第二章　范数理论及其应用

2.1 向量范数

2.2 矩阵范数

2.3 范数的一些应用

2.1 向量范数及其性质

一. 向量范数的概念及l_p 范数

设向量序列$\{x^{(k)}\}\in R^n$, $x^{(k)}=(\xi_1^{(k)},\xi_2^{(k)},\cdots,\xi_n^{(k)})(k=1,2,\cdots)$

如果 $$\lim_{k\to\infty}\xi_i^{(k)}=\xi_i$$

记$x=(\xi_1,\xi_2,\cdots,\xi_n)$, 则称向量序列$\{x^{(k)}\}$有**极限$x$**,
或称$\{x^{(k)}\}$**收敛于x**, 简称$\{x^{(k)}\}$收敛。

记为 $\lim_{k\to\infty}x^{(k)}=x$ 或 $x^{(k)}\to x$

不收敛的向量序列$\{x^{(k)}\}$称为是发散的.

例：向量序列

$$x^{(k)}=\begin{pmatrix}\dfrac{1}{2^k}\\[2ex]\dfrac{\sin k}{k}\end{pmatrix}(k=1,2,3,\cdots)$$

当 $k\to\infty$ 时，$\dfrac{1}{2^k}\to 0,\quad \dfrac{\sin k}{k}\to 0,$

所以 $\lim\limits_{k\to\infty}\boldsymbol{x}^{(k)}=\begin{pmatrix}0\\0\end{pmatrix}$，故 $\{\boldsymbol{x}^{(k)}\}$ 收敛.

例：向量序列

$$x^{(k)} = \begin{pmatrix} \sum_{i=1}^{k} \frac{1}{2^i} \\ \sum_{i=1}^{k} \frac{1}{i} \end{pmatrix} \quad (k = 1, 2, 3, \cdots)$$

因为 $\sum_{i=1}^{k} \frac{1}{2^i} = \frac{1}{2} \frac{1-(1/2)^k}{1-1/2} \to 1,$

$$\sum_{i=1}^{k} \frac{1}{i} \to \infty,$$

所以是发散的.

显然，当$\boldsymbol{x}^{(k)} \to \boldsymbol{x}$时，向量$\boldsymbol{x}^{(k)} - \boldsymbol{x}$的欧式长度

$$\left|\boldsymbol{x}^{(k)} - \boldsymbol{x}\right| = \sqrt{\left(\xi_1^{(k)} - \xi_1\right)^2 + \cdots + \left(\xi_n^{(k)} - \xi_n\right)^2} \to 0$$

反之亦成立.

所以, 向量的长度可用来刻画收敛的性质.

对于一般的线性空间，如何定义向量的长度呢？

范数!!

定义 2.1 如果V是数域K上的线性空间，且对于V的任意一向量 x，对应一个实值函数，它满足以下三个条件：

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

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

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

则称$||x||$为V上向量x的范数，简称向量范数.

例2.1 复向量$\boldsymbol{x}=(\xi_1,\xi_2,\ldots,\xi_n)\in \boldsymbol{C}^n$, 长度

$$\|\mathbf{x}\|=\sqrt{|\xi_1|^2+|\xi_2|^2+\cdots+|\xi_n|^2} \quad (2.1.1)$$

是否是范数?

解: （1）对于 $\|\boldsymbol{x}\|=\sqrt{|\xi_1|^2+|\xi_2|^2+\cdots+|\xi_n|^2}$

当$x\neq o$时，$||x||>0$;当$x=o$时，$||x||=0.$

（2）$\forall a\in C, \|ax\|=\sqrt{|a\xi_1|^2+|a\xi_2|^2+\ldots+|a\xi_n|^2}=|a|\|x\|$

（3）对于$x, y \in C^n$

$$\| x+y\|^2 = (x+y, x+y) = (x,x) + 2\,\mathrm{Re}\,(x,y) + (y,y)$$

因为 $\mathrm{Re}\,(x,y) \le |(x,y)| \le \sqrt{(x,x)(y,y)} = \| x \| \| y \|$

所以 $\|x+y\|^2 \le \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$

即 $\|x+y\| \le \|x\| + \|y\|$

因此，式(2. 1. 1)是C^n上的一种范数。称这种范数为2-范数，记作$\|x\|_2$.

