Bevezetés az algebrába 1

Egyenletrendszerek numerikus módszerei

Wettl Ferenc diáinak felhasználásával

LU-felbontás

Motiváció: Δ-mátrixokkal könnyebb

egyenletrendszert megoldani (behelyettesítésekkel),

determinánst számolni,

inverzet számolni (szimultán egyenletrendszerként).

Látni fogjuk, hogy sok mátrix felírható (alsó Δ) $\cdot$ (felső Δ) alakban.

A Δ-mátrix fogalmát kiterjeszthetjük nem négyzetes mátrixra is.

D Egy $A \in K^{m \times n}$ mátrix felső Δ-mátrix, ha $a_{ij} = 0 \ \forall j < i$,

alsó Δ-mátrix, ha $a_{ij} = 0 \ \forall j > i$,

normált Δ-mátrix, ha emellett a főátlóban csupa 1 áll, azaz $a_{ii} = 1 \ \forall i$.

P Minden lépcsős mátrix felső Δ-mátrix, de fordítva nem következik.

$$\begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix} \qquad \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 3 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix} \qquad \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

mind felső Δ-mátrixok, a harmadik normált is, viszont a második nem lépcsős.

D **LU-felbontás**

Egy $A \in K^{m\times n}$ mátrix LU-felbontása:
$A = LU$, ahol L normált alsó Δ-mátrix, U pedig felső Δ-mátrix.
(L lower, U upper triangular)

P Nem minden mátrixnak van LU-felbontása:
$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$-nak nincs: $\begin{bmatrix} x & 0 \\ y & z \end{bmatrix} \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} = \begin{bmatrix} xa & xb \\ ya & yb + zc \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
esetén $xa = 0 \Rightarrow xb = 0$ vagy $ya = 0$. $\lightning$

D Egy $A \in K^{m\times n}$-es mátrix k. főminorja a bal fölső $k \times k$-as sarokaldeterminánsa (ahol $k \leqslant \min\{m, n\}$), azaz

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & & \dots & a_{2k} \\ \vdots & & \dots & \vdots \\ a_{k1} & & \dots & a_{kk} \end{vmatrix}.$$

Pl. $\begin{bmatrix} 2 & -1 & 3 & 0 \\ 1 & 1 & 2 & -1 \\ 0 & -1 & 4 & 7 \end{bmatrix}$ főminorjai 2, 3 és 13.

Az LU-felbontás kiszámítása

Gauss-eliminációval felső Δ-alakra hozzuk a mátrixot úgy, hogy csak $s_i \mapsto s_i - cs_j$ alakú műveleteket végzünk, ahol $i > j$ (lefelé nullázás).

Ezekhez a sorműveletekhez tartozó elemi mátrixok mind normált alsó háromszögmátrixok, így a szorzatuk is ilyen.

$$(E_t \cdots E_1)A = U \;\Rightarrow\; A = (E_t \cdots E_1)^{-1}U = LU$$

Ha a szokásos módon, balról jobbra haladva nullázzuk az oszlopok alját, akkor az $L = E_1^{-1} \cdots E_t^{-1} I$ mátrix elemeit egyszerűen fel tudjuk írni: ha szerepelt egy $s_i \mapsto s_i - cs_j$ sorművelet a Gauss-eliminációban, akkor $(L)_{ij} = c$, ha nem volt (i, j)-hez tartozó sorművelet, akkor pedig 0.

Ugyanis az $I \to \to \;\cdots\; \to E_1^{-1} \cdots E_t^{-1} I = L$ elemi sorműveleteknél, amikor az $s_i \mapsto s_i + cs_j$ inverz sorműveletet végrehajtjuk, addig csak alatta levő sorokat módosítottunk, tehát az i. sorban csak az i. elem 1, a többi 0, és így csak az i. oszlopban történik változás.

P Határozzuk meg az alábbi mátrix LU-felbontását!

$$A = \begin{bmatrix} 4 & 8 & 4 & 8 \\ 2 & 6 & 4 & 4 \\ 1 & 3 & 2 & 4 \end{bmatrix}$$

m

$$A = \begin{bmatrix} 4 & 8 & 4 & 8 \\ 2 & 6 & 4 & 4 \\ 1 & 3 & 2 & 4 \end{bmatrix} \to \begin{bmatrix} 4 & 8 & 4 & 8 \\ 0 & 2 & 2 & 0 \\ 0 & 1 & 1 & 2 \end{bmatrix} \to \begin{bmatrix} 4 & 8 & 4 & 8 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} = U$$

és $L = \begin{bmatrix} 1 & 0 & 1 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{4} & \frac{1}{2} & 1 \end{bmatrix}$, ugyanis $\begin{array}{l} s_2 \mapsto s_2 - \frac{1}{2}s_1 \\ s_3 \mapsto s_3 - \frac{1}{4}s_1 \\ s_3 \mapsto s_3 - \frac{1}{2}s_2 \end{array}$

az elvégzett sorműveletek.

Néhány tudnivaló az LU-algoritmusról:

1. Csak $s_i \mapsto s_i - cs_j$ $(i > j)$ típusú elemi sorműveleteket végezzünk a j növekvő sorrendjében!
2. Nem kell feltétlenül lépcsős alakra törekedni, csak felső Δ-alakra.
3. Ha a soron következő nullázó elem 0, és alatta van nemnulla elem, akkor elakad az algoritmus.

