BUDAPESTI MŰSZAKI
MATEMATIKA
ÉS GAZDASÁGTUDOMÁNYI
INTÉZET
EGYETEM

Alkalmazott algebra

BMETE90MX57 (FELSŐBB MATEMATIKA INFORMATIKUSOKNAK)

Szinguláris értékek

2018-11-05 EIC

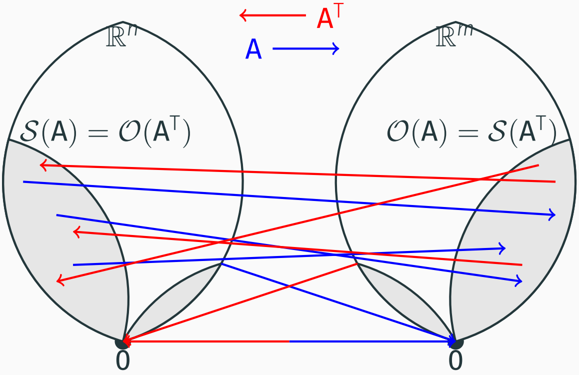
Wettl Ferenc
ALGEBRA TANSZÉK

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

Keressünk két ONB-t



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).

Diagonális két ONB-ban

m Az $\mathbf{A} \in \mathbb{R}^{m\times n}$ mátrixhoz olyan ortonormált $\{\mathbf{v}_1, \mathbf{v}_2 \dots, \mathbf{v}_n\} \subset \mathbb{R}^n$ és $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\} \subset \mathbb{R}^m$ 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

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

- $\mathbf{A}^\mathsf{T}\mathbf{A}$ szimmetrikus és pozitív szemidefinit $\rightsquigarrow$ ortogonálisan diagonalizálható, sajátértékei nem negatívak
- sv-okból kiválasztható $\mathbb{R}^n$ egy ONB-a, a 0 sé-hez tartozók a $\mathcal{N}(\mathbf{A}^\mathsf{T}\mathbf{A}) = \mathcal{N}(\mathbf{A})$ tér ONB-át adják ($\mathcal{N}(\mathbf{A}) = \{\mathbf{0}\}$ esetén a bázis $\emptyset$)
- a többi $\mathcal{N}(\mathbf{A})^\perp = \mathcal{S}(\mathbf{A})$ 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 $\mathbf{A}^\mathsf{T}\mathbf{A}$ mátrix $\lambda_i > 0$ sé-hez tartozó egységnyi sv-a, akkor $|\mathbf{A}\mathbf{v}_i| = \sqrt{\lambda_i}$, ui. $|\mathbf{A}\mathbf{v}_i|^2 = (\mathbf{A}\mathbf{v}_i)^\mathsf{T}(\mathbf{A}\mathbf{v}_i) = \mathbf{v}_i^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{v}_i = \lambda_i|\mathbf{v}_i|^2 = \lambda_i$.
- 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
- $\mathbf{A}\mathbf{v}_i = \sigma_i\mathbf{u}_i$, $\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{v}_i = \sigma_i^2\mathbf{v}_i \rightsquigarrow \mathbf{A}^\mathsf{T}\mathbf{u}_i = \frac{1}{\sigma_i}\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{v}_i = \frac{1}{\sigma_i}\sigma_i^2\mathbf{v} = \sigma_i\mathbf{v}_i$:

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

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

  $$\mathbf{A}\mathbf{v} = \sigma\mathbf{u}, \quad \mathbf{A}^\mathsf{T}\mathbf{u} = \sigma\mathbf{v}$$

Szinguláris értékek és vektorok

D Az r rangú $\mathbf{A} \in \mathbb{R}^{m\times n}$ mátrix szinguláris értékének nevezzük azt a nemnegatív σ valóst, melyhez van olyan $\mathbf{v} \in \mathbb{R}^n$ és $\mathbf{u} \in \mathbb{R}^m$ két nemzérus vektor, hogy

$$\mathbf{A}\mathbf{v} = \sigma\mathbf{u}, \quad \mathbf{A}^\mathsf{T}\mathbf{u} = \sigma\mathbf{v}.$$

