第十二讲 范数理论及其应用

一、向量范数

1. 向量范数定义：设 V 为数域 K 上的向量空间，若对于 V 的任一向量 x，对应一个实值函数$\|x\|$，并满足以下三个条件：

（1）非负性 $\|x\| \geq 0$，等号当且仅当 x=0 时成立；

（2）齐次性 $\|\alpha x\| = |\alpha|\|x\|, \alpha \in k, x \in V$;

（3）三角不等式 $\|x+y\| \leq \|x\| + \|y\|, x, y \in V$。

则称$\|x\|$为 V 中向量 x 的范数，简称为向量范数。

例 1. $x \in C^n$，它可表示成$x = [\xi_1 \quad \xi_2 \quad \cdots \quad \xi_n]^T$，$\xi_i \in C$，

$$\|x\|_2 \stackrel{\Delta}{=} \left(\sum_{i=1}^{n} |\xi_i|^2\right)^{1/2}$$ 就是一种范数

证明：

（i）非负性 $\|x\|_2=\left(\sum_{i=1}^{n}|\xi_i|^2\right)^{1/2}\geq 0$，当且仅当$\xi_i=0(i=1,2,\cdots,n)$时，即 $x=0$ 时，$\|x\|_2=0$

（ii）齐次性 $\|\alpha x\|_2=\left(\sum_{i=1}^{n}|\alpha\xi_i|^2\right)^{1/2}=|\alpha|\left(\sum_{i=1}^{n}|\xi_i|^2\right)^{1/2}=|\alpha|\|x\|_2$

（iii）$y=[\eta_1 \quad \eta_2 \quad \cdots \quad \eta_n]^T$ ，$\eta_i\in C$

$$x+y=[\xi_1+\eta_1 \quad \xi_2+\eta_2 \quad \cdots \quad \xi_n+\eta_n]^T,\quad \|x+y\|_2^2=\sum_{i=1}^{n}|\xi_i+\eta_i|^2$$

$$|\xi_i+\eta_i|^2=|\xi_i|^2+|\eta_i|^2+2\mathrm{Re}\left(\overline{\xi_i}\eta_i\right)\leq|\xi_i|^2+|\eta_i|^2+2|\xi_i||\eta_i|$$

$$\|x+y\|_2^2\leq\|x\|_2^2+\|y\|_2^2+2\sum_{i=1}^{n}|\xi_i||\eta_i|$$

$$\left(\|x\|_2+\|y\|_2\right)^2=\|x\|_2^2+\|y\|_2^2+2\|x\|_2\|y\|_2$$

根据 Hölder 不等式：

$$\sum_{i=1}^{n} a_i b_i \le \left(\sum_{i=1}^{n} a_i^p\right)^{1/p} \left(\sum_{i=1}^{n} b_i^q\right)^{1/q}, \quad p,q>1, \frac{1}{p}+\frac{1}{q}=1, a_i, b_i>0$$

$$\|x\|_2 \|y\|_2 = \left(\sum_{i=1}^{n} |\xi_i|^2\right)^{1/2} \left(\sum_{i=1}^{n} |\eta_i|^2\right)^{1/2} \ge \sum_{i=1}^{n} |\xi_i||\eta_i|$$

$$\therefore \|x+y\|_2 \le \|x\|_2 + \|y\|_2$$ [证毕]

2. p-范数（l_p范数）

$$\|x\|_p = \left(\sum_{i=1}^{n} |\xi_i|^p\right)^{1/p} \quad （p\ge 1）$$

- **1-范数：** $\|x\|_1 = \sum_{i=1}^{n} |\xi_i|$
- **2-范数：** $\|x\|_2 = \left(\sum_{i=1}^{n} |\xi_i|^2\right)^{\frac{1}{2}}$
- **∞-范数（chebyshev范数）：** $\|x\|_\infty = \max |\xi_i| = \lim_{p\to\infty} \|x\|_p$

证明：$\|\mathbf{x}\|_p$显然满足非负性和齐次性

$$\mathbf{y}=[\eta_1 \quad \eta_2 \quad \cdots \quad \eta_n]^T$$

$$\|\mathbf{x}\|_p=\left(\sum_{i=1}^{n}|\xi_i|^p\right)^{1/p},\quad \|\mathbf{y}\|_p=\left(\sum_{i=1}^{n}|\eta_i|^p\right)^{1/p},\quad \|\mathbf{x}+\mathbf{y}\|_p=\left(\sum_{i=1}^{n}|\xi_i+\eta_i|^p\right)^{1/p}$$

