Bevezetés az algebrába 2 – Vektor- és mátrixnorma

Wettl Ferenc
Algebra Tanszék

BUDAPESTI MŰSZAKI
MATEMATIKA
ÉS GAZDASÁGTUDOMÁNYI
INTÉZET
EGYETEM

2016. április 29.

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

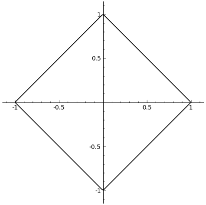
$$\|\mathbf{x}\|_\infty = \lim_{p\to\infty} \|\mathbf{x}\|_p = \lim_{p\to\infty} \left(\sum_{i=1}^n |x_i|^p \right)^{1/p} = \max_i |x_i|.$$

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

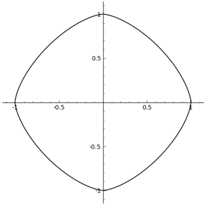
$$1 \le \sum_{i=1}^n |x_i/x_{\max}|^p \le n.$$

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

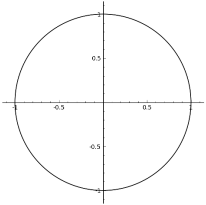
$$|x_{\max}| \le |x_{\max}| \left(\sum_{i=1}^n \left| \frac{x_i}{x_{\max}} \right|^p \right)^{1/p} \le |x_{\max}| n^{1/p},$$



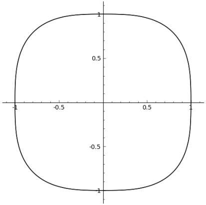
$p = 1$



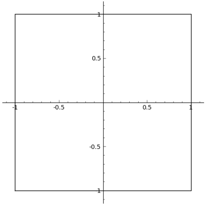
$p = \frac{3}{2}$



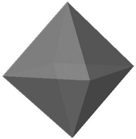
$p = 2$



$p = 3$

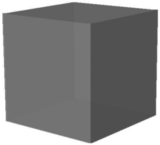


$p = \infty$



$p = 1$

$p = 2$



$p = \infty$

Á Az előző normák alaptulajdonságai:

- $\|\mathbf{x}\| \geq 0$
- $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$ ($\rightsquigarrow$ 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 $f : \mathbb{R}^n \to \mathbb{R}$, vagy $f : \mathbb{C}^n \to \mathbb{R}$ 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:

$$\|\mathbf{x}+\mathbf{y}\|_p \le \|\mathbf{x}\|_p + \|\mathbf{y}\|_p \,. \tag{1}$$

- A Hölder-egyenlőtlenség a CBS-egyenlőtlenség általánosítása:

$$|\mathbf{x}^{\mathsf{H}}\mathbf{y}| \le \|\mathbf{x}\|_p \, \|\mathbf{y}\|_q \,, \text{ ahol } \frac{1}{p}+\frac{1}{q}=1. \tag{2}$$

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

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 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.

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:

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

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

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

D $\|.\| : \mathbb{K}^{n\times n} \to \mathbb{R}$ 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

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

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

m a definíció ekvivalens alakjai

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

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

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

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

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

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

$$\left\|\mathbf{A}^{-1}\right\|_2 = \max_{\|\mathbf{x}\|_2=1} \frac{1}{\|\mathbf{A}\mathbf{x}\|_2} = \frac{1}{\min\limits_{\|\mathbf{x}\|_2=1} \|\mathbf{A}\mathbf{x}\|_2} = \frac{1}{\sigma_n}, \tag{6}$$

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.

B $p=1$: ha $\|\mathbf{x}\|_1 = 1$, akkor

$$\|\mathbf{Ax}\|_1 = \sum_{i=1}^{m}\left|\sum_{j=1}^{n} a_{ij}x_j\right| \le \sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}||x_j| = \sum_{j=1}^{n}|x_j|\sum_{i=1}^{m}|a_{ij}|$$
$$\le \left(\sum_{j=1}^{n}|x_j|\right)\max_j \sum_{i=1}^{m}|a_{ij}| = \max_j \sum_{i=1}^{m}|a_{ij}|$$

E max elérhető, ha k-adik oszlopössz. a max: $\|\mathbf{Ae}_k\|_1 = \max_j \sum_{i=1}^{m}|a_{ij}|$.