T (1) Ha az A mátrix minden bal fölső sarokaldeterminánsa, azaz főminorja nemnulla, akkor A-nak van LU-felbontása.

(2) Ha A invertálható, és van LU-felbontása, akkor az egyértelmű.

B Vegyük észre, hogy ha csak lefelé változtató sorműveleteket végzünk, akkor a főminorok nem változnak:

$s_i \mapsto s_i - cs_j$ $(i > j)$ sorműveletnél

$k < i$-re a $k \times k$-as bal fölső sarokaldetermináns változatlan marad;

$k \geqslant i$-re a $k \times k$-as főminoron megengedett (determinánst nem változtató) sorműveletet hajtunk végre, mert $k \geqslant i, j$.

(1): Belátjuk, hogy a Gauss-eliminació során az soron következő i. oszlop főátlóbeli eleme mindig nemnulla lesz, így lehet vele lefelé nullázni. Tegyük fel ugyanis, hogy az i. oszlop az első, ahol elakadunk.

$$\left[\begin{array}{ccc|c|c} * & & & & \\ 0 & \ddots & * & * & * \\ & & * & & \\ \hline & & & 0 & * \\ & 0 & & * & * \\ & & & * & * \end{array}\right]$$

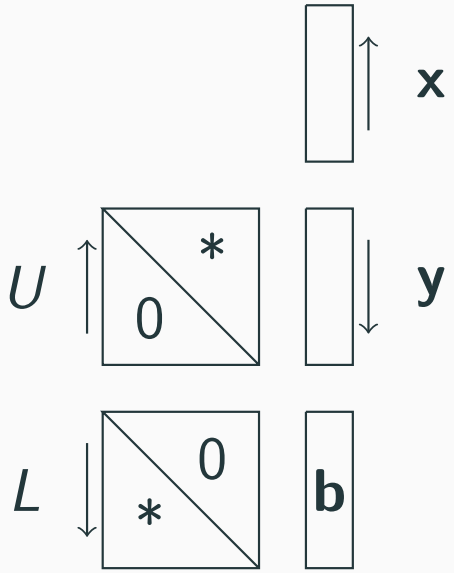
Ekkor az $(i+1) \times (i+1)$-es főminornak van nulla sora, tehát az értéke nulla, holott az eddigi sorműveletek nem változtatták a főminorokat. ↯

(2): Tfh $A = L_1 U_1 = L_2 U_2$. Ekkor $|A| \neq 0 \Rightarrow |L_i|, |U_i| \neq 0$.
Így $L_2^{-1} L_1 = U_2 U_1^{-1}$ alsó és felső Δ, és normált, tehát I.

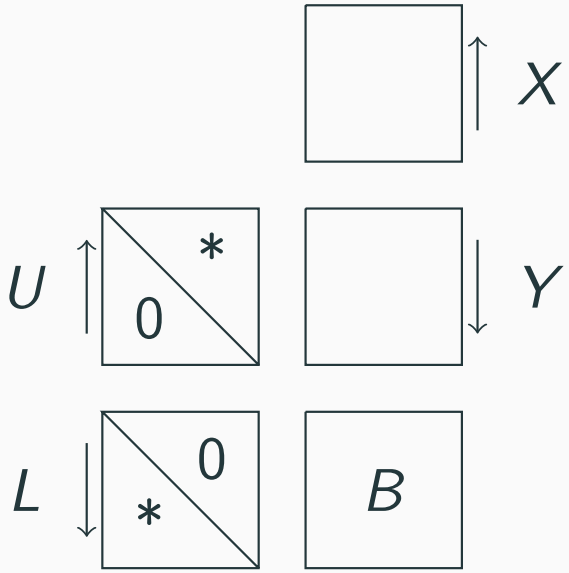
Egyenletrendszer megoldása és inverz kiszámolása LU-felbontással

$LU\mathbf{x} = \mathbf{b} \Leftrightarrow \exists \mathbf{y}\colon\ L\mathbf{y} = \mathbf{b}$ és $U\mathbf{x} = \mathbf{y}$

Invertálható $A = LU$ mátrixra használjuk a módszert:



$A\mathbf{x} = \mathbf{b}$ er.



$AX = B$ mátrixegyenlet

Inverzszámítás: $AX = I$ megoldása

P Határozzuk meg az A mátrix LU-felbontását, és ezzel oldjuk meg az $A\mathbf{x} = \mathbf{b}$ egyenletrendszert, ahol

$$A = \begin{bmatrix} 4 & 8 & 8 \\ 2 & 6 & 4 \\ 1 & 3 & 4 \end{bmatrix} \qquad \text{és} \qquad \mathbf{b} = \begin{bmatrix} 8 \\ 0 \\ 4 \end{bmatrix}$$

m

$$\begin{bmatrix} 4 & 8 & 8 \\ 2 & 6 & 4 \\ 1 & 3 & 4 \end{bmatrix} \to \begin{bmatrix} 4 & 8 & 8 \\ 0 & 2 & 0 \\ 0 & 1 & 2 \end{bmatrix} \to \begin{bmatrix} 4 & 8 & 8 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = U$$

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/4 & 1/2 & 1 \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} 8 \\ -4 \\ 4 \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix}$$

Mj Az LU-felbontás műveletigénye a Gauss-eliminációéval azonos, de kevesebb memóriára van szükség hozzá (az A eredeti helyén tárolhatjuk az L és U lényeges elemeit: az L átló alatti elemeit beírhatjuk az U éppen kinullázott elemeinek a helyére).

