2 \\
 0 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_1^3 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_2^3 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_3^3 & & \\
 0 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_1^4 & & & & & & \\
 0 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_1^5 & & & & & &
 \end{array} \quad (10.4)$$

A leghosszabb blokk mérete 4, ami $\mathbf{A} - \lambda\mathbf{I}$ ismeretében úgy kapható meg, hogy 4 az a legkisebb s kitevő, melyre $(\mathbf{A} - \lambda\mathbf{I})^s = \mathbf{O}$. Általában is igaz, a legnagyobb blokk mérete az a legkisebb s , melyre $(\mathbf{A} - \lambda\mathbf{I})^s = \mathbf{O}$. A mátrix hatványának zérus volta is invariáns, így hasonló mátrixokra a leghosszab lánc hossza is azonos.

Legyen a λ sajátértékhez tartozó i -hosszú Jordan-láncok száma m_i . A 10.4. diagramon $m_1 = 2$, $m_2 = 0$, $m_3 = 1$, $m_4 = 2$. Látható, hogy $(\mathbf{A} - \lambda\mathbf{I})$ hatványainak rangjából megmondható, hogy vektor nem futott még a nullvektorba, innen pedig az m_i értékek is kiszámolhatók. Esetünkben

$$\begin{aligned}
 m_4 &= r((\mathbf{A} - \lambda\mathbf{I})^3) = 2 \\
 m_3 + 2m_4 &= r((\mathbf{A} - \lambda\mathbf{I})^2) = 5 \\
 m_2 + 2m_3 + 3m_4 &= r(\mathbf{A} - \lambda\mathbf{I}) = 8 \\
 m_1 + m_2 + m_3 + m_4 &= n - r(\mathbf{A} - \lambda\mathbf{I}) = 13 - 8 = 5
 \end{aligned}$$

és ez az egyenletrendszer egyértelműen megoldható. Általában

$$\begin{aligned}
 m_s &= r \left((\mathbf{A} - \lambda \mathbf{I})^{s-1} \right) \\
 m_{s-1} + 2m_s &= r \left((\mathbf{A} - \lambda \mathbf{I})^{s-2} \right) \\
 m_{s-2} + 2m_{s-1} + 3m_s &= r \left((\mathbf{A} - \lambda \mathbf{I})^{s-2} \right) \\
 &\vdots \\
 m_2 + 2m_3 + \cdots + (s-1)m_s &= r(\mathbf{A} - \lambda \mathbf{I}) \\
 m_1 + m_2 + m_3 + \cdots + m_{s-1} + m_s &= n - r(\mathbf{A} - \lambda \mathbf{I})
 \end{aligned}$$

Ennek az egyenletrendszernek a jobb oldalán a hasonlóságra nézve invariáns értékek vannak, az együttthatómátrix háromszög alakú, így egyszerű visszahelyettesítéssel megoldható, főátlójában csupa egyes van, melynek következtében a megoldás egyértelmű, másrészt a megoldások mindegyike egész szám.

Ha a mátrixnak nem csak egy sajátértéke van, akkor sajátértékenként egy ilyen egyenletrendszert kapunk, mely annyiban változik az előzőhöz képest, hogy ha λ multiplicitása $m(\lambda)$, akkor $\mathbf{A} - \lambda \mathbf{I}$ hatványainak rangjához az összes λ -tól különböző sor eggyel hozzájárul, így az egyenletrendszerek jobb oldalából $(n - m(\lambda))$ -t ki kell vonni. A legnagyobb blokk mérete ekkor $\mathbf{A} - \lambda \mathbf{I}$ -nek az a legkisebb s hatványa, melynek rangja először éri el $n - m(\lambda)$ -t. Így az összes esetre általánosan érvényes egyenletrendszer:

$$\begin{aligned}
 m_s &= m(\lambda) - n + r \left((\mathbf{A} - \lambda \mathbf{I})^{s-1} \right) \\
 m_{s-1} + 2m_s &= m(\lambda) - n + r \left((\mathbf{A} - \lambda \mathbf{I})^{s-2} \right) \\
 m_{s-2} + 2m_{s-1} + 3m_s &= m(\lambda) - n + r \left((\mathbf{A} - \lambda \mathbf{I})^{s-2} \right) \\
 &\vdots \\
 m_2 + 2m_3 + \cdots + (s-1)m_s &= m(\lambda) - n + r(\mathbf{A} - \lambda \mathbf{I}) \\
 m_1 + m_2 + m_3 + \cdots + m_{s-1} + m_s &= n - r(\mathbf{A} - \lambda \mathbf{I})
 \end{aligned}$$