- $p=\infty$: ha $\|\mathbf{x}\|_\infty = 1$, akkor

$$\|\mathbf{Ax}\|_\infty = \max_i \left|\sum_{j=1}^{n} a_{ij}x_j\right| \le \max_i \sum_{j=1}^{n}|a_{ij}||x_j| \le \max_i \sum_{j=1}^{n}|a_{ij}|.$$

Ez a maximum el is érhető: ha a k-adik sor a max, akkor az

$$\mathbf{x} = \left(\frac{\overline{a_{k1}}}{|a_{k1}|}, \frac{\overline{a_{k2}}}{|a_{k2}|}, \dots, \frac{\overline{a_{kn}}}{|a_{kn}|}\right)$$

vektorra $\|\mathbf{x}\|_\infty = 1$ és $\|\mathbf{Ax}\|_\infty = \max_i \sum_{j=1}^{n}|a_{ij}|$.

- $p = 2$: A CBS-egyenlőtlenség szerint $|\mathbf{y}^H\mathbf{A}\mathbf{x}| \le \|\mathbf{y}\|_2 \|\mathbf{A}\mathbf{x}\|_2$, így

$$\max_{\|\mathbf{x}\|_2=1} \max_{\|\mathbf{y}\|_2=1} |\mathbf{y}^H\mathbf{A}\mathbf{x}| \le \max_{\|\mathbf{x}\|_2=1} \|\mathbf{A}\mathbf{x}\|_2 = \|\mathbf{A}\|_2 .$$

Legyen $\mathbf{x}_0$ az a vektor, melyben $\|\mathbf{A}\mathbf{x}\|_2$ a max

$$\|\mathbf{A}\mathbf{x}_0\|_2 = \max_{\|\mathbf{x}\|_2=1} \|\mathbf{A}\mathbf{x}\|_2 = \|\mathbf{A}\|_2 , \quad \mathbf{y}_0 = \frac{\mathbf{A}\mathbf{x}_0}{\|\mathbf{A}\mathbf{x}_0\|_2} = \frac{\mathbf{A}\mathbf{x}_0}{\|\mathbf{A}\|_2}.$$

$$\mathbf{y}_0^H\mathbf{A}\mathbf{x}_0 = \frac{\mathbf{x}_0^H\mathbf{A}^H\mathbf{A}\mathbf{x}_0}{\|\mathbf{A}\|_2} = \frac{\|\mathbf{A}\mathbf{x}_0\|_2^2}{\|\mathbf{A}\|_2} = \frac{\|\mathbf{A}\|_2^2}{\|\mathbf{A}\|_2} = \|\mathbf{A}\|_2 .$$

- Az $\|\mathbf{A}\|_2 = \sigma_1$ igazolásához a következő maximumot keressük:

$$\|\mathbf{A}\|_2^2 = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{Ax}\|_2^2}{\|\mathbf{x}\|_2^2} = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\mathbf{x}^{\mathsf{H}}\mathbf{A}^{\mathsf{H}}\mathbf{Ax}}{\mathbf{x}^{\mathsf{H}}\mathbf{x}}.$$

