BUDAPESTI MŰSZAKI
MATEMATIKA
ÉS GAZDASÁGTUDOMÁNYI
INTÉZET
EGYETEM

Bevezetés az algebrába 2

BMETE91AM37

Szinguláris érték, SVD

H607 – 2018-04-16

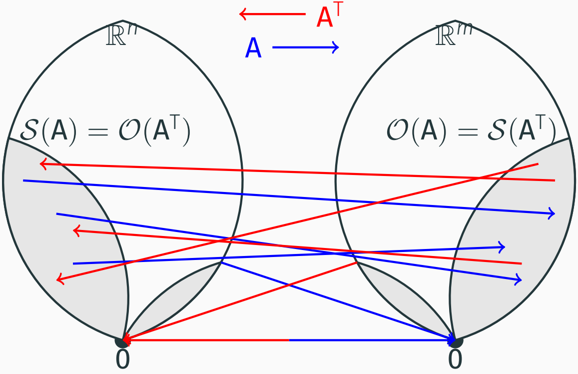
Wettl Ferenc

ALGEBRA TANSZÉK

SVD

Keressünk két ONB-t

- K Hogyan válasszunk bázist, ha a mátrix téglalap alakú? Mit tehetünk még, ha a mátrix nem diagonalizálható?
- m A $\mathcal{V} \to \mathcal{V}$ lineáris transzformációk ortogonális diagonalizálását fogjuk általánosítani $\mathcal{V}_1 \to \mathcal{V}_2$ lineáris leképezésekre.



Ötlet

m Az ortogonális diagonalizációt igyekszünk megvalósítani a kitüntetett alterek ONB-aival, azzal az „apró” módosítással, hogy a sortér bázisának minden vektora az oszloptér bázisának pozitív konstansszorosába menjen (a nulltérbeli bázisvektorok a $\mathbf{0}$-ba).

m Az $\mathbf{A} \in \mathbb{R}^{m \times n}$ mátrixhoz olyan $\{\mathbf{v}_1, \mathbf{v}_2 \ldots, \mathbf{v}_n\} \subset \mathbb{R}^n$ és $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \subset \mathbb{R}^m$ ONB-okat keresünk, melyekben $\mathbf{A}$ mátrixa diagonális.

m $\mathbf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i$, ahol $1 \leqslant i \leqslant \min(m, n)$

Á Ha az $\mathbb{R}^n$-beli $\mathbf{x} \perp \mathbf{y}$ vektorok legalább egyike az $\mathbf{A}^\mathsf{T}\mathbf{A}$ sajátvektora, akkor $\mathbf{A}\mathbf{x} \perp \mathbf{A}\mathbf{y}$

B $\mathbf{x} \cdot \mathbf{y} = 0$, $\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{y} = \lambda \mathbf{y} \rightsquigarrow$

$\mathbf{A}\mathbf{x} \cdot \mathbf{A}\mathbf{y} = (\mathbf{A}\mathbf{x})^\mathsf{T}\mathbf{A}\mathbf{y} = \mathbf{x}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{y} = \mathbf{x} \cdot (\lambda \mathbf{y}) = \lambda(\mathbf{x} \cdot \mathbf{y}) = 0$

Komplexben $\mathbf{A}^\mathsf{H}\mathbf{A}$-val!

Szinguláris érték, szinguláris vektor

- $A\colon \mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ 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 ONB-a $\rightsquigarrow$ $\mathbf{A}\mathbf{v}_i$ OB az oszloptérben
- $|\mathbf{A}\mathbf{v}_i|^2 = \mathbf{v}_i^T\mathbf{A}^T\mathbf{A}\mathbf{v}_i = \lambda_i|\mathbf{v}_i|^2 = \lambda_i$, $\sigma_i = \sqrt{\lambda_i}$, $\mathbf{u}_i = \frac{\mathbf{A}\mathbf{v}_i}{\sigma_i}$ $\rightsquigarrow$ $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ ONB