Ennek egyértelmű megoldhatósága bizonyítja állításunkat. $\square$

10.9. PÉLDA (JORDAN-BLOKKOK MÉRETE). Egy 10×10 -es $\mathbf{A}$ mátrixnak λ 10-szeres algebrai multiplicitású sajátértéke. $\mathbf{A} - \lambda \mathbf{I}$ hatványainak rangja rendre 5, 2, 1, 0. Írjuk fel a Jordan normálalakját!

MEGOLDÁS. A blokkok száma, ami megegyezik a Jordan-láncok számával 5, mivel $n - r(\mathbf{A}) = 10 - 5 = 5$. A leghosszabb lánc hossza 4, mivel $\mathbf{A} - \lambda \mathbf{I}$ legkisebb zérusmátrixot adó hatványa a 4-dik. Az

egyenletrendszer és megoldása, valamint a $\mathbf{J}$ Jordan-mátrix:

$$\begin{array}{rcl}
 m_4 = 1 & m_4 = 1 \\
 m_3 + 2m_4 = 2 & m_3 = 0 \\
 m_2 + 2m_3 + 3m_4 = 5 & m_2 = 2 \\
 m_1 + m_2 + m_3 + m_4 = 5 & m_1 = 2
 \end{array} \Rightarrow \mathbf{J} = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \lambda & 1 & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}$$

□

10.10. PÉLDA (JORDAN-BLOKKOK MÉRETE). Egy 14×14 -es $\mathbf{A}$ mátrix karakterisztikus polinomja $(3 - \lambda)^5(2 - \lambda)^5(1 - \lambda)^4$.

$\mathbf{A} - 3\mathbf{I}$ hatványainak rangja rendre: 12, 11, 10, 9;

$\mathbf{A} - 2\mathbf{I}$ hatványainak rangja rendre: 12, 10, 9;

$\mathbf{A} - \mathbf{I}$ hatványainak rangja rendre: 11, 10.

Írjuk fel a Jordan normálalakját!

MEGOLDÁS. Most $n = 14$, $m(3) = 5$, $m(2) = 5$, $m(1) = 4$.

$\lambda = 3$ esetén a blokkok (Jordan-láncok) száma $n - r(\mathbf{A} - 3\mathbf{I}) = 14 - 12 = 2$. A leghosszabb lánc hossza $s = 4$, ugyanis ez a legkisebb s , melyre $r((\mathbf{A} - 3\mathbf{I})^s) = 14 - 5 = 9$. Az egyenletrendszer és megoldása

$$\begin{array}{rcl}
 m_4 = 5 - 14 + 10 = 1 & m_4 = 1 \\
 m_3 + 2m_4 = 5 - 14 + 11 = 2 & m_3 = 0 \\
 m_2 + 2m_3 + 3m_4 = 5 - 14 + 12 = 3 & m_2 = 0 \\
 m_1 + m_2 + m_3 + m_4 = 14 - 12 = 2 & m_1 = 1
 \end{array} \Rightarrow$$

Hasonlóan járunk el a többi sajátérték esetén is (gondoljuk végig):

$$\begin{array}{rcl}
 m_3 = 5 - 14 + 10 = 1 & m_3 = 1 \\
 m_2 + 2m_3 = 5 - 14 + 12 = 3 & m_2 = 1 \\
 m_1 + m_2 + m_3 = 14 - 12 = 2 & m_1 = 0 \\
 \\
 m_2 = 4 - 14 + 11 = 1 & m_2 = 1 \\
 m_1 + m_2 = 14 - 11 = 3 & m_1 = 2
 \end{array} \Rightarrow$$

