BUDAPESTI MŰSZAKI MATEMATIKA ÉS GAZDASÁGTUDOMÁNYI INTÉZET EGYETEM

Haladó lineáris algebra

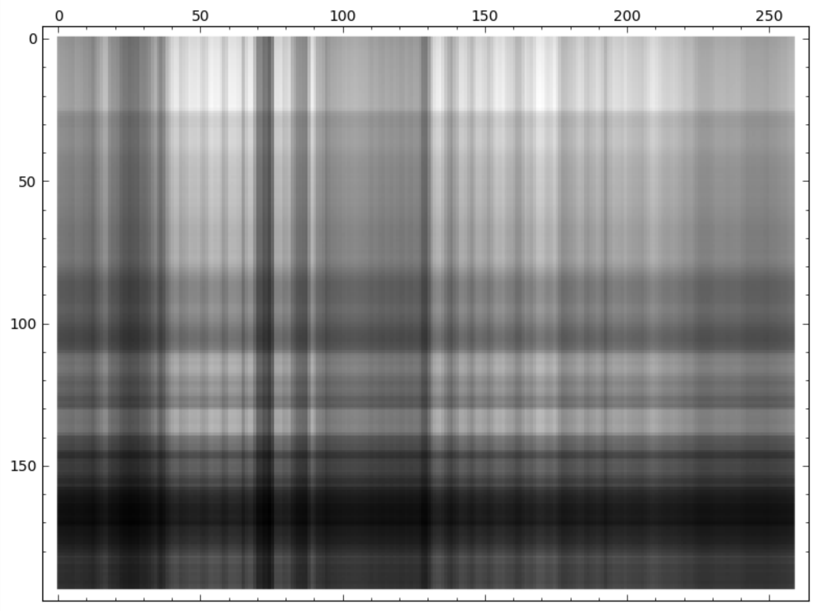
BMETE90MX54 (FELSŐBB MATEMATIKA VILLAMOSMÉRNÖKÖKNEK)