A $\mathbf{v}$ vektort jobb, az $\mathbf{u}$ vektort bal szinguláris vektornak nevezzük (ui. $\mathbf{u}^\mathsf{T}\mathbf{A} = \sigma\mathbf{v}^\mathsf{T}$). A σ szinguláris érték multiplicitása s, ha legfeljebb s olyan független $\mathbf{v}_1, \ldots, \mathbf{v}_s$ jobb és $\mathbf{u}_1, \ldots, \mathbf{u}_s$ bal szinguláris vektor található, hogy minden $i = 1, 2 \ldots, s$ indexre $\mathbf{A}\mathbf{v}_i = \sigma\mathbf{u}_i, \quad \mathbf{A}^\mathsf{T}\mathbf{u}_i = \sigma\mathbf{v}_i$.

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

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

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

J Jelölések:

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

$$\mathbf{U}_1 = [\mathbf{u}_1 \mid \cdots \mid \mathbf{u}_r]$$

$$\mathbf{V}_1 = [\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_r]$$

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

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

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

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

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

$\mathbb{R}$-ben V_2 oszlopvektorai $\perp \mathcal{S}$ ($\mathbb{C}$-ben $\perp \bar{\mathcal{S}}$-re) $\rightsquigarrow$ $r < i \le n$ esetén $\mathsf{Av}_i = \mathbf{0}$.

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

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

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

Szinguláris (érték szerinti) felbontás, SVD

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

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

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

- Redukált szinguláris felbontás: $\mathbf{A} = \mathbf{U}_1\Sigma_1\mathbf{V}_1^\mathsf{H}$, 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: $\mathbf{A} = \sigma_1\mathbf{u}_1\mathbf{v}_1^\mathsf{T} + \sigma_2\mathbf{u}_2\mathbf{v}_2^\mathsf{T} + \dots + \sigma_r\mathbf{u}_r\mathbf{v}_r^\mathsf{T}$, ahol $\{\mathbf{u}_1, \dots, \mathbf{u}_r\} \subset \mathbb{K}^m$, $\{\mathbf{v}_1, \dots, \mathbf{v}_r\} \subset \mathbb{K}^n$ ONR, $\sigma_1 \geqslant \sigma_2 \geqslant \dots \geqslant \sigma_r > 0$.

Példa a SVD három alakjára

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

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

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

Szinguláris érték

Az SVD meghatározása

SVD meghatározása $A^H A$-ból

m Ötlet: $\mathbf{A}^H\mathbf{A} = (\mathbf{U}\Sigma\mathbf{V}^H)^H\mathbf{U}\Sigma\mathbf{V}^H = \mathbf{V}\Sigma^H\mathbf{U}^H\mathbf{U}\Sigma\mathbf{V}^H = \mathbf{V}\Sigma^H\Sigma\mathbf{V}^H$
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. $\Sigma^H\Sigma = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2, 0, \dots, 0)$.

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 szinguláris 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 VV_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 Ötlet: $\mathbf{AA}^\mathsf{H} = \mathbf{U}\Sigma\mathbf{V}^\mathsf{H}(\mathbf{U}\Sigma\mathbf{V}^\mathsf{H})^\mathsf{H} = \mathbf{U}\Sigma\mathbf{V}^\mathsf{H}\mathbf{V}\Sigma^\mathsf{H}\mathbf{U}^\mathsf{H} = \mathbf{U}\Sigma\Sigma^\mathsf{H}\mathbf{U}^\mathsf{H}$

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. $\Sigma\Sigma^\mathsf{H} = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_r^2, 0, \ldots, 0)$.