$$\left(\|\mathbf{x}+\mathbf{y}\|_p\right)^p=\sum_{i=1}^{n}|\xi_i+\eta_i|^p=\sum_{i=1}^{n}|\xi_i+\eta_i|^{p-1}|\xi_i+\eta_i|$$

$$\le\sum_{i=1}^{n}|\xi_i+\eta_i|^{p-1}|\xi_i|+\sum_{i=1}^{n}|\xi_i+\eta_i|^{p-1}|\eta_i|$$

$$\le\left(\sum_{i=1}^{n}|\xi_i|^p\right)^{\frac{1}{p}}\cdot\left(\sum_{i=1}^{n}|\xi_i+\eta_i|^{(p-1)\cdot\frac{p}{p-1}}\right)^{1-\frac{1}{p}}+\left(\sum_{i=1}^{n}|\eta_i|^p\right)^{\frac{1}{p}}\cdot\left(\sum_{i=1}^{n}|\xi_i+\eta_i|^{(p-1)\cdot\frac{p}{p-1}}\right)^{1-\frac{1}{p}}$$

$$=\left[\left(\sum_{i=1}^{n}|\xi_i|^p\right)^{\frac{1}{p}}+\left(\sum_{i=1}^{n}|\eta_i|^p\right)^{\frac{1}{p}}\right]\cdot\left(\sum_{i=1}^{n}|\xi_i+\eta_i|^p\right)^{1-\frac{1}{p}}$$

$\therefore \left(\sum_{i=1}^{n}|\xi_i+\eta_i|^p\right)^{1/p}\le\left(\sum_{i=1}^{n}|\xi_i|^p\right)^{1/p}+\left(\sum_{i=1}^{n}|\eta_i|^p\right)^{1/p}$，即$\|\mathbf{x}+\mathbf{y}\|_p\le\|\mathbf{x}\|_p+\|\mathbf{y}\|_p$ [证毕]

∞-范数:

设 $\max_i |\xi_i| = |\xi_{i0}|$

$$\|x\|_p = \left(\sum_{i=1}^{n} |\xi_{i0}|^p \cdot \frac{|\xi_i|^p}{|\xi_{i0}|^p} \right)^{\frac{1}{p}} = |\xi_{i0}| \cdot \left(\sum_{i=1}^{n} \frac{|\xi_i|^p}{|\xi_{i0}|^p} \right)^{\frac{1}{p}}$$

$$\because \quad |\xi_{i0}|^p \le \sum_{i=1}^{n} |\xi_i|^p \le n|\xi_{i0}|^p \qquad \therefore \quad 1 \le \left(\sum_{i=1}^{n} \frac{|\xi_i|^p}{|\xi_{i0}|^p} \right)^{\frac{1}{p}} \le n^{\frac{1}{p}}$$

$$\text{又}\because \quad \lim_{p\to\infty} n^{1/p} = 1 \qquad \therefore \quad \lim_{p\to\infty} \|x\|_p = |\xi_{i0}| = \|x\|_\infty$$

3. 向量范数的等价性

定理: 设$\|\ \|_\alpha$、$\|\ \|_\beta$为C^n的两种向量范数，则必定存在正数 m、M，使得 $m\|x\|_\alpha \le \|x\|_\beta \le M\|x\|_\alpha$，（m、M 与 x 无关），它被称为向量范数的等价性。

二、矩阵范数

1. 矩阵范数定义： 设 $k^{m\times n}(k=c或R)$表示数域 k 上全体$m\times n$阶矩阵的集合。若对于$k^{m\times n}$中任一矩阵 A，均对应一个实值函数，并满足以下四个条件：

（1）非负性：$\|A\|\geq 0$ ，等号当且仅当 $A=0$ 时成立；

（2）齐次性：$\|\alpha A\|=|\alpha|\|A\|,\alpha\in k$；

（3）三角不等式：$\|A+B\|\leq\|A\|+\|B\|, A,B\in k^{m\times n}$，则称$\|A\|$为广义矩阵范数；

（4）相容性：$\|AB\|\leq\|A\|\|B\|$，则称$\|A\|$为矩阵范数。

2. 矩阵范数和向量范数

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

$$\|Ax\|_V\leq\|A\|_M\|x\|_V,\quad \forall A\in C^{m\times n},\forall x\in C^n$$

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

例：已知 $\mathbf{A}=(a_{ij})\in C^{n\times n}$ ，试证明以下两个函数都是矩阵范数：

$$\|\mathbf{A}\|_{m_1}=\sum_{i,j=1}^{n}|a_{ij}| \qquad \|\mathbf{A}\|_{m_\infty}=n\cdot\max_{i,j}|a_{ij}|$$