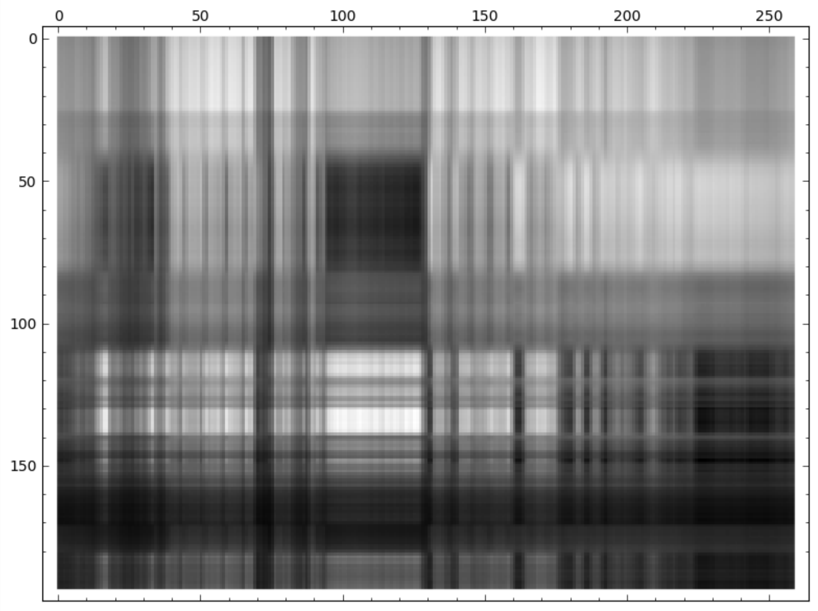
Szinguláris értékek

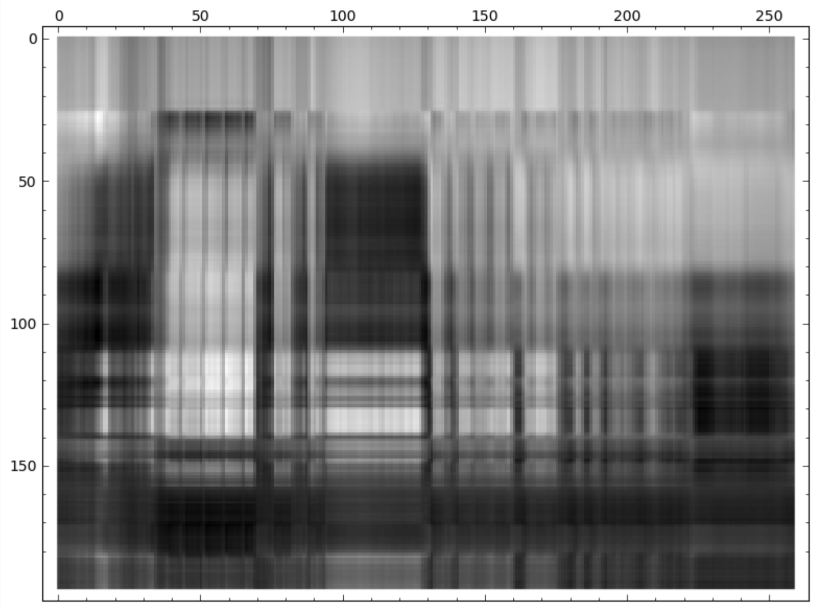
SZINGULÁRIS ÉRTÉK SZERINTI FELBONTÁS, VEKTOR- ÉS MÁTRIXNORMA

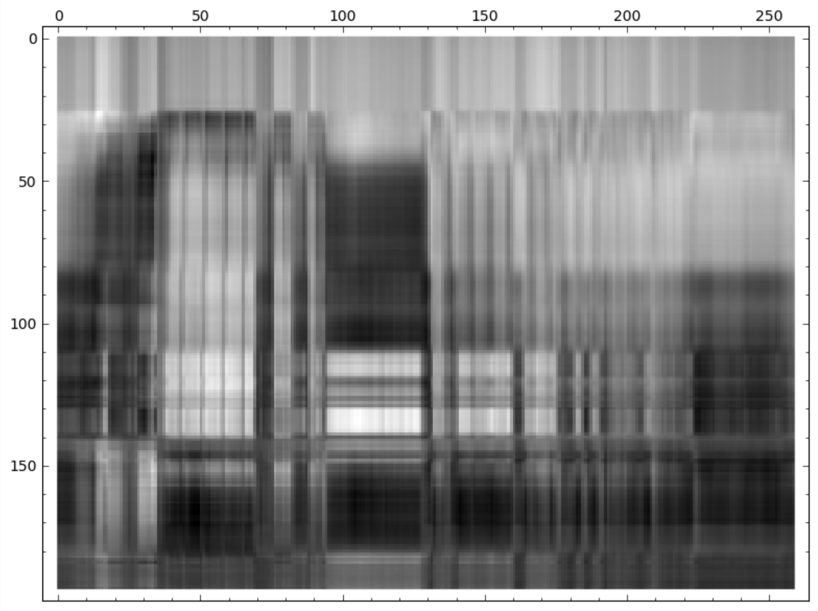
Wettl Ferenc

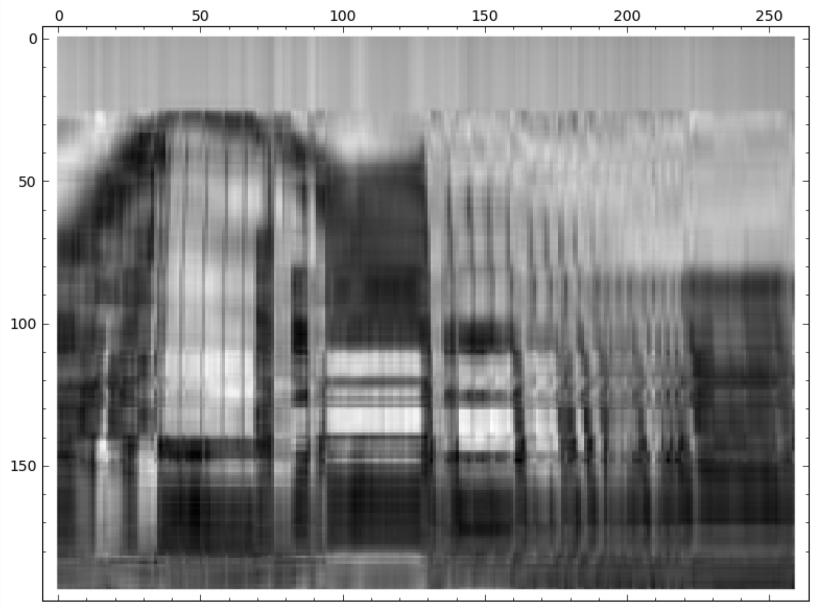
ALGEBRA TANSZÉK

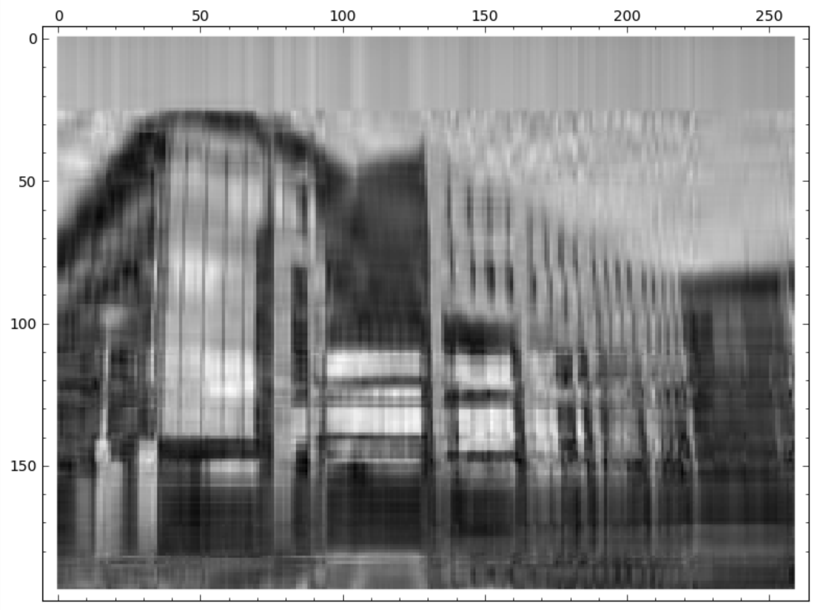

0
50
100
150
200
250
0
50
100
150


0
50
100
150
200
250
0
50
100
150


0
50
100
150
200
250
0
50
100
150


0
50
100
150
200
250
0
50
100
150


0
50
100
150
200
250
0
50
100
150


0
50
100
150
200
250
0
50
100
150


0
50
100
150
200
250
0
50
100
150


0
50
100
150
200
250
0
50
100
150


0
50
100
150
200
250
0
50
100
150

Tartalom

Ismeretek, képességek, célok

- Szinguláris érték, jobb és bal szinguláris vektor meghatározása, szinguláris felbontás és geometriai interpretációja
- SVD alkalmazásai: pszeudoinverz, polárfelbontás, kis rangú approximáció (Eckart–Young-tétel)
- Norma, mátrixnorma, indukált norma

Szinguláris érték, SVD

Az ortogonális diagonalizáció újraértelmezése

- A négyzetes valós szimmetrikus (komplex normális) mátrixok ortogonális (unitér) diagonalizálását újraértelmezve általánosítjuk tetszőleges valós (komplex) mátrixra:
    - sajátértékek $\twoheadrightarrow$ szinguláris értékek
    - sajátfelbontás $\twoheadrightarrow$ szinguláris érték szerinti felbontás (SVD)
- A feltételeken lazítunk: egy helyett két ortonormált bázis
    - $\mathcal{V} \to \mathcal{V}$ lineáris trafó $\twoheadrightarrow$ $\mathcal{V}_1 \to \mathcal{V}_2$ lineáris leképezés
    - $\mathcal{V}$ ONB $\twoheadrightarrow$ $\mathcal{V}_1$ és $\mathcal{V}_2$ egy-egy ONB
- alkalmazásai: statisztika, információtömörítés, képfeldolgozás (főkomponensanalízis képekre: eigenfaces), videomásolat igazolás (azonos kép–különböző forrás, különböző kép–azonos forrás), LCD hibakeresés, szkennelt 3D pontok alapján felületi normális pl. árnyékoláshoz, animáció generálása, az egy.rsz.-ek megoldásához használt leghatékonyabbak algoritmusok,...

Diagonalizálás két ONB-ban

m Az $\mathbf{A} \in \mathbb{R}^{m\times n}$ mátrixhoz olyan ortonormált $\{\mathbf{v}_1, \mathbf{v}_2 \ldots, \mathbf{v}_n\} \subset \mathbb{R}^n$ és $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \subset \mathbb{R}^m$ ortonormált bázisokat keresünk, melyekben $\mathbf{A}$ mátrixa diagonálissá válik. Ez azt jelenti, hogy léteznek olyan σ_i valósok, hogy $\mathbf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i$, ahol $1 \leqslant i \leqslant \min(m, n)$.

Á Ha $\mathbf{A} \in \mathbb{R}^{m\times n}$ és az egymásra merőleges $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ vektorok legalább egyike az $\mathbf{A}^\mathsf{T}\mathbf{A}$ Gram-mátrix sajátvektora, akkor az $\mathbf{A}\mathbf{x}, \mathbf{A}\mathbf{y} \in \mathbb{R}^m$ vektorok is merőlegesek egymásra.

B A feltételek szerint $\mathbf{x} \cdot \mathbf{y} = 0$, és legyen például $\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{y} = \lambda \mathbf{y}$. Ekkor

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

\- komplexben: $\mathbf{A} \in \mathbb{C}^{m\times n}$, $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, $\mathbf{A}^\mathsf{H}\mathbf{A}$ sajátvektora,

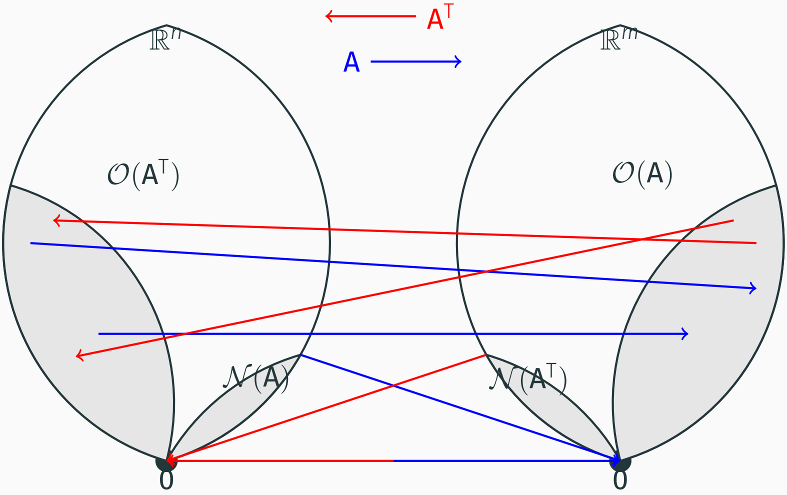
$$\mathbf{A}\mathbf{x} \cdot \mathbf{A}\mathbf{y} = (\mathbf{A}\mathbf{x})^\mathsf{H}\mathbf{A}\mathbf{y} = \mathbf{x}^\mathsf{H}\mathbf{A}^\mathsf{H}\mathbf{A}\mathbf{y} = \mathbf{x} \cdot (\lambda \mathbf{y}) = \lambda(\mathbf{x} \cdot \mathbf{y}) = 0.$$

