矩阵论　第九次作业

第 2 章　范数理论及其应用

2.1　向量范数及其性质

定义

定义 2.1　如果 V 是数域 K 上的线性空间，对任意的 $\boldsymbol{x} \in V$，定义一个实值函数 $\|\boldsymbol{x}\|$，它满足以下三个条件：

(1) 非负性：当 $\boldsymbol{x} \neq \mathbf{0}$ 时，$\|\boldsymbol{x}\| > 0$；当 $\boldsymbol{x} = \mathbf{0}$ 时，$\|\boldsymbol{x}\| = 0$；

(2) 齐次性：$\|a\boldsymbol{x}\| = |a|\,\|\boldsymbol{x}\|\ \ (a \in K, \boldsymbol{x} \in V)$

(3) 三角不等式：$\|\boldsymbol{x}+\boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}\|\ \ (\boldsymbol{x}, \boldsymbol{y} \in V)$

则称 $\|\boldsymbol{x}\|$ 为 V 上向量 $\boldsymbol{x}$ 的范数，简称**向量范数**.

定义 2.2　满足

$$c_1 \|\boldsymbol{x}\|_\beta \leqslant \|\boldsymbol{x}\|_\alpha \leqslant c_2 \|\boldsymbol{x}\|_\beta \qquad (\forall \boldsymbol{x} \in V)$$

的范数是等价的. 即有限维空间上的不同范数是等价的.

定理

定理 2.1　设 $\|\boldsymbol{x}\|_\alpha$ 和 $\|\boldsymbol{x}\|_\beta$ 为有限维线性空间 V 上的任意两种向量范数 (他们不限于 p-范数)，则存在两个与向量 $\boldsymbol{x}$ 无关的正常数 c_1 和 c_2，使满足

$$c_1 \|\boldsymbol{x}\|_\beta \leqslant \|\boldsymbol{x}\|_\alpha \leqslant c_2 \|\boldsymbol{x}\|_\beta \qquad (\forall \boldsymbol{x} \in V)$$

定理 2.2　C^n 中的向量序列

$$\boldsymbol{x}^{(k)} = (\xi_1^{(k)}, \xi_2^{(k)}, \cdots, \xi_n^{(k)}) \quad (k = 1, 2, 3, \cdots)$$

收敛到向量 $\boldsymbol{x} = (\xi_1, \xi_2, \cdots, \xi_n)$ 的充要条件是对任意一种向量范数 $\|\cdot\|$，数列 $\{\|\boldsymbol{x}^{(k)} - \boldsymbol{x}\|\}$ 收敛于零.

例题

例 2.1　在 n 维酉空间 C^n 上，复向量 $\boldsymbol{x} = (\xi_1, \xi_2, \cdots, \xi_n)$ 的长度

$$\|\boldsymbol{x}\| = \sqrt{|\xi_1|^2 + |\xi_2|^2 + \cdots + |\xi_n|^2}$$

就是一种范数.

解 为了说明这里的 $|x|$ 时范数，只需验证它满足范数的三个条件

(1) 当 $\boldsymbol{x} \neq \boldsymbol{0}$ 时，显然 $\|\boldsymbol{x}\| > 0$；当 $\boldsymbol{x} = \boldsymbol{0}$ 时，有 $\|\boldsymbol{x}\| = 0$.

(2) 对任意的复数 a，因为

$$a\boldsymbol{x} = (a\xi_1, a\xi_2, \cdots, a\xi_n)$$

所以

$$\|a\boldsymbol{x}\| = |a|\sqrt{|\xi_1|^2 + |\xi_2|^2 + \cdots + |\xi_n|^2} = |a|\,\|\boldsymbol{x}\|$$

(3) 对于任意两个复向量 $\boldsymbol{x} = (\xi_1, \xi_2, \cdots, \xi_n)$, $\boldsymbol{y} = (\eta_1, \eta_2, \cdots, \eta_n)$，有

$$\boldsymbol{x} + \boldsymbol{y} = (\xi_1 + \eta_1, \xi_2 + \eta_2, \cdots, \xi_n + \eta_n)$$

