

Szinguláris értékek

Wettl Ferenc

2016. április 12.

- 1 Szinguláris érték
- 2 Norma
- 3 Mátrixnorma
- 4 Alkalmazások

1 Szinguláris érték

- Az SVD meghatározása
- Szinguláris érték szerinti felbontás létezése
- Szinguláris felbontás geometriai interpretációja

2 Norma

- Euklideszi vektornorma és p -norma
- A norma általános fogalma
- Vektornormák ekvivalenciája

3 Mátrixnorma

- Vektornorma mátrixokon
- A mátrixnorma általános fogalma
- Az 1-, 2- és ∞ -norma mátrixokra

4 Alkalmazások

- Információtömörítés
- Mögöttes tartalom analízise

Szinguláris értékek és vektorok

- D** A pozitív $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ számok az r -rangú valós (komplex) **A** mátrix **szinguláris értékei**, ha van olyan $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ ONB-a a sortérnek (konjugáltjának) és $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ ONB-a az oszloptérnek, hogy

$$\mathbf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i, \quad i = 1, \dots, r.$$

A $\mathbf{v}_i$ vektorokat **jobb**, az $\mathbf{u}_i$ vektorokat **bal szinguláris vektoroknak** nevezzük.

- m** $|\mathbf{u}_i| = 1 \ (i = 1, 2, \dots, r) \rightsquigarrow |\mathbf{A}\mathbf{v}_i| = \sigma_i.$

Lehetne így: $k = \min(m, n)$, és $\sigma_{r+1} = \sigma_{r+2} = \dots = \sigma_k = 0.$

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

$$\begin{bmatrix} -4/13 & 6 \\ 11/13 & -4 \end{bmatrix} \begin{bmatrix} 4/5 \\ -3/5 \end{bmatrix} = 10 \begin{bmatrix} -5/13 \\ 12/13 \end{bmatrix}, \quad \begin{bmatrix} -4/13 & 6 \\ 11/13 & -4 \end{bmatrix} \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix} = 5 \begin{bmatrix} 12/13 \\ 5/13 \end{bmatrix},$$

$$\rightsquigarrow \sigma_1 = 10, \sigma_2 = 5.$$

J Jelölések

$$\Sigma_1 = \text{diag}(\sigma_1, \dots, \sigma_r) = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_r \end{bmatrix}$$

$$\mathbf{U}_1 = \{\mathbf{u}_1, \dots, \mathbf{u}_r\}$$

$$\mathbf{V}_1 = \{\mathbf{v}_1, \dots, \mathbf{v}_r\}$$

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

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

$\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ -ből $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ -ből ONB:

$$\mathbf{V}_2 = \begin{bmatrix} \mathbf{v}_{r+1} & \dots & \mathbf{v}_n \end{bmatrix}$$

$$\mathbf{U}_2 = \begin{bmatrix} \mathbf{u}_{r+1} & \dots & \mathbf{u}_m \end{bmatrix}$$

$\mathbb{R}$ -ben $\mathbf{V}_2$ oszlopvektorai $\perp \mathcal{S}$ ($\mathbb{C}$ -ben $\perp \bar{\mathcal{S}}$ -re) $\rightsquigarrow r < i \leq n$ esetén $\mathbf{A}\mathbf{v}_i = \mathbf{0}$.

$$\begin{aligned}\mathbf{A}\mathbf{V} &= [\mathbf{A}\mathbf{v}_1 \quad \dots \quad \mathbf{A}\mathbf{v}_r \mid \mathbf{A}\mathbf{v}_{r+1} \quad \dots \quad \mathbf{A}\mathbf{v}_n] \\ &= [\sigma_1 \mathbf{u}_1 \quad \dots \quad \sigma_r \mathbf{u}_r \mid \mathbf{0} \quad \dots \quad \mathbf{0}]\end{aligned}$$

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

$$= \mathbf{U}\Sigma,$$

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

$$\mathbf{A} [\mathbf{V}_1 \mid \mathbf{V}_2] = [\mathbf{U}_1 \mid \mathbf{U}_2] \left[\begin{array}{c|c} \Sigma_1 & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array} \right].$$

Szinguláris (érték szerinti) felbontás, SVD

D Szinguláris érték szerinti felbontás, vagy szinguláris felbontás:

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$$

Redukált szinguláris felbontás: $\mathbf{A} = \mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^H$

Szinguláris érték szerinti diadikus felbontás vagy a szinguláris felbontás diadikus alakja

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

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

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

SVD kiszámítása:

- 1 meghatározzuk $\mathbf{A}^T \mathbf{A}$ sajátértékeit (λ_i) és sajátvektorait ($\mathbf{x}_i$),
- 2 a sajátértékek gyökei a szinguláris értékek ($\sigma_i = \sqrt{\lambda_i}$),
- 3 a jobb szinguláris vektorok a sajátvektorok ($\mathbf{v}_i = \mathbf{x}_i / |\mathbf{x}_i|$),
- 4 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 alternatíva: a bal szinguláris vektorok megegyeznek az $\mathbf{A} \mathbf{A}^T$ sajátvektoraival, innen $\mathbf{v}_i = \mathbf{A}^T \mathbf{u}_i / \sigma_i$,
- 5 a szinguláris vektorokból fölírható $\mathbf{V}_1$ és $\mathbf{U}_1$, a redukált szinguláris felbontás és a diadikus alak,
- 6 a $\mathbf{V}$ -hez ki kell egészíteni a $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ vektorokat ONB-sá (hasonlóképp $\mathbf{U}$ -hoz)

T A szinguláris értékek létezése és egyértelműsége: Minden r -rangú komplex $\mathbf{A}$ mátrixnak létezik r (pozitív) szinguláris értéke. Ezek megegyeznek $\mathbf{A}^H \mathbf{A}$ (illetve az $\mathbf{A} \mathbf{A}^H$) pozitív sajátértékeinek négyzetgyökeivel. A szinguláris értékek monoton csökkenő sorozata egyértelmű.

B $\mathbf{A}^H \mathbf{A}$ önadjungált: $(\mathbf{A}^H \mathbf{A})^H = \mathbf{A}^H (\mathbf{A}^H)^H = \mathbf{A}^H \mathbf{A}$

$\rightsquigarrow$ minden sajátértéke valós, így ortogonálisan diagonalizálható,

$\mathbf{A}^H \mathbf{A}$ pozitív szemidefinit, így minden sajátértéke nemnegatív,

ui. $\mathbf{x}^H \mathbf{A}^H \mathbf{A} \mathbf{x} = (\mathbf{A} \mathbf{x})^H (\mathbf{A} \mathbf{x}) = |\mathbf{A} \mathbf{x}|^2 \geq 0$.

$r(\mathbf{A}^H \mathbf{A}) = r(\mathbf{A}) = r$, így $(\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ rendezés után), $\lambda_i > 0$, ha $1 \leq i \leq r$, és $\lambda_i = 0$, ha $r < i \leq n$

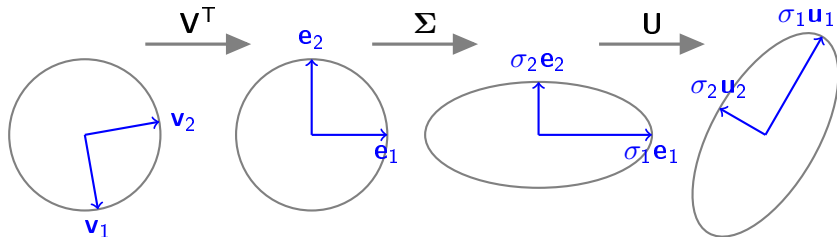
$\sigma_i = \sqrt{\lambda_i} > 0$, ha $1 \leq i \leq r$

$\mathbf{A}^H \mathbf{A}$ sajátértékei egyértelműek, ezért $\mathbf{A}$ szinguláris értékei is azok

$\mathbf{A}^H \mathbf{A}$ és $\mathbf{A} \mathbf{A}^H$ 0-tól különböző sajátértékei megegyeznek

P SVD hatása egységgörön 2-rangú 2×2 -es mátrixszal: $\mathbf{V}^T \mathbf{v}_i = \mathbf{e}_i$, $\Sigma \mathbf{e}_i = \sigma_i \mathbf{e}_i$, $\mathbf{U} \sigma \mathbf{e}_i = \sigma \mathbf{U} \mathbf{e}_i = \sigma \mathbf{u}_i$, azaz $\mathbf{V}^T$ a $\{\mathbf{v}_i\}$ bázist a standardba viszi ortogonális leképezéssel (forgatás vagy tükrözés), ott Σ tengelyirányban nyújtja/összenyomja, végül az ortogonális $\mathbf{U}$ hat rá.