D Az r rangú $\mathbf{A} \in \mathbb{R}^{m\times n}$ mátrix szinguláris értékeinek nevezzük az $\mathbf{A}^T\mathbf{A}$ pozitív sajátértékeinek négyzetgyökeit ($\sigma_1 \geqslant \sigma_2 \geqslant \ldots \geqslant \sigma_r > 0$). Az $\mathbf{A}^T\mathbf{A}$ mátrix valamely σ^2 sajátértékéhez tartozó egységnyi $\mathbf{v}$ sajátvektorát az $\mathbf{A}$ jobb szinguláris vektorának, az $\mathbf{u} = \frac{1}{\sigma}\mathbf{A}\mathbf{v}$ egységvektort pedig a $\mathbf{v}$-hez tartozó bal szinguláris vektorának nevezzük.

- $|\mathbf{u}_i| = 1$ $(i = 1, 2, \ldots, r) \rightsquigarrow |\mathsf{A}\mathbf{v}_i| = \sigma_i$
- úgy is definiálható, hogy ha $k = \min(m, n)$, akkor a szinguláris értékek száma k, és értékük nem negatív, így $\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_k = 0$.
- Ha σ t-szeres multiplicitású szinguláris értéke A-nak, akkor σ^2 az $\mathsf{A}^\mathsf{T}\mathsf{A}$-nak t-szeres algebrai (és t-szeres geometriai) multiplicitású sajátértéke, így a σ-hoz tartozó jobb szinguláris vektorok a sortérben, a hozzá tartozó bal szinguláris vektorok pedig az oszloptérben t-dimenziós alteret feszítenek ki.
- Ha $\mathsf{A}\mathbf{v} = \sigma\mathbf{u}$, akkor $\mathsf{A}^\mathsf{T}\mathbf{u} = \frac{1}{\sigma}\mathsf{A}^\mathsf{T}\mathsf{A}\mathbf{v} = \frac{1}{\sigma}\sigma^2\mathbf{v} = \sigma\mathbf{v}$, tehát az
$$\mathsf{A}\mathbf{v} = \sigma\mathbf{u}, \quad \mathsf{A}^\mathsf{T}\mathbf{u} = \sigma\mathbf{v} \tag{1}$$
összefüggések párba állítják a jobb és bal szinguláris vektorokat.
- Ha $\mathsf{A} \in \mathbb{C}^{m\times n}$, akkor a sortér helyett a sortér konjugáltjával, azaz az $\mathcal{O}(\mathsf{A}^\mathsf{H})$ altérrel, transzponálás helyett adjungálással, tehát $\mathsf{A}^\mathsf{T}\mathsf{A}$ helyett az $\mathsf{A}^\mathsf{H}\mathsf{A}$ mátrixszal igazak maradnak a fentiek.

P Határozzuk meg

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

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

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

$\mathsf{A}^\mathsf{T}\mathsf{A}$ 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

- $\Sigma_1 = \operatorname{diag}(\sigma_1, \dots, \sigma_r)$ $\mathbf{U}_1 = [\mathbf{u}_1|\dots|\mathbf{u}_r]$, $\mathbf{V}_1 = [\mathbf{v}_1|\dots|\mathbf{v}_r]$
  $\mathbf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i \rightsquigarrow \mathbf{A}\mathbf{V}_1 = \mathbf{U}_1\Sigma_1$, azaz

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

$\rightsquigarrow$ $\mathbf{A} = \mathbf{U}_1\Sigma_1\mathbf{V}_1^T$ (mert $\mathbf{V}_1$ oszlopvektorai $\mathcal{S}(\mathbf{A})$ ONB-a, így $\mathbf{A}\mathbf{V}_1\mathbf{V}_1^T = \mathbf{A}$)

