Лекция 3. Нормы векторов и матриц

Панкратов Владимир Александрович

Московский Государственный Технический Университет имени Н.Э.Баумана

Москва, 2022

Нормы векторов и матриц

Ключевые слова

Линейное пространство. Скалярное произведение. Евклидово пространство. Норма. Матричная норма. Подчиненная норма. Октаэдрическая норма. Кубическая норма. Норма Гильберта-Шмидта. p-норма.

Множество $\mathcal{L}$ называется линейным пространством, а его элементы — векторами, если
1) задано сложение элементов, т. е. закон или правило, которое любым $x, y \in \mathcal{L}$ ставит в соответствие элемент $z = x + y \in \mathcal{L}$, называемый суммой элементов x и y;
2) задано умножение на скаляр, т. е. закон или правило, которое любым $x \in \mathcal{L}$ и $\lambda \in \mathbb{R}$ ставит в соответствие элемент $z = \lambda x \in \mathcal{L}$, называемый произведением x на скаляр λ;

Множество $\mathcal{L}$ называется линейным пространством, а его элементы — векторами, если
3) указанные операции подчиняются следующим аксиомам линейного пространства:
$\forall x, y, z \in \mathcal{L}$, $\forall \lambda, \mu \in \mathcal{L}$

1. $x + y = y + x$
2. $x + (y + z) = (x + y) + z$
3. $\exists 0 \in \mathcal{L}$: $x + 0 = 0$
4. $\exists (-x)$: $x + (-x) = 0$
5. $1 \cdot x = x$
6. $\lambda(\mu x) = (\lambda\mu)x$
7. $\lambda(x + y) = \lambda x + \lambda y$
8. $(\lambda + \mu)x = \lambda x + \mu x$

Линейное пространство $\mathcal{E}$ называется *евклидовым пространством*, если

1. в этом пространстве задано *скалярное умножение*, т. е. определён закон или правило, по которому любым $x, y \in \mathcal{E}$ ставится в соответствие действительное число $(x, y) \in \mathbb{R}$, называемое *скалярным произведением*;
2. при этом выполняются следующие аксиомы
   - 2.1 $(x, y) = (y, x)$ — симметричность;
   - 2.2 $(x + y, z) = (x, z) + (y, z)$ — линейность;
   - 2.3 $(\lambda x, y) = \lambda(x, y)$ — однородность;
   - 2.4 $(x, x) \geq 0$, причём $(x, x) = 0 \Leftrightarrow x = 0$.

Функцию, заданную на линейном пространстве $\mathcal{L}$, которая каждому вектору $x \in \mathcal{L}$ ставит в соответствие действительное число $\|x\|$, называют *нормой вектора*, если она удовлетворяет следующим *аксиомам нормы*

$\forall x, y \in \mathcal{L}$, $\lambda \in \mathbb{R}$

1. $\|x\| = 0 \Leftrightarrow x = 0$;
2. $\|\lambda x\| = |\lambda|\|x\|$ — однородность;
3. $\|x + y\| \leqslant \|x\| + \|y\|$ — неравенство треугольника.

Теорема

Для любого x: $\|x\| \geqslant 0$.

$\triangleleft$ Поскольку $0 = \|0\| = \|x + (-x)\| \leqslant \|x\| + \|x\| = 2\|x\|$, то $\|x\| \geqslant 0$. $\triangleright$

Нормы векторов и матриц

Норма

Функцию, заданную на линейном пространстве $\mathcal{L}$, которая каждому вектору $x \in \mathcal{L}$ ставит в соответствие действительное число $\|x\|$, называют *нормой вектора*, если она удовлетворяет следующим *аксиомам нормы*
$\forall x, y \in \mathcal{L}$, $\lambda \in \mathbb{R}$

1. $\|x\| = 0 \Leftrightarrow x = 0$;
2. $\|\lambda x\| = |\lambda| \|x\|$ — однородность;
3. $\|x + y\| \leqslant \|x\| + \|y\|$ — неравенство треугольника.

Теорема

*let $\mathcal{E}$ — произвольное евклидово пространство $\Rightarrow \forall x \in \mathcal{E}$:
$\|x\| = \sqrt{(x, x)}$ — индуцированная скалярным произведением норма*

$\triangleleft$ Доказать самостоятельно. $\triangleright$

В $\mathbb{R}^n$ наиболее часто используют следующие нормы

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|\,,$$

$$\|x\|_2 = \left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}},$$

$$\|x\|_\infty = \max_{i=\overline{1,n}} |x_i|\,.$$

Доказать самостоятельно, что для введённых выше норм выполнены следующие оценки:

$$\|x\|_\infty \le \|x\|_2 \le \|x\|_1 \le n\|x\|_\infty.$$

Норма Гильберта-Шмидта или *p-норма*

Определённая выражением $\|x\|_p = \left(\sum\limits_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$ $(p \geqslant 0)$ норма.

