# Numerikus módszerek 2B.

## 8–9. előadás: Szinguláris felbontás, általánosított inverz

Dr. Bozsik József

ELTE IK

2020. november 5–12.

# Tartalomjegyzék

# Tartalomjegyzék

**Szemléltetés:**

- Szimmetrikus $A$ mátrix esetén $\exists$ $U$ unitér és $D$ diagonális mátrix, melyre $A = UDU^* \quad \Leftrightarrow \quad AU = UD$.
  Ez szemléletesen azt jelenti, hogy az $A$ mátrix az $u_1, \ldots, u_n$ ONR-t a $\lambda_1 u_1, \ldots, \lambda_n u_n$ vektorokba viszi.
  A kép vektorok ortogonalitása továbbra is megmarad.

- Nem szimmetrikus esetben ez nem lehetséges. Azt szeretnénk elérni, hogy $A$ az értelmezési tartomány ($\mathbb{R}^n$) $v_1, \ldots, v_r$ ortonormált vektorait (ahol $r = rang(A)$) a képhalmaz ($\mathbb{R}^m$) $\sigma_1 u_1, \ldots, \sigma_r u_r$ ortogonális vektoraiba vigye. Ekkor

$$A \cdot \begin{bmatrix} \vdots & \vdots & \vdots \\ v_1 & & v_r \\ \vdots & \vdots & \vdots \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & \vdots \\ u_1 & & u_r \\ \vdots & \vdots & \vdots \end{bmatrix} \cdot \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{bmatrix}.$$

**Tétel: Szinguláris felbontás**

Tetszőleges $A \in \mathbb{C}^{m \times n}$ esetén $\exists$ $U \in \mathbb{C}^{m \times m}$, $V \in \mathbb{C}^{n \times n}$ unitér mátrix és $D \in \mathbb{C}^{m \times n}$, particionálva $D = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}$, hogy

$$A = UDV^*,$$

ahol $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_r)$, $r = \operatorname{rang}(A)$ és $\sigma_1 \geq \ldots \geq \sigma_r > 0$.

**Definíció:** Szinguláris értékek

A szinguláris felbontásbeli $\sigma_1, \ldots, \sigma_r > 0$ értékeket az $A$ *szinguláris értékeinek* nevezzük.

$$\sigma_i = \sqrt{\lambda_i(A^*A)} > 0, \quad (i = 1, \ldots, r)$$

**Megjegyzések:**

- Az $A^*A \in \mathbb{C}^{n\times n}$ és $AA^* \in \mathbb{C}^{m\times m}$ pozitív sajátértékei azonosak.
- $A^*A$ sajátvektorai a $V$ oszlopai és
- $AA^*$ sajátvektorai az $U$ oszlopai.

**Speciális esetben:** $m = n$ esetén

- $\sigma_1 = \max_{i=1}^n \sigma_i = \|A\|_2 = \|D\|_2$,
- $\|A\|_F^2 = \|D\|_F = \sum_{i=1}^n \sigma_i^2$.

**Tétel: Lineáris transzformáció jellemzése**

Tetszőleges $A \in \mathbb{C}^{m\times n}$ mátrix és $\sigma_1 \geq \ldots \geq \sigma_r > 0$ szinguláris értékek esetén $\operatorname{rang}(A) = r$, továbbá $\ker A = \operatorname{span}\{u_1 \ldots, u_r\}$ és $\operatorname{Im} A = \operatorname{span}\{v_{r+1}, \ldots, v_n\}$.

**Alkalmazási területei:**

1. Többváltozós statisztikában a szóródási mátrix vizsgálata, a variancia tömörítése.
2. Nagy dokumentumhalmazokban való keresés és tematizálás (Főkomponens analízis, adatbányászat).
3. Geofizikai mérések elemzésénél.

**Tétel:** Eckart–Young-tétel

Ha $\text{rang}\,(A) = r$, akkor a $k$ rangú legjobb közelítése előállítható a szinguláris felbontásból

$$A_k = \sum_{i=1}^{k} \sigma_i u_i v_i^T,$$

ahol $\sigma_i$ az $i$. szinguláris értéket, $u_i$ és $v_i$ a bal és jobb szinguláris vektort jelöli. Ez a közelítés a 2-es és a Frobenius normában is a legjobb, a hibák:

$$\|A - A_k\|_F = \sqrt{\sum_{i=k+1}^{r} \sigma_i^2}$$

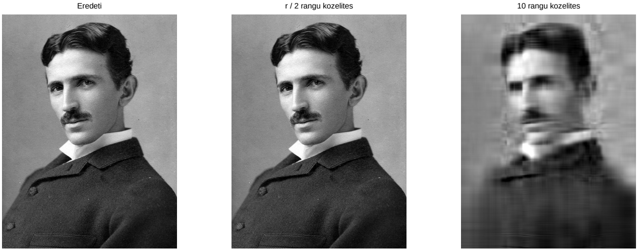
$$\|A - A_k\|_2 = \sigma_{k+1}.$$

**Példa:**

Vegyünk egy pixelformátumú szürkeárnyalatos fényképet és készítsünk belőle egy mátrixot, melynek minden eleme a kép egy pontjának árnyalatát adja meg. Legyen a mátrix rangja $r$.

- Írjuk fel a mátrix szinguláris felbontását,
- majd abból készítsünk két olyan kép-mátrixot, hogy egyiknek $r/2$, másiknak 10 a rangja, és Frobenius-normában a legközelebb vannak az eredeti kép mátrixához.

# Alkalmazás jelfeldolgozásban



## Tartalomjegyzék

**Definíció:** Általánosított inverz

Az $A \in \mathbb{C}^{m \times n}$ mátrix *Moore-Penrose-féle általánosított (pszeudo) inverze* az $A^+ \in \mathbb{C}^{n \times m}$ mátrix, ha

1. $AA^+$ önadjungált,
2. $A^+A$ önadjungált,
3. $AA^+A = A$,
4. $A^+AA^+ = A^+$.

**Tétel:**

$A^+$ egyértelmű.

**Tétel:** Diagonális mátrix általánosított inverze

A $D \in \mathbb{C}^{m \times n}$ diagonális mátrix általánosított inverze a $D^+ \in \mathbb{C}^{n \times m}$ diagonális mátrix, ahol $d_{ii}^+ = \frac{1}{d_{ii}}$, ha $d_{ii} \neq 0$, különben 0 az értéke.

**Biz.:** Az 1)-4) tulajdonságokat kell ellenőrizni.

**Tétel:** Az általánosított inverz előállítása

A szinguláris felbontás felhasználásával az $A \in \mathbb{C}^{m \times n}$ mátrix általánosított inverze a következő alakban állítható elő:

$$A^+ = VD^+U^*, \text{ ahol } D^+ \in \mathbb{C}^{n \times m} \text{ lásd az előző tételben.}$$

**Biz.:** Következik az előző tételből, az 1)-4) tulajdonságokat kell ellenőrizni.

**Tétel:** Túlhatározott teljes rangú eset

Legyen $A \in \mathbb{C}^{m \times n}$, $m > n$ és $r = \operatorname{rang}(A) = n$. Ekkor

$$A^+ = (A^*A)^{-1}A^*.$$

**Biz.:** Az 1)-4) tulajdonságokat kell ellenőrizni.

**Tétel:** Alulhatározott teljes rangú eset

Legyen $A \in \mathbb{C}^{m \times n}$, $m < n$ és $r = \operatorname{rang}(A) = m$. Ekkor

$$A^+ = A^*(AA^*)^{-1}.$$

**Biz.:** Az 1)-4) tulajdonságokat kell ellenőrizni.

# Az általánosított inverz approximációs tulajdonsága

**Definíció:** Általánosított megoldás

Legyen $A \in \mathbb{C}^{m \times n}$, $b \in \mathbb{C}^m$, ekkor az $x^+ := A^+ b \in \mathbb{C}^n$ vektort az $Ax = b$ *LER általánosított megoldásának* nevezzük.

**Következmények:**

- Ha $A$ túlhatározott és teljes rangú, akkor

$$x^+ = A^+ b = (A^* A)^{-1} A^* b \;\Leftrightarrow\; A^* A\, x^+ = A^* b.$$

  Az $A^* A\, x = A^* b$ LER-t Gauss-féle normálegyenleteknek nevezik, $x^+$ a megoldása.

- Ha $A$ alulhatározott és teljes rangú, akkor

$$x^+ = A^+ b = A^* \underbrace{(AA^*)^{-1} b}_{=:y} \;\Leftrightarrow\; AA^* y = b \;\text{ és }\; x^+ = A^* y.$$

- A fenti két esetben nem szükséges szinguláris felbontás az általánosított inverz előállításához.