可以证明范数$\|x\|_2$还满足不等式

$$\big|\|x\| - \|y\|\big| \le \|x-y\|$$

两边长度之差不大于第三边的长度.

由此，可以证明向量范数是其分量的连续函数！

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

证 当$x \neq o$时，有$||x|| = \max_i |\xi_i| > 0;$

当$x = o$时，有$||x|| = 0.$

又$\forall a \in C$, 有 $\|ax\| = \max_i |a\xi_i| = |a| \max_i |\xi_i| = |a| \|x\|$

对$\forall x, y \in C^n$, 其中$x = (\xi_1, \xi_2, \cdots \xi_n)$
和$y = (\eta_1, \eta_2, \dots, \eta_n)$有

$$\| x + y \| = \max_i |\xi_i + \eta_i| \le \max_i |\xi_i| + \max_i |\eta_i| = \| x \| + \| y \|$$

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

称例2.2中的范数为$\infty -$范数，记为$||x||_\infty$

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

证 (1), (2) 略;

对$\forall x, y \in C^n$，有

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

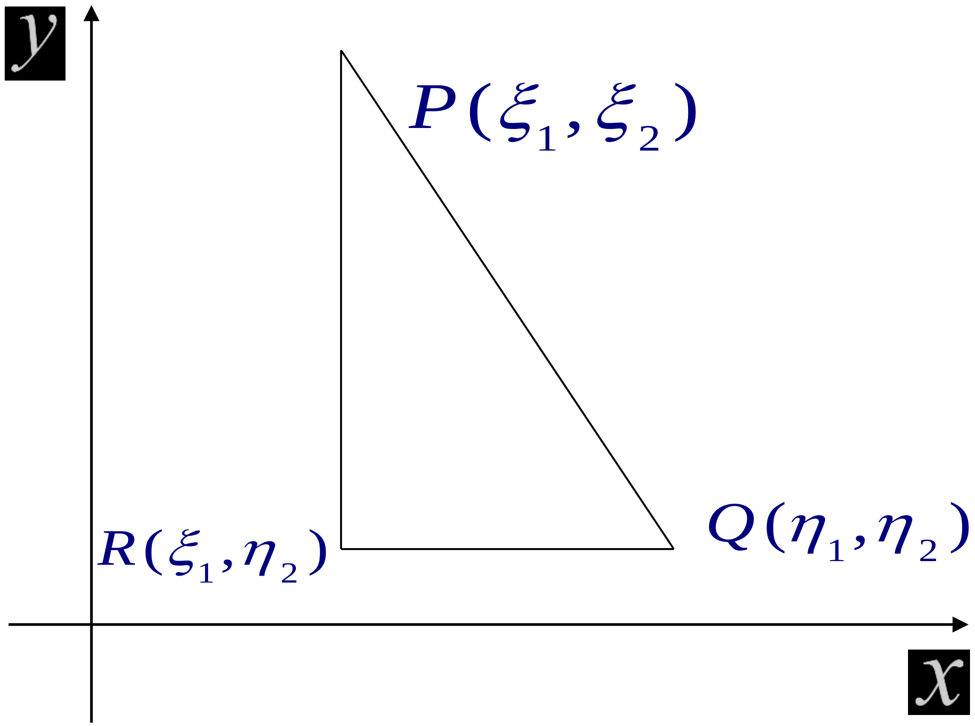
因此, $||x|| = \sum_{i=1}^{n} |\xi_i|$是$\boldsymbol{C^n}$上的一种范数.
称例2. 3中的范数为**1** $-$范数，记为$||x||_1$

***PQ*距离:**

$$\left|\xi_1-\eta_1\right|+\left|\xi_2-\eta_2\right|$$

$$\sqrt{\left(\xi_1-\eta_1\right)^2+\left(\xi_2-\eta_2\right)^2}$$

$$\max\left\{\left|\xi_1-\eta_1\right|,\left|\xi_2-\eta_2\right|\right\}$$



可以证明 $\|x\| = (\sum_{i=1}^{n} |\xi_i|^p)^{1/p}, 1 \le p < +\infty$ 是向量范数.