Speciális, pl. ritka mátrixokra van gyorsabb algoritmus is az LU-felbontásra.

Mj Bár nem minden mátrixnak van LU-felbontása, mindegyiknek van PLU-felbontása, ahol P permutációmátrix:

a szükséges sorcseréket előre elvégezhetjük az A-n, és ezután végigfut az LU-algoritmus.

$$(E_t \cdots E_1)PA = U \;\Rightarrow\; A = P^T(E_t \cdots E_1)^{-1}U$$

(Menetközben is végrehajthatjuk a sorcseréket, és alkalmasan módosíthatjuk az L-et, nem kell kétszer végigmenni.)

F Oldjuk meg az $A\mathbf{x} = \mathbf{b}$ egyenletrendszert LU-felbontás segítségével, ahol

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{bmatrix} \qquad \text{és} \qquad \mathbf{b} = \begin{bmatrix} 5 \\ -1 \\ 3 \\ 7 \end{bmatrix}$$

P Számítsuk ki a PLU-felbontást:

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

Ellentmondásos egyenletrendszer közelítő megoldása

D **Merőleges vetület**

Legyen $\mathcal{W} \leqslant \mathbb{R}^n$ és $\mathbf{v} \in \mathbb{R}^n$.

$\mathbf{v} \in \mathbb{R}^n$ merőleges vetülete $\mathcal{W}$-re az a $\mathbf{v}'$ vektor, amelyre

$$\mathbf{v}' \in \mathcal{W} \text{ és } \mathbf{v} - \mathbf{v}' \in \mathcal{W}^\perp.$$

(Tudjuk: $\mathbb{R}^n$ minden vektora egyértelműen felírható egy $\mathcal{W}$-beli és egy $\mathcal{W}^\perp$-beli vektor összegeként, tehát $\exists!\, \mathbf{v}'$.)

Á Legyen $\mathcal{W} \leqslant \mathbb{R}^n$, $\{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ bázis $\mathcal{W}$-ben, és $B = [\mathbf{b}_1 \mid \ldots \mid \mathbf{b}_k]$ ($\mathbf{b}_i$-k a bázisvektorok oszlopokként írva).

Ekkor a $\mathcal{W}$-re való merőleges vetítés $\mathbf{v} \mapsto B(B^T B)^{-1} B^T \mathbf{v}$.

B $B^T B$ invertálható:

Elég belátni, hogy a $B^T B\mathbf{x} = \mathbf{0}$ er.-nek csak triviális megoldása van.

Ez igaz, mert ha $B^T B\mathbf{x} = \mathbf{0}$, akkor

$0 = \mathbf{x}^T\mathbf{0} = \mathbf{x}^T B^T B\mathbf{x} = (B\mathbf{x})^T(B\mathbf{x}) = |B\mathbf{x}|^2 \Rightarrow B\mathbf{x} = \mathbf{0}$,

de B oszlopai ftlnek $\Rightarrow \mathbf{x} = \mathbf{0}$.

$B(B^T B)^{-1}B^T\mathbf{v} \in \mathcal{O}(B) = \mathcal{W}$,

mert a B-t szorozzuk jobbról egy vektorral.

$\mathbf{v} - B(B^T B)^{-1}B^T\mathbf{v} \in \mathcal{W}^\perp$,

mert $\mathcal{W}^\perp = \mathcal{O}(B)^\perp = \mathcal{S}(B^T)^\perp = \mathcal{N}(B^T)$, és

$B^T(\mathbf{v} - B(B^T B)^{-1}B^T\mathbf{v}) = B^T\mathbf{v} - B^T B(B^T B)^{-1}B^T\mathbf{v} =$
$B^T\mathbf{v} - IB^T\mathbf{v} = B^T\mathbf{v} - B^T\mathbf{v} = \mathbf{0}$.

P Adjuk meg a $\mathbf{v} = (2, 1, 0, 3)$ vektor $\mathcal{W} = \operatorname{span}((1, 2, -1, 0),\ (1, 0, 1, 1))$ altérre való merőleges vetületét $\mathbb{R}^4$-ben!

m $$B^T B = \begin{bmatrix} 1 & 2 & -1 & 0 \\ 1 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ -1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 0 \\ 0 & 3 \end{bmatrix} \Rightarrow (B^T B)^{-1} = \frac{1}{6} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

$$\frac{1}{6} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & -1 & 0 \\ 1 & 0 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 2 \\ 1 \\ 0 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 \\ 2 & 0 \\ -1 & 1 \\ 0 & 1 \end{bmatrix} \quad \frac{1}{6} \cdot \begin{bmatrix} 1 & 2 \\ 2 & 0 \\ -1 & 2 \\ 0 & 2 \end{bmatrix} \frac{1}{6} \cdot \begin{bmatrix} 3 & 2 & 1 & 2 \\ 2 & 4 & -2 & 0 \\ 1 & -2 & 3 & 2 \\ 2 & 0 & 2 & 2 \end{bmatrix} \frac{1}{6} \cdot \begin{bmatrix} 14 \\ 8 \\ 6 \\ 10 \end{bmatrix} = \begin{bmatrix} 7/3 \\ 4/3 \\ 1 \\ 5/3 \end{bmatrix} = \mathbf{v}'$$

T **A merőleges vetítés jellemzései**

(1) Egy $f : \mathbb{R}^n \to \mathbb{R}^n$ lineáris transzformáció a $\mathcal{W} \leqslant \mathbb{R}^n$ altérre való merőleges vetítés $\Leftrightarrow \begin{cases} f(\mathbf{w}) = \mathbf{w} \ \forall \mathbf{w} \in \mathcal{W} \\ f(\mathbf{u}) = \mathbf{0} \ \forall \mathbf{u} \in \mathcal{W}^\perp \end{cases}$

(2) $A \in \mathbb{R}^{n \times n}$ egy merőleges vetítés mátrixa $\Leftrightarrow \begin{cases} A^T = A \text{ és} \\ A^2 = A \end{cases}$

(és ekkor A az $\mathcal{O}(A)$-ra vetít).

