5.3 向量和矩阵的范数

朱升峰

华东师范大学 数学科学学院

2022 年 11 月

目录

- 向量范数
- 矩阵范数

向量范数

为研究线性方程组近似解的误差估计和迭代法的收敛性，对 $\mathbb{R}^n$ 中向量 (或 $\mathbb{R}^{n\times n}$ 中矩阵) 的“大小”引进某种度量——向量 (或矩阵) 范数的概念. 向量范数的概念是二维、三维欧氏空间中向量长度的推广.

首先将向量长度概念推广到 $\mathbb{R}^n$(或 $\mathbb{C}^n$) 中

定义 1

设 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}, \boldsymbol{y}=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}\in\mathbb{R}^n$ (或 $\mathbb{C}^n$. 将实数 $(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{x}=\sum\limits_{i=1}^{n}x_iy_i$ (或复数 $(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{y}^{\mathrm{H}}\boldsymbol{x}=\sum\limits_{i=1}^{n}x_i\bar{y}_i$) 称为向量 $\boldsymbol{x},\boldsymbol{y}$ 的数量积. 将非负实数 $\|\boldsymbol{x}\|_2=(\boldsymbol{x},\boldsymbol{x})^{\frac{1}{2}}=\left(\sum\limits_{i=1}^{n}x_i^2\right)^{\frac{1}{2}}$ (或 $\|\boldsymbol{x}\|_2=(\boldsymbol{x},\boldsymbol{x})^{\frac{1}{2}}=(\sum\limits_{i=1}^{n}|x_i|^2)^{\frac{1}{2}}$) 称为向量 $\boldsymbol{x}$ 的欧氏范数.

下述定理可在线性代数书中找到.

定理 2

设 $x, y \in \mathbb{R}^n$ (或$\mathbb{C}^n$), 则

(1) $(\boldsymbol{x}, \boldsymbol{x}) \geqslant 0$, 等号当且仅当 $\boldsymbol{x} = 0$ 时成立;

(2) $(\alpha \boldsymbol{x}, \boldsymbol{y}) = \alpha(\boldsymbol{x}, \boldsymbol{y}), \alpha$ 为实数 *(*或 $(\boldsymbol{x}, \alpha \boldsymbol{y}) = \bar{\alpha}(\boldsymbol{x}, \boldsymbol{y}), \alpha$ 为复数*)*;

(3) $(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{y}, \boldsymbol{x})($ 或 $(\boldsymbol{x}, \boldsymbol{y}) = (\overline{\boldsymbol{y}, \boldsymbol{x}}))$;

(4) $(\boldsymbol{x}_1 + \boldsymbol{x}_2, \boldsymbol{y}) = (\boldsymbol{x}_1, \boldsymbol{y}) + (\boldsymbol{x}_2, \boldsymbol{y})$;

(5) Cauchy-Schwarz

$$|(\boldsymbol{x}, \boldsymbol{y})| \leqslant \|\boldsymbol{x}\|_2 \|\boldsymbol{y}\|_2,$$

等式当且仅当 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 线性相关时成立;

(6) 三角不等式

$$\|\boldsymbol{x} + \boldsymbol{y}\|_2 \leq \|\boldsymbol{x}\|_2 + \|\boldsymbol{y}\|_2$$

向量范数的定义

定义 3 (向量的范数)

如果向量 $\boldsymbol{x} \in \mathbb{R}^n$(或 $\mathbb{C}^n$) 的某个实值函数 $N(\boldsymbol{x}) = \|\boldsymbol{x}\|$, 满足条件:
(1) $\|\boldsymbol{x}\| \geqslant 0(\|\boldsymbol{x}\| = 0$ 当且仅当 $\boldsymbol{x} = \mathbf{0})$ (正定条件),
(2) $\|\alpha \boldsymbol{x}\| = |\alpha| \|\boldsymbol{x}\|, \forall \alpha \in \mathbb{R}$ (或 $\alpha \in \mathbb{C}$),
(3) $\|\boldsymbol{x} + \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ (三角不等式),
则称 $N(\boldsymbol{x})$ 是 $\mathbb{R}^n$ (或 $\mathbb{C}^n$) 上的一个向量范数(或模). 且由 (3) 可推出不等式
(4) $|\|\boldsymbol{x}\| - \|\boldsymbol{y}\|| \leqslant \|\boldsymbol{x} - \boldsymbol{y}\|$.

常用的向量范数

(1) 向量的 ∞-范数 (最大范数):

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

容易验证这样定义的向量 $\boldsymbol{x}$ 的函数 $N(\boldsymbol{x}) := \|\boldsymbol{x}\|_\infty$ 满足向量范数的三个条件.