只证明三角不等式

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

其中 $y = (\eta_1, \eta_2, \cdots, \eta_n) \in C^n$

注：Holder不等式：

$$\sum_{i=1}^{n} |a_i b_i| \le \left(\sum_{i=1}^{n} |a_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |b_i|^q \right)^{\frac{1}{q}}$$

其中 $\frac{1}{p} + \frac{1}{q} = 1, p > 1, q > 1.$

因为

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

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

$$(\sum_{i=1}^{n}|\eta_i|^p)^{1/p}(\sum_{i=1}^{n}|\xi_i+\eta_i|^{(p-1)\frac{p}{p-1}})^{1-\frac{1}{p}}$$

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

所以，$(\sum_{i=1}^{n}|\xi_i+\eta_i|^p)^{1/p} \le (\sum_{i=1}^{n}|\xi_i|^p)^{1/p}+(\sum_{i=1}^{n}|\eta_i|^p)^{1/p}$

称 $(\sum_{i=1}^{n}|\xi_i|^p)^{\frac{1}{p}}$为向量$\boldsymbol{x}$的$\boldsymbol{p}$ −范数或$\boldsymbol{l_p}$范数，记为$\|x\|_p$.

当p=1,2, 便得$\|x\|_1, \|x\|_2$，并且 $\|x\|_\infty = \lim\limits_{p\to\infty}\|x\|_p$

证: 设 $\left|\xi_{i_0}\right| = \max\limits_i \left|\xi_i\right|, \quad \|x\|_\infty = \max\limits_i \left|\xi_i\right| = \left|\xi_{i_0}\right|.$

又 $\|x\|_p = \left(\sum_{i=1}^n \left|\xi_{i_0}\right|^p \frac{\left|\xi_i\right|^p}{\left|\xi_{i_0}\right|^p}\right)^{1/p} = \left|\xi_{i_0}\right| \left(\sum_{i=1}^n \frac{\left|\xi_i\right|^p}{\left|\xi_{i_0}\right|^p}\right)^{1/p}$

$$\left|\xi_{i_0}\right|^p \le \sum_{i=1}^n \left|\xi_i\right|^p \le n\left|\xi_{i_0}\right|^p$$

所以 $1 \le \left(\sum_{i=1}^n \frac{\left|\xi_i\right|^p}{\left|\xi_{i_0}\right|^p}\right)^{1/p} \le n^{1/p}$, 从而有 $\lim\limits_{p\to\infty}\|x\|_p = \left|\xi_{i_0}\right| = \|x\|_\infty$

例2.4 设A 是任意一个n阶对称正定矩阵，列向量 $x \in R^n$, 则函数 $\|x\|_A = \left(x^T Ax\right)^{1/2}$ 是一种向量范数，称为加权范数或椭圆范数.

证： 因为A 正定，所以当$x=0$时， $\|x\|_A = 0$
当$x \neq 0$时， $||x||_A > 0$, 即$||x||_A$具有非负性

又对任意$\forall \boldsymbol{a} \in \boldsymbol{R}$, 有

$$\|a\boldsymbol{x}\|_{\boldsymbol{A}} = \sqrt{(a\boldsymbol{x})^{\mathrm{T}} \boldsymbol{A}(a\boldsymbol{x})} = |a|\sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}} = |a|\|\boldsymbol{x}\|_{\boldsymbol{A}}$$

所以$||\boldsymbol{x}||_{\boldsymbol{A}}$具有齐次性.

由于$\boldsymbol{A}$ 为正定，所以存在非奇异矩阵$\boldsymbol{P}$使

$$\boldsymbol{A}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}.$$

于是 $\|\boldsymbol{x}\|_{\boldsymbol{A}}=\sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}}=\sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x}}=\|\boldsymbol{P}\boldsymbol{x}\|_2$

从而 $\|\boldsymbol{x}+\boldsymbol{y}\|_{\boldsymbol{A}}=\left\|\boldsymbol{P}\left(\boldsymbol{x}+\boldsymbol{y}\right)\right\|_2\le\|\boldsymbol{P}\boldsymbol{x}\|_2+\|\boldsymbol{P}\boldsymbol{y}\|_2\le\|x\|_A+\|y\|_A$

所以, $||x||_A$是一种向量范数.

例2.5 设 $\forall f(x)\in C[a,b]$，可以验证

$$\|f(t)\|_p=\left[\int_a^b |f(t)|^p dt\right]^{1/p},1\le p<\infty$$

$$\|f(t)\|_\infty=\max_{t\in[a,b]}|f(t)|$$

是线性空间 C[*a*,*b*]的范数.