可得

$$\begin{aligned} \|\boldsymbol{x} + \boldsymbol{y}\| \ &= \sqrt{|\xi_1 + \eta_1|^2 + |\xi_2 + \eta_2|^2 + \cdots + |\xi_n + \eta_n|^2} \\ \|\boldsymbol{x} + \boldsymbol{y}\|^2 &= (\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{x} + \boldsymbol{y}) = (\boldsymbol{x}, \boldsymbol{x}) + 2Re(\boldsymbol{x}, \boldsymbol{y}) + (\boldsymbol{y}, \boldsymbol{y}) \end{aligned}$$

因为

$$Re(\boldsymbol{x}, \boldsymbol{y}) \leqslant |(\boldsymbol{x}, \boldsymbol{y})| \leqslant= \sqrt{(\boldsymbol{x}, \boldsymbol{x}), (\boldsymbol{y}, \boldsymbol{y})} = \|\boldsymbol{x}\|\,\|\boldsymbol{y}\|$$

所以

$$\|\boldsymbol{x} + \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + 2\,\|\boldsymbol{x}\|\,\|\boldsymbol{y}\| = (\|\boldsymbol{x} + \boldsymbol{y}\|)^2$$

即 $\|\boldsymbol{x} + \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\|\,\|\boldsymbol{y}\|$.

例 2.2 证明 $\|\boldsymbol{x} + \boldsymbol{y}\| = \max\|\xi_i|$ 是 C^n 上的一种范数，这里 $\boldsymbol{x} = (\xi_1, \xi_2, \cdots, \xi_n) \in C^n$.

证 当 $\boldsymbol{x} \neq 0$ 时，有 $\|\boldsymbol{x}\| = \max_i |\xi_i| > 0$；当 $\boldsymbol{x} = \boldsymbol{0}$ 时，显然有 $\|\boldsymbol{x}\| = 0$.
又对任意的 $a \in C$，有

$$\|a\boldsymbol{x}\| = \max_i |a\xi_i| = |a| \max_i |\xi_i| = |a|\,\|\boldsymbol{x}\|$$

对 C^n 的任意两个向量 $\boldsymbol{x} = (\xi_1, \xi_2, \cdots, \xi_n)$, $\boldsymbol{y} = (\eta_1, \eta_2, \cdots, \eta_n)$，有

$$\begin{aligned} \|\boldsymbol{x} + \boldsymbol{y}\| = &\max_i |\xi_i + \eta_i| \leqslant \\ &\max_i |\xi_i| + \max_i |\eta_i| = \|\boldsymbol{x}\| + \|\boldsymbol{y}\| \end{aligned}$$

因此，$\|\boldsymbol{x}\| = \max_i |\xi_i|$ 是 C^n 上的一种范数.

例 2.3 证明 $\|\boldsymbol{x}\| = \sum_{i=1}^n |\xi_i|$ 是 C^n 上的一种范数，其中 $\boldsymbol{x} = (\xi_1, \xi_2, \cdots, \xi_n) \in C^n$.

证 当 $\boldsymbol{x} \neq \boldsymbol{0}$ 时，显然 $\|\boldsymbol{x}\| = \sum_{i=1}^{n} |\xi_i| > 0$；当 $\boldsymbol{x} = \boldsymbol{0}$ 时，有 $\|\boldsymbol{x}\| = 0$.

又对于任意 $a \in C$，有

$$\|a\boldsymbol{x}\| = |a| \sum_{i=1}^{n} |\xi_i| = |a| \, \|\boldsymbol{x}\|$$

对于任意两个向量 $\boldsymbol{x}, \boldsymbol{y} \in C^n$，有

$$\|\boldsymbol{x} + \boldsymbol{y}\| = \sum_{i=1}^{n} |\xi_i + \eta_i| \leqslant \sum_{i=1}^{n} (|\xi_i| + |\eta|) =$$
$$\sum_{i=1}^{n} |\xi_i| + \sum_{i=1}^{n} |\eta_i| = \|\boldsymbol{x} + \boldsymbol{y}\|$$

于是 $\|\boldsymbol{x}\| = \sum_{i=1}^{n} |\xi_i|$ 是 C^n 上的一种范数.

习题

习题 2.1.1 求向量 $\boldsymbol{e} = (1, 1, \cdots, 1)$ 的 l_1, l_2 及 l_∞ 范数.

