

# Bevezetés az algebra 2 – Szinguláris érték, SVD

Wettl Ferenc  
Algebra Tanszék



2016. április 29.

# Ötlet

- K** Hogyan válasszunk bázist, ha a mátrix téglalap alakú? Mit tehetünk még, ha a mátrix nem diagonalizálható?
- m** A  $\mathcal{V} \rightarrow \mathcal{V}$  lineáris transzformációk ortogonális diagonalizálását fogjuk általánosítani  $\mathcal{V}_1 \rightarrow \mathcal{V}_2$  lineáris leképezésekre.
- m** Az  $\mathbf{A} \in \mathbb{R}^{m \times n}$  mátrixhoz olyan  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\} \subset \mathbb{R}^n$  és  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\} \subset \mathbb{R}^m$  ONB-okat keresünk, melyekben  $\mathbf{A}$  mátrixa diagonális.
- m**  $\mathbf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i$ , ahol  $1 \leq i \leq \min(m, n)$
- Á** Ha az  $\mathbb{R}^n$ -beli  $\mathbf{x} \perp \mathbf{y}$  vektorok legalább egyike az  $\mathbf{A}^T \mathbf{A}$  sajátvektora, akkor  $\mathbf{A}\mathbf{x} \perp \mathbf{A}\mathbf{y}$
- B**  $\mathbf{x} \cdot \mathbf{y} = 0$ ,  $\mathbf{A}^T \mathbf{A}\mathbf{y} = \lambda \mathbf{y} \rightsquigarrow$   
 $\mathbf{A}\mathbf{x} \cdot \mathbf{A}\mathbf{y} = (\mathbf{A}\mathbf{x})^T \mathbf{A}\mathbf{y} = \mathbf{x}^T \mathbf{A}^T \mathbf{A}\mathbf{y} = \mathbf{x} \cdot (\lambda \mathbf{y}) = \lambda(\mathbf{x} \cdot \mathbf{y}) = 0$   
Komplexben  $\mathbf{A}^H \mathbf{A}$ -val!

# Szinguláris érték, szinguláris vektor

- $A: \mathbf{x} \mapsto \mathbf{Ax}$  leképezés kölcsönösen egyértelmű a sortér és az oszloptér között (komplex függvény esetén  $\mathcal{O}(\mathbf{A}^H)$  és  $\mathcal{O}(\mathbf{A})$  között).
- $\mathbf{A}^T \mathbf{A}$  szimmetrikus és pozitív szemidefinit  $\rightsquigarrow$  ortogonálisan diagonalizálható és a sajátértékei nem negatívak
- sajátvektorokból ONB, a pozitív sajátértékekhez tartozók kifeszítik a sorteret ( $\mathcal{N}(\mathbf{A}^T \mathbf{A}) = \mathcal{N}(\mathbf{A})$  merőlegesét)
- $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$  a sortér bázisa  $\rightsquigarrow \mathbf{Av}_i$  OB az oszloptérben
- $|\mathbf{Av}_i|^2 = \mathbf{v}_i^T \mathbf{A}^T \mathbf{Av}_i = \lambda_i |\mathbf{v}_i|^2 = \lambda_i$
- L!  $\sigma_i = \sqrt{\lambda_i}$ ,  $\mathbf{u}_i = \mathbf{Av}_i / \sigma_i \rightsquigarrow \{\mathbf{u}_1, \dots, \mathbf{u}_r\}$  ONB
- D Az  $r$  rangú  $\mathbf{A} \in \mathbf{R}^{m \times n}$  mátrix **szinguláris értékeinek** nevezzük az  $\mathbf{A}^T \mathbf{A}$  pozitív sajátértékeinek négyzetgyökeit ( $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ ).  
Az  $\mathbf{A}^T \mathbf{A}$  mátrix valamely  $\sigma^2$  sajátértékéhez tartozó egységnyi  $\mathbf{v}$  sajátvektorát az  $\mathbf{A}$  **jobb szinguláris vektorának**, az  $\mathbf{u} = \frac{1}{\sigma} \mathbf{Av}$  egységvektort pedig a  $\mathbf{v}$ -hez tartozó **bal szinguláris vektorának** nevezzük.