Очевидно, что нормы $\|x\|_1 = \sum\limits_{i=1}^{n} |x_i|$ и $\|x\|_2 = \left(\sum\limits_{i=1}^{n} x_i^2\right)^{\frac{1}{2}}$ являются частными, а норма $\|x\|_\infty = \max\limits_{i=\overline{1,n}} |x_i|$ — предельным (при $p \to +\infty$) случаями

Нормы $\|x\|_\infty$ называют *кубической*, $\|x\|_1$ — *октаэдрической*[1], а $\|x\|_2$ — *евклидовой*.

Кубическая норма также называется *нормой Чебышёва*.

[1] По названиям тел, которыми являются множества $\|x\| \leqslant 1$ при $x \in \mathbb{R}^3$

Норма матрицы, подчинённая векторной норме $\|x\|_*$

$\forall A \in \mathbb{R}^{m\times n}$ величина

$$\|A\|_* = \sup_{x\neq 0} \frac{\|Ax\|_*}{\|x\|_*}.$$

Теорема

Для подчиненной нормы матрицы A выполняются все аксиомы нормы.

$\forall A, B \in \mathbb{R}^{m\times n}$, $\lambda \in \mathbb{R}$:

1. $\|A\| = 0 \Leftrightarrow A = 0$;
2. $\|\lambda A\| = |\lambda| \|A\|$;
3. $\|A + B\| \leqslant \|A\| + \|B\|$.

Нормы векторов и матриц

Матричные нормы

$\triangleleft$ $\forall A, B \in \mathbb{R}^{m\times n}$, $\lambda \in \mathbb{R}$:

1. $\|A\| = \sup\limits_{x\neq 0} \dfrac{\|Ax\|}{\|x\|} \geqslant \left\{ \begin{array}{l} \|Ax\| \geqslant 0 \\ \|x\| \geqslant 0 \end{array} \right\} \geqslant 0 \Rightarrow \|A\| = 0 \Leftrightarrow$
   $\forall x \neq 0\colon \|Ax\| = 0 \Leftrightarrow \forall x \neq 0\colon\ Ax = 0 \Leftrightarrow A = 0;$
2. $\|\lambda A\| = \sup\limits_{x\neq 0} \dfrac{\|\lambda Ax\|}{\|x\|} = \sup\limits_{x\neq 0} \dfrac{|\lambda|\|Ax\|}{\|x\|} = |\lambda|\|A\|;$
3. $$\|A+B\| = \sup_{x\neq 0} \frac{\|(A+B)x\|}{\|x\|} = \sup_{x\neq 0} \frac{\|Ax+Bx\|}{\|x\|} \leqslant$$
   $$\leqslant \sup_{x\neq 0} \frac{\|Ax\|+\|Bx\|}{\|x\|} \leqslant \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|} + \sup_{x\neq 0} \frac{\|Bx\|}{\|x\|}.$$

Т. о. подчиненния матричная норма представляет собой норму на линейном пространстве $\mathbb{R}^{m\times n}$ $\triangleright$

Теорема

Дополнительно для подчиненной нормы матрицы A выполняются свойства:

- *субмультипликативности* $\forall A, B \in \mathbb{R}^{n\times n}$*:* $\|AB\| \leqslant \|A\|\|B\|$*;*
- *согласованности* $\forall A \in \mathbb{R}^{m\times n}$*,* $x \in \mathbb{R}^n$*:* $\|Ax\| \leqslant \|A\|\|x\|$*.*

◁ **Субмультипликативность.** $\forall A, B \in \mathbb{R}^{n\times n}$:

$$
\begin{aligned}
\|AB\| &= \sup_{x\neq 0} \frac{\|(A\cdot B)x\|}{\|x\|} = \\
&= \left\{ \begin{array}{l} \sup\limits_{x\neq 0} \left(\dfrac{\|A(B\cdot x)\|}{\|(B\cdot x)\|} \cdot \dfrac{\|B\cdot x\|}{\|x\|} \right), \text{ при } Bx \neq 0; \\ 0, \text{ при } Bx = 0 \end{array} \right\} \leqslant \\
&\leqslant \sup_{y\neq 0} \frac{\|Ay\|}{\|y\|} \cdot \sup_{x\neq 0} \frac{\|Bx\|}{\|x\|} = \|A\|\|B\|
\end{aligned}
$$