B (1): $\mathbb{R}^n = \mathcal{W} \oplus \mathcal{W}^\perp$

$\Rightarrow$: $\mathbf{w} \in \mathcal{W}$-re $\mathbf{w} = \mathbf{w} + \mathbf{0} \in \mathcal{W} + \mathcal{W}^\perp$ felbontásból $f(\mathbf{w}) = \mathbf{w}$.

$\mathbf{u} \in \mathcal{W}^\perp$-re $\mathbf{u} = \mathbf{0} + \mathbf{u} \in \mathcal{W} + \mathcal{W}^\perp$ felbontásból $f(\mathbf{u}) = \mathbf{0}$.

$\Leftarrow$: $\mathbf{v} = \mathbf{w} + \mathbf{u} \in \mathcal{W} + \mathcal{W}^\perp$ felbontásra

$f(\mathbf{v}) = f(\mathbf{w}) + f(\mathbf{u}) = \mathbf{w} + \mathbf{0} = \mathbf{w}$.

(2): $\Rightarrow$: Ellenőrizzük a feltételeket az $A = B(B^TB)^{-1}B^T$ mátrixra, ahol B oszlopai a $\mathcal{W}$ bázisát alkotják.

$$A^T = B\left((B^TB)^{-1}\right)^T B^T = B(B^TB)^{-1}B^T = A$$

$$A^2 = \left(B(B^TB)^{-1}B^T\right)\left(B(B^TB)^{-1}B^T\right) = B(B^TB)^{-1}B^T = A$$

$\Leftarrow$: $\mathbf{w} \in \mathcal{W} := \mathcal{O}(A)$-ra $\mathbf{w} = A\mathbf{x}$ vmely $\mathbf{x}$-re $\Rightarrow A\mathbf{w} = AA\mathbf{x} = A\mathbf{x} = \mathbf{w}$, és $\mathbf{u} \in \mathcal{W}^\perp = \mathcal{N}(A^T)$-ra $A\mathbf{u} = A^T\mathbf{u} = \mathbf{0}$,
így az (1) rész miatt A az $\mathcal{O}(A)$-ra való vetítés mátrixa.

P A korábbi példában kiszámoltuk, hogy a $\mathcal{W} = \text{span}((1,2,-1,0),(1,0,1,1))$ altérre való merőleges vetítés mátrixa

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

Látható, hogy szimmetrikus, azaz $A^T = A$, és ellenőrizhető, hogy $A^2 = A$.

Ebből levezethetők:

1. Vektorra való vetítés mátrixa:

$\mathbf{0} \neq \mathbf{b} \in \mathbb{R}^n$, $\mathcal{W} = \text{span}(\mathbf{b})$, $B = \mathbf{b}$. f a $\mathbf{b}$-re való vetítés.

$[f]_{\mathcal{E}} = \mathbf{b}(\mathbf{b}^T\mathbf{b})^{-1}\mathbf{b}^T = \mathbf{b}\frac{1}{|\mathbf{b}|^2}\mathbf{b}^T = \frac{1}{|\mathbf{b}|^2}\mathbf{b}\mathbf{b}^T$, ahogy ezt korábban láttuk.

2. A $\mathcal{W}$-re való merőleges vetítés és a $\mathcal{W}^\perp$-ra való merőleges vetítés összege az identikus leképezés, ui. $\mathbf{v} = \mathbf{w} + \mathbf{u}$ ($\mathbf{w} \in \mathcal{W}$, $\mathbf{u} \in \mathcal{W}^\perp$) felbontás esetén $\mathbf{v}$ vetülete az elsőre $\mathbf{w}$, a másodikra $\mathbf{u}$. Tehát ha $\mathcal{W}^\perp$ a kisebb dimenziós, érdemes a $\mathcal{W}^\perp$-ra vetítés B mátrixából kifejezni a $\mathcal{W}$-re való vetítését mint $I - B$-t.

3. Hipersíkra való vetítés: Ha $\mathcal{W}$ a $c_1x_1 + \ldots + c_nx_n = 0$ hipersík, akkor $\mathcal{W}^\perp = \text{span}(\mathbf{c})$, tehát a 2. pont szerint a $\mathcal{W}$-re való merőleges vetítés mátrixa $I - \frac{1}{|\mathbf{c}|^2}\mathbf{c}\mathbf{c}^T$.

4. Hipersíkra való tükrözés: Ha $\mathbf{v}''$ az $\mathbf{v}$ merőleges vetülete a $\mathbf{c}$ normálvektorra, akkor a tükörképe $\mathbf{v} - 2\mathbf{v}''$, és a tükrözés mátrixa $I - \frac{2}{|\mathbf{c}|^2}\mathbf{c}\mathbf{c}^T$.