- $|\mathbf{u}_i| = 1$  ( $i = 1, 2, \dots, r$ )  $\rightsquigarrow |\mathbf{A}\mathbf{v}_i| = \sigma_i$
- úgy is definiálható, hogy ha  $k = \min(m, n)$ , akkor a szinguláris értékek száma  $k$ , és értékük nem negatív, így  $\sigma_{r+1} = \sigma_{r+2} = \dots = \sigma_k = 0$ .
- Ha  $\sigma$   $t$ -szeres multiplicitású szinguláris értéke  $\mathbf{A}$ -nak, akkor  $\sigma^2$  az  $\mathbf{A}^\top \mathbf{A}$ -nak  $t$ -szeres algebrai (és  $t$ -szeres geometriai) multiplicitású sajátértéke, így a  $\sigma$ -hoz tartozó jobb szinguláris vektorok a sortérben, a hozzá tartozó bal szinguláris vektorok pedig az oszloptérben  $t$ -dimenziós alteret feszítenek ki.
- Ha  $\mathbf{A}\mathbf{v} = \sigma\mathbf{u}$ , akkor  $\mathbf{A}^\top \mathbf{u} = \frac{1}{\sigma} \mathbf{A}^\top \mathbf{A}\mathbf{v} = \frac{1}{\sigma} \sigma^2 \mathbf{v} = \sigma \mathbf{v}$ , tehát az

$$\mathbf{A}\mathbf{v} = \sigma\mathbf{u}, \quad \mathbf{A}^\top \mathbf{u} = \sigma\mathbf{v} \quad (1)$$

összefüggések párba állítják a jobb és bal szinguláris vektorokat.

- Ha  $\mathbf{A} \in \mathbb{C}^{n \times n}$ , akkor a sortér helyett a sortér konjugáltjával, azaz az  $\mathcal{O}(\mathbf{A}^H)$  altérrel, transzponálás helyett adjungálással, tehát  $\mathbf{A}^\top \mathbf{A}$  helyett az  $\mathbf{A}^H \mathbf{A}$  mátrixszal igazak maradnak a fentiek.

**P** Határozzuk meg

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

szinguláris értékeit, és keressünk megfelelő ONB-t a sor és oszloptérben.

**M**  $\mathbf{A}^T \mathbf{A} = \begin{bmatrix} 73 & -36 \\ -36 & 52 \end{bmatrix}$ ,  $\rightsquigarrow \chi(x) = x^2 - 125x + 2500 \rightsquigarrow \lambda_1 = 100$ ,  
 $\lambda_2 = 25 \rightsquigarrow \sigma_1 = 10, \sigma_2 = 5$ .

$\mathbf{A}^T \mathbf{A}$  sajátpárjai:  $(100, (4/5, -3/5))$ ,  $(25, (3/5, 4/5))$ .

jobb szinguláris vektorok:  $\mathbf{v}_1 = (4/5, -3/5)$ ,  $\mathbf{v}_2 = (3/5, 4/5)$

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

$\rightsquigarrow \mathbf{u}_1 = (-5/13, 12/13)$  és  $\mathbf{u}_2 = (12/13, 5/13)$  a hozzájuk tartozó bal szinguláris vektorok.

# Redukált szinguláris felbontás és annak diadikus alakja

- $\mathbf{\Sigma}_1 = \text{diag}(\sigma_1, \dots, \sigma_r)$   $\mathbf{U}_1 = \{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ ,  $\mathbf{V}_1 = \{\mathbf{v}_1, \dots, \mathbf{v}_r\}$   
 $\mathbf{A}\mathbf{v}_i = \sigma\mathbf{u}_i \rightsquigarrow \mathbf{A}\mathbf{V}_1 = \mathbf{U}_1\mathbf{\Sigma}_1$ , azaz

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

$\rightsquigarrow \mathbf{A} = \mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^T$  (mert  $\mathbf{V}_1$  oszlopvektorai  $\mathcal{S}(\mathbf{A})$  ONB-a, így  $\mathbf{A}\mathbf{V}\mathbf{V}^T = \mathbf{A}$ )

**D** Redukált szinguláris felbontás:  $\mathbf{A} = \mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^T$  ( $\mathbf{A} = \mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^H$ ), ahol  $\mathbf{V}_1$ ,  $\mathbf{U}_1$  szemiortogonális („szemiunitér”),  $\mathbf{\Sigma}_1$  négyzetes diagonális pozitív főátlóval