Согласованность. $\forall A \in \mathbb{R}^{m\times n}$, $x \in \mathbb{R}^n$:

$$\|Ax\| = \left\{ \begin{array}{l} \dfrac{\|Ax\|}{\|x\|} \cdot \|x\|, \ \text{при } x \neq 0; \\ 0, \ \text{при } x = 0 \end{array} \right\} \leqslant$$

$$\leqslant \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|} \cdot \|x\| = \|A\|\|x\|. \quad \triangleright$$

Нормы векторов и матриц

Формулы для вычисления основных матричных норм

Теорема

Векторным нормам $\|x\|_1$, $\|x\|_2$ и $\|x\|_\infty$ подчинены, соответственно, следующие матричные нормы

$$\|A\|_1 = \max_{j=\overline{1,n}} \sum_{i=1}^{n} |a_{ij}|,$$

$$\|A\|_2 = \max_{i=\overline{1,n}} \sqrt{\lambda_i(A^T A)},$$

$$\|A\|_\infty = \max_{i=\overline{1,n}} \sum_{j=1}^{n} |a_{ij}|,$$

где $\lambda_i(A^T A)$ — собственные значения матрицы $A^T A$.

$\triangleleft$ Матричная норма подчиненная **октаэдрической векторной норме**.

$\forall A \in \mathbb{R}^{m\times n}$, $\forall x \in \mathbb{R}^n$: $\|x\|_1 = \sum\limits_{i=1}^{n} |x_i| \Rightarrow$

$$\|Ax\|_1 = \sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij}x_j \right| \leqslant \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}||x_j| =$$

$$= \sum_{j=1}^{n} |x_j| \sum_{i=1}^{m} |a_{ij}| \leqslant \|x\|_1 \max_{j=\overline{1,n}} \sum_{i=1}^{m} |a_{ij}| \Rightarrow$$

$\dfrac{\|Ax\|_1}{\|x\|_1} \leqslant \max\limits_{j=\overline{1,n}} \sum\limits_{i=1}^{m} |a_{ij}|$ и осталось показать, что $\max\limits_{j=\overline{1,n}} \sum\limits_{i=1}^{m} |a_{ij}|$ — **точная** верхняя грань.

Нормы векторов и матриц

Формулы для вычисления основных матричных норм

let $j_{\max}$ — номер столбца матрицы, сумма модулей элементов которого максимальна: $\max\limits_{j=\overline{1,n}} \sum\limits_{i=1}^{m} |a_{ij}| = \sum\limits_{i=1}^{m} |a_{ij_{\max}}|$.

Рассмотрим вектор $z = (0, \ldots, 0, \underbrace{1}_{j_{\max}}, 0 \ldots, 0)^T$:

$$\frac{\|Az\|_1}{\|z\|_1} = \frac{\sum\limits_{i=1}^{m} \left| \sum\limits_{j=1}^{n} a_{ij} z_j \right|}{1} = \sum_{i=1}^{m} |a_{ij_{\max}}| = \max_{j=\overline{1,n}} \sum_{i=1}^{m} |a_{ij}| \Rightarrow$$

$\max\limits_{j=\overline{1,n}} \sum\limits_{i=1}^{m} |a_{ij}|$ не только является верхней гранью $\dfrac{\|Ax\|_1}{\|x\|_1}$, но и достигается на $z \neq 0 \Rightarrow \|A\|_1 = \sup\limits_{x \neq 0} \dfrac{\|Ax\|_1}{\|x\|_1} = \max\limits_{j=\overline{1,n}} \sum\limits_{i=1}^{n} |a_{ij}|$.

Матричная норма подчиненная **кубической векторной норме**.
$\forall A \in \mathbb{R}^{m\times n}$, $\forall x \in \mathbb{R}^n$: $\|x\|_\infty = \max\limits_{i=\overline{1,n}} |x_i| \Rightarrow$

$$\|Ax\|_\infty = \max_{i=\overline{1,m}} \left| \sum_{j=1}^{n} a_{ij} x_j \right| \leqslant \max_{i=\overline{1,m}} \sum_{j=1}^{n} |a_{ij}| \max_{j=\overline{1,n}} |x_j| =$$

$$= \|x\|_\infty \max_{i=\overline{1,m}} \sum_{j=1}^{n} |a_{ij}| \Rightarrow$$

т. е. величина $\max\limits_{i=\overline{1,m}} \sum\limits_{j=1}^{n} |a_{ij}|$ — верхняя (возможно не точная) грань $\dfrac{\|Ax\|_\infty}{\|x\|_\infty}$ и осталось показать, что $\max\limits_{j=\overline{1,n}} \sum\limits_{i=1}^{m} |a_{ij}|$ — **точная** верхняя грань.

let $i_{\max}$ — номер строки матрицы, сумма модулей элементов которой максимальна: $\max\limits_{i=\overline{1,m}} \sum\limits_{j=1}^{m} |a_{ij}| = \sum\limits_{i=1}^{m} |a_{i_{\max}j}|$.

