Numerical analysis
Vector norms and matrix norms

Kazuaki Tanaka

Iterative methods in numerical analysis need the knowledge of vector norms and matrix norms. Here, we consider the norms only on $\mathbb{R}^n$ but this can be extended to that on $\mathbb{C}^n$ immediately.

1 Vector norms

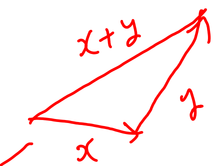
axiom of norms (公理)

Definition 1.1. *A real valued function $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ is called a vector norm, if it satisfies the following properties for all $x, y \in \mathbb{R}^n$:*

(N1) **(Positivity)** $\|x\| \geq 0$, *and* $\|x\| = 0 \Leftrightarrow x = 0$ $= \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$

(N2) **(Homogeneity)** $\|\alpha x\| = |\alpha| \, \|x\|$ *for all* $\alpha \in \mathbb{R}$

(N3) **(Triangle inequality)** $\|x + y\| \leq \|x\| + \|y\|$



1.1 l^p norm

The most important vector norm in numerical analysis is the following l^p norm

$\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}, \tag{1}$$

$x = \begin{pmatrix} -1 \\ 2 \\ -3 \end{pmatrix}$, $\|x\|_1 = 6$

where $1 \leq p < \infty$. The l^∞ norm (also called the max norm) is defined by $\|x\|_\infty = 3$

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i| \,. \tag{2}$$

In practice, we usually consider the case where $p = 1, 2$, or ∞.

Proposition 1.2. *l^1, l^2, and l^∞ norms satisfy the properties* (N1)-(N3).

Practice 1.3. *Prove Proposition 1.2.*
Note: The proof for l^1 or l^∞ is easy. For l^2, (N1) *and* (N2) *also can be confirmed easily. The difficulty is to prove* (N3) *with l^2; use the Cauchy-Schwarz inequality:*

$$\sum_{i=1}^{n} a_i b_i \leq \left(\sum_{i=1}^{n} a_i^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} b_i^2 \right)^{\frac{1}{2}}, a_i > 0, b_i > 0, i = 1, 2, \cdots, n \tag{3}$$

Definition 1.4. *Let $\|\cdot\|_A$ and $\|\cdot\|_B$ be norms in $\mathbb{R}^n$. $\|\cdot\|_A$ and $\|\cdot\|_B$ are called* **"equivalent"** *when there exist positive numbers M and M' satisfying*

$$M' \|x\|_{\underset{B}{A}} \leq \|x\|_{\underset{A}{B}} \leq M \|x\|_{\underset{B}{A}} \,. \tag{4}$$

In fact, any norms in finite-dimensional spaces are equivalent (the proof is not easy).

Proposition 1.5. *For $x \in \mathbb{R}^n$, we have*

$$\|x\|_2 \overset{②}{\leq} \|x\|_1 \overset{③}{\leq} \sqrt{n} \, \|x\|_2 \,, \tag{5}$$

$$\|x\|_\infty \underset{\text{① and ②}}{\leq} \|x\|_1 \underset{\text{ok}}{\leq} n \, \|x\|_\infty \,, \tag{6}$$

$$\|x\|_\infty \underset{①}{\leq} \|x\|_2 \underset{\text{ok}}{\leq} \sqrt{n} \, \|x\|_\infty \,. \tag{7}$$

① $\|x\|_\infty \le \|x\|_2$

$x = \begin{pmatrix} x_1 \\ \vdots\, x_k \\ x_n \end{pmatrix} \in \mathbb{R}^n$, $\|x\|_\infty \left(:= \max_i |x_i|\right) = |x_k|$

$$\|x\|_\infty = |x_k| = \left(|x_k|^2\right)^{1/2}$$
$$= \left(0^2 + \cdots + 0^2 + |x_k|^2 + 0^2 + \cdots + 0^2 + 0^2\right)^{1/2}$$
$$\le \left(|x_1|^2 + \cdots + |x_k|^2 + \cdots + |x_n|^2\right)^{1/2}$$
$$= \|x\|_2$$

② $\|x\|_2 \le \|x\|_1$

$$\|x\|_2^{\cancel{2}} = \sum_i |x_i|^2 = \sum_i \underbrace{|x_i|}_{\text{maximize}} |x_i|$$
$$\le \underbrace{\max_i |x_i|}_{\text{using ① } \|x\|_\infty} \cdot \underbrace{\sum_i |x_i|}_{\|x\|_1}$$
$$\le \cancel{\|x\|_2}\, \|x\|_1$$