**D** Szinguláris felbontás diadikus alakja:

$$\mathbf{A} = \sigma_1\mathbf{u}_1\mathbf{v}_1^T + \sigma_2\mathbf{u}_2\mathbf{v}_2^T + \dots + \sigma_r\mathbf{u}_r\mathbf{v}_r^T$$

# Szinguláris felbontás, SVD

- Egészítsük ki  $\mathbf{V}_1$ -et ortogonális  $\mathbf{V}$  mátrixszá:  $n > r$  esetén

$$\mathbf{V}_2 = [\mathbf{v}_{r+1} \quad \dots \quad \mathbf{v}_n], \quad \mathbf{V} = [\mathbf{V}_1 | \mathbf{V}_2]$$

$\mathbf{V}_2$  oszlopvektorai kifeszítik a nullteret  $\leadsto r < i \leq n$  esetén  $\mathbf{A}\mathbf{v}_i = \mathbf{0}$   
 $\leadsto \mathbf{A}\mathbf{V} = [\mathbf{U}_1 | \mathbf{0}]$

- Egészítsük ki  $\mathbf{U}_1$ -et ortogonális  $\mathbf{U}$  mátrixszá:  $m > r$  esetén

$$\mathbf{U}_2 = [\mathbf{u}_{r+1} \quad \dots \quad \mathbf{u}_m], \quad \mathbf{V} = [\mathbf{U}_1 | \mathbf{U}_2]$$

- Egészítsük ki  $\Sigma_1$ -et  $m \times n$ -es mátrixszá:  $\Sigma = \begin{bmatrix} \Sigma_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$

$\leadsto \mathbf{A}\mathbf{V} = \mathbf{U}\Sigma$ , ugyanis  $\mathbf{A}\mathbf{V} = \mathbf{A}[\mathbf{V}_1 | \mathbf{V}_2] = [\mathbf{U}_1 | \mathbf{0}] = [\mathbf{U}_1 | \mathbf{U}_2] \begin{bmatrix} \Sigma_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} = \mathbf{U}\Sigma$

$\leadsto \mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T$

- D Szinguláris felbontás:**  $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T$  (komplexben  $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^H$ ), ahol  $\mathbf{V}$  és  $\mathbf{U}$  ortogonális (unitér),  $\Sigma$  valós diagonális (monoton csökkenő nemnegatív diagonális elemekkel)

# Létezés és egyértelműség

**T** Minden  $\mathbf{A} \in \mathbb{R}^{m \times n}$  ( $\in \mathbb{C}^{m \times n}$ ) mátrixnak létezik szinguláris és redukált szinguláris felbontása, melyben a szinguláris értékek monoton csökkenő sorozata egyértelmű.

**m** Maga a felbontás nem egyértelmű!

**m** A mátrixok méreteit is kiírva  $\mathbf{A}_{m \times n} = \mathbf{U}_{m \times m} \mathbf{\Sigma}_{m \times n} \mathbf{V}_{n \times n}^T$

**B**  $\mathbf{A}^T \mathbf{A} \in \mathbb{R}^{n \times n}$  szimmetrikus ( $\mathbf{A}^H \mathbf{A}$  önadjungált), pozitív szemidefinit

~> ortogonálisan (unitéren) diag-ható és a diag. elemek nemneg. valósok

~> sajátvektorokból  $\rightarrow$  ONB  $\rightarrow \mathbf{V}$  ortogonális (unitér).

( $\mathbf{A}\mathbf{V}$  nemnulla oszlopvektorai)  $\cup$  ( $\mathcal{N}(\mathbf{A}^T)$  egy ONB-a)  $\rightarrow \mathbf{U}$

~>  $\exists \mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$  ( $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ ),  $\mathbf{\Sigma}$ -ben a sajátértékek négyzetgyökei.

■ ha  $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$  egy felbontás, akkor

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

~>  $\mathbf{\Sigma}$  elemei a sajátértékeinek gyökei, ugyanis

$\mathbf{\Sigma}^H \mathbf{\Sigma} = \mathbf{\Sigma}^T \mathbf{\Sigma} = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2, 0, \dots, 0)$  és a sajátértékek

egyértelműek, így a szinguláris értékek is.