# Az általánosított inverz approximációs tulajdonsága

**Tétel:** Az általánosított inverz approximációs tulajdonsága

1. $\|Ax - b\|_2 \geq \|Ax^+ - b\|_2, \quad \forall\, x \in \mathbb{C}^n.$
2. $H := \{x \in \mathbb{C}^n |\ \|Ax - b\|_2 = \|Ax^+ - b\|_2\}$, akkor

$$\|x^+\|_2 < \|x\|_2 \ \forall\, x \in \mathbb{C}^n,\ \ x \neq x^+.$$

**Megjegyzések:**

1. Ha nem létezik megoldás, akkor a $\|\cdot\|_2$ norma szerint $b$-hez legközelebbi pontot adó $x^+$ az általánosított megoldás.
2. Ha létezik megoldás, akkor az origóhoz legközelebbi megoldás az általánosított megoldás.

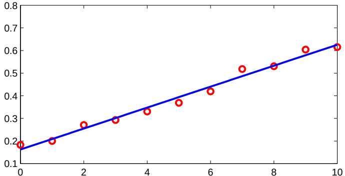
# Tartalomjegyzék

**Definíció:** A legkisebb négyzetek módszerének alapfeladata

Adottak az $x_1, \ldots, x_N \in [a; b]$ különböző alappontok, $y_1, \ldots, y_N \in \mathbb{R}$ függvényértékek vagy mérési eredmények. Olyan $p_n \in P_n$ polinomot keresünk $(n+1 \leq N$, általában $N \gg n)$, melyre

$$\sum_{i=1}^{N} (y_i - p_n(x_i))^2 \quad \text{minimális.}$$

A $p_n$ polinomot *négyzetesen legjobban közelítő polinomnak* nevezzük.



Írjuk fel a $p_n(x_i) = y_i$, $(i = 1, \ldots, N)$ LER-t mátrix alakban, ahol

$$p_n(x) := a_n x^n + \ldots + a_1 x + a_0.$$

Vezessük be hozzá a következő jelöléseket:

$$A := \begin{bmatrix} 1 & x_1 & \ldots & x_1^n \\ 1 & x_2 & \ldots & x_2^n \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 1 & x_N & \ldots & x_N^n \end{bmatrix}_{N\times(n+1)} \qquad a := \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}_{n+1\times 1} \qquad y := \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ y_N \end{bmatrix}_{N\times 1} .$$

A kapott $A \cdot a = y$ LER klasszikus értelemben nem oldható meg, különböző alappontok esetén rang $(A) = n + 1$, vagyis teljes rangú LER-t kaptunk, amit általánosított értelemben meg tudunk oldani.

Túlhatározott teljes rangú esetben az általánosított megoldásra a Gauss-féle normálegyenleteket kapjuk:

$$a^+ = A^+ y = (A^T A)^{-1} \cdot A^T y \quad \Leftrightarrow \quad A^T A \cdot a^+ = A^T y.$$

A szimmetrikus LER alakja:

$$\begin{bmatrix} N & \sum x_i & \cdots & \sum x_i^n \\ \sum x_i & \sum x_i^2 & \cdots & \sum x_i^{n+1} \\ \vdots & & & \vdots \\ \sum x_i^n & \sum x_i^{n+1} & \cdots & \sum x_i^{2n} \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \vdots \\ \sum x_i^n y_i \end{bmatrix}.$$

Alkalmazzuk az approximációs tulajdonságot a LER-re:

$$\|A \cdot a^+ - y\|_2 \leq \|A \cdot a - y\|_2, \quad \forall\, a \in \mathbb{R}^{n+1}.$$

$$\|A \cdot a - y\|_2^2 = \sum_{i=1}^{N} (p_n(x_i) - y_i)^2 \;\; \to \;\; \text{minimalizálása},$$

így $a^+$ a négyzetesen legjobban közelítő polinom együtthatóit adja.

**Megjegyzések:**

- Ha $A$ teljes rangú, akkor $A^T A$ mindig szimmetrikus és invertálható.
- A LER alakja $n = 1$ esetben:
$$\begin{bmatrix} N & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}.$$
- Ha $n = 1$ esetben $\sum x_i = 0$, akkor diagonális LER-t kapunk, melynek megoldása:
$$a_0 = \frac{1}{N} \sum y_i, \quad a_1 = \frac{\sum x_i y_i}{\sum x_i^2}.$$
A közgazdászok előszeretettel használják (lásd statisztika, regressziószámítás).
- A négyzetesen legjobban közelítő egyenes mindig átmegy az $\left(\frac{1}{N} \sum x_i, \frac{1}{N} \sum y_i\right)$ (átlagokból álló) ponton. ($n = 1$-re a Gauss-féle normálegyenletek első sora épp ezt jelenti.)