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 mint fönt ($\mathbf{V}_1$, $\mathbf{U}_1$),

5 mint fönt (a nullterek bázisaival kiegészítés ONB-sá).

Szinguláris érték

Szinguláris érték szerinti felbontás létezése

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, és unitéren diagonalizálható,

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

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

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

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

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

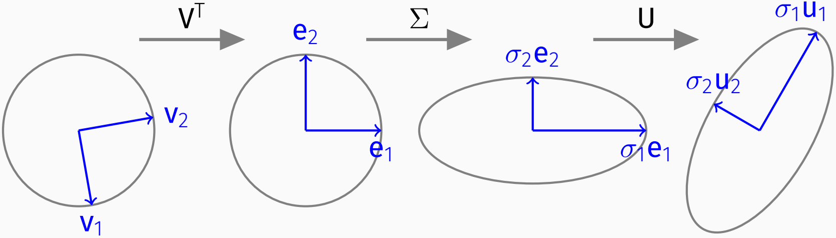
$\mathbf{A}^H\mathbf{A}$ és $\mathbf{A}\mathbf{A}^H$ 0-tól különböző sajátértékei megegyeznek

Szinguláris érték

Geometriai interpretáció

P SVD hatása egységkörön 2-rangú 2×2-es mátrixszal: $\mathbf{V}^\mathsf{T}\mathbf{v}_i = \mathbf{e}_i$, $\Sigma\mathbf{e}_i = \sigma_i\mathbf{e}_i$, $\mathbf{U}\sigma\mathbf{e}_i = \sigma\mathbf{U}\mathbf{e}_i = \sigma\mathbf{u}_i$, azaz $\mathbf{V}^\mathsf{T}$ 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.

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



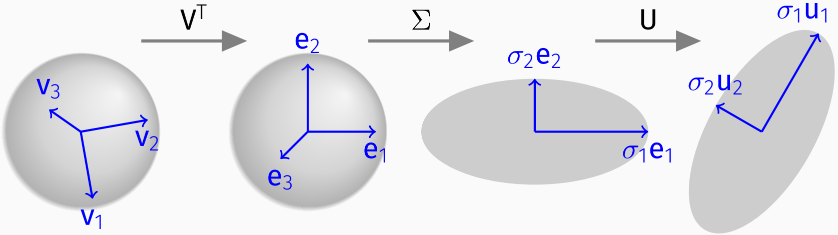
P 2×3-as, valós, 2-rangú mátrix: $\mathbf{V}^T$ ort. $\rightsquigarrow \mathbf{V}^T\mathbf{v}_i = \mathbf{e}_i$ $(i = 1, 2, 3)$.

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

A kép nem egy ellipszisvonal, hanem ellipszistartomány.

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

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



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

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 Polárfelbontáson egy négyzetes mátrixnak egy pozitív szemidefinit és egy ortogonális mátrix szorzatára való felbontását értjük.

T Bármely komplex (valós) négyzetes A mátrix előáll

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

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

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

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

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

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

Pszeudoinverz

T SVD: $\mathsf{A} = \mathsf{U}_1\Sigma_1\mathsf{V}_1^\mathsf{T}$, illetve $\mathsf{A} = \mathsf{U}\Sigma\mathsf{V}^\mathsf{T}$. Ekkor

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

B U_1 teljes oszloprangú, $\Sigma_1\mathsf{V}_1^\mathsf{T}$ teljes sorrangú, így

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

Ebből

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

Norma

Norma

Euklideszi vektornorma és p-norma

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 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 $\|\mathbf{x}\|_\infty = \lim_{p\to\infty} \|\mathbf{x}\|_p$.

m 1-norma = rácsnorma = Manhattan-norma

maximum norma = ∞-norma

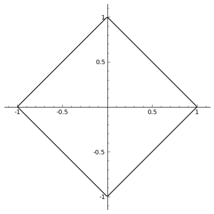
$$\|\mathsf{x}\|_\infty = \lim_{p\to\infty} \|\mathsf{x}\|_p = \lim_{p\to\infty} \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} = \max_i |x_i|.$$