(2) 向量的 1-范数:

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

同样可证 $N(\boldsymbol{x}) = \|\boldsymbol{x}\|_1$ 是 $\mathbb{R}^n$ 上的向量范数.

(3) 向量的 2-范数:

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

由定理 2 知 $N(\boldsymbol{x}) = \|\boldsymbol{x}\|_2$ 是 $\mathbb{R}^n$ 上一个向量范数, 称为向量 $\boldsymbol{x}$ 的欧氏范数.

(4) 向量的 p-范数:

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

其中 $p \in [1, \infty)$, 可以证明向量函数 $N(\boldsymbol{x}) \equiv \|\boldsymbol{x}\|_p$ 是 $\mathbb{R}^n$ 上向量的范数, 上述三种范数是 p-范数的特殊情况 ($\|\boldsymbol{x}\|_\infty = \lim_{p\to\infty} \|\boldsymbol{x}\|_p$).

定义 4

设 $\{\boldsymbol{x}^{(k)}\}$ 为 $\mathbb{R}^n$ 中一向量序列，$\boldsymbol{x}^* \in \mathbb{R}^n$，记 $\boldsymbol{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})^{\mathrm{T}}$，$\boldsymbol{x}^* = (x_1^*, x_2^*, \cdots, x_n^*)^{\mathrm{T}}$, 如果 $\lim\limits_{k\to\infty} x_i^{(k)} = x_i^* (i = 1, 2, \cdots, n)$, 则称 $\boldsymbol{x}^{(k)}$ 收敛于向量 $\boldsymbol{x}^*$, 记为

$$\lim_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}^*$$

定理 5 ($N(\boldsymbol{x})$ 的连续性)

设非负函数 $N(\boldsymbol{x}) = \|\boldsymbol{x}\|$ 为 $\mathbb{R}^n$ 上任一向量范数，则 $N(\boldsymbol{x})$ 是 $\boldsymbol{x}$ 的分量 $x_1, x_2, \cdots, x_n$ 的连续函数.

定理 6 (向量范数的等价性)

设 $\|\boldsymbol{x}\|_s$,$\|\boldsymbol{x}\|_t$ 为 $\mathbb{R}^n$ 上向量的任意两种范数，则存在常数 $c_1, c_2 > 0$ 使得对一切 $\boldsymbol{x} \in \mathbb{R}^n$ 有

$$c_1\|\boldsymbol{x}\|_s \leq \|\boldsymbol{x}\|_t \leq c_2\|\boldsymbol{x}\|_s$$

注 定理 ??不能推广到无穷维空间，由定理 ??可得到结论: 如果在一种范数意义下向量序列收敛时，则在任何一种范数意义下该向量序列均收敛.

定理 7

$\lim\limits_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}^* \Leftrightarrow \lim\limits_{k\to\infty} \left\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^*\right\| = 0$, 其中 $\|\cdot\|$ 为向量的任一范数.

矩阵范数

将向量范数推广到矩阵，视 $\mathbb{R}^{n\times n}$ 中的矩阵为 $\mathbb{R}^{n^2}$ 中的向量，则由 $\mathbb{R}^{n^2}$ 上的 2-范数可以得到 $\mathbb{R}^{n\times n}$ 中的矩阵的一种范数

$$F(\boldsymbol{A}) := \|\boldsymbol{A}\|_{\mathrm{F}} := \left(\sum_{i,j=1}^{n} a_{ij}^2\right)^{1/2}$$

称为 $\boldsymbol{A}$ 的 Frobenius 范数. 显然 $\|\boldsymbol{A}\|_{\mathrm{F}}$ 满足正定性、齐次性及三角不等式. 下面给出矩阵范数的一般定义

定义 8 (矩阵的范数)

如果矩阵 $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ 的某个非负的实值函数 $N(\boldsymbol{A}) := \|\boldsymbol{A}\|$, 满足条件:
(1) $\|\boldsymbol{A}\| \geqslant 0(\|\boldsymbol{A}\| = 0 \Leftrightarrow \boldsymbol{A} = \mathbf{0}$)(正定条件);
(2) $\|c\boldsymbol{A}\| = |c|\|\boldsymbol{A}\|, c$ 为实数 (齐次条件);
(3) $\|\boldsymbol{A}+\boldsymbol{B}\| \leqslant \|\boldsymbol{A}\| + \|\boldsymbol{B}\|$ (三角不等式);
(4) $\|\boldsymbol{A}\boldsymbol{B}\| \leqslant \|\boldsymbol{A}\|\|\boldsymbol{B}\|$.
则称 $N(\boldsymbol{A})$ 是 $\mathbf{R}^{n\times n}$ 上的一个矩阵范数 (或模).