- $\mathbf{A}^T\mathbf{A}$ szimmetrikus ($\mathbf{A}^H\mathbf{A}$ önadjungált), pozitív szemidefinit $\rightsquigarrow$ ortogonálisan (unitéren) diag-ható, sajátértékei nem negatívak
- A s.v.-okból kiválasztható $\mathbb{R}^n$ egy ONB-a, a 0 s.é.-hez tartozók a $\mathcal{N}(\mathbf{A}^T\mathbf{A}) = \mathcal{N}(\mathbf{A})$ tér ONB-át adják, $\mathcal{N}(\mathbf{A}) = \{\mathbf{0}\}$ esetén a bázis $\emptyset$.
- a többi $\mathcal{N}(\mathbf{A})^\perp = \mathcal{O}(\mathbf{A}^T)$ ONB-a: $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$, ahol r az $\mathbf{A}$ rangja
- ezek $\mathbf{A}$ általi képei páronként merőlegesek egymásra $\rightsquigarrow$ az $\mathbf{A}\mathbf{v}_i$ vektorok ortogonális bázist alkotnak az oszloptérben.
- ha $\mathbf{v}_i$ az $\mathbf{A}^T\mathbf{A}$ mátrix $\lambda_i > 0$ sé-hez tartozó egységnyi sv-a, akkor $\|\mathbf{A}\mathbf{v}_i\| = \sqrt{\lambda_i}$, ui. $\|\mathbf{A}\mathbf{v}_i\|^2 = (\mathbf{A}\mathbf{v}_i)^T(\mathbf{A}\mathbf{v}_i) = \mathbf{v}_i^T\mathbf{A}^T\mathbf{A}\mathbf{v}_i = \lambda_i\|\mathbf{v}_i\|^2 = \lambda_i$.
- L! $\sigma_i = \sqrt{\lambda_i}$, így $\mathbf{A}\mathbf{v}_i = \sigma_i\mathbf{u}_i$, $\|\mathbf{u}_i\| = 1 \rightsquigarrow \{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ ONB $\mathcal{O}(\mathbf{A})$-ban
- $\mathbf{A}\mathbf{v}_i = \sigma_i\mathbf{u}_i$, $\mathbf{A}^T\mathbf{A}\mathbf{v}_i = \sigma_i^2\mathbf{v}_i \rightsquigarrow \mathbf{A}^T(\sigma_i\mathbf{u}_i) = \sigma_i^2\mathbf{v}_i \rightsquigarrow$

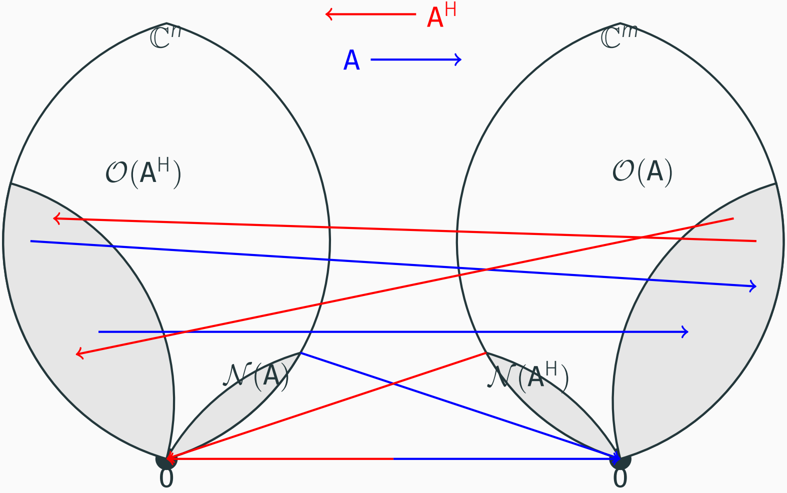
  $$\boxed{\mathbf{A}\mathbf{v}_i = \sigma_i\mathbf{u}_i,\ \mathbf{A}^T\mathbf{u}_i = \sigma_i\mathbf{v}_i}$$

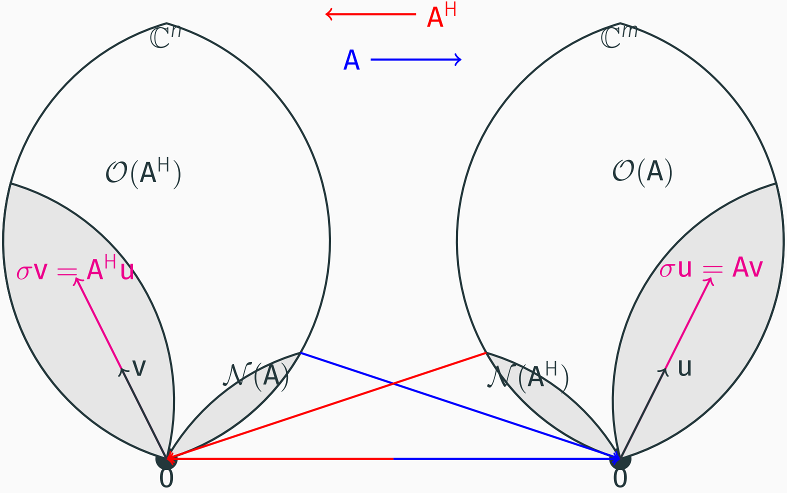
  párba állítják a $\mathbf{v}_i$ és $\mathbf{u}_i$ vektorokat
- $\mathbf{v} \in \mathcal{N}(\mathbf{A})$, $\mathbf{u} \in \mathcal{N}(\mathbf{A}^T)$, $\sigma = 0$: $\boxed{\mathbf{A}\mathbf{v} = \sigma\mathbf{u},\ \mathbf{A}^T\mathbf{u} = \sigma\mathbf{v}}$

A négy kitüntetett altérben



A négy kitüntetett altérben




A^H
A
$\mathbb{C}^n$
$\mathbb{C}^m$
$\mathcal{O}(\mathsf{A}^\mathsf{H})$
$\mathcal{O}(\mathsf{A})$
$\sigma\mathsf{v} = \mathsf{A}^\mathsf{H}\mathsf{u}$
$\sigma\mathsf{u} = \mathsf{A}\mathsf{v}$
v
u
$\mathcal{N}(\mathsf{A})$
$\mathcal{N}(\mathsf{A}^\mathsf{H})$
0
0

Szinguláris értékek és vektorok: $(\mathbf{u}, \sigma, \mathbf{v})$

D Az r rangú $\mathbf{A} \in \mathbb{K}^{m\times n}$ mátrix ($\mathbb{K} = \mathbb{R}$ vagy $\mathbb{C}$) **szinguláris értékének** nevezzük azt a nemnegatív σ valóst, melyhez van olyan nemzérus $\mathbf{v} \in \mathbb{K}^n$ és $\mathbf{u} \in \mathbb{K}^m$ vektor, hogy $\boxed{\mathbf{A}\mathbf{v} = \sigma\mathbf{u},\ \mathbf{A}^{\mathsf{H}}\mathbf{u} = \sigma\mathbf{v}.}$ A $\mathbf{v}$ vektort **jobb**, az $\mathbf{u}$-t **bal** szinguláris vektornak nevezzük (ui. $\mathbf{u}^{\mathsf{H}}\mathbf{A} = \sigma\mathbf{v}^{\mathsf{H}}$).

Á A $\sigma \neq 0$-hoz tartozó jobb szinguláris vektorok $\mathbb{K}^n$-ben, és a bal szinguláris vektorok $\mathbb{K}^m$-ben a nullvektorral mindketten s-dimenziós alteret alkotnak ($\sigma \neq 0$ multiplicitása s).

m $(\mathbf{u}, \sigma, \mathbf{v})$, $\|\mathbf{v}\| = \|\mathbf{u}\| = 1 \rightsquigarrow \|\mathbf{A}\mathbf{v}\| = \|\mathbf{A}^{\mathsf{H}}\mathbf{u}\| = \sigma$.