# SVD kiszámítása

**P** 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!

$$\mathbf{M} \quad \mathbf{A}^T \mathbf{A} = \begin{bmatrix} 20 & -14 & 4 \\ -14 & 17 & -10 \\ 4 & -10 & 8 \end{bmatrix}.$$

$\chi(x) = x^3 - 45x^2 + 324x$ , gyökei 36, 9 és 0  $\rightsquigarrow$  a szinguláris értékek 6 és 3

**■**  $\mathbf{A}^T \mathbf{A}$  egységnyi sajátvektorai:

$$\lambda_1 = 36$$

$$\mathbf{v}_1 = (2/3, -2/3, 1/3)$$

$$\lambda_2 = 9$$

$$\mathbf{v}_2 = (2/3, 1/3, -2/3)$$

$$\lambda_3 = 0$$

$$\mathbf{v}_3 = (1/3, 2/3, 2/3).$$



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

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

■  $\mathbf{u}_3$  így nem számolható ( $\mathbf{A}\mathbf{v}_3 = \mathbf{0}$ ), így az  $\{\mathbf{u}_1, \mathbf{u}_2\}$  rendszert nekünk kell kiegészítenünk  $\mathbb{R}^3$  bázisává.

■  $\mathcal{O}(\mathbf{A})^\perp$  ONB-át keressük:

$$\blacksquare \text{span}(\mathbf{u}_1, \mathbf{u}_2)^\perp: \begin{bmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -2 \end{bmatrix} \rightarrow \mathbf{u}_3 = \frac{1}{3} \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$$

$$\blacksquare \mathcal{O}(\mathbf{A})^\perp: \begin{bmatrix} 0 & -2 & 4 \\ -2 & 3 & -2 \\ 2 & -2 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -2 \end{bmatrix} \rightarrow \mathbf{u}_3 = \frac{1}{3} \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$$

$$\blacksquare \mathbf{u}_3 = \mathbf{u}_1 \times \mathbf{u}_2 = (-2/3, -2/3, -1/3).$$

$$\rightsquigarrow \mathbf{U} = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ -2/3 & 1/3 & 2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix} \quad \mathbf{V} = \begin{bmatrix} 2/3 & 2/3 & 1/3 \\ -2/3 & 1/3 & 2/3 \\ 1/3 & -2/3 & 2/3 \end{bmatrix}$$

■ Szinguláris felbontás:

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

■ Redukált szinguláris felbontás:

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

■ Szinguláris felbontás diadikus alakja:

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

P Számítsuk ki az  $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}$  szinguláris érték szerinti felbontását!

- $\mathbf{A}^T \mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ , karakterisztikus polinomja  $\lambda^2 - 4\lambda + 3 = (\lambda - 3)(\lambda - 1)$
- $\mathbf{A}$  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  $\mathbf{u}_i = \mathbf{A}\mathbf{v}_i/\sigma_i$ :  $\mathbf{u}_1 = \frac{1}{\sqrt{6}}(1, 2, 1)$ ,  $\mathbf{u}_2 = \frac{1}{\sqrt{2}}(1, 0, -1)$ .
- ( $\mathbf{u}_3 = \mathbf{u}_1 \times \mathbf{u}_2$  is megy)  $\mathbf{A}^T$  nullterének bázisát keressük:  
 $\mathbf{A}^T \mathbf{x} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \mathbf{x} = \mathbf{0}$  megoldása:  $\mathbf{u}_3 = \frac{1}{\sqrt{3}}(1, -1, 1)$
- Tehát a szinguláris felbontás

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

## $\mathbf{A}^T \mathbf{A}$ és $\mathbf{A} \mathbf{A}^T$

- m** A számolás kezdhető  $\mathbf{U}$  meghatározásával is.  $\mathbf{A}^T \mathbf{A}$  és  $\mathbf{A} \mathbf{A}^T$  nem zérus sajátértékei megegyeznek,

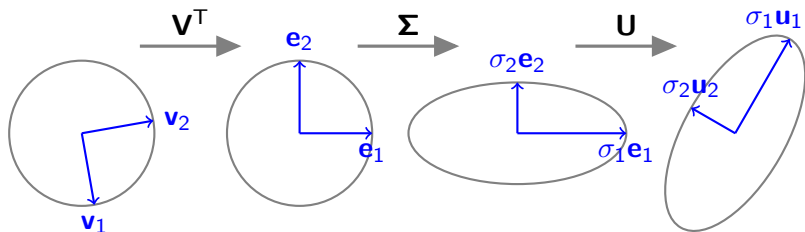
$$\mathbf{A} \mathbf{A}^T = \mathbf{U} \Sigma \mathbf{V}^T (\mathbf{U} \Sigma \mathbf{V}^T)^T = \mathbf{U} \Sigma \mathbf{V}^T \mathbf{V} \Sigma^T \mathbf{U}^T = \mathbf{U} \Sigma \Sigma^T \mathbf{U}^T.$$