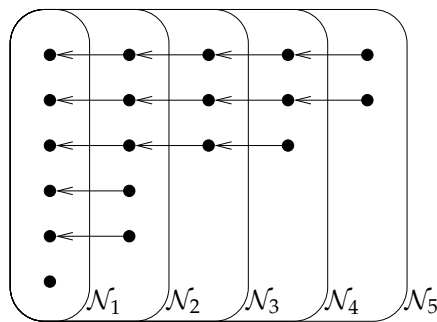
Összefoglalva:

$$J = \begin{bmatrix} 3 & 1 & & & & & \\ & 3 & 1 & & & & \\ & & 3 & 1 & & & \\ & & & 3 & & & \\ & & & & 3 & & \\ & & & & & 2 & 1 \\ & & & & & 2 & 1 \\ & & & & & & 2 \\ & & & & & 2 & 1 \\ & & & & & & 2 \\ & & & & & 1 & 1 \\ & & & & & & 1 \\ & & & & & & 1 \\ & & & & & & 1 \end{bmatrix}$$

az A mátrix Jordan normálalakja. □

A Jordan-bázis konstrukciója A Jordan-tétel bizonyításában megkonstruált bázist kis mátrixokra kézzel is ki lehet számolni. Az itt ismertetendő egyszerű naív algoritmus nem számítógépes megvalósításra való, ennél hatékonyabb is létezik.

Példaként tekintsünk egy 18-adrendű A mátrixot, melynek λ 18-szoros sajátértéke, és amelyre a Jordan-láncok hossza rendre $m_5 = 2$, $m_4 = 1$, $m_3 = 0$, $m_2 = 1$, $m_1 = 2$. Ennek alapján az $A - \lambda I$ hatása a Jordan-bázison meghatározható, amit a 10.1. ábra szemlélt. Jelölje $(A - \lambda I)^k$ nullterét $\mathcal{N}_k$. $\mathcal{N}_5$ a λ -hoz tartozó teljes általánosított sajátaltér – e példában $\mathbb{C}^{18}$. A nulltér meghatározása egy homogén lineáris egyenletrendszer megoldását jelenti, ami nem okoz nehézséget, de az már nem teljesen mindegy, hogy a bázist hogy választjuk benne.



10.1. ábra: Az $A - \lambda I$ hatványainak nullterei, és hatása az A mátrix Jordan-bázisán

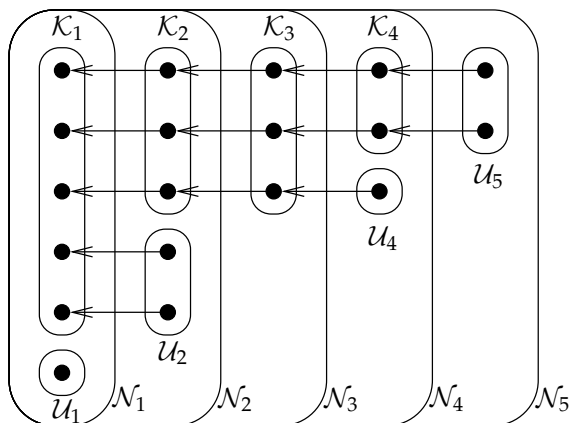
Az algoritmus – melyet a 10.2. ábra is szemléltet – lényege, hogy a nagyobb indexű terek felől indulva minden lépésben kiterjesztjük $\mathcal{N}_{i-1}$ bázisát $\mathcal{N}_i$ bázisává, majd a kiterjesztésnek vesszük az $A - \lambda I$ általi képét, ami már $\mathcal{N}_{i-1}$ -be esik, és ezt $\mathcal{N}_{i-2}$ bázisához adva azt kiterjesztjük $\mathcal{N}_{i-1}$ új bázisává. . . . A konvenció az lesz, hogy a vektortér

egy generátorának vektorait egyetlen mátrixba tesszük, és azt azonos, de félkövér betűvel jelöljük, tehát pl. $\mathcal{N}$ generátormátrixát $\mathbf{N}$ jelöli. Részletezve:

- Meghatározzuk $\mathcal{N}_i$ egy tetszőleges $\mathbf{N}_i$ bázismátrixát ($i = 1, 2, 3, 4, 5$).
- Kiegészítjük $\mathbf{N}_4$ -et $\mathcal{N}_5$ bázisává, az új báziselemek mátrixát jelölje $\mathbf{U}_5$, tehát $\mathcal{N}_5$ bázisa most $[\mathbf{N}_4 | \mathbf{U}_5]$.
- Legyen $\mathbf{K}_4$ az új elemek $\mathbf{A} - \lambda \mathbf{I}$ általi képe, azaz $\mathbf{K}_4 = (\mathbf{A} - \lambda \mathbf{I})\mathbf{U}_5$. Mivel $\mathcal{K}_4 \subset \mathcal{N}_4$, de független az $\mathcal{N}_3$ altértől, ezért alkalmas arra, hogy bázisvektorait a Jordan-bázis $\mathcal{N}_4 \setminus \mathcal{N}_3$ -ba eső elemei közé vegyük.
- Ha szükséges, egészítsük ki a $[\mathbf{N}_3 | \mathbf{K}_4]$ -et $\mathcal{N}_4$ bázisává új elemek hozzávételével, így $\mathcal{N}_4$ bázisa most $[\mathbf{N}_3 | \mathbf{K}_4 | \mathbf{U}_4]$.
- Vegyük a $[\mathbf{K}_4 | \mathbf{U}_4]$ mátrix képét, azaz legyen $\mathbf{K}_3 = (\mathbf{A} - \lambda \mathbf{I})[\mathbf{K}_4 | \mathbf{U}_4]$, és ha szükséges, egészítsük ki $[\mathbf{N}_2 | \mathbf{K}_3]$ -at $\mathcal{N}_3$ bázisává új elemek hozzávételével (azaz $\mathcal{N}_3$ bázisa most $[\mathbf{N}_2 | \mathbf{K}_3 | \mathbf{U}_3]$ – most $\mathcal{U}_3 = \emptyset$). Hasonlóan folytatjuk $\mathcal{N}_i$ indexét csökkentve: $\mathbf{K}_2 = (\mathbf{A} - \lambda \mathbf{I})[\mathbf{K}_3 | \mathbf{U}_3]$, új elemek hozzávételével előállítjuk $\mathcal{N}_2$ bázisát: $[\mathbf{N}_1 | \mathbf{K}_2 | \mathbf{U}_2]$.
- Végül legyen $\mathbf{K}_1 = (\mathbf{A} - \lambda \mathbf{I})[\mathbf{K}_2 | \mathbf{U}_2]$, amit $\mathbf{U}_1$ -gyel kibővítünk $\mathcal{N}_1$ bázisává. A tér Jordan-bázisa a kép- és az új elemek egyesítése:

$$[\mathbf{K}_1 | \mathbf{K}_2 | \mathbf{K}_3 | \mathbf{K}_4 | \mathbf{U}_1 | \mathbf{U}_2 | \mathbf{U}_3 | \mathbf{U}_4 | \mathbf{U}_5].$$

- Végül a Jordan-bázis elemeit úgy rendezzük, hogy a láncokat egymás után, minden láncot a sajátvektorból indulva felsorolunk.



10.2. ábra: A Jordan-bázist megkonstruáló algoritmus

Ezek után lássunk egy konkrét példát, majd az algoritmust általánosan. A számolás kivitelezéséhez két megjegyzés:

► Ha $\mathcal{U} \subset \mathcal{V}$ két altér, és $\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$, illetve $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ ($m < n$) a bázisuk, akkor elemi sorműveletekkel konstruálhatunk $\mathcal{V}$ -nek egy olyan $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ bázist, hogy $\{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ az $\mathcal{U}$ bázisa legyen. Ehhez írjuk $\mathcal{U}$, majd $\mathcal{V}$ bázisvektorait egyetlen

$$[\mathbf{u}_1 \ \dots \ \mathbf{u}_m \mid \mathbf{v}_1 \ \dots \ \mathbf{v}_n]$$

mátrixba. Ennek lépcsős alakjában vezéregyesek lesznek az első m oszlopban, és $n - m$ további oszlopban. A nekik megfelelő vektorok az eredeti bázisokban (tehát $\mathcal{V}$ összes vektora és $\mathcal{U}$ -nak $n - m$ vektora) adják az új bázist.

► Emlékeztetünk rá, hogy ha egy mátrix redukált lépcsős alakja $[\mathbf{I} \mid \mathbf{S}]$ alakú, akkor $\begin{bmatrix} -\mathbf{S} \\ \mathbf{I} \end{bmatrix}$ vagy $\begin{bmatrix} \mathbf{S} \\ -\mathbf{I} \end{bmatrix}$ oszlopvektorai a mátrix nulltere bázisát adják. Vigyázzunk, ha $\mathbf{I}$ nem az első oszlopokban van, akkor $\begin{bmatrix} -\mathbf{S} \\ \mathbf{I} \end{bmatrix}$ sorait megfelelően permutálni kell.