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

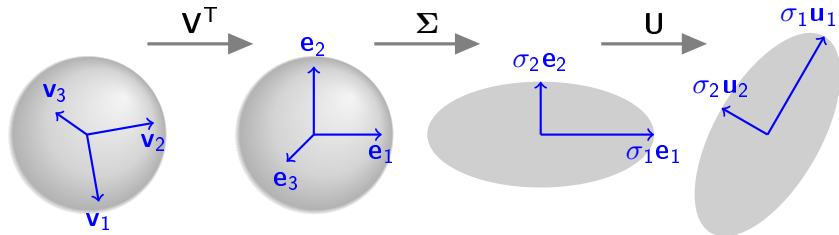


P 2×3 -as, valós, 2-rangú mátrix: $\mathbf{V}^T$ ortogonális, így hatása: $\mathbf{V}^T \mathbf{v}_i = \mathbf{e}_i$ ($i = 1, 2, 3$).

Σ a két első tengely irányában nyújt/összenyom: $\Sigma \mathbf{e}_i = \sigma_i \mathbf{e}_i$ ($i = 1, 2$), a harmadik tengely irányával párhuzamosan vetít: $\Sigma \mathbf{e}_3 = \mathbf{0}$.

A kép nem egy ellipsziszvonal, hanem a teljes általa határolt tartomány.

Végül az ortogonális $\mathbf{U}$ ezt elforgatja vagy tükrözi egy egyenesre.



- T** **A** egy r -rangú, $m \times n$ -es valós mátrix. Az $\mathbf{x} \mapsto \mathbf{Ax}$ leképezés $\mathbb{R}^n$ az $\mathbf{e}^T \mathbf{e} = 1$ egyenletet kielégítő, egységgömb felületén lévő pontjait, $\mathbb{R}^m$ egy r -dimenziós altere
- egy ellipszoidjának felületére képi, ha $r = n$, és
 - egy ellipszoidja által határolt tartományára képi, ha $r < n$.

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** **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 SVD: $\mathbf{A} = \mathbf{U}_1 \mathbf{\Sigma}_1 \mathbf{V}_1^T$, illetve $\mathbf{A} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T$. Ekkor

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

B $\mathbf{U}_1$ teljes oszloprangú, $\mathbf{\Sigma}_1 \mathbf{V}_1^T$ teljes sorrangú, így

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

Ebből

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

- 1 Szinguláris érték
 - Az SVD meghatározása
 - Szinguláris érték szerinti felbontás létezése
 - Szinguláris felbontás geometriai interpretációja
- 2 Norma
 - Euklideszi vektornorma és p -norma
 - A norma általános fogalma
 - Vektornormák ekvivalenciája
- 3 Mátrixnorma
 - Vektornorma mátrixokon
 - A mátrixnorma általános fogalma
 - Az 1-, 2- és ∞ -norma mátrixokra
- 4 Alkalmazások
 - Információtömörítés
 - Mögöttes tartalom analízise

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

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

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

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

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

m 1-norma = rácsnorma = Manhattan-norma

maximum norma = ∞ -norma

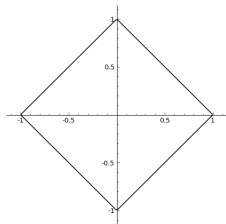
$$\|\mathbf{x}\|_{\infty} = \lim_{p \rightarrow \infty} \|\mathbf{x}\|_p = \lim_{p \rightarrow \infty} \left(\sum_{i=1}^n |x_i|^p \right)^{1/p} = \max_i |x_i|.$$

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

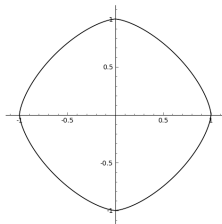
$$1 \leq \sum_{i=1}^n |x_i/x_{\max}|^p \leq n.$$

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

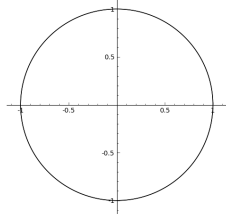
$$|x_{\max}| \leq |x_{\max}| \left(\sum_{i=1}^n \left| \frac{x_i}{x_{\max}} \right|^p \right)^{1/p} \leq |x_{\max}| n^{1/p},$$



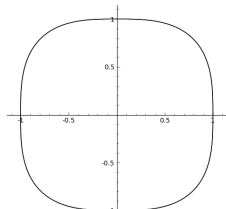
$$p = 1$$



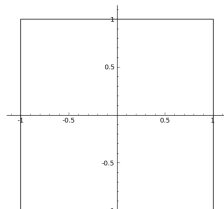
$$p = \frac{3}{2}$$



$$p = 2$$



$$p = 3$$



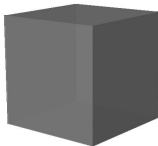
$$p = \infty$$



$$p = 1$$



$$p = 2$$



$$p = \infty$$

Á Az előző normák alaptulajdonságai:

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

D Az $f: \mathbb{R}^n \rightarrow \mathbb{R}$, vagy $f: \mathbb{C}^n \rightarrow \mathbb{R}$ függvény **norma**, ha