D Redukált szinguláris felbontás: $\mathbf{A} = \mathbf{U}_1\Sigma_1\mathbf{V}_1^T$ ($\mathbf{A} = \mathbf{U}_1\Sigma_1\mathbf{V}_1^H$), ahol $\mathbf{V}_1$, $\mathbf{U}_1$ szemiortogonális („szemiunitér”), Σ_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 $\mathbf{V}_2 = \begin{bmatrix} \mathbf{v}_{r+1} & \dots & \mathbf{v}_n \end{bmatrix}$, $\mathbf{V} = [\mathbf{V}_1|\mathbf{V}_2]$
  $\mathbf{V}_2$ oszlopvektorai kifeszítik a nullteret $\rightsquigarrow$ $r < i \leqslant n$ esetén $\mathbf{A}\mathbf{v}_i = \mathbf{0}$
  $\rightsquigarrow$ $\mathbf{A}\mathbf{V} = [\mathbf{U}_1\Sigma_1|\mathbf{O}]$
- Egészítsük ki $\mathbf{U}_1$-et ortogonális $\mathbf{U}$ mátrixszá: $m > r$ esetén $\mathbf{U}_2 = \begin{bmatrix} \mathbf{u}_{r+1} & \dots & \mathbf{u}_m \end{bmatrix}$, $\mathbf{U} = [\mathbf{U}_1|\mathbf{U}_2]$
- Egészítsük ki Σ_1-et $m \times n$-es mátrixszá: $\Sigma = \left[\begin{smallmatrix} \Sigma_1 & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{smallmatrix}\right]$

$\rightsquigarrow$ $\mathbf{A}\mathbf{V} = \mathbf{U}\Sigma$, ugyanis $\mathbf{A}\mathbf{V} = \mathbf{A}[\mathbf{V}_1|\mathbf{V}_2] = [\mathbf{U}_1\Sigma_1|\mathbf{O}] = [\mathbf{U}_1|\mathbf{U}_2]\left[\begin{smallmatrix} \Sigma_1 & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{smallmatrix}\right] = \mathbf{U}\Sigma$

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

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

Létezés és egyértelműség

T Minden $\mathsf{A} \in \mathbb{R}^{m\times n}$ ($\in \mathbb{C}^{m\times n}$) 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 $\mathsf{A}_{m\times n} = \mathsf{U}_{m\times m}\Sigma_{m\times n}\mathsf{V}^{\mathsf{T}}_{n\times n}$

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

$\rightsquigarrow$ ortog/unitéren diag-ható és a diag. elemek $\geqslant 0$ valósok

$\rightsquigarrow$ sajátvektorokból $\to$ ONB $\to$ V ortogonális/unitér.
(AV nemnulla oszlopvektorai) $\cup$ ($\mathcal{N}(\mathsf{A}^{\mathsf{T}})$ egy ONB-a) $\to$ U

$\rightsquigarrow$ $\exists\ \mathsf{A} = \mathsf{U}\Sigma\mathsf{V}^{\mathsf{T}}$ ($\mathsf{A} = \mathsf{U}\Sigma\mathsf{V}^{\mathsf{H}}$), Σ-ben a sajátértékek négyzetgyökei.

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

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

$\rightsquigarrow$ $\Sigma^{\mathsf{H}}\Sigma = \Sigma^{\mathsf{T}}\Sigma = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2, 0, \dots, 0)$ és a sajátértékek egyértelműek, így a szinguláris értékek is.

SVD kiszámítása

P Számítsuk ki az

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

mátrix szinguláris értékeit, és írjuk fel szinguláris felbontását!

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

$\chi(x) = x^3 - 45x^2 + 324x$, gyökei 36, 9, 0 $\rightsquigarrow$ a szing. ért.: 6, 3 (0)

- $A^T A$ egységnyi sajátvektorai:

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

- $\mathsf{Av}_i = \sigma_i \mathbf{u}_i \rightsquigarrow \mathbf{u}_i = \mathsf{Av}_i/\sigma_i \rightsquigarrow \mathbf{u}_1 = \dfrac{\mathsf{Av}_1}{\sigma_1} = \dfrac{(2,-4,4)}{6} = \left(\dfrac{1}{3}, -\dfrac{2}{3}, \dfrac{2}{3}\right)$