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

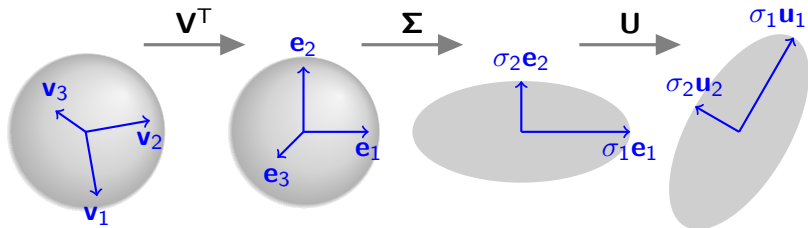
- érdemes  $\mathbf{A} \mathbf{A}^T$ -t használni, ha  $m < n$ , mert ekkor csak  $m$ -dimenziós vektorokkal kell számolni.

# Szinguláris felbontás geometriai interpretációja

- **L!**  $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$   $2 \times 2$ -es 2-rangú



- **L!**  $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$   $2 \times 3$ -as 2-rangú



# Egységgömb képe

**T**  $\mathbf{A} \in \mathbb{R}^{m \times n}$   $r$ -rangú, az  $\mathbf{x} \mapsto \mathbf{Ax}$  leképezés az  $\mathbf{e}^T \mathbf{e} = 1$  egyenletet kielégítő pontokat az  $\mathbb{R}^m$  egy  $r$ -dimenziós altere

- 1 egy ellipszoidjának felületére képzi, ha  $r = n$ , és
- 2 egy ellipszoidja által határolt tartományára képzi, ha  $r < n$ .

**B** A szinguláris felbontás diadikus alakja:

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

$$|\mathbf{e}| = 1 \rightsquigarrow |\mathbf{V}^T \mathbf{e}|^2 = (\mathbf{v}_1^T \mathbf{e})^2 + \cdots + (\mathbf{v}_n^T \mathbf{e})^2 = 1, \text{ hisz } \mathbf{V} \text{ ortogonális}$$

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

$$\text{ahol } x_i = \sigma_i \mathbf{v}_i^T \mathbf{e} \ (i = 1, 2, \dots, r).$$

$$\blacksquare \text{ L! } \mathbf{x} = (x_1, x_2, \dots, x_r, 0, \dots, 0) \rightsquigarrow \mathbf{Ae} = \mathbf{Ux} \rightsquigarrow |\mathbf{Ae}| = |\mathbf{Ux}| = |\mathbf{x}|$$

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

Eszerint az egyenlet

$$\begin{aligned} \left(\frac{x_1}{\sigma_1}\right)^2 + \left(\frac{x_2}{\sigma_2}\right)^2 + \dots + \left(\frac{x_r}{\sigma_r}\right)^2 &= 1, \quad \text{ha } r = n, \\ \left(\frac{x_1}{\sigma_1}\right)^2 + \left(\frac{x_2}{\sigma_2}\right)^2 + \dots + \left(\frac{x_r}{\sigma_r}\right)^2 &\leq 1, \quad \text{ha } r < n. \end{aligned}$$



# Polárfelbontás

- m**  $re^{i\varphi}$ : egy nemnegatív nyújtási tényező ( $r$ ) és egy egységnyi abszolút értékű komplex szám ( $e^{i\varphi}$ , ami a komplex síkon  $\varphi$ -vel való forgatás) szorzata.
- D** **Polárfelbontáson** egy négyzetes mátrixnak egy pozitív szemidefinit és egy ortogonális mátrix szorzatára való felbontását értjük.
- T** Bármely komplex (valós) négyzetes **A** mátrix előáll

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

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

**B**  $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^H\mathbf{U}\mathbf{V}^H = (\mathbf{U}\mathbf{\Sigma}\mathbf{U}^H)(\mathbf{U}\mathbf{V}^H)$ , ahonnan  $\mathbf{P} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^H$ ,  $\mathbf{Q} = \mathbf{U}\mathbf{V}^H$ .