10.11. PÉLDA (JORDAN-BÁZIS ELŐÁLLÍTÁSA). Határozzuk meg az

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & -2 & 1 & -1 \\ 3 & -3 & 6 & -2 & 4 \\ 4 & -5 & 9 & -3 & 5 \\ 4 & -5 & 8 & -2 & 5 \\ -1 & 1 & -1 & 0 & 0 \end{bmatrix}$$

mátrix Jordan normálalakját és Jordan-bázisát!

MEGOLDÁS. A karakterisztikus polinom

$$\det(\mathbf{A} - \lambda \mathbf{I}) = -\lambda^5 + 4\lambda^4 - 6\lambda^3 + 4\lambda^2 - \lambda = -\lambda(1 - \lambda)^4.$$

A 0-hoz tartozó sajátvektor a redukált lépcsős alakból kiszámítva:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 \\ -3 \\ -4 \\ -4 \\ 1 \end{bmatrix}$$

Mivel itt az algebrai és geometriai multiplicitás egyaránt 1, ez a Jordan-lánc egyelemű. A $\lambda = 1$ esetén a geometriai multiplicitás 2, ugyanis $\mathbf{A} - \mathbf{I}$ redukált lépcsős alakja és abból a nulltér bázisa:

$$\text{rref}(\mathbf{A} - \mathbf{I}) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \mathbf{N}_1 = \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$(\mathbf{A} - 1 \cdot \mathbf{I})^2$ nullterét gyorsabb úgy számolni, ha nem $\mathbf{A} - \mathbf{I}$ -t szorozzuk önmagával, hanem lépcsős alakját jobbról, és annak a szorzatnak vesszük a lépcsős alakját. A lépcsős alak kiszámolása ugyanis csak elemi mátrixokkal való balról szorzást jelent, így

$$(\text{rref}(\mathbf{A} - \mathbf{I}))(\mathbf{A} - \mathbf{I}) = \mathbf{E}(\mathbf{A} - \mathbf{I})(\mathbf{A} - \mathbf{I}) = \mathbf{E}(\mathbf{A} - \mathbf{I})^2$$

vagyis a szorzat az $(\mathbf{A} - \mathbf{I})^2$ -en végrehajtott elemi sorműveletek eredménye. Sokkal kevesebb viszont a számolnivaló, mivel a 0-sorok elhagyhatók:

$$\left\{ \begin{array}{l} (\text{rref}(\mathbf{A} - \mathbf{I}))(\mathbf{A} - \mathbf{I}) \\ \text{a 0-sorok nélkül} \end{array} \right\} = \begin{bmatrix} -1 & 1 & -2 & 1 & -1 \\ 0 & 0 & -1 & 1 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -1 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 & 0 \end{bmatrix}$$

ahonnan a bázis vektorai:

$$\mathbf{N}_2 = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Hasonlóan határozzuk meg $(\mathbf{A} - \mathbf{I})^3$ bázisát!

$$\begin{bmatrix} 1 & -1 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 & 0 \end{bmatrix} (\mathbf{A} - \mathbf{I}) = \begin{bmatrix} 1 & -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

ami már lépcsős alakú, és ahonnan a bázismátrix

$$\mathbf{N}_3 = \begin{bmatrix} 1 & -1 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Ezután határozzuk meg $\mathbf{U}_3$ vektorait, vagyis azokat, amelyek $\mathcal{N}_2$ bázisát $(\mathbf{N}_2\text{-t})$ $\mathcal{N}_3$ bázisává egészítik ki. Ehhez az $[\mathbf{N}_2|\mathbf{N}_3]$ mátrixot kell redukált lépcsős alakra hozni, $\mathcal{U}_3$ elemei a $\mathbf{N}_3$ azon oszlopai lesznek, melyek függetlenek $\mathbf{N}_2$ -től, azaz melyek redukált lépcsős alakjában végegyes van.

$$[\mathbf{N}_2|\mathbf{N}_3] = \left[\begin{array}{ccc|ccc} 1 & -1 & -1 & 1 & -1 & 0 & -1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array} \right] \Rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

Tehát $\mathbf{U}_3 = [-1 \ 0 \ 1 \ 0 \ 0]^T$, és ebből $\mathbf{K}_2 = (\mathbf{A} - \mathbf{I})\mathbf{U}_3 = [-1 \ 3 \ 4 \ 4 \ 0]^T$.