- $f(x) \geq 0$ minden x vektorra, és $f(x) = 0$ pontosan akkor áll fenn, ha $x = \mathbf{0}$,
- $f(cx) = |c|f(x)$ minden x vektorra,
- $f(x + y) \leq f(x) + f(y)$.

m $\|x\| = \|-x\|$ bármely $\|\cdot\|$ normára igaz, hisz
 $\|-x\| = |-1| \|x\| = \|x\|$.

m háromszögeyenlőtlenség másik alakja:

$$\|z - x\| \geq \left| \|z\| - \|x\| \right|$$

m A háromszögegyenlőtlenség p -normánál: **Minkowski-egyenlőtlenség:**

$$\|\mathbf{x} + \mathbf{y}\|_p \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p. \quad (1)$$

m A **Hölder-egyenlőtlenség** a CBS-egyenlőtlenség általánosítása:

$$|\mathbf{x}^H \mathbf{y}| \leq \|\mathbf{x}\|_p \|\mathbf{y}\|_q, \text{ ahol } \frac{1}{p} + \frac{1}{q} = 1. \quad (2)$$

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

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

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

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

m Minden norma folytonos függvény.

- D** A $\|\cdot\|_a$ és $\|\cdot\|_b$ normák ekvivalensek, ha van olyan c és d pozitív valós szám, hogy $\|\cdot\|_a \leq c \|\cdot\|_b$ és $\|\cdot\|_b \leq d \|\cdot\|_a$.
- m** Az 1-, 2- és ∞ -normák ekvivalensek
- T** A $\mathbb{K}^n$ téren értelmezett bármely két norma ekvivalens ($\mathbb{K} = \mathbb{R}$ vagy $\mathbb{C}$).
- m** Ez csak véges dimenziós terekben igaz, itt tehát a konvergenciakérdésekhez bármelyik norma jó.

- 1 Szinguláris érték
 - Az SVD meghatározása
 - Szinguláris érték szerinti felbontás létezése
 - Szinguláris felbontás geometriai interpretációja
- 2 Norma
 - Euklideszi vektornorma és p -norma
 - A norma általános fogalma
 - Vektornormák ekvivalenciája
- 3 Mátrixnorma
 - Vektornorma mátrixokon
 - A mátrixnorma általános fogalma
 - Az 1-, 2- és ∞ -norma mátrixokra
- 4 Alkalmazások
 - Információtömörítés
 - Mögöttes tartalom analízise

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

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

T Frobenius-norma ekvivalens alakjai:

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

B $[\mathbf{A}^H \mathbf{A}]_{jj} = \|\mathbf{A}_{*j}\|_2^2$

nyom = sajátértékek összege

Á Bármely $\mathbf{x} \in \mathbb{C}^n$ vektorra és $\mathbf{A} \in \mathbb{C}^{m \times n}$ mátrixra $\|\mathbf{A}\mathbf{x}\|_2 \leq \|\mathbf{A}\|_F \|\mathbf{x}\|_2$.

B Cauchy–Bunyakovszkij–Schwarz-egyenlőtlenségből:

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

D $\|\cdot\| : \mathbb{K}^n \rightarrow \mathbb{R}$ mátrixnorma, ha

- $\|\mathbf{A}\| \geq 0$, 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 $\|\cdot\|$ egy tetszőleges vektornorma. Ekkor az

$$\|\mathbf{A}\| = \max_{\|\mathbf{x}\|=1} \|\mathbf{Ax}\|$$

egyenlőséggel definiált függvényt a vektornorma által **indukált** mátrixnormának nevezzük.

m Jelölés: $\|\mathbf{A}\|_p = \max_{\|\mathbf{x}\|_p=1} \|\mathbf{Ax}\|_p$

m Ha lineáris leképezésre értelmezzük, **operátornormáról** beszélünk.

m A normák ekvivalenciájából $\leadsto$ bármely normában az egységgömb korlátos és zárt $\leadsto$ a $\mathbf{x} \mapsto \mathbf{Ax}$ függvénynek van maximuma és minimuma

m a definíció ekvivalens alakjai

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

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

$$\|\mathbf{A}\|_1 = \max_j \sum_{i=1}^n |a_{ij}| = \text{legnagyobb abszolút oszlopösszeg}, \quad (3)$$

$$\|\mathbf{A}\|_\infty = \max_i \sum_{j=1}^m |a_{ij}| = \text{legnagyobb abszolút sorösszeg}, \quad (4)$$

$$\|\mathbf{A}\|_2 = \left\| \mathbf{A}^H \right\|_2 = \max_{\|\mathbf{x}\|_2=1} \max_{\|\mathbf{y}\|_2=1} |\mathbf{y}^H \mathbf{A} \mathbf{x}| = \sigma_1, \quad (5)$$