解

(1) $\|\boldsymbol{e}\|_1 = \sum_{i=1}^{n} |1| = n$;

(2) $\|\boldsymbol{e}\|_2 = \sqrt{|1|^2 + |1|^2 + \cdots + |1|^2} = \sqrt{n}$;

(3) $\|\boldsymbol{e}\|_\infty = \max_i |1| = 1$.

2.1 矩阵的范数

定义

定义 2.3 设 $\boldsymbol{A} = \boldsymbol{C}^{m \times n}$，定义一个实值函数 $\|\boldsymbol{A}\|$，它满足以下三个条件：

(1) 非负性：当 $\boldsymbol{A} \neq \boldsymbol{O}$ 时，$\|\boldsymbol{A}\| > 0$；当 $\boldsymbol{A} = \boldsymbol{O}$ 时，$\|\boldsymbol{A}\| = 0$;

(2) 齐次性：$\|a\boldsymbol{A}\| = |a| \, \|\boldsymbol{A}\| \quad (a \in C)$;

(3) 三角不等式：$\|\boldsymbol{A} + \boldsymbol{B}\| \leqslant \|\boldsymbol{A}\| + \|\boldsymbol{B}\| \quad (\boldsymbol{B} \in \boldsymbol{C}^{m \times n})$ 则称 $\|\boldsymbol{A}\|$ 为 $\boldsymbol{A}$ 的**广义矩阵范数**.

若对 $\boldsymbol{C}^{m \times n}, \boldsymbol{C}^{n \times l}$ 及 $\boldsymbol{C}^{m \times l}$ 上的同类广义矩阵范数 $\|\cdot\|$，还满足下面一个条件：

(4) 相容性：$\|\boldsymbol{A}\boldsymbol{B}\| \leqslant \|\boldsymbol{A}\| \, \|\boldsymbol{B}\| \qquad (\boldsymbol{B} \in \boldsymbol{C}^{n \times l})$ 则称 $\|\boldsymbol{A}\|$ 为 $\boldsymbol{A}$ 的**矩阵范数**.

定义 2.4 对于 $C^{m \times n}$ 上的矩阵范数 $\|\cdot\|_M$ 和 $\boldsymbol{C}^m$ 与 $\boldsymbol{C}^n$ 上的同类向量范数 $\|\cdot\|_V$，如果

$$\|\boldsymbol{A}\boldsymbol{x}\|_V \leqslant \|\boldsymbol{A}\|_M \, \|\boldsymbol{x}\|_V \qquad (\forall \boldsymbol{A} \in C^{m \times n}, \ \forall bmx \in C^n)$$

则称矩阵范数 $\|\cdot\|_M$ 与向量范数 $\|\cdot\|_V$ 是相容的.

定理

定理 2.4 已知 C^m 和 C^n 上的同类向量范数 $\|\cdot\|$，设 $\boldsymbol{A} \in \boldsymbol{C}^{m\times n}$，则函数

$$\|\boldsymbol{A}\| = \max_{\|\boldsymbol{x}\|=1} \|\boldsymbol{A}\boldsymbol{x}\|$$

是 $\boldsymbol{C}^{m\times n}$ 上的矩阵范数，且与已知的向量范数相容.

定理 2.5 设 $\boldsymbol{A} = (a_{ij})_{m\times n} \in \boldsymbol{C}^{m\times n}$, $\boldsymbol{x} = (\xi_1, \xi_2, \cdots, \xi_n)^T \in \boldsymbol{C}^n$，则从属于向量 $\boldsymbol{x}$ 的三种范数 $\|\boldsymbol{x}\|_1, \|\boldsymbol{x}\|_2, \|\boldsymbol{x}\|_\infty$ 的矩阵范数计算公式依次为

(1) $\|\boldsymbol{A}\|_1 = \max_j \sum_{i=1}^m |a_{ij}|$;

(2) $\|\boldsymbol{A}\|_2 = \sqrt{\lambda_1}$，$\lambda_1$ 为 $\boldsymbol{A}^H\boldsymbol{A}$ 的最大特征值;

(3) $\|\boldsymbol{A}\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|$

通常称 $\|\boldsymbol{A}\|_1, \|\boldsymbol{A}\|_2$ 及 $\|\boldsymbol{A}\|_\infty$ 依次为**列和范数、谱范数**及**行和范数**.