P $\{\mathbf{v}_1,\mathbf{v}_2\}=\{(\frac{4}{5}, -\frac{3}{5}), (\frac{3}{5}, \frac{4}{5})\}$, $\{\mathbf{u}_1,\mathbf{u}_2\}=\{(-\frac{5}{13},\frac{12}{13}), (\frac{12}{13},\frac{5}{13})\}$, $\sigma_1=10$, $\sigma_2=5$

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

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

- J r-rangú $A_{m\times n}$ mátrix esetén $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = 0$

$$\Sigma_1 = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_r \end{bmatrix}, \quad \mathbf{U}_1 = [\mathbf{u}_1 \mid \cdots \mid \mathbf{u}_r], \quad \mathbf{V}_1 = [\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_r]$$

- $\mathbf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i \rightsquigarrow \mathbf{A}\mathbf{V}_1 = \mathbf{U}_1\Sigma_1$

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

- $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ és $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ kiterjesztése ONB-sá, majd

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

- $\mathbf{V}_2$ oszlopvektorai $\perp \mathcal{O}(\mathbf{A}^{\mathsf{H}}) \rightsquigarrow r < i \leq n$ esetén $\mathbf{A}\mathbf{v}_i = \mathbf{0}$.
- $\mathbf{U}_2$ oszlopvektorai $\perp \mathcal{O}(\mathbf{A}) \rightsquigarrow r < i \leq m$ esetén $\mathbf{A}^{\mathsf{H}}\mathbf{u}_i = \mathbf{0}$.

$$\mathbf{A}\mathbf{V} = \left[\mathbf{A}\mathbf{v}_1 \ \ \dots \ \ \mathbf{A}\mathbf{v}_r \mid \mathbf{A}\mathbf{v}_{r+1} \ \ \dots \ \ \mathbf{A}\mathbf{v}_n\right] = \left[\sigma_1\mathbf{u}_1 \ \ \dots \ \ \sigma_r\mathbf{u}_r \mid \mathbf{0} \ \ \dots \ \ \mathbf{0}\right]$$

$$= \left[\mathbf{u}_1 \ \ \dots \ \ \mathbf{u}_r \mid \mathbf{u}_{r+1} \ \ \dots \ \ \mathbf{u}_m\right] \left[\begin{array}{cccc|ccc} \sigma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \sigma_r & 0 & \dots & 0 \\ \hline 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{array}\right] = \mathbf{U}\Sigma,$$

azaz $\mathbf{A}_{m\times n}\mathbf{V}_{n\times n} = \mathbf{U}_{m\times m}\Sigma_{m\times n}$, blokkmátrix alakban

$$\mathbf{A}\left[\mathbf{V}_1 \mid \mathbf{V}_2\right] = \left[\mathbf{U}_1 \mid \mathbf{U}_2\right] \left[\begin{array}{c|c} \Sigma_1 & \mathbf{O} \\ \hline \mathbf{O} & \mathbf{O} \end{array}\right].$$

$\mathbf{V}^{\mathsf{H}}$-tal jobbról szorozva ($\mathbf{V}\mathbf{V}^{\mathsf{H}} = \mathbf{I}$):

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^{\mathsf{H}} = \left[\mathbf{U}_1 \mid \mathbf{U}_2\right] \left[\begin{array}{c|c} \Sigma_1 & \mathbf{O} \\ \hline \mathbf{O} & \mathbf{O} \end{array}\right] \left[\begin{array}{c} \mathbf{V}_1^{\mathsf{H}} \\ \hline \mathbf{V}_2^{\mathsf{H}} \end{array}\right] = \mathbf{U}_1\Sigma_1\mathbf{V}_1^{\mathsf{H}}$$

Szinguláris (érték szerinti) felbontás, SVD

D L! $\mathbf{A} \in \mathbb{K}^{m \times n}$ ($\mathbb{K} = \mathbb{R}$ vagy $\mathbb{C}$). Szinguláris érték szerinti felbontás, vagy szinguláris felbontás: $\boxed{\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^{\mathsf{H}}}$, ahol $\mathbf{U}$ és $\mathbf{V}$ unitér (ortogonális) és Σ diagonális, monoton csökkenő nemnegatív elemekkel a főátlóban.

- Redukált szinguláris felbontás: $\boxed{\mathbf{A} = \mathbf{U}_1\Sigma_1\mathbf{V}_1^{\mathsf{H}}}$, ahol $\mathbf{V}_1$, $\mathbf{U}_1$ szemiunitér (szemiortogonális) és Σ_1 négyzetes, diagonális, monoton csökkenő pozitív elemekkel a főátlóban.
- Szinguláris érték szerinti diadikus felbontás vagy a szinguláris felbontás diadikus alakja: $\mathbf{A} = \sigma_1\mathbf{u}_1\mathbf{v}_1^{\mathsf{H}} + \sigma_2\mathbf{u}_2\mathbf{v}_2^{\mathsf{H}} + \cdots + \sigma_r\mathbf{u}_r\mathbf{v}_r^{\mathsf{H}}$, ahol $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\} \subset \mathbb{K}^m$, $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\} \subset \mathbb{K}^n$ ONR, $\sigma_1 \geqslant \sigma_2 \geqslant \ldots \geqslant \sigma_r > 0$.

Példa a SVD három alakjára

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

Szinguláris érték, SVD

Az SVD meghatározása

SVD meghatározása A^HA-ból

m $\mathbf{A}^H\mathbf{A} = (\mathbf{U}\Sigma\mathbf{V}^H)^H\mathbf{U}\Sigma\mathbf{V}^H = \mathbf{V}\Sigma^H\mathbf{U}^H\mathbf{U}\Sigma\mathbf{V}^H = \mathbf{V}\Sigma^H\Sigma\mathbf{V}^H$
Ez $\mathbf{A}^H\mathbf{A}$ sajátfelbontása: $\mathbf{V}$ unitér (ortogonális), $\Sigma^H\Sigma$ diagonális ($\sigma_i \geq 0$) $\rightsquigarrow$ $\mathbf{A}^H\mathbf{A}$ sajátértékei épp a szinguláris értékek négyzetei, ui. $\Sigma^H\Sigma = \operatorname{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_r^2, 0, \ldots, 0)$.
SVD kiszámítása:

1. kiszámoljuk $\mathbf{A}^H\mathbf{A}$ poz.sé-eit ($\lambda_i > 0$) és sv-ai közül kiválasztunk egy ONR-t (ez $\mathcal{O}(\mathbf{A}^H)$ ONB-a: $\{\mathbf{v}_i\}_{i=1}^r$), ezek jobb sztng. vektorok,
2. a sajátértékek gyökei a szinguláris értékek ($\sigma_i = \sqrt{\lambda_i}$),
3. bal szinguláris vektorok: mivel $\mathbf{A}\mathbf{v}_i = \sigma_i\mathbf{u}_i$, ezért $\mathbf{u}_i = \mathbf{A}\mathbf{v}_i/\sigma_i$,
4. a szinguláris vektorokból fölírható $\mathbf{V}_1$ és $\mathbf{U}_1$, a redukált szinguláris felbontás és a diadikus alak,
5. az $\mathcal{N}(\mathbf{A})$, illetve $\mathcal{N}(\mathbf{A}^H)$ egy-egy tetszőleges ONB-ának vektoraival kiegészítve $\mathbf{V}_1$ és $\mathbf{U}_1$ oszlopait megkapjuk a $\mathbf{V}$ és $\mathbf{U}$ mátrixokat, és a teljes SVD-t.

Alternatíva $m < n$ esetére, ha a mátrix $m \times n$-es

m ha $m < n$ lehet így is:
$\mathbf{AA}^{\mathsf{H}} = \mathbf{U}\Sigma\mathbf{V}^{\mathsf{H}}(\mathbf{U}\Sigma\mathbf{V}^{\mathsf{H}})^{\mathsf{H}} = \mathbf{U}\Sigma\mathbf{V}^{\mathsf{H}}\mathbf{V}\Sigma^{\mathsf{H}}\mathbf{U}^{\mathsf{H}} = \mathbf{U}\Sigma\Sigma^{\mathsf{H}}\mathbf{U}^{\mathsf{H}}$
Ez $\mathbf{AA}^{\mathsf{H}}$ sajátfelbontása: $\mathbf{U}$ unitér (ortogonális), $\Sigma\Sigma^{\mathsf{H}}$ diagonális ($\sigma_i \geq 0$) $\rightsquigarrow$ $\mathbf{AA}^{\mathsf{H}}$ sajátértékei épp a szinguláris értékek négyzetei, ui. $\Sigma\Sigma^{\mathsf{H}} = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2, 0, \dots, 0)$.
SVD kiszámítása:

1 meghatározzuk $\mathbf{AA}^{\mathsf{H}}$ pozitív sajátértékeit ($\lambda_i > 0$) és sajátvektorai közül kiválasztunk egy ONR-t ($\mathbf{u}_i$), ezek bal szinguláris vektorok,
2 a sajátértékek gyökei a szinguláris értékek ($\sigma_i = \sqrt{\lambda_i}$),
3 $\mathbf{v}_i = \mathbf{A}^{\mathsf{H}}\mathbf{u}_i/\sigma_i$, a jobb szinguláris vektorok
4 $\mathbf{V}_1$, $\mathbf{U}_1$, a redukált SVD és a diadikus alak felírása
5 a nullterek bázisaival kiegészítés ONB-sá, teljes SVD

P Határozzuk meg az

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

mátrix szinguláris értékeit, vektorait, és felbontását.