Mivel $\mathbf{K}_2$ egyetlen vektorból áll, és $\mathcal{N}_3$ és $\mathcal{N}_2$ dimenzióinak különbsége

is 1, ezért itt nem kell számolnunk semmit, $\mathcal{U}_2 = \{\mathbf{0}\}$, azaz $\mathbf{U}_2$ üres. $\mathbf{K}_1 = (\mathbf{A} - \mathbf{I})\mathbf{K}_2 = [0 \ 1 \ 1 \ 1 \ 0]^T$, és mivel e vektor benn van $\mathcal{N}_1$ bázisában, $\mathbf{U}_1 = [0 \ 1 \ 0 \ 0 \ 1]^T$ teret a másik bázisvektor generálja. A Jordan normálalak felírásához a Jordan-láncok vektorait egymás után fel kell sorolni, a belőlük képzett $\mathbf{P}$ mátrixszal lesz $\mathbf{J} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. E két mátrix

$$\mathbf{P} = \left[\begin{array}{c|cccc|c} 0 & 0 & -1 & -1 & & 1 \\ 1 & 1 & 3 & 0 & & -3 \\ 0 & 1 & 4 & 1 & & -4 \\ 0 & 1 & 4 & 0 & & -4 \\ 1 & 0 & 0 & 0 & & 1 \end{array} \right] \quad \mathbf{J} = \left[\begin{array}{c|cccc|c} 1 & 0 & 0 & 0 & & 0 \\ 0 & 1 & 1 & 0 & & 0 \\ 0 & 0 & 1 & 1 & & 0 \\ 0 & 0 & 0 & 1 & & 0 \\ 0 & 0 & 0 & 0 & & 0 \end{array} \right] \quad \square$$

Az algoritmus általánosan:

- Input: $\mathbf{A}$, $p(x)$ karakterisztikus polinom lineáris tényezőkre bontva,
- Minden λ sajátértékre
 - Határozzuk meg a leghosszabb lánc s hosszát, és $(\mathbf{A} - \lambda\mathbf{I})^i$ nullterét $\mathcal{N}_i$, $i = 1, 2, \dots, s$. Legyen $\mathcal{U}_{s+1} = \mathcal{K}_{s+1} = \mathcal{N}_0 = \{\mathbf{0}\}$, azaz e terek bázisa az üreshalmaz.
 - Minden i -re s -től 1-ig haladva:
 - * Legyen $\mathbf{K}_i = (\mathbf{A} - \lambda\mathbf{I})[\mathbf{K}_{i+1}|\mathbf{U}_{i+1}]$.
 - * Határozzuk meg $\mathbf{U}_i$ -t úgy, hogy $[\mathbf{N}_{i-1}|\mathbf{K}_i|\mathbf{U}_i]$ bázisa legyen $\mathcal{N}_i$ -nek. Ehhez az $[\mathbf{N}_{i-1}|\mathbf{K}_i|\mathbf{N}_i]$ mátrix redukált lépcsős alakja alapján válasszuk a vezéroszlopokat $\mathbf{N}_i$ -ből az $\mathbf{U}_i$ -be.
- Tegyük a Jordan-láncok vektorait balról jobbra egymás mellé minden láncot a sajátvektorral kezdve, az így kapott $\mathbf{P}$ mátrixszal $\mathbf{J} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$.

10.12. DEFINÍCIÓ (MÁTRIX EXPONENCIÁLIS FÜGGVÉNYE). Az $\mathbf{A}$ négyzetes mátrix exponenciális függvényét a következő sor definiálja:

$$e^{\mathbf{A}} = \mathbf{I} + \mathbf{A} + \frac{\mathbf{A}^2}{2!} + \cdots + \frac{\mathbf{A}^n}{n!} + \cdots$$