P A korábbi példabeli $\mathcal{W} = \operatorname{span}((1,2,-1,0),(1,0,1,1))$ altérre való merőleges vetítés mátrixából számítsuk ki a $\mathcal{W}^\perp$ altérre való merőleges vetítés mátrixát. Adjuk meg $\mathcal{W}^\perp$ egy bázisát.

m A $\mathcal{W}$-re való merőleges vetítés A mátrixa (amit korábban kiszámoltunk), és ebből a $\mathcal{W}^\perp$-ra való merőleges vetítés B mátrixa

$$A = \frac{1}{6}\begin{bmatrix} 3 & 2 & 1 & 2 \\ 2 & 4 & -2 & 0 \\ 1 & -2 & 3 & 2 \\ 2 & 0 & 2 & 2 \end{bmatrix} \quad \text{és} \quad B = I - A = \frac{1}{6}\begin{bmatrix} 3 & -2 & -1 & -2 \\ -2 & 2 & 2 & 0 \\ -1 & 2 & 3 & -2 \\ -2 & 0 & -2 & 4 \end{bmatrix}$$

$\mathcal{W}^\perp = \mathcal{O}(B)$, és tudjuk, hogy kétdimenziós, ezért B bármely két ftln oszlopa kifeszíti, pl. $\{(3,-2,-1,-2),(-2,2,2,0)\}$ bázisa $\mathcal{W}^\perp$-nek.

P A $2x + y - 3z = 0$ síkra való tükrözés mátrixa

$$I - \tfrac{1}{7}\begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}\begin{bmatrix} 2 & 1 & -3 \end{bmatrix} = \tfrac{1}{7}\begin{bmatrix} 3 & -2 & 6 \\ -2 & 6 & 3 \\ 6 & 3 & -2 \end{bmatrix}$$

Vegyük észre, hogy az oszlopok merőleges, 1 hosszúságú vektorok, mert $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ képei egy egybevágóságnál.

Á Egy $\mathbf{v} \in \mathbb{R}^n$ vektorhoz egyetlen legközelebbi vektor a $\mathcal{W}$ altérben (azaz $\mathbf{w} \in \mathcal{W}$, amelyre $|\mathbf{v} - \mathbf{w}|$ minimális) a $\mathbf{v}$ merőleges vetülete $\mathcal{W}$-re.

B Legyen $\mathbf{v} = \mathbf{v}' + \mathbf{v}''$, ahol $\mathbf{v}' \in \mathcal{W}$ és $\mathbf{v}'' \in \mathcal{W}^\perp$

(azaz $\mathbf{v}'$ a $\mathbf{v}$ merőleges vetülete $\mathcal{W}$-re).

Ekkor a Pitagorasz-tétel miatt tetszőleges $\mathbf{w} \in \mathcal{W}$-re

$$|\mathbf{v} - \mathbf{w}|^2 = |(\mathbf{v}' - \mathbf{w}) + \mathbf{v}''|^2 = |\mathbf{v}' - \mathbf{w}|^2 + |\mathbf{v}''|^2 \geqslant |\mathbf{v}''|^2 = |\mathbf{v} - \mathbf{v}'|^2,$$

és $=$ csak akkor van, ha $|\mathbf{v}' - \mathbf{w}| = 0$, azaz $\mathbf{w} = \mathbf{v}'$.

Ellentmondásos egyenletrendszer optimális közelítő megoldása

Tfh. az $A\mathbf{x} = \mathbf{b}$ er. konstans tagjait mérések eredményeként kaptuk. A mérések nem feltétlenül pontosak, ezért lehet, hogy az er.-nek nincs megoldása. Keressük azt a $\mathbf{b}$-hez legközelebbi $\mathbf{b}'$-t, amelyre már van megoldás, és a hozzá tartozó $\mathbf{x}$ vektorokat.

$$A\mathbf{x} = \mathbf{b} \text{ megoldható} \Leftrightarrow \mathbf{b} \in \mathcal{O}(A),$$

tehát a keresett $\mathbf{b}'$ a $\mathbf{b}$-hez legközelebbi vektor $\mathcal{O}(A)$-ban, azaz $\mathbf{b}'$ merőleges vetülete $\mathcal{O}(A)$-ra.

A vetületképlet alkalmazásához ki kell választani egy bázist $\mathcal{O}(A)$-ban,. . .

Egyszerűbben is lehet:

Á $A\mathbf{x} = \mathbf{b}'$ megoldásai az $A^T A\mathbf{x} = A^T\mathbf{b}$ megoldásai,

(ahol $\mathbf{b}'$ a $\mathbf{b}$ merőleges vetülete $\mathcal{O}(A)$-ra).

B $\Rightarrow$: $A\mathbf{x} = \mathbf{b}' \Rightarrow A^T A\mathbf{x} = A^T\mathbf{b}' = A^T\mathbf{b} + A^T(\mathbf{b}' - \mathbf{b})$.

Mivel $\mathbf{b}'$ a $\mathbf{b}$ merőleges vetülete $\mathcal{O}(A)$-ra, $(\mathbf{b}' - \mathbf{b}) \perp \mathcal{O}(A) = \mathcal{S}(A^T)$

$\Rightarrow \mathbf{b}' - \mathbf{b} \in \mathcal{N}(A^T) \Rightarrow A^T(\mathbf{b}' - \mathbf{b}) = \mathbf{0} \Rightarrow A^T A\mathbf{x} = A^T\mathbf{b}$.