M Kiszámítjuk az

$$\mathsf{A}^\mathsf{T}\mathsf{A} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

sajátértékeit és sajátvektorait.

- sajátértékek: $\lambda_{1,2} = 2$, $\lambda_3 = 0 \rightsquigarrow$
  szinguláris értékek: $\sigma_{1,2} = \sqrt{2}$, $\sigma_3 = 0$.

- $A^T A$ sajátvektorai láthatóan lehetnek a standard bázis elemei: $\mathbf{v}_1 = \mathbf{e}_1$, $\mathbf{v}_2 = \mathbf{e}_2$, $\mathbf{v}_3 = \mathbf{e}_3$ (az *xy*-sík bármely két merőleges vektora lehetne).
- $\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i$ képlettel:

$$\mathbf{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix},$$

$$\mathbf{u}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}.$$

- $\mathrm{span}(\mathbf{u}_1, \mathbf{u}_2)^\perp$ meghatározása:

$$\mathbf{u}_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \quad \mathbf{u}_4 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}.$$

- Innen az teljes és a redukált SVD:

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

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

Szinguláris érték, SVD

SVD-re vonatkozó állítások

T **A szinguláris értékek létezése és egyértelműsége**

Minden r-rangú komplex vagy valós $\mathbf{A}$ mátrixnak létezik r pozitív szinguláris értéke, melyek monoton csökkenő sorozata egyértelmű.

m Általában a sajátértékek és a szinguláris értékek négyzetes mátrixok esetén is különböznek. Például az $\left[\begin{smallmatrix}1 & 1\\ 0 & 0\end{smallmatrix}\right]$ mátrix sajátértékei 1, 0, míg szinguláris értékei $\sqrt{2}$, 0.

Á Normális mátrix szinguláris értékei a sajátértékeinek abszolút értékeivel egyeznek meg.

B Ha normális, akkor unitéren diagonalizálható $\rightsquigarrow$ $\mathbf{A}^{\mathsf{H}}\mathbf{A} = (\mathbf{Q}\Lambda\mathbf{Q}^{\mathsf{H}})^{\mathsf{H}}\mathbf{Q}\Lambda\mathbf{Q}^{\mathsf{H}} = \mathbf{Q}\Lambda^{\mathsf{H}}\Lambda\mathbf{Q}^{\mathsf{H}} \rightsquigarrow \sigma_i^2 = \lambda_i\bar{\lambda}_i$, azaz $\sigma_i = |\lambda_i|$.

Á Ha $\mathbf{A}$ pozitív szemidefinit, akkor sajátértékei megegyeznek szinguláris értékeivel, és sajátfelbontása szinguláris felbontása is (ha a sajátértékeket csökkenő sorrendben tettük a Λ-ba).

B $\mathbf{A} = \mathbf{Q}\Lambda\mathbf{Q}^{\mathsf{T}}$: $\mathbf{U} = \mathbf{V} = \mathbf{Q}$, $\Sigma = \Lambda$.

Szimmetrikus mátrix szinguláris felbontása

m Ha $\mathbf{A}$ szimmetrikus (önadjungált), akkor sajátfelbontásából megkapható az SVD: legyenek Λ-ban a sajátértékek abszolútértékük szerint monoton csökkenő sorrendben, és képezzük $\mathbf{E}$-t az $\mathbf{I}$-ből úgy, hogy a Λ-beli negatív sajátértékek helyén $\mathbf{E}$-ben -1 legyen. Ekkor a sajátfelbontásból az $\mathbf{U} = \mathbf{QE}$, $\Sigma = \mathbf{E}\Lambda$ és $\mathbf{V} = \mathbf{Q}$ helyettesítéssel szinguláris felbontást kapunk:

$$\mathbf{A} = \mathbf{Q}\Lambda\mathbf{Q}^{\mathsf{T}} = \mathbf{QEE}\Lambda\mathbf{Q}^{\mathsf{T}} = (\mathbf{QE})(\mathbf{E}\Lambda)\mathbf{Q}^{\mathsf{T}} = \mathbf{U}\Sigma\mathbf{V}^{\mathsf{T}}.$$

($\mathbf{QE}$ a $\mathbf{Q}$ oszlopait szorozza -1-gyel)

P Egy szimmetrikus mátrix saját- és szinguláris felbontása:

$$\begin{bmatrix} -7 & 24 \\ 24 & 7 \end{bmatrix} = \mathbf{Q}\Lambda\mathbf{Q}^{\mathsf{T}} = \begin{bmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix} \begin{bmatrix} 25 & 0 \\ 0 & -25 \end{bmatrix} \begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{bmatrix}$$

$$\begin{bmatrix} -7 & 24 \\ 24 & 7 \end{bmatrix} = \mathbf{U}\Sigma\mathbf{V}^{\mathsf{T}} = \begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{bmatrix} \begin{bmatrix} 25 & 0 \\ 0 & 25 \end{bmatrix} \begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{bmatrix}$$

Szinguláris érték, SVD

Geometriai interpretáció

$\mathbb{R}^n$ standard bázisa $\mathcal{E}_1$, másik bázisa $\mathcal{B}_1 = \{\mathsf{v}_1, \ldots, \mathsf{v}_n\}$
$\mathbb{R}^m$ standard bázisa $\mathcal{E}_2$, másik bázisa $\mathcal{B}_2 = \{\mathsf{u}_1, \ldots, \mathsf{u}_m\}$

$$\begin{tikzcd}
{[\mathsf{x}]_{\mathcal{B}_1}} \arrow[r, "\Sigma_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}"] \arrow[d, "\mathsf{V}_{\mathcal{E}_1 \leftarrow \mathcal{B}_1}"'] & {[\Sigma\mathsf{x}]_{\mathcal{B}_2}} \arrow[d, "\mathsf{U}_{\mathcal{E}_2 \leftarrow \mathcal{B}_2}"] \\
{[\mathsf{x}]_{\mathcal{E}_1}} \arrow[r, "\mathsf{A}_{\mathcal{E}_2 \leftarrow \mathcal{E}_1}"'] & {[\mathsf{A}\mathsf{x}]_{\mathcal{E}_2}}
\end{tikzcd}$$

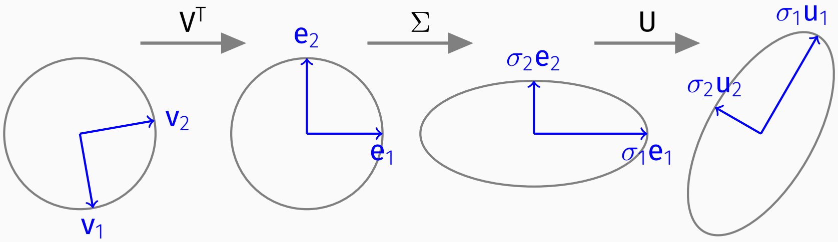
$$\begin{aligned}
[\mathsf{A}\mathsf{x}]_{\mathcal{E}_2} &= \mathsf{A}_{\mathcal{E}_2 \leftarrow \mathcal{E}_1}[\mathsf{x}]_{\mathcal{E}_1} \\
&= \mathsf{U}_{\mathcal{E}_2 \leftarrow \mathcal{B}_2}\Sigma_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}\mathsf{V}^{-1}_{\mathcal{B}_1 \leftarrow \mathcal{E}_1}[\mathsf{x}]_{\mathcal{E}_1} \quad \rightsquigarrow \quad \mathsf{A} = \mathsf{U}\Sigma\mathsf{V}^\mathsf{T} \\
&= \mathsf{U}_{\mathcal{E}_2 \leftarrow \mathcal{B}_2}\Sigma_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}\mathsf{V}^\mathsf{T}_{\mathcal{B}_1 \leftarrow \mathcal{E}_1}[\mathsf{x}]_{\mathcal{E}_1}
\end{aligned}$$

P SVD hatása egységkörön 2-rangú 2×2-es mátrixszal:

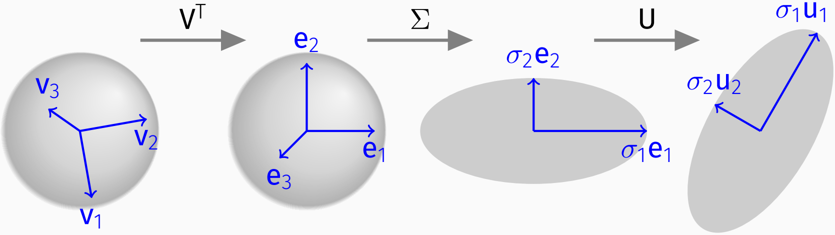
$$\mathbf{V}^\mathsf{T}\mathbf{v}_i = \mathbf{e}_i, \quad \Sigma\mathbf{e}_i = \sigma_i\mathbf{e}_i, \quad \mathbf{U}\sigma_i\mathbf{e}_i = \sigma_i\mathbf{U}\mathbf{e}_i = \sigma_i\mathbf{u}_i, \quad (i = 1, 2)$$

azaz $\mathbf{V}^\mathsf{T}$ a $\{\mathbf{v}_i\}$ bázist a standardba viszi ortogonális leképezéssel (forgatás vagy tükrözés), ott Σ tengelyirányban nyújt/összenyom, végül az ortogonális $\mathbf{U}$ hat rá.