证明：对于函数 $\|\mathbf{A}\|_{m_1}$，其非负性和齐次性显然。下面证明三角不等式和相容性。

$$\|\mathbf{A}+\mathbf{B}\|_{m_1}=\sum_{i,j=1}^{n}|a_{ij}+b_{ij}|\le\sum_{i,j=1}^{n}(|a_{ij}|+|b_{ij}|)$$

$$=\sum_{i,j=1}^{n}|a_{ij}|+\sum_{i,j=1}^{n}|b_{ij}|=\|\mathbf{A}\|_{m_1}+\|\mathbf{B}\|_{m_1}$$

$$\|\mathbf{AB}\|_{m_1}=\sum_{i,j=1}^{n}|a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{in}b_{nj}|\le\sum_{i,j=1}^{n}(|a_{i1}||b_{1j}|+|a_{i2}||b_{2j}|+\cdots+|a_{in}||b_{nj}|)$$

$$\le\sum_{i=1}^{n}(|a_{i1}|+|a_{i2}|+\cdots+|a_{in}|)\times\sum_{j=1}^{n}(|b_{1j}|+|b_{2j}|+\cdots+|b_{nj}|)$$

$$=\sum_{i,j=1}^{n}|a_{ij}|\cdot\sum_{i,j=1}^{n}|b_{ij}|=\|\mathbf{A}\|_{m_1}\cdot\|\mathbf{B}\|_{m_1}$$

证毕

3. 常用的矩阵范数

（1）Frobenius 范数（F-范数）和导出性范数

F-范数： $\|A\|_F=\left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right)^{1/2}=(tr(A^H A))^{1/2}$

（2）导出性范数（从属范数）：设$\|x\|$为数域 K 上 n 维向量空间k^n（K=R 或 C）的一种向量范数。可定义矩阵范数为：

$$\|A\|=\max_{\|x\|=1}\|Ax\|=\max_{x\neq 0}\left(\frac{\|Ax\|}{\|x\|}\right)$$

从属于向量范数$\|x\|_1,\|x\|_2,\|x\|_\infty$的矩阵范数依次是：

$$\|A\|_1=\max_{1\le j\le n}\sum_{i=1}^{m}|a_{ij}|$$ 列（和）范数

$$\|A\|_2=\max_{1\le i\le n}\sqrt{\lambda_i(A^H A)}$$ 谱范数

$$\|A\|_\infty=\max_{1\le i\le m}\sum_{j=1}^{n}|a_{ij}|$$ 行（和）范数

三、范数的应用

逼近和误差估计是矩阵范数应用的主要领域。

定义：设矩阵A可逆，称

$$\operatorname{cond}(A)=\|A\|\|A^{-1}\|$$

为A的条件数。

由相容性可知：$\|A\|\|A^{-1}\|\ge\|AA^{-1}\|=\|I\|$

$\|x\|=\|Ix\|\le\|I\|\|x\|\Rightarrow\|I\|\ge 1$

对于导出性范数 $\|I\|=1$

$\therefore\quad \operatorname{cond}(A)\ge 1$

条件数反映了误差放大的程度，条件数越大，矩阵越病态。

1. 矩阵的摄动

定理 1：设$A\in C^{n\times n}$,且对某种矩阵范数有$\|A\|<1$，则矩阵$I-A$非奇异，且

$$\|(I-A)^{-1}\|\le\frac{\|I\|}{1-\|A\|}$$

证明：设矩阵范数$\|A\|$与向量范数$\|x\|_V$相容，假设$I-A$为奇异矩阵，则齐次方程$(I-A)x=0$有非零解x_0，即有

$$(I-A)x_0=0,\qquad x_0\ne 0$$

则：

$$\|x_0\|_V=\|Ax_0\|_V\le\|A\|\cdot\|x_0\|_V$$

与$\|A\|<1$相矛盾，故$I-A$非奇异。又$(I-A)(I-A)^{-1}=I$

所以，

$$(I-A)^{-1}=I+A(I-A)^{-1}$$

$$\|(I-A)^{-1}\|\le\|I\|+\|A\|\cdot\|(I-A)^{-1}\|\quad 即\ \|(I-A)^{-1}\|\le\frac{\|I\|}{1-\|A\|}$$