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

$$\left\| \mathbf{A}^{-1} \right\|_2 = \max_{\|\mathbf{x}\|_2=1} \frac{1}{\|\mathbf{A} \mathbf{x}\|_2} = \frac{1}{\min_{\|\mathbf{x}\|_2=1} \|\mathbf{A} \mathbf{x}\|_2} = \frac{1}{\sigma_n}, \quad (6)$$

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

m Az 1-, a ∞ - és a 2-normára szokásos másik elnevezés: **oszlopnorma**, **sornorma** és **spektrálnorma**.

- 1 Szinguláris érték
 - Az SVD meghatározása
 - Szinguláris érték szerinti felbontás létezése
 - Szinguláris felbontás geometriai interpretációja
- 2 Norma
 - Euklideszi vektornorma és p -norma
 - A norma általános fogalma
 - Vektornormák ekvivalenciája
- 3 Mátrixnorma
 - Vektornorma mátrixokon
 - A mátrixnorma általános fogalma
 - Az 1-, 2- és ∞ -norma mátrixokra
- 4 Alkalmazások
 - Információtömörítés
 - Mögöttes tartalom analízise

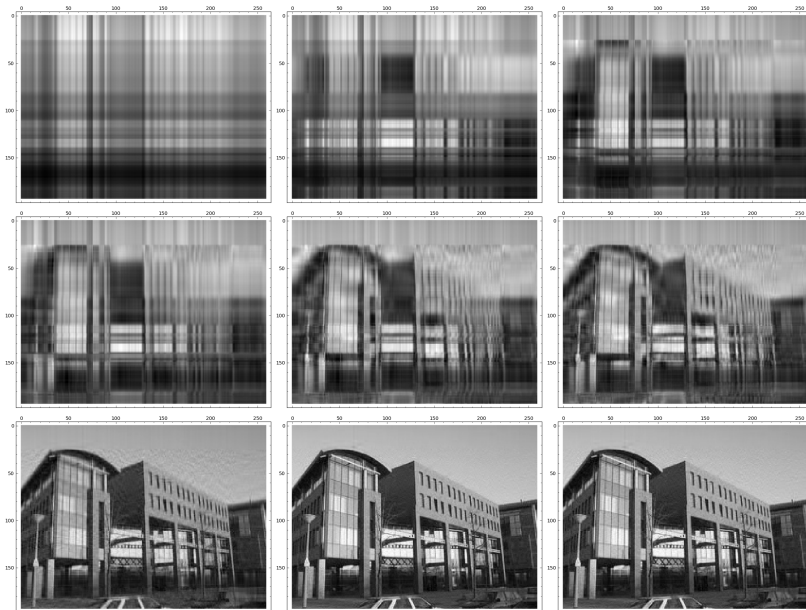
T Kis rangú approximáció tétele – Eckart–Young-tétel **A** r -rangú, k -adik szinguláris értéke σ_k , jobb és bal szinguláris vektora $\mathbf{v}_k$, illetve $\mathbf{u}_k$.
Legyen

$$\mathbf{A}_k = \sum_{i=1}^k \sigma_i \mathbf{u}_i \mathbf{v}_i^T.$$

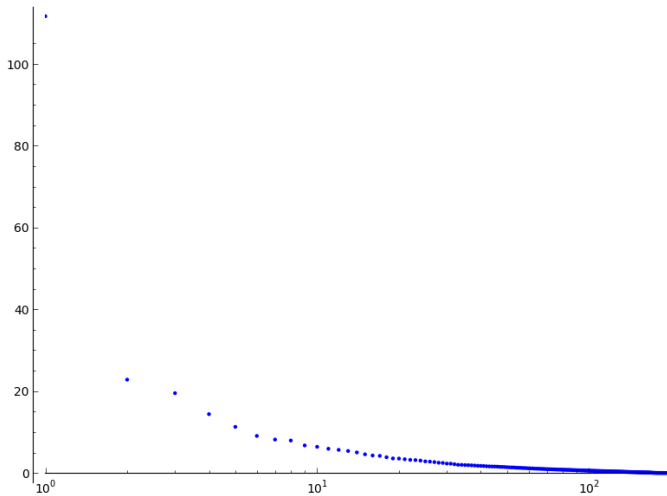
Ekkor $\mathbf{A}_k$ az $\mathbf{A}$ mátrix legjobb legfölbbebb k -rangú közelítése, azaz

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

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



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