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^\mathsf{T} : \mathbf{x} \mapsto \mathbf{U}\Sigma\mathbf{V}^\mathsf{T}\mathbf{x} :$$



P 2×3-as, valós, 2-rangú mátrix: $\mathbf{V}^\mathsf{T}$ ort. $\rightsquigarrow \mathbf{V}^\mathsf{T}\mathbf{v}_i = \mathbf{e}_i$ $(i = 1, 2, 3)$.



- Σ a két első tengely irányában nyújt/összenyom: $\Sigma\mathbf{e}_i = \sigma_i\mathbf{e}_i$ $(i = 1, 2)$, a harmadik tengely irányban vetít: $\Sigma\mathbf{e}_3 = \mathbf{0}$.
- A kép nem egy ellipszisvonal, hanem ellipszistartomány.
- Végül az ortogonális $\mathbf{U}$ ezt elforgatja vagy egyenesre tükrözi.

$$\mathbf{A}_{2\times 3} = \mathbf{U}\Sigma\mathbf{V}^\mathsf{T} : \mathbf{x}_{3\times 1} \mapsto \mathbf{U}_{2\times 2}\Sigma_{2\times 3}\mathbf{V}^\mathsf{T}_{3\times 3}\mathbf{x}_{3\times 1}$$

$$= [\mathbf{u}_1 \mid \mathbf{u}_2] \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 \\ \hline \mathbf{v}_2 \\ \hline \mathbf{v}_3 \end{bmatrix} \mathbf{x}_{3\times 1}$$

T $\mathbf{A}$ egy r-rangú, $m \times n$-es valós mátrix. Az $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ leképezés $\mathbb{R}^n$ az $\mathbf{e}^\mathsf{T}\mathbf{e} = 1$ egyenletet kielégítő, egységgömb felületén lévő pontjait, $\mathbb{R}^m$ egy r-dimenziós altere

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

Szinguláris érték, SVD

Polárfelbontás, pszeudoinverz

Polárfelbontás

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

D **Polárfelbontás**

egy négyzetes valós/komplex mátrix pozitív szemidefinit és ortogonális/unitér mátrixok szorzatára való bontása.

T **A polárfelbontás létezése és egyértelműsége**

Bármely komplex/valós négyzetes $\mathbf{A}$ mátrix előáll $\boxed{\mathbf{A} = \mathbf{PQ}}$ alakban, ahol $\mathbf{P}$ pozitív szemidefinit önadjungált/szimmetrikus mátrix, $\mathbf{Q}$ unitér/ortogonális. Ha $\mathbf{A}$ invertálható, akkor $\mathbf{P}$ pozitív definit, és a felbontás egyértelmű.

- B $A = U\Sigma V^H = U\Sigma U^H U V^H = (U\Sigma U^H)(UV^H)$, ahonnan $\boxed{P = U\Sigma U^H,\ Q = UV^H.}$
- P önadjungált, hisz $(U\Sigma U^H)^H = U\Sigma^H U^H = U\Sigma U^H$.
- A P pozitív szemidefinit, hisz hasonló a pozitív szemidefinit Σ mátrixhoz (ha A invertálható, akkor Σ pozitív definit)
- Q unitér (ortogonális), hisz két unitér (ortogonális) mátrix szorzata.
- A P egyértelmű (nem csak akkor, ha pozitív definit), ugyanis

$$AA^H = PQQ^H P^H = PP^H = P^2,$$

azaz $P = \sqrt{AA^H}$, és pozitív szemidefinit önadjungált mátrix négyzetgyöke egyértelmű az önadjungált pozitív szemidefinit mátrixok körében.
- Ha P pozitív definit, akkor invertálható, így $Q = P^{-1}A$ 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 Számítsuk ki $A = \begin{bmatrix} 0 & -2 & 2 \\ -2 & 3 & -2 \\ 4 & -2 & 0 \end{bmatrix}$ polárfelbontásait!

M1 $P = U\Sigma U^T$, $Q = UV^T$, $\hat{P} = V\Sigma V^T$

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

M2 Megoldás „ránézésre”:

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

T **Pszeudoinverz SVD-ből**

L! $\mathbf{A} = \mathbf{U}_1\Sigma_1\mathbf{V}_1^H$, illetve $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^H$. Ekkor

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

B $\mathbf{U}_1$ teljes oszloprangú, $\Sigma_1\mathbf{V}_1^H$ teljes sorrangú, így

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

Ebből

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

Pszeudoinverz kiszámítása az SVD-ből

P Számítsuk ki az alábbi A mátrix szinguláris érték szerinti felbontását, és abból a pszeudoinverzét!

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

M $\mathsf{A}^\mathsf{T}\mathsf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, sajátpárjai $(3, (1, 1))$, $(1, (-1, 1)) \rightsquigarrow$

$\sigma_1 = \sqrt{3}$, $\sigma_2 = 1$, $\mathbf{v}_1 = \frac{1}{\sqrt{2}}(1, 1)$, $\mathbf{v}_2 = \frac{1}{\sqrt{2}}(-1, 1)$,

\- $\mathbf{u}_1 = \frac{1}{\sqrt{3}}\mathsf{A}\mathbf{v}_1 = \frac{1}{\sqrt{6}}(1, 1, 2)$, $\mathbf{u}_2 = \frac{1}{1}\mathsf{A}\mathbf{v}_2 = \frac{1}{\sqrt{2}}(-1, 1, 0)$,

\- $\mathsf{A} = \mathsf{U}_1\Sigma_1\mathsf{V}_1^\mathsf{T} = \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & 0 \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$

Pszeudoinverz kiszámítása az SVD-ből

- Innen

$$A^+ = V_1 \Sigma_1^{-1} U_1^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}$$

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

Norma

Norma

Vektornorma

D Az $\mathbf{x}$ vektor euklideszi normája vagy más néven abszolút értéke

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{\mathbf{x}^\mathsf{T}\mathbf{x}}, \qquad \text{ha } \mathbf{x} \in \mathbb{R}^n,$$

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} = \sqrt{\mathbf{x}^\mathsf{H}\mathbf{x}}, \qquad \text{ha } \mathbf{x} \in \mathbb{C}^n.$$

m Manhattan ($|x| + |y|$), képméretezés ($\max\{|x|, |y|\}$).