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

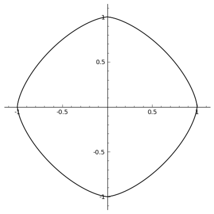
$$1 \leq \sum_{i=1}^n |x_i/x_{\max}|^p \leq n.$$

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

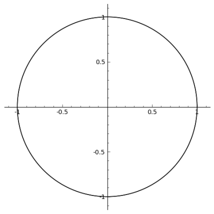
$$|x_{\max}| \leq |x_{\max}| \left(\sum_{i=1}^n \left|\frac{x_i}{x_{\max}}\right|^p\right)^{1/p} \leq |x_{\max}| n^{1/p},$$



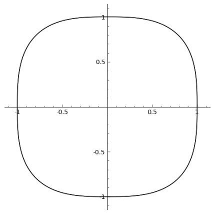
$p = 1$



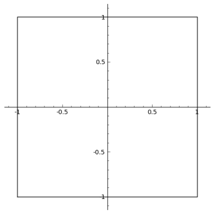
$p = \frac{3}{2}$



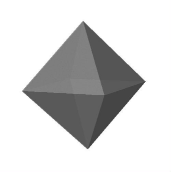
$p = 2$



$p = 3$

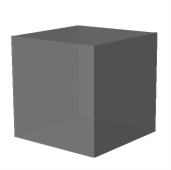


$p = \infty$



$p = 1$

$p = 2$



$p = \infty$

Norma

A norma 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 Az $f\colon \mathbb{R}^n \to \mathbb{R}$, vagy $f\colon \mathbb{C}^n \to \mathbb{R}$ függvény norma, ha

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

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

m háromszögegyenlőtlenség másik alakja:

$$\|\mathbf{z} - \mathbf{x}\| \geq \Big|\|\mathbf{z}\| - \|\mathbf{x}\|\Big|$$

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

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

- 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

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

- $\max_i\{|x_i|\} \le \sqrt{|x_1|^2 + \cdots + |x_n|^2} \le |x_1| + \cdots + |x_n|$ azaz

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

- Másrészt

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

- Minden norma folytonos függvény.

Norma

Vektornormák ekvivalenciája

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 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.

Mátrixnorma

Mátrixnorma

Vektornorma mátrixokon

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:

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

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

nyom = sajátértékek összege

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

B Cauchy–Bunyakovszkij–Schwarz-egyenlőtlenségből:

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

Mátrixnorma

A mátrixnorma általános fogalma

D $\|.\| : \mathbb{K}^n \to \mathbb{R}$ 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 $\|.\|$ egy tetszőleges vektornorma. Ekkor az

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

a vektornorma által indukált mátrixnorma.

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

Ha lineáris leképezésre értelmezzük, operátornormáról beszélünk.

m Normák ekvivalenciája $\rightsquigarrow$ az egységgömb korlátos és zárt bármely normában $\rightsquigarrow$ $\mathbf{x} \mapsto \mathbf{Ax}$ függvénynek van maximuma és minimuma

m a definíció ekvivalens alakjai

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

Mátrixnorma

Az 1-, 2- és ∞-norma mátrixokra

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

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

$$\|\mathbf{A}\|_\infty = \max_i \sum_{j=1}^{m} |a_{ij}| = \text{legnagyobb abszolút sorösszeg}, \tag{2}$$

$$\|\mathbf{A}\|_2 = \left\|\mathbf{A}^{\mathrm{H}}\right\|_2 = \max_{\|\mathbf{x}\|_2=1} \max_{\|\mathbf{y}\|_2=1} |\mathbf{y}^{\mathrm{H}}\mathbf{A}\mathbf{x}| = \sigma_1, \tag{3}$$