例2.6 给定 V^n 的基$x_1, x_2,\dots, x_n$,设$x\in V^n$在该基下的坐标向量为$\tilde{x}=(\xi_1,\xi_2,\cdots,\xi_n)$, 那么$\|x\|_p=\|\tilde{x}\|_p$, $1\le p\le\infty$,是V^n上的范数，也称为x的p-范数.

二. 向量范数的等价性

定理2.1 设 $\|x\|_\alpha$和$\|x\|_\beta$为有限线性空间V的任意两种向量范数，则存在两个与向量无关的正常数 c_1和c_2 使下面不等式成立

$$c_1\|x\|_\beta \le \|x\|_\alpha \le c_2\|x\|_\beta$$

证 只需证每种范数与**2-范数**等价即可.

设V是n维的， $x_1, x_2, \cdots, x_n$ **是它的一个基，于是，**

对于 $\forall \boldsymbol{x} \in V$ 有 $\boldsymbol{x} = \xi_1 \boldsymbol{x}_1 + \xi_2 \boldsymbol{x}_2 + \cdots + \xi_n \boldsymbol{x}_n$,从而

$$\|\boldsymbol{x}\|_\alpha = \|\xi_1 \boldsymbol{x}_1 + \xi_2 \boldsymbol{x}_2 + \cdots + \xi_n \boldsymbol{x}_n\|_\alpha$$

令 $\phi(\xi_1, \xi_2, \cdots, \xi_n) = \|x\|_\alpha$, **证明它是连续函数**

设 $x' = \xi_1' x_1 + \xi_2' x_2 + \cdots + \xi_n' x_n \in V$, **则**

$$\left|\phi(\xi_1', \xi_2', \cdots, \xi_n') - \phi(\xi_1, \xi_2, \cdots, \xi_n)\right|$$

$$= \left|\|x'\|_\alpha - \|x\|_\alpha\right| \le \|x' - x\|_\alpha$$

$$= \|(\xi_1' - \xi_1)x_1 + (\xi_2' - \xi_2)x_2 + \cdots + (\xi_n' - \xi_n)x_n\|_\alpha$$

$$\le |\xi_1' - \xi_1|\|x_1\|_\alpha + |\xi_2' - \xi_2|\|x_2\|_\alpha + \cdots + |\xi_n' - \xi_n|\|x_n\|_\alpha$$

由于$\left\|x_i\right\|_\alpha (i=1,2,\ldots,n)$是常数，所以当$\xi_i' \to \xi_i$时

有$\phi(\xi_1',\xi_2',\ldots,\xi_n') \to \phi(\xi_1,\xi_2,\ldots,\xi_n)$

根据连续函数的性质可知，在有界闭集

$$S=\left\{(\xi_1,\ \xi_2,\cdots,\ \xi_n)\middle|\left|\xi_1\right|^2+\left|\xi_2\right|^2+\cdots+\left|\xi_n\right|^2=1\right\}$$

上，函数$\phi(\xi_1,\ \xi_2,\cdots,\ \xi_n)$可以达到最大值$\boldsymbol{c_1}$和最小值$\boldsymbol{c_2}$.

设向量 $y=\dfrac{\xi_1}{\|x\|_2}x_1+\dfrac{\xi_2}{\|x\|_2}x_2+\cdots+\dfrac{\xi_n}{\|x\|_2}x_n$

从而有

$$0 < c_1 \le \|y\|_\alpha = \phi\left(\frac{\xi_1}{\|x\|_2}, \frac{\xi_2}{\|x\|_2}, \cdots \frac{\xi_n}{\|x\|_2}\right) \le c_2$$

但 $y = \dfrac{x}{\|x\|_2}$, $\quad c_1 \le \dfrac{\|x\|_\alpha}{\|x\|_2} \le c_2$,

故 $$c_1\|x\|_2 \le \|x\|_\alpha \le c_2\|x\|_2$$ 证毕

例如: $$\|x\|_\infty \le \|x\|_1 \le n\|x\|_\infty$$

$$\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\|x\|_\infty$$

定义2.2 满足不等式(2. 1. 9)的两种范数称为是等价的.