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

- $\mathbf{u}_3$ így nem számolható ($\mathsf{Av}_3 = \mathbf{0}$), így az $\{\mathbf{u}_1, \mathbf{u}_2\}$ rendszert nekünk kell kiegészítenünk $\mathbb{R}^3$ bázisává.
- $\mathcal{O}(\mathsf{A})^\perp$ ONB-át keressük:
- $\operatorname{span}(\mathbf{u}_1, \mathbf{u}_2)^\perp$: $\begin{bmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -2 \end{bmatrix} \to \mathbf{u}_3 = \frac{1}{3} \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$
- $\mathcal{O}(\mathsf{A})^\perp$: $\begin{bmatrix} 0 & -2 & 4 \\ -2 & 3 & -2 \\ 2 & -2 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -2 \end{bmatrix} \to \mathbf{u}_3 = \frac{1}{3} \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$
- $\mathbf{u}_3 = \mathbf{u}_1 \times \mathbf{u}_2 = (-2/3, -2/3, -1/3)$.

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

- 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:

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

$$\left(= \begin{bmatrix} 4/3 & -4/3 & 2/3 \\ -8/3 & 8/3 & -4/3 \\ 8/3 & -8/3 & 4/3 \end{bmatrix} + \begin{bmatrix} -4/3 & -2/3 & 4/3 \\ 2/3 & 1/3 & -2/3 \\ 4/3 & 2/3 & -4/3 \end{bmatrix}. \right)$$

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

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

A^TA és AA^T

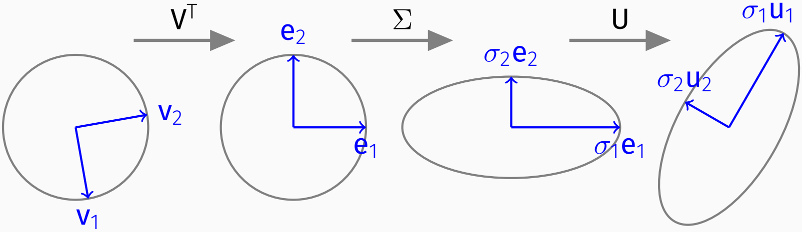
m A számolás kezdhető U meghatározásával is. A^TA és AA^T nem zérus sajátértékei megegyeznek,

$$AA^T = U\Sigma V^T(U\Sigma V^T)^T = U\Sigma V^T V\Sigma^T U^T = U\Sigma\Sigma^T U^T.$$

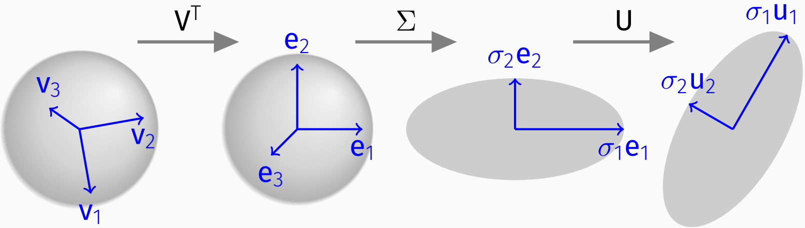
$A^T u_i = \sigma_i v_i$, azaz $v_i = \dfrac{A^T u_i}{\sigma_i}$.

- érdemes AA^T-t használni, ha $m < n$, mert ekkor csak m-dimenziós vektorokkal kell számolni.

- L! $A = U\Sigma V^T$ 2×2-es 2-rangú



- L! $A = U\Sigma V^T$ 2×3-as 2-rangú



Egységgömb képe

T $\mathbf{A} \in \mathbb{R}^{m\times n}$ r-rangú, az $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ leképezés az $\mathbf{e}^\mathsf{T}\mathbf{e} = 1$ 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:

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

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

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

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

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

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

  Eszerint az egyenlet

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

P Határozzuk meg a következő mátrix teljes és redukált szinguláris felbontását!

$$\mathsf{A} = \begin{bmatrix} 1 & 1-\mathrm{i} \\ 1+\mathrm{i} & 2 \end{bmatrix}$$