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

$$\left\|\mathbf{A}^{-1}\right\|_2 = \max_{\|\mathbf{x}\|_2=1} \frac{1}{\|\mathbf{A}\mathbf{x}\|_2} = \frac{1}{\min_{\|\mathbf{x}\|_2=1} \|\mathbf{A}\mathbf{x}\|_2} = \frac{1}{\sigma_n}, \tag{4}$$

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

m a legkisebb pontosan akkor pozitív, ha $\mathbf{A}$ invertálható

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

Alkalmazások

Alkalmazások

Információtömörítés

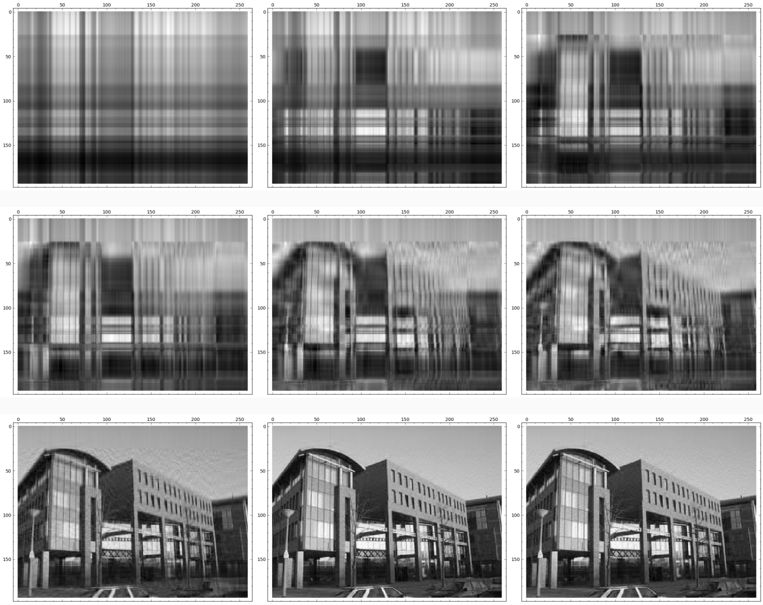
T Kis rangú approximáció tétele – Eckart–Young-tétel $\mathbf{A}$ r-rangú, k-adik szinguláris értéke σ_k, jobb és bal szinguláris vektorta $\mathbf{v}_k$, illetve $\mathbf{u}_k$. Legyen

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

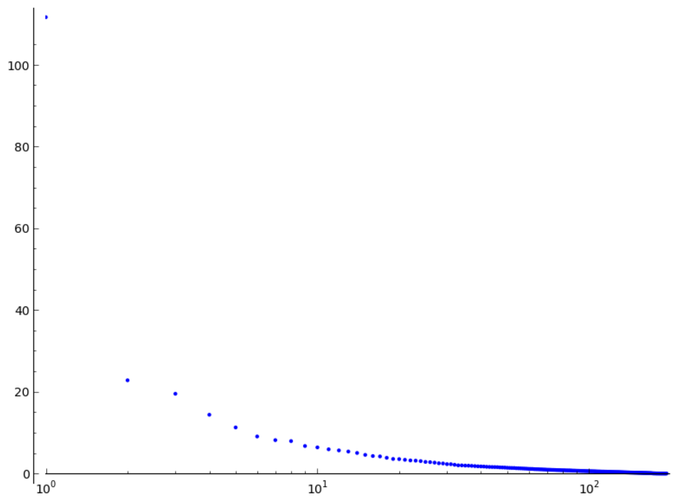
Ekkor $\mathbf{A}_k$ az $\mathbf{A}$ mátrix legjobb legföljebb k-rangú közelítése, azaz

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

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



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


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

Alkalmazások

Mögöttes tartalom analízise

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.

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

U_k, ill. V_k oszlopainak vektorterében a szavak, ill. dokumentumok kapcsolatát a hozzájuk tartozó vektorok helyzete jellemzi