$\Leftarrow$: $A^T A\mathbf{x} = A^T\mathbf{b} \Rightarrow A^T(A\mathbf{x} - \mathbf{b}) = \mathbf{0} \Rightarrow (A\mathbf{x} - \mathbf{b}) \in \mathcal{N}(A^T)$

$\Rightarrow (A\mathbf{x} - \mathbf{b}) \perp \mathcal{S}(A^T) = \mathcal{O}(A)$, és $A\mathbf{x} \in \mathcal{O}(A)$,

így $\mathbf{b}$-nek az $\mathcal{O}(A)$-ra vett merőleges vetülete $A\mathbf{x}$, azaz $A\mathbf{x} = \mathbf{b}'$.

D **Normálegyenlet**

$A^T A\mathbf{x} = A^T\mathbf{b}$ az $A\mathbf{x} = \mathbf{b}$ er.-hez tartozó normálegyenlet.

Ennek a kibővített mátrixa $A^T[A \mid \mathbf{b}]$.

K1 Megkaphatjuk a **minimális abszolút értékű** optimális közelítő megoldást a normálegyenlet **sortérbe eső** megoldásaként.

K2 Egy $\mathbf{b}$ vektor $\mathcal{W}$ altérre vett merőleges vetületét kiszámíthatjuk az optimális közelítő megoldás segítségével is, és ekkor $\mathcal{W}$-nek elég csak egy generátorrendszerét ismerni, nem kell feltétlenül bázist választani belőle (bár lehet, hogy érdemes):

Ha $\mathcal{W} = \operatorname{span}(\mathbf{a}_1, \ldots, \mathbf{a}_k)$, $A = [\mathbf{a}_1 \mid \ldots \mid \mathbf{a}_k]$, és az $A\mathbf{x} = \mathbf{b}$ egyenletrendszer egyik optimális közelítő megoldása $\mathbf{x}$, akkor $\mathbf{b}' = A\mathbf{x}$.

Mj Ha az $A\mathbf{x} = \mathbf{b}$ egyenletrendszer konzisztens, akkor az optimális közelítő megoldások nyilván az igazi megoldások ($\mathbf{b}' = \mathbf{b}$), tehát a normálegyenlet megoldáshalmaza megegyezik az eredetiével.

P Keressük meg az alábbi egyenletrendszer optimális közelítő megoldásait, és ezek közül a legkisebb abszolút értékűt!

$$\begin{array}{rcrcrcr} & & y & + & z & = & 3 \\ x & + & y & + & 2z & = & 2 \\ x & & & + & z & = & 2 \end{array}$$

m

$$[A \mid \mathbf{b}] = \left[\begin{array}{ccc|c} 0 & 1 & 1 & 3 \\ 1 & 1 & 2 & 2 \\ 1 & 0 & 1 & 2 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 3 \\ 0 & 1 & 1 & 0 \end{array}\right] \text{ ellentmondásos.}$$

$$\underbrace{\begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix}}_{A^T} \underbrace{\left[\begin{array}{ccc|c} 2 & 1 & 3 & 4 \\ 1 & 2 & 3 & 5 \\ 3 & 3 & 6 & 9 \end{array}\right]}_{[A^TA \mid A^T\mathbf{b}]} \to \left[\begin{array}{ccc|c} 1 & 2 & 3 & 5 \\ 0 & -3 & -3 & -6 \\ 0 & -3 & -3 & -6 \end{array}\right] \to$$

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right] \Rightarrow \mathbf{x} = \begin{bmatrix} 1-t \\ 2-t \\ t \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \left(\Rightarrow A\mathbf{x} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} = \mathbf{b}' \right)$$

(folyt.) Min. abszolút értékű közelítő megoldás:

$A^T A$ sortere $\operatorname{span}((-1,-1,1)^\perp$.

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ -1 & -1 & 1 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 3 & 3 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right] \Rightarrow \mathbf{x} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

F Számítsuk ki a $\mathbf{v} = (2,1,0,3)$ vektor merőleges vetületét a $W = \operatorname{span}((1,2,-1,0),(1,0,1,1))$-re ezzel a módszerrel!

m

$$[A \mid \mathbf{b}] = \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 2 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & 3 \end{array}\right]$$

$$\underbrace{\begin{bmatrix} 1 & 2 & -1 & 0 \\ 1 & 0 & 1 & 1 \end{bmatrix}}_{A^T} \underbrace{\left[\begin{array}{cc|c} 6 & 0 & 4 \\ 0 & 3 & 5 \end{array}\right]}_{[A^T\mathbf{x} \mid A^T\mathbf{b}]} \to \left[\begin{array}{cc|c} 1 & 0 & 2/3 \\ 0 & 1 & 5/3 \end{array}\right], \quad \mathbf{b}' = A\mathbf{x} = \begin{bmatrix} 7/3 \\ 4/3 \\ 1 \\ 5/3 \end{bmatrix}$$

Moore–Penrose-féle pszeudoinverz

Invertálható A mátrixra $A\mathbf{x} = \mathbf{b}$ megoldása $\mathbf{x} = A^{-1}\mathbf{b}$.

Ha A nem invertálható, szeretnénk egy olyan A^+ „általánosított inverzet”, amelyre $A^+\mathbf{b}$ megad egy megoldást, vagy ha az er. inkonzisztens, egy optimális közelítő megoldást.