M A mátrix első sorát $1+\mathrm{i}$-vel szorozva a második sort kapjuk, tehát $\mathcal{S}(\mathsf{A}) = \operatorname{span}((1, 1-\mathrm{i})) \rightsquigarrow \mathcal{O}(\mathsf{A}^{\mathsf{H}}) = \overline{\mathcal{S}(\mathsf{A})} = \operatorname{span}((1, 1+\mathrm{i})) \rightsquigarrow$ $\mathbf{v}_1 = \frac{1}{\sqrt{3}}(1, 1+\mathrm{i}) \rightsquigarrow$ felhasználva az $\mathsf{A}\mathbf{v}_1 = \sigma_1 \mathbf{u}_1$ képletet $\sigma_1 = 3$, ui.

$$\begin{bmatrix} 1 & 1-\mathrm{i} \\ 1+\mathrm{i} & 2 \end{bmatrix} \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1+\mathrm{i} \end{bmatrix} = \frac{1}{\sqrt{3}} \begin{bmatrix} 3 \\ 3+3\mathrm{i} \end{bmatrix} = 3 \left(\frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1+\mathrm{i} \end{bmatrix} \right) = 3\mathbf{u}_1$$

- a merőleges egységvektorok: $\mathbf{v}_2 = \mathbf{u}_2 = \frac{1}{\sqrt{3}}(-1+\mathrm{i}, 1)$.
- SVD: $\mathsf{A} = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1+\mathrm{i}}{\sqrt{3}} \end{bmatrix} 3 \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1-\mathrm{i}}{\sqrt{3}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{-1+\mathrm{i}}{\sqrt{3}} \\ \frac{1+\mathrm{i}}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1-\mathrm{i}}{\sqrt{3}} \\ \frac{-1-\mathrm{i}}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}$

Ha a mátrix normális

Á Ha a mátrix normális (szimmetrikus), az unitér (ortogonális) diagonalizálásából a szinguláris felbontás megkapható. Ha a mátrix pozitív szemidefinit, akkor van olyan sajátfelbontása, mely egyúttal szinguláris felbontás is!

P Határozzuk meg az $\mathbf{A} = \begin{bmatrix} 1 & 1-\mathrm{i} \\ 1+\mathrm{i} & 0 \end{bmatrix}$ szinguláris felbontását!

M $\mathbf{A}$ sajátpárjai: $(2, \frac{1}{\sqrt{3}}(1-\mathrm{i}, 1))$, $(-1, \frac{1}{\sqrt{6}}(1-\mathrm{i}, -2)) \rightsquigarrow$ sajátfelbontás:

$$\begin{bmatrix} 1 & 1-\mathrm{i} \\ 1+\mathrm{i} & 0 \end{bmatrix} = \begin{bmatrix} \frac{1-\mathrm{i}}{\sqrt{3}} & \frac{1-\mathrm{i}}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{-2}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \frac{1+\mathrm{i}}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1+\mathrm{i}}{\sqrt{6}} & \frac{-2}{\sqrt{6}} \end{bmatrix}$$

- A negatív sajátértéket és egy hozzá tartozó sajátvektort -1-gyel szorozva SVD-t kapunk:

$$\begin{bmatrix} 1 & 1-\mathrm{i} \\ 1+\mathrm{i} & 0 \end{bmatrix} = \begin{bmatrix} \frac{1-\mathrm{i}}{\sqrt{3}} & \frac{-1+\mathrm{i}}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1+\mathrm{i}}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1+\mathrm{i}}{\sqrt{6}} & \frac{-2}{\sqrt{6}} \end{bmatrix}$$

SVD

Polárfelbontás, pszeudoinverz

Polárfelbontás

m $r\mathrm{e}^{\mathrm{i}\varphi}$: egy nemnegatív nyújtási tényező (r) és egy egységnyi abszolút értékű komplex szám ($\mathrm{e}^{\mathrm{i}\varphi}$, ami a komplex síkon φ-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