10.13. ÁLLÍTÁS (EXPONENCIÁLIS FÜGGVÉNY KISZÁMÍTÁSA). Ha $\mathbf{J} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ az $\mathbf{A}$ mátrix Jordan normálalakja, akkor

$$e^{\mathbf{A}} = \mathbf{P}e^{\mathbf{J}}\mathbf{P}^{-1},$$

ahol $e^{\mathbf{J}}$ blokkonként számolható, és

$$\mathbf{J}_\lambda = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda \end{bmatrix} \text{ esetén } e^{\mathbf{J}_\lambda} = e^\lambda \begin{bmatrix} 1 & 1 & \frac{1}{2!} & \cdots & \frac{1}{(n-1)!} \\ 0 & 1 & 1 & \cdots & \frac{1}{(n-2)!} \\ 0 & 0 & 1 & \cdots & \frac{1}{(n-3)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

Mátrixfüggvények Általában mondható, hogy

$$f\left(\begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \lambda \end{bmatrix}\right) = \begin{bmatrix} \frac{f(\lambda)}{0!} & \frac{f'(\lambda)}{1!} & \frac{f''(\lambda)}{2!} & \cdots & \frac{f^{(n-1)}(\lambda)}{(n-1)!} \\ 0 & \frac{f(\lambda)}{0!} & \frac{f'(\lambda)}{1!} & \cdots & \frac{f^{(n-2)}(\lambda)}{(n-2)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{f(\lambda)}{0!} \end{bmatrix}.$$

10.14. PÉLDA (MÁTRIX EXPONENCIÁLIS FÜGGVÉNYE). Legyen

$$\begin{pmatrix} -3 & 2 & 1 \\ 1 & -2 & 1 \\ -1 & -2 & -5 \end{pmatrix}$$

Határozzuk meg az $e^{\mathbf{A}}$ mátrixot!

MEGOLDÁS. $\mathbf{A}$ karakterisztikus polinomja

$$x^3 + 10x^2 + 32x + 32 = (x + 2)(x + 4)^2,$$

így

$$\mathbf{J} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -4 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & -1 & -2 \end{bmatrix},$$

$$e^{\mathbf{A}} = \mathbf{P}e^{\mathbf{J}}\mathbf{P}^{-1} = \begin{bmatrix} \frac{1}{2} \frac{e^2+1}{e^4} & \frac{e^2-1}{e^4} & \frac{1}{2} \frac{e^2-1}{e^4} \\ \frac{1}{2} \frac{e^2-1}{e^4} & e^2 & \frac{1}{2} \frac{e^2-1}{e^4} \\ -\frac{1}{2} \frac{e^2-1}{e^4} & -\frac{e^2-1}{e^4} & -\frac{1}{2} \frac{e^2-3}{e^4} \end{bmatrix} \quad \square$$

A Jordan normálalak használata a differenciálegyenletrendszerek megoldásában

Irodalomjegyzék

Wolf Holzmann. Uniqueness of reduced row echelon form. <http://www.cs.uleth.ca/~holzmann/notes/reduceduniq.pdf>, 2002.

Tárgymutató

p -norma, 417
általánosított sajátvektor, 430

algebrai multiplicitása, 378
altér
 invariáns, 391

diadikus felbontás
 szinguláris érték szerinti, 406
diagonalizálhatóság, 386

euklideszi norma, 416

geometriai multiplicitás, 378

Hölder-egyenlőtlenség, 419, 426

illeszkedő normák, 422
indukált mátrixnorma, 422
invariáns altér, 391

Jordan normálalak, 433

Jordan-blokk, 431
Jordan-lánc, 430

karakterisztikus egyenlet, 372
karakterisztikus polinom, 372
konzisztens normák, 422
kvadratikus forma, 398

lineáris transzformáció
 karakterisztikus polinomja, 386
 sajátértéke, 384
 sajátértékei, 386

másodfokú tag, 397
mátrixnorma, 422
Minkowski-egyenlőtlenség, 419, 426

negatív (szemi)definit, 400
normálás, 416
norma, 418

operátornorma, 423

ortogonális diagonalizálás, 395

polárfelbontás, 411
polinom
 homogén másodfokú, 397
pozitív (szemi)definit, 400

redukált szinguláris felbontás, 406

sajátérték, 370
 lineáris transzformációé, 384
sajátaltér, 370
sajátpár, 370
sajátvektor, 370
szinguláris felbontás, 406
szinguláris érték, 404
szinguláris felbontás
 diadikus alak, 406
szinguláris vektor, 404

unitér diagonalizálás, 397