D **p-norma**

A $p \geq 1$ valósra az $\mathbf{x} \in \mathbb{C}^n$ vektor p-normája $\|\mathbf{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$, míg ennek határértéke a ∞-norma, azaz $\|\mathbf{x}\|_\infty = \lim_{p\to\infty} \|\mathbf{x}\|_p$. A 2-norma = euklideszi-norma.

m 1-norma = rácsnorma = Manhattan-norma

maximum norma = ∞-norma

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

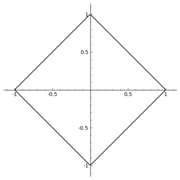
B* a legnagyobb abszolút értékű koordináta $x_{\max}$ ($|x_i|/|x_{\max}| \leq 1$)

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

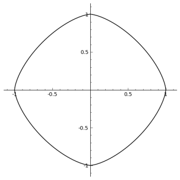
Mindegyik kifejezést $1/p$-edik hatványra emelve, majd $|x_{\max}|$-szal beszorozva kapjuk, hogy

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

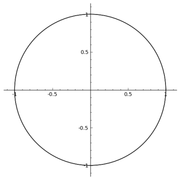
Egységkörök különböző normákban



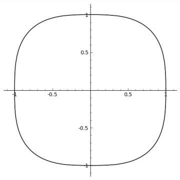
$p = 1$



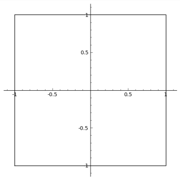
$p = \frac{3}{2}$



$p = 2$

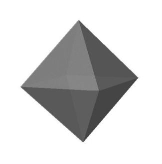


$p = 3$



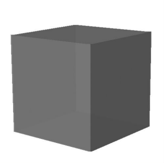
$p = \infty$

Egységgömbök különböző normákban



$p = 1$

$p = 2$



$p = \infty$

A vektornorma általános fogalma

Á Az előző normák alaptulajdonságai:

- $\|\mathbf{x}\| \geq 0$
- $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$ ($\rightsquigarrow$ a $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$ távolságfüggvény szeparálja a pontokat, azaz két különböző pont távolsága sosem 0)
- $\|c\mathbf{x}\| = |c|\|\mathbf{x}\|$ (pozitív homogenitás)
- $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (háromszögegyenlőtlenség)

D **Vektornorma**

Az $f \colon \mathbb{K}^n \to \mathbb{R}$ ($\mathbb{K} = \mathbb{R}$ vagy $\mathbb{K} = \mathbb{C}$) függvény norma, ha minden $\mathbf{x}, \mathbf{y} \in \mathbb{K}^n$ vektorra és minden $c \in \mathbb{K}$ konstansra

- $f(\mathbf{x}) \geq 0$, és $f(\mathbf{x}) = 0 \iff \mathbf{x} = \mathbf{0}$,
- $f(c\mathbf{x}) = |c| f(\mathbf{x})$,
- $f(\mathbf{x} + \mathbf{y}) \leq f(\mathbf{x}) + f(\mathbf{y})$.

A norma néhány tulajdonsága

J Ha f norma, az $f(\mathbf{x})$-et rendszerint $\|\mathbf{x}\|$ jelöli.

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

T A háromszögegyenlőtlenség p-normánál: Minkowski-egyenlőtlenség: $\|\mathbf{x}+\mathbf{y}\|_p \le \|\mathbf{x}\|_p + \|\mathbf{y}\|_p$.

T A Hölder-egyenlőtlenség a CBS-egyenlőtlenség általánosítása:

$$|\mathbf{x}^{\mathsf{H}}\mathbf{y}| \le \|\mathbf{x}\|_p \, \|\mathbf{y}\|_q \,, \text{ ahol } \frac{1}{p} + \frac{1}{q} = 1.$$

Á Minden norma folytonos függvény.

m Ha $\mathbf{x} \mapsto \|\mathbf{x}\|$ egy norma, és A egy egy-egy értelmű lineáris leképezés, akkor az $\mathbf{x} \mapsto \|A\mathbf{x}\|$ leképezés is norma és a következő is

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

Vektornormák ekvivalenciája

m $\max_i\{|x_i|\} \leq \sqrt{|x_1|^2 + \cdots + |x_n|^2} \leq |x_1| + \cdots + |x_n|$ azaz

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1 .$$

m Másrészt

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

D A $\|.\|_a$ és $\|.\|_b$ normák ekvivalensek, ha van olyan c és d pozitív valós szám, hogy $\|.\|_a \leq c\,\|.\|_b$ és $\|.\|_b \leq d\,\|.\|_a$.

m Az 1-, 2- és ∞-normák ekvivalensek

T **Vektornormák ekvivalenciája**

Legyen $\mathbb{K} = \mathbb{R}$ vagy $\mathbb{K} = \mathbb{C}$. A $\mathbb{K}^n$ téren értelmezett bármely két norma ekvivalens.

m Ez csak véges dimenziós terekben igaz, itt tehát a konvergenciakérdésekhez bármelyik norma jó.

Norma

Vektornorma mátrixokon

Frobenius-norma

D

Frobenius-norma

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

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

T

Frobenius-norma ekvivalens alakjai

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

B $[\mathbf{A}^H \mathbf{A}]_{jj} = \|\mathbf{A}_{*j}\|_2^2$, és nyom = sajátértékek összege

Á $\forall\, \mathbf{x} \in \mathbb{C}^n$, $\mathbf{A} \in \mathbb{C}^{m \times n}$ esetén $\|\mathbf{A}\mathbf{x}\|_2 \leq \|\mathbf{A}\|_F \, \|\mathbf{x}\|_2$.

B Cauchy–Bunyakovszkij–Schwarz-egyenlőtlenségből:

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

Norma

A mátrixnorma általános fogalma

D

Mátrixnorma

$\|.\| : \mathbb{K}^{m\times n} \to \mathbb{R}$ mátrixnorma, ha $\forall\, \mathbf{A}, \mathbf{B} \in \mathbb{K}^{m\times n}, \mathbf{C} \in \mathbb{K}^{n\times k}, c \in \mathbb{K}$:

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

D

Indukált mátrixnorma, *p*-norma, operátornorma

$\|.\|$ egy tetszőleges vektornorma. Ekkor az

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

a vektornorma által indukált mátrixnorma.

p-norma mátrixokra: $\|\mathbf{A}\|_p = \max_{\|\mathbf{x}\|_p=1} \|\mathbf{Ax}\|_p$

Operátornorma: ha lineáris leképezésre értelmezzük.

T

Az 1-, 2- és ∞-mátrixnorma kiszámítása

Legyen $\mathbf{A} \in \mathbb{C}^{m \times n}$, ekkor

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

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

$$\|\mathbf{A}\|_2 = \left\|\mathbf{A}^{\mathsf{H}}\right\|_2 = \max_{\|\mathbf{x}\|_2=1} \max_{\|\mathbf{y}\|_2=1} |\mathbf{y}^{\mathsf{H}}\mathbf{A}\mathbf{x}| = \sigma_1,$$

Ha az $\mathbf{A} \in \mathbb{C}^{n \times n}$ invertálható, akkor

$$\left\|\mathbf{A}^{-1}\right\|_2 = \max_{\|\mathbf{x}\|_2=1} \frac{1}{\|\mathbf{A}\mathbf{x}\|_2} = \frac{1}{\min\limits_{\|\mathbf{x}\|_2=1} \|\mathbf{A}\mathbf{x}\|_2} = \frac{1}{\sigma_n},$$

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

- Normák ekvivalenciája $\rightsquigarrow$ az egységgömb korlátos és zárt bármely normában $\rightsquigarrow$ $\mathsf{x} \mapsto \|\mathsf{A}\mathsf{x}\|$ függvénynek van maximuma és minimuma
- az indukált norma definíciójának ekvivalens alakjai

$$\|\mathsf{A}\| = \sup_{\|\mathsf{x}\| \neq 0} \frac{\|\mathsf{A}\mathsf{x}\|}{\|\mathsf{x}\|} = \max_{\|\mathsf{x}\| \neq 0} \frac{\|\mathsf{A}\mathsf{x}\|}{\|\mathsf{x}\|}.$$

- a legkisebb szinguláris érték pontosan akkor pozitív, ha A invertálható, azaz ha a 0 nem szinguláris érték.
- Az 1-, a ∞- és a 2-normára szokásos másik elnevezés: oszlopnorma, sornorma és spektrálnorma.

Példák

P Számítsuk ki a Frobenius-, 1-, 2- és ∞-normát!

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

M $\|A\|_F = 5$, $\|A\|_1 = 6$, $\|A\|_2 = 5$, $\|A\|_\infty = 6$.
$\|B\|_F = 5$, $\|B\|_1 = 7$, $\|B\|_2 = 5$, $\|B\|_\infty = 4$.
$\|C\|_F = \sqrt{13}$, $\|C\|_1 = 4$, $\|C\|_2 = 3$, $\|C\|_\infty = 4$.

P Számítsuk ki a Frobenius-, 1-, 2- és ∞-normát!