$$\mathsf{A} = \mathsf{PQ}$$

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

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

- $\mathbf{P}$ önadjungált, hisz $(\mathbf{U}\Sigma\mathbf{U}^H)^H = \mathbf{U}\Sigma^H\mathbf{U}^H = \mathbf{U}\Sigma\mathbf{U}^H$.
- A $\mathbf{P}$ pozitív szemidefinit, hisz hasonló a pozitív szemidefinit Σ mátrixhoz (ha $\mathbf{A}$ invertálható, akkor Σ pozitív definit)
- $\mathbf{Q}$ unitér (ortogonális), hisz két unitér (ortogonális) mátrix szorzata.
- 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 determinánson is: $\det \mathsf{P} = r$, $\det \mathsf{Q} = \mathrm{e}^{\mathrm{i}\varphi}$ (hisz Q unitér), $\det \mathsf{A} = r\mathrm{e}^{\mathrm{i}\varphi}$.
- Fordított sorrend: azonos unitér (ortogonális) mátrixszal:

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

- Valós polárfelbontás geometriai jelentése: minden mátrixleképezés két olyan leképezés kompozíciója, amelyekből az egyik forgatja vagy forgatva tükrözi a teret (Q), a másik pedig egy ortonormált bázis tengelyei mentén nyújtja/összenyomja a teret minden tengelyirányban egy-egy nemnegatív tényezővel.

P Polárfelbontásait számítsuk ki

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

B $P = U\Sigma U^\top$, $Q = UV^\top$, $\hat{P} = V\Sigma V^\top$

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

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

E példában könnyen találunk egy másik polárfelbontást is:

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

Pszeudoinverz

T A valós mátrix és $\mathsf{A} = \mathsf{U}_1\Sigma_1\mathsf{V}_1^\mathsf{T} = \mathsf{U}\Sigma\mathsf{V}^\mathsf{T}$. Ekkor

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

B Az $\mathsf{A} = \mathsf{U}_1(\Sigma_1\mathsf{V}_1^\mathsf{T})$ felbontásban U_1 teljes oszloprangú, $\Sigma_1\mathsf{V}_1^\mathsf{T}$ teljes sorrangú:

$$\begin{aligned}\mathsf{A}^+ &= (\Sigma_1\mathsf{V}_1^\mathsf{T})^\mathsf{T}\left(\Sigma_1\mathsf{V}_1^\mathsf{T}(\Sigma_1\mathsf{V}_1^\mathsf{T})^\mathsf{T}\right)^{-1}\left(\mathsf{U}_1^\mathsf{T}\mathsf{U}_1\right)^{-1}\mathsf{U}_1^\mathsf{T} = \mathsf{V}_1\Sigma_1\Sigma_1^{-2}\mathsf{U}_1^\mathsf{T}\\ &= \mathsf{V}_1\Sigma_1^{-1}\mathsf{U}_1^\mathsf{T}.\end{aligned}$$

Ebből következik a másik egyenlőség is, mivel

$$\mathsf{V}\Sigma^+\mathsf{U}^\mathsf{T} = \begin{bmatrix}\mathsf{V}_1 & \mathsf{V}_2\end{bmatrix}\begin{bmatrix}\Sigma_1^{-1} & \mathsf{O}\\ \mathsf{O} & \mathsf{O}\end{bmatrix}\begin{bmatrix}\mathsf{U}_1^\mathsf{T}\\ \mathsf{U}_2^\mathsf{T}\end{bmatrix} = \mathsf{V}_1\Sigma_1^{-1}\mathsf{U}_1^\mathsf{T}.$$

- Komplex: $\mathsf{A} = \mathsf{U}\Sigma\mathsf{V}^\mathsf{H} = \mathsf{U}_1\Sigma_1\mathsf{V}_1^\mathsf{H} \rightsquigarrow \mathsf{A}^+ = \mathsf{V}\Sigma^+\mathsf{U}^\mathsf{H} = \mathsf{V}_1\Sigma_1^{-1}\mathsf{U}_1^\mathsf{H}$.