例题

例 2.8 设 $\boldsymbol{A} = (a_{ij})_{m\times n} \in C^{m\times n}$，证明函数

$$\|\boldsymbol{A}\|_F = (\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|)^{\frac{1}{2}} = (tr(\boldsymbol{A}^H\boldsymbol{A}))^{\frac{1}{2}}$$

是 $C^{m\times n}$ 上的矩阵范数，且与向量范数 $\|\cdot\|_2$ 相容.

证 显然 $\|\boldsymbol{A}\|_F$ 具有非负性与齐次性. 设 $\boldsymbol{B} \in \boldsymbol{C}^{m\times n}$，且 $\boldsymbol{A}$ 的第 j 列分别为 $\boldsymbol{a}_j, \boldsymbol{b}_j (j = 1, 2, \cdots, n)$，则有

$$\begin{aligned}\|\boldsymbol{A}+\boldsymbol{B}\|_F^2 = \|\boldsymbol{a}_1+\boldsymbol{b}_1\|_2^2 + \cdots + \|\boldsymbol{a}_n+\boldsymbol{b}_n\|_2^2 \leqslant \\ (\|\boldsymbol{a_1}\|_2 + \|\boldsymbol{b}_1\|_2) + \cdots + (\|\boldsymbol{a_n}\|_2 + \|\boldsymbol{b}_n\|_2)^2 = \\ (\|\boldsymbol{a_1}\|_2^2 + \cdots + \|\boldsymbol{a}_n\|_2^2) + \\ 2(\|\boldsymbol{a_1}\|_2 \|\boldsymbol{b_1}\|_2 + \cdots + \|\boldsymbol{a_n}\|_2 \|\boldsymbol{b_n}\|_2) \\ (\|\boldsymbol{b_1}\|_2^2 + \cdots + \|\boldsymbol{b}_n\|_2^2)\end{aligned}$$

可得

$$\|\boldsymbol{A}+\boldsymbol{B}\|_F^2 \leqslant \|\boldsymbol{A}\|_F^2 + 2\|\boldsymbol{A}\|_F \|\boldsymbol{B}\|_F + \|\boldsymbol{B}\|_F^2 = (\|\boldsymbol{A}\|_F + \|\boldsymbol{B}\|_F)^2$$

即三角不等式成立.

再设 $\boldsymbol{B} = (b_{ij})n \times l \in \boldsymbol{C}^{n\times l}$，则 $\boldsymbol{AB} = (\sum_{k=1}^n a_{ik}b_{kj})_{m\times l} \in \boldsymbol{C}^{m\times l}$，于是有

$$\|\boldsymbol{AB}\|_F^2 = \sum_{i=1}^m \sum_{j=1}^l \left|\sum_{k=1}^n a_{ik}b_{kj}\right|^2 \leqslant \sum_{i=1}^m \sum_{j=1}^l (\sum_{k=1}^n |a_{ik}| |b_{kj}|)^2$$

可得

$$\|\boldsymbol{AB}\|_F^2 \leqslant = \sum_{i=1}^{m}\sum_{j=1}^{l}[(\sum_{k=1}^{n}|a_{ik}|^2)(\sum_{k=1}^{n}|b_{kj}|^2)] =$$
$$(\sum_{i=1}^{m}\sum_{k=1}^{n}|a_{ik}|^2)(\sum_{j=1}^{l}\sum_{k=1}^{n}|b_{kj}|^2) = \|\boldsymbol{A}\|_F^2\|\boldsymbol{B}\|_F^2$$

即 $\|\boldsymbol{A}\|_F$ 是 $\boldsymbol{A}$ 的矩阵范数.

取 $\boldsymbol{B} = \boldsymbol{x} \in \boldsymbol{C}^{n\times l}$，则有

$$\|\boldsymbol{Ax}\|_2 = \|\boldsymbol{AB}\|_F \leqslant \|\boldsymbol{A}\|_F\|\boldsymbol{B}\|_F = \|\boldsymbol{A}\|_F\|\boldsymbol{x}\|_2$$

即矩阵范数 $\|\cdot\|_F$ 与向量范数 $\|\cdot\|_2$ 相容.