定理 2： 设 $A\in C^{n\times n}$,且对某种矩阵范数有 $\|A\|<1$，则

$$\left\|I-(I-A)^{-1}\right\|\le\frac{\|A\|}{1-\|A\|}$$

证明：因为 $\|A\|<1$，所以 $(I-A)^{-1}$ 存在。又

$$(I-A)-I=-A$$

可得 $$I-(I-A)^{-1}=-A(I-A)^{-1}$$

取范数 $$\left\|I-(I-A)^{-1}\right\|\le\|A\|\cdot\left\|(I-A)^{-1}\right\|$$

即 $$\left\|I-(I-A)^{-1}\right\|\le\frac{\|A\|}{1-\|A\|}$$

定理 3： 设 A 可逆，δA 为摄动矩阵，且 $\|A^{-1}\delta A\|<1$，则

（1）$A+\delta A$ 为可逆矩阵；

（2）$(A+\delta A)^{-1}=(I-F)A^{-1}$，其中，$\|F\|\le\dfrac{\|A^{-1}\delta A\|}{1-\|A^{-1}\delta A\|}$；

（3）$\dfrac{\|A^{-1}-(A+\delta A)^{-1}\|}{\|A^{-1}\|}\le\dfrac{\|A^{-1}\delta A\|}{1-\|A^{-1}\delta A\|}$

推论： 若 $\|A^{-1}\|\cdot\|\delta A\|<1$, 则

$$\|F\|\le\frac{cond(A)\dfrac{\|\delta A\|}{\|A\|}}{1-cond(A)\dfrac{\|\delta A\|}{\|A\|}},\qquad \frac{\|A^{-1}-(A+\delta A)^{-1}\|}{\|A^{-1}\|}\le\frac{cond(A)\dfrac{\|\delta A\|}{\|A\|}}{1-cond(A)\dfrac{\|\delta A\|}{\|A\|}}$$

例： $A=\begin{bmatrix}2 & 6\\ 2 & 6.00001\end{bmatrix}$， $\delta A=\begin{bmatrix}0 & 0\\ 0 & -0.00002\end{bmatrix}$

计算得： $A^{-1}=\begin{bmatrix}300000.5 & -300000\\ -100000 & 100000\end{bmatrix}$

$$(A+\delta A)^{-1}=\begin{bmatrix}-299999.5 & -300000\\ 100000 & -100000\end{bmatrix}$$

其条件数 $cond(A)=\|A\|_2\cdot\|A^{-1}\|_2\approx 1105$

2. 线性方程组的摄动

对于方程 $Ax=b \quad \to x=A^{-1}b$

(1)b 存在误差 Δb，求出的 x 存在误差 Δx， $\Delta x=A^{-1}\Delta b$

$\|\Delta x\|\le\|A^{-1}\|\|\Delta b\|$,考察相对误差，求 $\dfrac{\|\Delta x\|}{\|x\|}$

$$\|b\|\le\|A\|\|x\|\to\|x\|\ge\frac{\|b\|}{\|A\|} \qquad \therefore \qquad \frac{\|\Delta x\|}{\|x\|}\le\|A\|\|A^{-1}\|\frac{\|\Delta b\|}{\|b\|}$$

$$\frac{\|\Delta x\|}{\|x\|}\Big/\frac{\|\Delta b\|}{\|b\|}\le\|A\|\|A^{-1}\|=\mathrm{cond}(A)$$

(2)A 存在误差 ΔA，求出的解 x 存在误差 Δx

$$Ax=b \qquad (A+\Delta A)(x+\Delta x)=b \quad \to A\Delta x=-\Delta Ax-\Delta A\Delta x$$

忽略高阶小量得： $\|\Delta x\|\le\|A^{-1}\|\|\Delta A\|\|x\|$

$$\frac{\|\Delta x\|}{\|x\|}\le\|A^{-1}\|\|A\|\frac{\|\Delta A\|}{\|A\|}$$

$$\frac{\|\Delta x\|}{\|x\|}\Big/\frac{\|\Delta A\|}{\|A\|}\le\|A^{-1}\|\|A\|=\mathrm{cond}(A)$$

常用条件数用$\|\mathbf{A}\|_2$来考虑：

$$\|\mathbf{A}\|_2=\sqrt{\lambda_{\max}\left(\mathbf{A}^{\mathrm{H}}\mathbf{A}\right)}$$

$$\|\mathbf{A}^{-1}\|_2=\sqrt{\lambda_{\min}^{-1}\left(\mathbf{A}^{\mathrm{H}}\mathbf{A}\right)}$$

$$\mathbf{cond}(\mathbf{A})=\sqrt{\frac{\lambda_{\max}\left(\mathbf{A}^{\mathrm{H}}\mathbf{A}\right)}{\lambda_{\min}\left(\mathbf{A}^{\mathrm{H}}\mathbf{A}\right)}}$$

作业：**P275　1、2**