③ $\|x\|_1 \le \sqrt{n}\, \|x\|_2$

C-S ineq

$$\|x\|_1 = \sum_{i=1}^n |x_i| \cdot 1 \le \underbrace{\left(\sum_i |x_i|^2\right)^{1/2}}_{\|x\|_2} \underbrace{\left(\sum_i 1^2\right)^{1/2}}_{\sqrt{n}}$$
$$= \sqrt{n}\, \|x\|_2$$

2 Matrix norms (operator norms for matrices)

Definition 2.1. *Let $A \in \mathbb{R}^{n\times n}$ be ~~a~~ an $n \times n$ real matrix. The ~~fun~~ ra ctional defined by*

$$\|A\| := \max_{x\in\mathbb{R}^n\setminus\{0\}} \frac{\|Ax\|}{\|x\|} \tag{8}$$

vector norms

is called **the matrix norm induced by the vector norm**.

Proposition 2.2. *Let $\|\cdot\|$ is the matrix norm defined above. It follow that*

same

$$\|A\| = \max_{\|x\|=1} \|Ax\|. \tag{9}$$

$$\|A\| = \max_{x\in\mathbb{R}^n\setminus\{0\}} \frac{\|Ax\|}{\|x\|}$$

$$= \max_{x\in\mathbb{R}^n\setminus\{0\}} \Big\| A \cdot \overbrace{\frac{x}{\|x\|}}^{=y} \Big\| , \ \|y\|=1 \quad \left(\|y\| = \frac{\|x\|}{\|x\|} = 1\right)$$

$$= \max_{\substack{\cancel{x\in\mathbb{R}^n\setminus\{0\}} \\ \|y\|=1}} \|Ay\|$$

Proposition 2.3. *For all $A, B \in \mathbb{R}^{n\times n}$:*

same as vectors

1. $\|A\| \geq 0$, *and* $\|A\| = 0 \Leftrightarrow A = 0$
2. $\|\alpha A\| = |\alpha| \, \|A\|$ *for all* $\alpha \in \mathbb{R}$
3. $\|A + B\| \leq \|A\| + \|B\|$
4. $\|Ax\| \leq \|A\| \, \|x\|$ *for all* $x \in \mathbb{R}^n$ $\quad \frac{\|Ax\|}{\|x\|} \leq \|A\|$
5. $\|AB\| \leq \|A\| \, \|B\|$

Definition 2.4. *l^p norm for matrix $A \in \mathbb{R}^{n\times n}$ is defined by*

$$\|A\|_p := \max_{x\in\mathbb{R}^n\setminus\{0\}} \frac{\|Ax\|_p}{\|x\|_p} \tag{10}$$

Proposition 2.5. *Let $A = (a_{ij}) \in \mathbb{R}^{n\times n}$. We have*

$p=1$ $\qquad$ $p=\infty$

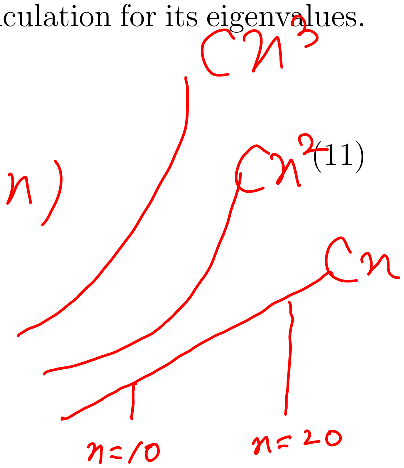
$$\|A\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{n} |a_{ij}|, \qquad \|A\|_\infty = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}|.$$

Proposition 2.6. *Let $A = (a_{ij}) \in \mathbb{R}^{n\times n}$. Let $\lambda_1,\lambda_2,\cdots,\lambda_n$ be the eigenvalues of A^TA (which are nonnegative), where A^T denotes the transposed matrix of A. Then, we have*

$$\|A\|_2 = \sqrt{\max_{1\leq i\leq n} \lambda_i}. \qquad \text{Time} = O(n^3)$$

In general, $\|A\|_2$ is difficult to compute because it requires the calculation for its eigenvalues. The Frobenius norm

$$\|A\|_F = \sqrt{\sum_{i,j=1}^{n} |a_{ij}|^2} \qquad O(n) \tag{11}$$



is useful for estimating $\|A\|_2$.

Proposition 2.7. $\|A\|_2 \leq \|A\|_F \leq \sqrt{n}\,\|A\|_2$ *for all* $A \in \mathbb{R}^{n\times n}$.

opnorm(A,2) $\qquad$ norm(A)