Ha A^+ ilyen, akkor $A(A^+\mathbf{b}) = \mathbf{b}'$ a $\mathbf{b}$ merőleges vetülete $\mathcal{O}(A)$-ra. Tehát AA^+ az $\mathcal{O}(A)$-ra való merőleges vetítés mátrixa,

így AA^+ szimmetrikus, továbbá minden $A\mathbf{x} \in \mathcal{O}(A)$ vektorra $AA^+A\mathbf{x} = A\mathbf{x}$, vagyis $AA^+A = A$.

Ha ez az A^+ ált. inverz olyan értelemben is hasonlít az inverzre, hogy $(A^+)^+ = A$, akkor az előző két tulajdonság fordított szereposztásban is teljesül.

D **Pszeudoinverz**

$A \in \mathbb{R}^{m\times n}$-re $A^+ \in \mathbb{R}^{n\times m}$ az A pszeudoinverze, ha

P1. $AA^+A = A$,

P2. $A^+AA^+ = A^+$,

P3. AA^+ szimmetrikus,

P4. A^+A szimmetrikus.

K A feltételekből következik, hogy ha A^+ az A pszeudoinverze, akkor

$$A \text{ pszeudoinverze } A^+\text{-nak},$$

$$(A^+)^T \text{ pszeudoinverze } A^T\text{-nak}.$$

K Ha $A \in \mathbb{R}^{n\times n}$ invertálható, akkor $A^+ = A^{-1}$

B P1–P4. teljesül, továbbá, ha A^+ pszeudoinverze A-nak, akkor $I = AA^{-1} = AA^+AA^{-1} = AA^+ \Rightarrow A^+ = A^{-1}$, mert A négyzetes.

T **Pszeudoinverz létezése**

Minden valós mátrixnak létezik pszeudoinverze:

(1) Ha B teljes oszloprangú, azaz $B \in \mathbb{R}^{m\times r}$, ahol $r(B) = r$, akkor $B^+ = (B^TB)^{-1}B^T$, és ez B-nek balinverze: $B^+B = I$.

(2) Ha C teljes sorrangú, azaz $C \in \mathbb{R}^{r\times m}$, ahol $r(C) = r$, akkor $C^+ = C^T(CC^T)^{-1}$, és ez a C-nek jobbinverze: $CC^+ = I$.

(3) Ha $A \in \mathbb{R}^{m\times n}$, és $A = BC$ az A-nak bázisfelbontása (azaz $r(A) = r(B) = r(C) = r$, $B \in \mathbb{R}^{n\times r}$, $C \in \mathbb{R}^{r\times m}$), akkor $A^+ = C^+B^+$, és itt B^+ az (1), C^+ a (2) alapján számítható ki.

B (1): Legyen $B^+ := (B^T B)^{-1} B^T$. Belátjuk, hogy teljesül rá a pszeudoinverz P1–P4 definiáló tulajdonsága.

B oszlopai ftln-ek $\Rightarrow$ láttuk, hogy $B^T B$ invertálható, és $B(B^T B)^{-1} B^T = BB^+$ az $\mathcal{O}(B)$-re való merőleges vetítés mátrixa, így BB^+ szimmetrikus (P3 teljesül).

$B^+ B = (B^T B)^{-1} B^T B = I \Rightarrow$ P4 teljesül, és B^+ balinverz.

$BB^+ B = BI = B$ és $B^+ BB^+ = IB^+ = B^+$ (P1,P2 teljesül).

(2): Alkalmazzuk az (1)-et $B = C^T$-ra. Ekkor $C^+ = (B^T)^+ =$
$= (B^+)^T = \Big((B^T B)^{-1} B^T)\Big)^T = \Big((CC^T)^{-1} C\Big)^T = C^T (CC^T)^{-1}$.

(3): (1) és (2) miatt $B^+ B = I$ és $CC^+ = I$, így $A^+ := C^+ B^+$-ra

P1: $AA^+ A = BCC^+ B^+ BC = BIIC = BC = A$.

P2: $A^+ AA^+ = C^+ B^+ BCC^+ B^+ = C^+ IIB^+ = C^+ B^+ = A^+$.

P3: $AA^+ = BCC^+ B^+ = BIB^+ = BB^+$ szimmetrikus,

P4: $A^+ A = C^+ B^+ BC = C^+ IC = C^+ C$ szimmetrikus,

felhasználva a B^+ és C^+ pszeudoinverz tulajdonságait is.

T **A pszeudoinverz egyértelműsége**

Minden $A \in \mathbb{R}^{m \times n}$-re egyetlen olyan A^+ van, amely kielégíti a P1–P4. tulajdonságokat.

B Ha X és Y is pszeudoinverze A-nak $\Rightarrow$

$$X \overset{P2}{=} XAX \overset{P3}{=} X(AX)^T = XX^TA^T \overset{P1}{=} XX^T(AYA)^T = XX^TA^TY^TA^T = X(AX)^T(AY)^T \overset{P3}{=} XAXAY \overset{P2}{=} XAY$$

$$Y \overset{P2}{=} YAY \overset{P4}{=} (YA)^TY = A^TY^TY \overset{P1}{=} (AXA)^TY^TY = A^TX^TA^TY^TY = (XA)^T(YA)^TY \overset{P4}{=} XAYAY \overset{P2}{=} XAY.$$

$$\Rightarrow X = XAY = Y.$$

F Számítsuk ki $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & -1 & 1 \end{bmatrix}$ pszeudoinverzét!