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 pszeudoinverzet megadó $\mathbf{A}^+ = \mathbf{V}_1\Sigma_1^{-1}\mathbf{U}_1^T$ képlettel

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

SVD

Alkalmazások

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

$$\|\mathbf{A}\|_F = \sqrt{\operatorname{trace}(\mathbf{A}^{\mathsf{H}}\mathbf{A})} = \sqrt{\sum_{i=1}^{\mathrm{r}(\mathbf{A})} \sigma_i^2}. \tag{2}$$

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

nyom = sajátértékek összege,

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

Kis rangú approximáció tétele – Eckart–Young-tétel

T L! 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 legföljebb k-rangú közelítése, azaz

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

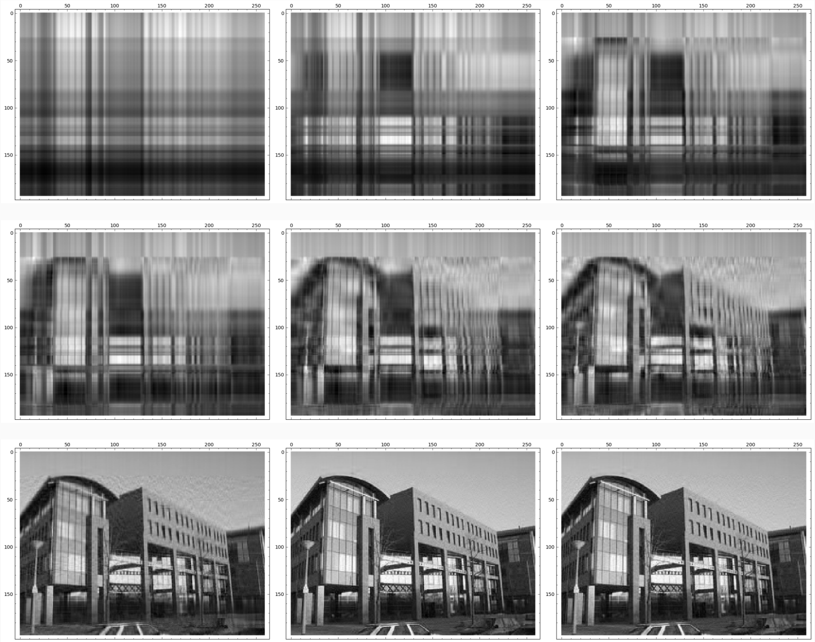
B (később)

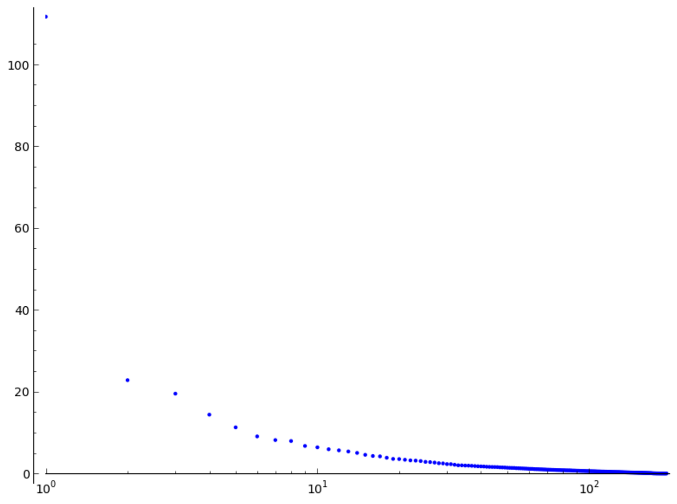
P Melyik 1-rangú mátrix van Frobenius-normában a legközelebb a $\mathbf{A} = \left[\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right]$ mátrixhoz?

M $\mathbf{A}$ szimmetrikus, pozitív definit $\rightsquigarrow$ szinguláris felbontása = sajátfelbontásával. Sajátpárok $(3, (1, 1))$, $(1, (1, -1))$ $\rightsquigarrow$ a szinguláris felbontás: $\mathbf{A} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \rightsquigarrow$

1-rangú közelítés $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} [3] \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{3}{2} & \frac{3}{2} \\ \frac{3}{2} & \frac{3}{2} \end{bmatrix}$

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




100
80
60
40
20
0
10^0
10^1
10^2

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.

$\mathbf{A}$ 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.

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

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