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

M $\|A\|_F = 9$, $\|A\|_1 = 8$, $\|A\|_2 = 8$, $\|A\|_\infty = 8$.
$\|B\|_F = 3\sqrt{3}$, $\|B\|_1 = 5$, $\|B\|_2 = 3$, $\|B\|_\infty = 5$.
$\|C\|_F = 3\sqrt{3}$, $\|C\|_1 = 5$, $\|C\|_2 = 5$, $\|C\|_\infty = 5$.

Mátrixnorma és spektrálsugár

T Tetszőleges mátrixnormára $\rho(\mathsf{A}) \leq \|\mathsf{A}\|$, ahol $\rho(\mathsf{A})$ az A spektrálsugara.

B Ha $(\lambda, \mathbf{x})$ sajátpár, azaz $\mathsf{A}\mathbf{x} = \lambda\mathbf{x}$, akkor

$$\mathsf{A}\mathbf{x}\mathbf{x}^{\mathsf{H}} = \lambda\mathbf{x}\mathbf{x}^{\mathsf{H}} \rightsquigarrow \|\mathsf{A}\| \left\|\mathbf{x}\mathbf{x}^{\mathsf{H}}\right\| \geq \left\|\mathsf{A}\mathbf{x}\mathbf{x}^{\mathsf{H}}\right\| = |\lambda| \left\|\mathbf{x}\mathbf{x}^{\mathsf{H}}\right\|,$$

és mivel $\mathbf{x} \neq \mathbf{0}$, így $\mathbf{x}\mathbf{x}^{\mathsf{H}} \neq \mathsf{O}$, azaz $\left\|\mathbf{x}\mathbf{x}^{\mathsf{H}}\right\| \neq 0$, vagyis leosztva vele $|\lambda| \leq \|\mathsf{A}\|$ adódik. $\rightsquigarrow \rho(\mathsf{A}) \leq \|\mathsf{A}\|$.

T Ha A normális ($\mathsf{A}^{\mathsf{H}}\mathsf{A} = \mathsf{A}\mathsf{A}^{\mathsf{H}}$), akkor $\|\mathsf{A}\|_2 = \rho(\mathsf{A})$.

B A normális $\iff$ unitéren diagonalizálható: $\mathsf{A} = \mathsf{U}\mathsf{D}\mathsf{U}^{\mathsf{H}}$.

- $\mathsf{A}^{\mathsf{H}}\mathsf{A} \sim \mathsf{D}^{\mathsf{H}}\mathsf{D}$ ui. $\mathsf{A}^{\mathsf{H}}\mathsf{A} = (\mathsf{U}\mathsf{D}\mathsf{U}^{\mathsf{H}})^{\mathsf{H}}(\mathsf{U}\mathsf{D}\mathsf{U}^{\mathsf{H}}) = \mathsf{U}\mathsf{D}^{\mathsf{H}}\mathsf{D}\mathsf{U}^{\mathsf{H}}$, tehát $\mathsf{A}^{\mathsf{H}}\mathsf{A}$ és $\mathsf{D}^{\mathsf{H}}\mathsf{D}$ sajátértékei megegyeznek.
- $\mathsf{D}^{\mathsf{H}}\mathsf{D}$ minden sajátértéke $|\lambda|^2$ alakú, ahol λ az A valamely sajátértéke $\rightsquigarrow \|\mathsf{A}\|_2 = \sigma_1$, azaz az $\mathsf{A}^{\mathsf{H}}\mathsf{A}$ legnagyobb sajátértékének gyöke = A legnagyobb sajátértékével = $\rho(\mathsf{A})$

Mátrix érzékenysége invertálásra – kondíciószám

P Az alábbi egyenletrendszer konstansait 0.01-dal megváltoztatjuk:

$$6.73x - 8.97y = 5.61$$

$$4.79x - 6.39y = 3.99$$

- A megoldás:	$x = 1.5,\ y = 0.5$
$6.73 \to 6.72$	$x \approx -2.26,\ y \approx -2.32$
$-8.97 \to -8.96$	$x \approx 4.35,\ y \approx 2.64$
$5.61 \to 5.60$	$x = 1.5,\ y = 0.5$
$4.79 \to 4.80,\ 6.39 \to 6.40$	inkonz., opt. sort.: $0.2998, -0.3997$
$3.99 \to 4.00$	∞ sok mo. (sortérbe $0.3, -0.4$)

- Együtthatómátrixra $\sigma_1 = 13.767$, $\sigma_2 = 0.002789$, $\sigma_1/\sigma_2 = 4935.7$.

D Egy $\mathbf{A}_{n\times n}$ mátrix kondíciószáma $\operatorname{cond}(\mathbf{A}) = \|\mathbf{A}\|\|\mathbf{A}^{-1}\| = \sigma_1/\sigma_n$.

- Ha e szám nagy, a mátrix és az $\mathbf{Ax} = \mathbf{b}$ rosszul kondicionált.
- Nem invertálható mátrixra ez ∞, ortogonális/unitér mátrixra 1.

Norma

Kis rangú approximáció

Kis rangú approximáció tétele – Eckart–Young-tétel

T A r-rangú, szinguláris érték szerinti felbontásának diadikus alakja

$$\mathsf{A} = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^{\mathsf{T}},$$

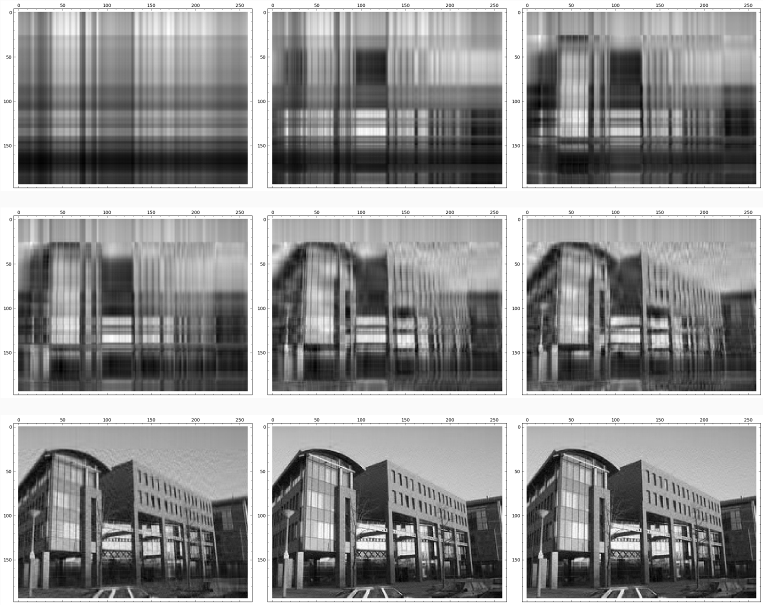
és legyen

$$\mathsf{A}_k = \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^{\mathsf{T}} \quad (k = 1, 2, \dots, r).$$

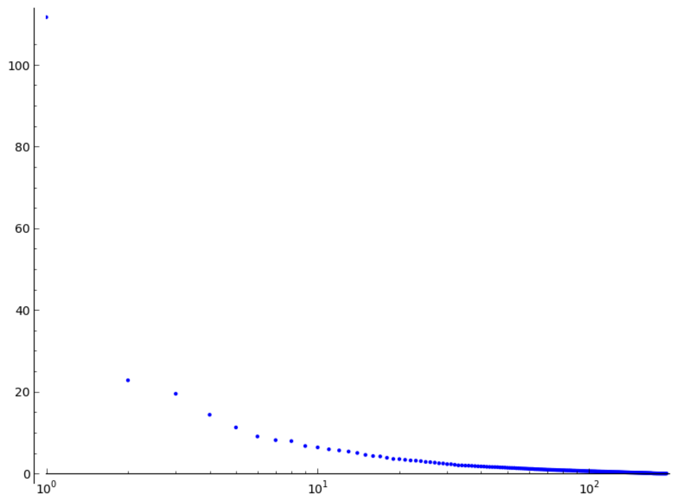
Ekkor A_k az A mátrix legjobb legföljebb k-rangú közelítése, azaz

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

$$\min_{\mathrm{r}(\mathsf{B}) \le k} \|\mathsf{A} - \mathsf{B}\|_2 = \|\mathsf{A} - \mathsf{A}_k\|_2 = \sigma_{k+1}.$$



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


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