Рассмотрим вектор $z = (\operatorname{sign} a_{i_{\max}1}, \ldots, \operatorname{sign} a_{i_{\max}n})^T$:

$$\frac{\|Az\|_\infty}{\|z\|_\infty} = \frac{\left|\sum\limits_{j=1}^{n} a_{i_{\max}j} \cdot \operatorname{sign} a_{i_{\max}j}\right|}{1} = \sum_{i=1}^{m} |a_{i_{\max}j}| = \max_{i=\overline{1,m}} \sum_{j=1}^{n} |a_{ij}| \Rightarrow$$

$\dfrac{\|Ax\|_\infty}{\|x\|_\infty}$ достигает $\max\limits_{i=\overline{1,m}} \sum\limits_{j=1}^{n} |a_{ij}|$ на $z \neq 0 \Rightarrow$

$$\|A\|_\infty = \sup_{x \neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty} = \max_{i=\overline{1,m}} \sum_{j=1}^{n} |a_{ij}| \underbrace{= \|A^T\|_1}_{remark!} .$$

Матричная норма подчиненная **евклидовой векторной норме** [2]. Воспользуемся связью евклидовой нормы и скалярного произведения $\forall x \in \mathbb{R}^n$: $\|x\|^2 = (x, x) = x^T x$, а также сингулярным разложением матрицы $A \in \mathbb{R}^{m\times n}$:

$$A = U\Sigma V^T\,,$$

$$U^T U = E_m\,, \quad V^T V = E_n\,, \quad \Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)\,.$$

[2] Доказательство выходит за рамки курса, поскольку основано на применении сингулярного разложения матрицы. Для интересующихся, про сингулярное разложение можно почитать методическое пособие: Логинов Н. В. Сингулярное разложение матриц. М.: МГАПИ. 1996.

Нормы векторов и матриц

Формулы для вычисления основных матричных норм

Рассмотрим

$$\frac{\|Ax\|^2}{\|x\|^2} = \frac{x^T A^T A x}{x^T x} = \frac{x^T A^T A x}{x^T x} = \frac{x^T V \Sigma U^T U \Sigma V^T x}{x^T x} =$$

$$= \frac{(V^T x)^T \Sigma^2 (V^T x)}{(V^T x)^T (V^T x)} = \frac{z^T \Sigma^2 z}{z^T z} = \frac{\|\Sigma z\|^2}{\|z\|^2},$$

где $z = V^T x$, $V^T V = E_n$.

Замечание

Заметим, если $x \in \mathbb{R}^n$ принимает все ненулевые значения, то и z принимает все значения из $\mathbb{R}^n$ отличные от нуля $\Rightarrow$

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{z \neq 0} \frac{\|\Sigma z\|}{\|z\|} = \|\Sigma\|.$$

Нормы векторов и матриц

Формулы для вычисления основных матричных норм

$$\frac{\|Ax\|^2}{\|x\|^2} = \frac{\|\Sigma z\|^2}{\|z\|^2} = \frac{\sum\limits_{i=1}^{n} \sigma_i^2 z_i^2}{\sum\limits_{i=1}^{n} z_i^2} \leqslant \sigma_{\max}^2 \frac{\sum\limits_{i=1}^{n} z_i^2}{\sum\limits_{i=1}^{n} z_i^2} = \sigma_{\max}^2 \,,$$

где $\sigma_{\max} = \max\limits_{i=\overline{1,n}} \sigma_i \Rightarrow \|A\| = \sup\limits_{x \neq 0} \dfrac{\|Ax\|}{\|x\|} \leqslant \sigma_{\max}$.

Очевидно, что данное значение достигается на векторе $z = (0, \ldots, 0, \underbrace{1}_{j_{\max}}, 0 \ldots, 0)^T$, в котором $j_{\max}$ — номер максимального сингулярного числа матрицы $A \Rightarrow$

$$\|A\| = \sigma_{\max} = \sqrt{\lambda_{\max}\left(A^T A\right)}\,. \quad \triangleright$$