m $r(A) = 2 \Rightarrow A$ teljes sorrangú $\Rightarrow A^+ = A^T(AA^T)^{-1}$

$$AA^T = \begin{bmatrix} 1 & 2 & 1 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 1 \\ 1 & 6 \end{bmatrix}, \quad (AA^T)^{-1} = \frac{1}{35} \begin{bmatrix} 6 & -1 \\ -1 & 6 \end{bmatrix}$$

$$A^+ = \begin{bmatrix} 1 & 2 \\ 2 & -1 \\ 1 & 1 \end{bmatrix} \frac{1}{35} \begin{bmatrix} 6 & -1 \\ -1 & 6 \end{bmatrix} = \frac{1}{35} \begin{bmatrix} 4 & 11 \\ 13 & -8 \\ 5 & 5 \end{bmatrix}$$

F Számítsuk ki $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix}$ pszeudoinverzét!

m $|A| = 0 \Rightarrow A$ nem teljes rangú. $\Rightarrow$ Kell a bázisfelbontás.

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix} = BC$$

$$B^+ = (B^T B)^{-1} B^T = \frac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \end{bmatrix}$$

$$C^+ = C^T (CC^T)^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 2 \end{bmatrix} \frac{1}{6} \begin{bmatrix} 5 & -2 \\ -2 & 2 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 5 & -2 \\ -2 & 2 \\ 1 & 2 \end{bmatrix}$$

$$A^+ = C^+ B^+ = \frac{1}{18} \begin{bmatrix} 12 & -9 & 3 \\ -6 & 6 & 0 \\ 0 & 3 & 3 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 4 & -3 & 1 \\ -2 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

Á Ha A^+ az A pszeudoinverze, akkor

(1) AA^+ az A oszlopterére való merőleges vetítés mátrixa.

(2) A^+A az A sorterére való merőleges vetítés mátrixa.

B (1): AA^+ szimmetrikus, és $(AA^+)^2 = AA^+AA^+ \overset{P1}{=} AA^+$, tehát AA^+ az $\mathcal{O}(AA^+)$-ra való merőleges vetítés mátrixa.

Másrészt $\mathcal{O}(AA^+) = \mathcal{O}(A)$, ugyanis $\leqslant$ nyilvánvaló, és $\mathcal{O}(A) = \mathcal{O}(AA^+A) \leqslant \mathcal{O}(AA^+)$.

(2): Az (1) miatt A^+A az A^+ oszlopterére való merőleges vetítés mátrixa. Azt kell még belátni, hogy $\mathcal{O}(A^+) = \mathcal{S}(A)$. Valóban:

$$\mathcal{O}(A^+) \overset{P2}{=} \mathcal{O}(A^+AA^+) \leqslant \mathcal{O}(A^+A) \overset{P4}{=} \mathcal{O}(A^T(A^+)^T) \leqslant \mathcal{O}(A^T) = \mathcal{S}(A),$$

$$\mathcal{S}(A) \overset{P1}{=} \mathcal{S}(AA^+A) \leqslant \mathcal{S}(A^+A) \overset{P4}{=} \mathcal{S}(A^T(A^+)^T) \leqslant \mathcal{S}((A^+)^T) = \mathcal{O}(A^+).$$

T **Legkisebb abszolút értékű optimális közelítő megoldás**

Ha A^+ az $A \in \mathbb{R}^{m\times n}$ mátrix pszeudoinverze, akkor $\forall \mathbf{b} \in \mathbb{R}^m$-re $A^+\mathbf{b}$ az $A\mathbf{x} = \mathbf{b}$ er. sortérbe eső, azaz legkisebb abszolút értékű (optimális közelítő) megoldása.

B Az előző tétel szerint $AA^+\mathbf{b}$ a $\mathbf{b}$ merőleges vetülete $\mathcal{O}(A)$-ra, tehát

ha az er. konzisztens, akkor $\mathbf{x} = A^+\mathbf{b}$-re $A\mathbf{x} = AA^+\mathbf{b} = \mathbf{b}$, így $\mathbf{x}$ megoldás,

ha inkonzisztens, akkor $A\mathbf{x} = AA^+\mathbf{b} = \mathbf{b}'$ miatt $\mathbf{x}$ optimális közelítő megoldás.

A sorterébe esik, mert $\mathbf{x} = A^+\mathbf{b} \in \mathcal{O}(A^+) = \mathcal{S}(A)$.

P Határozzuk meg az $A\mathbf{x} = \mathbf{b}$ er. sortérbe eső optimális közelítő megoldását az A korábban kiszámított pszeudoinverze segítségével, ahol

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix} \text{ és } \mathbf{b} = \begin{bmatrix} 1 \\ 4 \\ 5 \end{bmatrix}, \text{ illetve } \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$$

m Az elsőre $A^+\mathbf{b} = \frac{1}{6}\begin{bmatrix} 4 & -3 & 1 \\ -2 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 4 \\ 5 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix}$ megoldás: $A\mathbf{x} = \mathbf{b}$.

A másodikra $A^+\mathbf{b} = \frac{1}{6}\begin{bmatrix} 4 & -3 & 1 \\ -2 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ -\frac{1}{3} \\ \frac{1}{3} \end{bmatrix}$ optimális közelítő megoldás, $\mathbf{b}' = A\mathbf{x} = (\frac{4}{3}, \frac{1}{3}, \frac{5}{3})^T$ (az eltérés $\frac{1}{\sqrt{3}}$).