$\mathbf{P}$  önadjungált, hisz  $(\mathbf{U}\mathbf{\Sigma}\mathbf{U}^H)^H = \mathbf{U}\mathbf{\Sigma}^H\mathbf{U}^H = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^H$ .

$\mathbf{A}$   $\mathbf{P}$  pozitív szemidefinit, hisz hasonló a pozitív szemidefinit  $\mathbf{\Sigma}$  mátrixhoz (ha  $\mathbf{A}$  invertálható, akkor  $\mathbf{\Sigma}$  pozitív definit)

$\mathbf{Q}$  unitér (ortogonális), hisz két unitér (ortogonális) mátrix szorzata.

$\mathbf{A}$   $\mathbf{P}$  egyértelmű (nem csak akkor, ha pozitív definit), ugyanis

$$\mathbf{A}\mathbf{A}^H = \mathbf{P}\mathbf{Q}\mathbf{Q}^H\mathbf{P}^H = \mathbf{P}\mathbf{P}^H = \mathbf{P}^2,$$

azaz  $\mathbf{P} = \sqrt{\mathbf{A}\mathbf{A}^H}$ , és pozitív szemidefinit önadjungált mátrix négyzetgyöke egyértelmű az önadjungált pozitív szemidefinit mátrixok körében.

Ha  $\mathbf{P}$  pozitív definit, akkor invertálható, így  $\mathbf{Q} = \mathbf{P}^{-1}\mathbf{A}$  is egyértelmű.

- analógia a determinánsra is:  $\det \mathbf{P} = r$ ,  $\det \mathbf{Q} = e^{i\varphi}$  (hisz  $\mathbf{Q}$  unitér),  $\det \mathbf{A} = r e^{i\varphi}$ .
- Fordított sorrend: azonos unitér (ortogonális) mátrixszal:

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H = \mathbf{U}\mathbf{V}^H\mathbf{V}\mathbf{\Sigma}\mathbf{V}^H = (\mathbf{U}\mathbf{V}^H)(\mathbf{V}\mathbf{\Sigma}\mathbf{V}^H) = \mathbf{Q}\hat{\mathbf{P}},$$

- Valós polárfelbontás geometriai jelentése: minden mátrixleképezés két olyan leképezés kompozíciója, 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 tényezővel.

**P** Polárfelbontásait számítsuk ki

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

**B**  $\mathbf{P} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^T$ ,  $\mathbf{Q} = \mathbf{U}\mathbf{V}^T$ ,  $\hat{\mathbf{P}} = \mathbf{V}\mathbf{\Sigma}\mathbf{V}^H$

$$\begin{aligned} \mathbf{A} = \mathbf{P}\mathbf{Q} &= \begin{bmatrix} 2 & -2 & 0 \\ -2 & 3 & -2 \\ 0 & -2 & 4 \end{bmatrix} \begin{bmatrix} -4/9 & -8/9 & 1/9 \\ -4/9 & 1/9 & -8/9 \\ 7/9 & -4/9 & -4/9 \end{bmatrix} \\ &= \mathbf{Q}\hat{\mathbf{P}} = \begin{bmatrix} -4/9 & -8/9 & 1/9 \\ -4/9 & 1/9 & -8/9 \\ 7/9 & -4/9 & -4/9 \end{bmatrix} \begin{bmatrix} 4 & -2 & 0 \\ -2 & 3 & -2 \\ 0 & -2 & 2 \end{bmatrix}. \end{aligned}$$

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

# Pseudoinverz

**T**  $\mathbf{A}$  valós mátrix és  $\mathbf{A} = \mathbf{U}_1 \mathbf{\Sigma}_1 \mathbf{V}_1^T = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T$ . Ekkor

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

**B** Az  $\mathbf{A} = \mathbf{U}_1 (\mathbf{\Sigma}_1 \mathbf{V}_1^T)$  felbontásban  $\mathbf{U}_1$  teljes oszloprangú,  $\mathbf{\Sigma}_1 \mathbf{V}_1^T$  teljes sorrangú:

$$\begin{aligned} \mathbf{A}^+ &= (\mathbf{\Sigma}_1 \mathbf{V}_1^T)^T \left( \mathbf{\Sigma}_1 \mathbf{V}_1^T (\mathbf{\Sigma}_1 \mathbf{V}_1^T)^T \right)^{-1} \left( \mathbf{U}_1^T \mathbf{U}_1 \right)^{-1} \mathbf{U}_1^T = \mathbf{V}_1 \mathbf{\Sigma}_1 \mathbf{\Sigma}_1^{-2} \mathbf{U}_1^T \\ &= \mathbf{V}_1 \mathbf{\Sigma}_1^{-1} \mathbf{U}_1^T. \end{aligned}$$