  Legyenek $\mathbf{v}_1, \dots, \mathbf{v}_n$ a jobb szinguláris vektorok, ekkor bármely $\mathbf{x} = \sum_j c_j \mathbf{v}_j \neq \mathbf{0}$ vektorra

$$\lambda_1 = \frac{\mathbf{v}_1^{\mathsf{H}}\mathbf{A}^{\mathsf{H}}\mathbf{A}\mathbf{v}_1}{\mathbf{v}_1^{\mathsf{H}}\mathbf{v}_1},\ \lambda_1 - \frac{\mathbf{x}^{\mathsf{H}}\mathbf{A}^{\mathsf{H}}\mathbf{Ax}}{\mathbf{x}^{\mathsf{H}}\mathbf{x}} = \lambda_1 - \frac{\sum_j \lambda_j c_j^2}{\sum_j c_j^2} = \frac{\sum_j (\lambda_1 - \lambda_j) c_j^2}{\sum_j c_j^2} \geq 0,$$

  tehát $\|\mathbf{A}\|_2^2 = \lambda_1$, azaz $\|\mathbf{A}\|_2 = \sqrt{\lambda_1} = \sigma_1$.

- $$\frac{1}{\min\limits_{\|\mathbf{x}\|_2=1} \|\mathbf{Ax}\|_2} = \max_{\|\mathbf{x}\|_2=1} \frac{1}{\|\mathbf{Ax}\|_2} = \max_{\mathbf{y}\neq\mathbf{0}} \frac{1}{\left\|\mathbf{A}\frac{\mathbf{A}^{-1}\mathbf{y}}{\|\mathbf{A}^{-1}\mathbf{y}\|_2}\right\|_2} = \max_{\mathbf{y}\neq\mathbf{0}} \frac{\|\mathbf{A}^{-1}\mathbf{y}\|_2}{\|\mathbf{y}\|_2}$$

$$= \max_{\mathbf{y}\neq\mathbf{0}} \left\|\mathbf{A}^{-1}\left(\frac{\mathbf{y}}{\|\mathbf{y}\|_2}\right)\right\|_2 = \max_{\|\mathbf{x}\|_2=1} \left\|\mathbf{A}^{-1}\mathbf{x}\right\| = \left\|\mathbf{A}^{-1}\right\|_2.$$

  $\mathbf{A}^{-1}$ szinguláris értékei az $\mathbf{A}$ szinguláris értékeinek reciprokai $\rightsquigarrow$ $\mathbf{A}^{-1}$ legnagyobb szing.ért. az $\mathbf{A}$ legkisebb szing.értékének reciproka.

Eckart–Young-tétel

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

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

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

B 2-normára: $r(\mathbf{B}) \leqslant k$, így $\dim \mathcal{N}(\mathbf{B}) \geqslant n - k$.

Legyen $\mathcal{V} = \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{k+1})$ a $k+1$ legnagyobb szing.ért.hez tartozó jobb szinguláris vektorok által kifeszített altér.

Ha $k \geqslant r$, akkor kész vagyunk, **A** legjobb közelítése **A** és $\sigma_{k+1} = 0$.

Feltehető $r > k \rightsquigarrow \dim \mathcal{N}(\mathbf{B}) + \dim \mathcal{V} \geqslant (n-k) + (k+1) > n \rightsquigarrow$ $\mathcal{N}(\mathbf{B}) \cap \mathcal{V} \neq \emptyset$

$\mathbf{w} \in \mathcal{N}(\mathbf{B}) \cap \mathcal{V}$, $\|\mathbf{w}\|_2 = 1$. Ekkor

$$\begin{aligned}\|\mathbf{A} - \mathbf{B}\|_2^2 &\geqslant \|(\mathbf{A} - \mathbf{B})\mathbf{w}\|_2^2 = \|\mathbf{A}\mathbf{w}\|_2^2 \\ &= \sum_{i=1}^{k+1} \sigma_i^2 |\mathbf{v}_i^T \mathbf{w}|^2 \geqslant \sigma_{k+1}^2 \sum_{i=1}^{k+1} |\mathbf{v}_i^T \mathbf{w}|^2 = \sigma_{k+1}^2\end{aligned}$$

Másrészt $\|\mathbf{A} - \mathbf{A}_k\|_2 = \sigma_{k+1}$, így $\|\mathbf{A} - \mathbf{B}\|_2 \geqslant \|\mathbf{A} - \mathbf{A}_k\|_2$.