上面我们定义的 $F(\boldsymbol{A})=\|\boldsymbol{A}\|_{\mathrm{F}}$ 就是 $\mathbb{R}^{n\times n}$ 上的一个矩阵范数.

由于在大多数与估计有关的问题中，矩阵和向量会同时参与讨论, 所以希望引进一种矩阵范数, 使其与向量范数相联系. 如要求对任何向量 $\boldsymbol{x}\in\mathbb{R}^n$ 及 $\boldsymbol{A}\in\mathbb{R}^{n\times n}$ 都成立

$$\|\boldsymbol{A}\boldsymbol{x}\|\leqslant\|\boldsymbol{A}\|\|\boldsymbol{x}\|.$$

这时称矩阵范数和向量范数相容. 为此再引进一种矩阵的范数.

定义 9 (矩阵的算子范数)

设 $\boldsymbol{x}\in\mathbb{R}^n$, $\boldsymbol{A}\in\mathbb{R}^{n\times n}$, 给出一种向量范数 $\|\boldsymbol{x}\|_v$ (如 $v=1,2$ 或 ∞), 相应地定义一个矩阵的非负函数

$$\|\boldsymbol{A}\|_v=\max_{x\neq 0}\frac{\|\boldsymbol{A}\boldsymbol{x}\|_v}{\|\boldsymbol{x}\|_v}.$$

可验证 $\|\boldsymbol{A}\|_v$ 满足定义 ??, 所以 $\|\boldsymbol{A}\|_v$ 是 $\mathbb{R}^{n\times n}$ 上矩阵的一个范数, 称为 $\boldsymbol{A}$ 的算子范数，也称从属范数.

定理 10

设 $\|\boldsymbol{x}\|_v$ 是 $\mathbb{R}^n$ 上的一个向量范数，则 $\|\boldsymbol{A}\|_v$ 是 $\mathbb{R}^{n\times n}$ 上矩阵的范数，且满足相容性条件

$$\|\boldsymbol{A}\boldsymbol{x}\|_v \leqslant \|\boldsymbol{A}\|_v\|\boldsymbol{x}\|_v.$$

这种矩阵的范数 $\|\boldsymbol{A}\|_v$ 依赖于向量范数 $\|\boldsymbol{x}\|_v$ 的具体含义, 即当给出一种具体的向量范数 $\|\boldsymbol{x}\|_v$ 时, 相应的得到了一种矩阵范数 $\|\boldsymbol{A}\|_v$.

定理 11

设 $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, 则:

(1) $\|\boldsymbol{A}\|_\infty = \max\limits_{1\leq i\leq n} \sum\limits_{j=1}^{n} |a_{ij}|$ *(*称为 $\boldsymbol{A}$ 的行范数*);*

(2) $\|\boldsymbol{A}\|_1 = \max\limits_{1\leqslant j\leqslant n} \sum\limits_{i=1}^{n} |a_{ij}|$ *(*称为 $\boldsymbol{A}$ 的列范数*);*

(3) $\|\boldsymbol{A}\|_2 = \sqrt{\lambda_{\max}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)}$ *(*称为 $\boldsymbol{A}$ 的 *2-*范数*)*, 其中 $\lambda_{\max}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)$ 表示 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 的最大征值*.*

注 对于复矩阵 ($\boldsymbol{A} \in \mathbb{C}^{n\times n}$), 上述定理中的 (1),(2) 显然也成立, 对于 (3) 应改为

$$\|\boldsymbol{A}\|_2 = \max_{x\neq 0}\left(\frac{x^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}}\right)^{1/2} = \sqrt{\lambda_{\max}\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\right)}.$$

定理 12

对任何 $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\|\cdot\|$ 为一种算子范数, 则

$$\rho(\boldsymbol{A}) \leq \|\boldsymbol{A}\|, \qquad (\text{对}\|\boldsymbol{A}\|_{\mathbb{F}}\text{也成立.})$$

反之, 对任意实数 $\varepsilon > 0$, 至少存在一种算子范数 $\|\cdot\|_\varepsilon$, 使

$$\|\boldsymbol{A}\|_\varepsilon \leqslant \rho(\boldsymbol{A}) + \varepsilon.$$

定理 13

如果 $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ 为对称矩阵, 则 $\|\boldsymbol{A}\|_2 = \rho(A)$.

定理 14

如果 $\|\boldsymbol{B}\| < 1$, 则 $\boldsymbol{I} \pm \boldsymbol{B}$ 为非奇异矩阵, 且

$$\left\|(\boldsymbol{I} \pm \boldsymbol{B})^{-1}\right\| \leqslant \frac{1}{1 - \|\boldsymbol{B}\|},$$

其中 I 是单位矩阵, $\|\cdot\|$ 是矩阵的算子范数.