例 2.9 设 $\|\cdot\|_M$ 是矩阵范数，任取 C^n 中的非零列向量 $\boldsymbol{y}$，则函数

$$\|\boldsymbol{x}\|_V = \left\|\boldsymbol{x}\boldsymbol{y}^H\right\|_M \qquad (\forall \boldsymbol{x} \in \boldsymbol{C}^n)$$

是 $\boldsymbol{C}^n$ 上的向量范数，且矩阵范数 $\|\cdot\|_M$ 与向量范数 $\|\cdot\|_V$ 相容.

证 非负性. 当 $\boldsymbol{x} \neq \boldsymbol{0}$ 时，$\boldsymbol{x}\boldsymbol{y}^H \neq \boldsymbol{O}$，从而 $\|\boldsymbol{x}\|_V > 0$；当 $\boldsymbol{x} = \boldsymbol{0}$ 时，$\boldsymbol{x}\boldsymbol{y}^H = \boldsymbol{O}$，从而 $\|\boldsymbol{x}\|_V = 0$.

齐次性. 对任意 $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{C}^n$，有

$$\|k\boldsymbol{x}\|_V = \left\|k\boldsymbol{x}\boldsymbol{y}^H\right\|_M = |k|\left\|\boldsymbol{x}\boldsymbol{y}^H\right\|_M = \left\|k\boldsymbol{x}\boldsymbol{y}^H\right\|_M = |k|\,\|\boldsymbol{x}\|_V$$

三角不等式. 对任意 $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{C}^n$，有

$$\|\boldsymbol{x}_1 + \boldsymbol{x}_2\|_V = \left\|(\boldsymbol{x}_1 + \boldsymbol{x}_2)\boldsymbol{y}^H\right\|_M = \left\|\boldsymbol{x}_1\boldsymbol{y}^H + \boldsymbol{x}_2\boldsymbol{y}^H\right\|_M \leqslant$$
$$\left\|\boldsymbol{x}_1\boldsymbol{y}^H\right\|_M + \left\|\boldsymbol{x}_2\boldsymbol{y}^H\right\|_M = \|\boldsymbol{x}_1\|_V + \|\boldsymbol{x}_2\|_V$$

因此，$\|\boldsymbol{x}\|_V$ 是 $\boldsymbol{C}^n$ 上的向量范数. 当 $\boldsymbol{A} \in \boldsymbol{C}^{n\times n}$，$\boldsymbol{x} \in \boldsymbol{C}^n$ 时，有

$$\|\boldsymbol{Ax}\|_V = \left\|(\boldsymbol{Ax})\boldsymbol{y}^H\right\|_M = \left\|\boldsymbol{x}_1\boldsymbol{y}^H + \boldsymbol{x}_2\boldsymbol{y}^H\right\|_M \leqslant$$
$$\|A\|_M\left\|\boldsymbol{x}\boldsymbol{y}^H\right\|_M = \|\boldsymbol{A}\|_M\|\boldsymbol{x}\|_V$$

即矩阵范数 $\|\cdot\|_M$ 与向量范数 $\|\cdot\|_V$ 相容.

习题

习题 2.2.1 求矩阵 $\boldsymbol{A} = \begin{bmatrix}-1 & 2 & 1\end{bmatrix}$ 和 $\boldsymbol{B} = \begin{bmatrix}-j & 2 & 2\\ 1 & 0 & j\end{bmatrix}$ 的 $\|\cdot\|_1$，$\|\cdot\|_\infty$ 及 $\|\cdot\|_2$.

解

(1) $\|\boldsymbol{A}\|_1 = \max_j \sum_{i=1}^{m} |a_{ij}| = 2$；
$\|\boldsymbol{B}\|_1 = \max_j \sum_{i=1}^{m} |b_{ij}| = 4$；

(2) $\|\boldsymbol{A}\|_\infty = \max_i \sum_{i=j}^{m} |a_{ij}| = 4$；
$\|\boldsymbol{B}\|_\infty = \max_i \sum_{i=j}^{m} |b_{ij}| = 6$；

(3) $\|\boldsymbol{A}\|_2 = \sqrt{\lambda_1} = \sqrt{6}$；
$\|\boldsymbol{B}\|_2 = \sqrt{\lambda_1} = \sqrt{8+2\sqrt{13}}$；