Ebből következik a másik egyenlőség is, mivel

$$\mathbf{V} \mathbf{\Sigma}^+ \mathbf{U}^T = \begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_2 \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_1^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U}_1^T \\ \mathbf{U}_2^T \end{bmatrix} = \mathbf{V}_1 \mathbf{\Sigma}_1^{-1} \mathbf{U}_1^T.$$

■ Komplex:  $\mathbf{A} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^H = \mathbf{U}_1 \mathbf{\Sigma}_1 \mathbf{V}_1^H \rightsquigarrow \mathbf{A}^+ = \mathbf{V} \mathbf{\Sigma}^+ \mathbf{U}^H = \mathbf{V}_1 \mathbf{\Sigma}_1^{-1} \mathbf{U}_1^H$ .

**P** Számítsuk ki a pszeudoinverzét!

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

**M** Az  $\mathbf{A}$  mátrix redukált szinguláris felbontása

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

amiből a pszeudoinverzét megadó  $\mathbf{A}^+ = \mathbf{V}_1 \mathbf{\Sigma}_1^{-1} \mathbf{U}_1^T$  képlettel

$$\mathbf{A}^+ = \frac{1}{3} \begin{bmatrix} 2 & 2 \\ -2 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} \frac{1}{6} & 0 \\ 0 & \frac{1}{3} \end{bmatrix} \frac{1}{3} \begin{bmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{9} & 0 & \frac{2}{9} \\ -\frac{1}{9} & \frac{1}{9} & 0 \\ \frac{1}{6} & -\frac{1}{9} & -\frac{1}{9} \end{bmatrix}.$$

# Frobenius-norma

**D** Az  $\mathbf{A} \in \mathbb{C}^{m \times n}$  mátrix **Frobenius-normája**

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

**T** Frobenius-norma ekvivalens alakjai

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

**B**  $\mathbf{A}^H \mathbf{A}$  átlójának  $j$ -edik eleme éppen  $|\mathbf{A}_{*j}|^2$

nyom = sajátértékek összege,

az  $\mathbf{A}^H \mathbf{A}$  sajátértékei =  $\mathbf{A}$  szinguláris értékeinek négyzetei

## Kis rangú approximáció tétele – Eckart–Young-tétel

**T** L!  $\mathbf{A}$   $r$ -rangú  $\mathbf{A}_k = \sum_{i=1}^k \sigma_i \mathbf{u}_i \mathbf{v}_i^T$ . Ekkor  $\mathbf{A}_k$  az  $\mathbf{A}$  mátrix legjobb legfölbbebb  $k$ -rangú közelítése, azaz

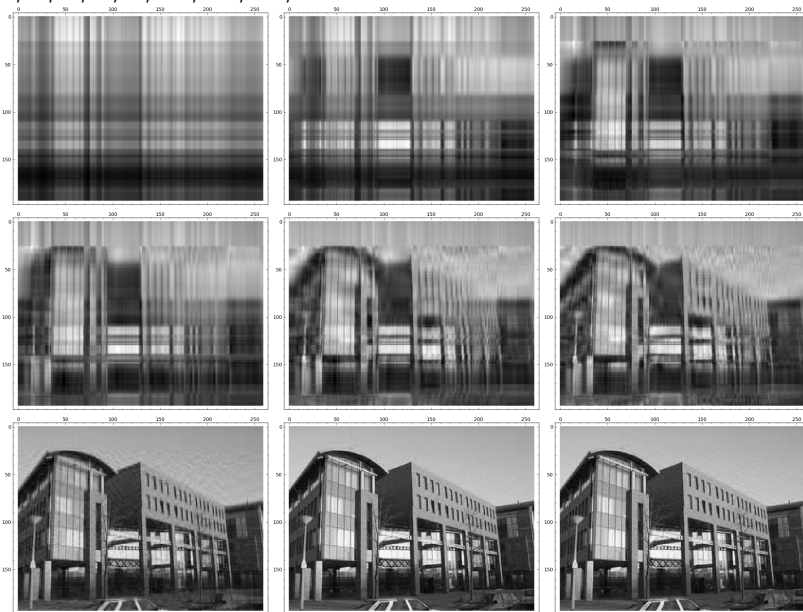
$$\min_{r(\mathbf{B}) \leq k} \|\mathbf{A} - \mathbf{B}\|_F = \|\mathbf{A} - \mathbf{A}_k\|_F = \sqrt{\sum_{i=k+1}^r \sigma_i^2}.$$

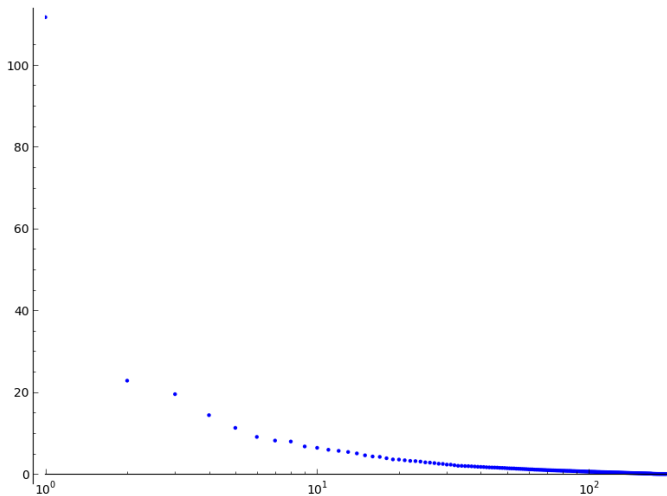
**B**  $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$  az SVD, és legyen  $\mathbf{C} = \mathbf{V}\mathbf{B}\mathbf{U}^T$ , azaz  $\mathbf{B} = \mathbf{U}\mathbf{C}\mathbf{V}^T$ ,  $\mathbf{C} = [c_{ij}]$ .

$$\begin{aligned} \|\mathbf{A} - \mathbf{B}\|_F^2 &= \|\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T - \mathbf{U}\mathbf{C}\mathbf{V}^T\|_F^2 = \|\mathbf{U}(\mathbf{\Sigma} - \mathbf{C})\mathbf{V}^T\|_F^2 \\ &= \text{trace} \left( (\mathbf{U}(\mathbf{\Sigma} - \mathbf{C})\mathbf{V}^T)^T (\mathbf{U}(\mathbf{\Sigma} - \mathbf{C})\mathbf{V}^T) \right) \\ &= \text{trace} \left( \mathbf{V}(\mathbf{\Sigma} - \mathbf{C})^T (\mathbf{\Sigma} - \mathbf{C})\mathbf{V}^T \right) = \text{trace} \left( (\mathbf{\Sigma} - \mathbf{C})^T (\mathbf{\Sigma} - \mathbf{C}) \right) \\ &= \|\mathbf{\Sigma} - \mathbf{C}\|_F^2 = \sum_i |\sigma_i - c_{ii}|^2 + \sum_{i \neq j} |c_{ij}|^2 \end{aligned}$$



1, 2, 3, 4, 8, 12, 40, 97, 194.





**A** latent semantic indexing (LSI) vagy latent semantic analysis (LSA)

**Ö** Az egy dokumentumban szereplő szavakat összekapcsolja a dokumentum tartalma, e kapcsolatokat – a szavak mögött lévő tartalmat – az SVD kiemeli, mint lényeges információt.

**A** sorai a szavakat, oszlopai a dokumentumokat reprezentálják.

$t_{ij}$  az  $i$ -edik szó gyakorisága a  $j$ -edik dokumentumban és  $T_i$  a teljes szöveggyűjteményben

$$a_{ij} = \left( 1 + \sum_{k=1}^n \frac{\frac{t_{ik}}{T_i} \log \frac{t_{ik}}{T_i}}{\log n} \right) \log(1 + t_{ij}).$$

az első tényező egy csak az  $i$ -edik szónak az egész gyűjteményhez való kapcsolatától függő globális súly, a második csak a lokális érték függvénye.

$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$  és közelítése  $\mathbf{A}_k = \mathbf{U}_k\mathbf{\Sigma}_k\mathbf{V}_k^T$

$\mathbf{U}_k$ , illetve  $\mathbf{V}_k$  oszlopainak vektorterében a szavak, illetve dokumentumok kapcsolatát a hozzájuk tartozó vektorok helyzete jellemzi