定理2.2 C^n中的$x^{(k)}=(\xi_1^{(k)},\xi_2^{(k)},\cdots,\xi_n^{(k)}),k=1,2,\cdots,$ 收敛到向量$x=(\xi_1,\xi_2,\cdots,\xi_n)$ 的充要条件是对任一种范数 $\|\bullet\|^*$，序列 $\{\|x^{(k)}-x\|_\infty\}$收敛于0.

证 取 $\|\cdot\|=\|\cdot\|_\infty$

充分性. 设 $\|x^{(k)}-x\|_\infty\to 0$ 即$\max\limits_i\left|\xi_i^{(k)}-\xi_i\right|\to 0$

但是 $\left|\xi_j^{(k)}-\xi_j\right|\le\max\limits_i\left|\xi_i^{(k)}-\xi_i\right|(j=1,2,\cdots,n)$

从而 $\left|\xi_j^{(k)}-\xi_j\right|\to 0\quad(j=1,2,\cdots,n)$, 即 $x^{(k)}\to x$

必要性. 设 $x^{(k)} \to x$ ，则 $x^{(k)} - x \to 0$ **即向量**

$\left(\xi_1^{(k)} - \xi_1, \xi_2^{(k)} - \xi_2, \cdots, \xi_n^{(k)} - \xi_n\right)$ **的每一个分量收敛**

到零，于是对$\forall \epsilon > 0$, $\exists k_i$使得当 $k > k_i$时，

有 $\left|\xi_i^{(k)} - \xi_i\right| < \epsilon$，取$N = \max\limits_i k_i$ 当$k > N$时

有$\left|\xi_i^{(k)} - \xi_i\right| < \epsilon(i = 1,2,\cdots,n)$，**故**

$\max\limits_{i}\left|\xi_i^{(k)} - \xi_i\right| < \varepsilon$ **即** $\left\|x^{(k)} - x\right\|_\infty < \varepsilon$

从而可知序列 $\left\{\left\|x^{(k)} - x\right\|\right\}$收敛于零.　　　　**证毕**

2.2 矩阵范数

2.2.1 定义与性质

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

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

当$\boldsymbol{A}=O$时, $\|\boldsymbol{A}\|=0$

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

(3) 三角不等式: $\|\boldsymbol{A}+\boldsymbol{B}\|\leq\|\boldsymbol{A}\|+\|\boldsymbol{B}\|, \boldsymbol{B}\in\mathrm{C}^{m\times n}$

则称 $\|A\|$ 为A 的广义矩阵范数.

矩阵之间的乘法运算？

若对 $\mathbf{C}^{m\times n}$, $\mathbf{C}^{n\times l}$ 与$\mathbf{C}^{m\times l}$上的同类广义矩阵范数$\|\bullet\|$有

(4) 相容性： $\|\boldsymbol{AB}\| \le \|\boldsymbol{A}\|\|\boldsymbol{B}\|, \boldsymbol{B} \in \mathrm{C}^{n\times l}$ （2.2.1）

则称 $\|\boldsymbol{A}\|$为$\boldsymbol{A}$ 的矩阵范数.

极限：

设矩阵序列$\{\boldsymbol{A}^{(k)}\}$，其中 $\boldsymbol{A}^{(k)} \in \mathbb{C}^{m\times n}, k=1,2,\cdots,$ 用$a_{ij}^{(k)}$记 $\boldsymbol{A}^{(k)}$ 的第i行第j列的元素，且 $a_{ij}^{(k)} \to a_{ij}$ 则称$\{\boldsymbol{A}^{(k)}\}$有极限$\boldsymbol{A}$，或称 $\boldsymbol{A}^{(k)}$ 收敛于$\boldsymbol{A}$：

$$\lim_{k\to\infty} \boldsymbol{A}^{(k)} = \boldsymbol{A} \text{或} \boldsymbol{A}^{(k)} \to \boldsymbol{A}$$

不收敛的矩阵序列称为发散的.

可以证明：(1) $\boldsymbol{A}^{(k)} \to \boldsymbol{A} \Leftrightarrow \left\|\boldsymbol{A}^{(k)} - \boldsymbol{A}\right\| \to 0$

(2) $\big|\|\boldsymbol{A}\| - \|\boldsymbol{B}\|\big| \le \|\boldsymbol{A} - \boldsymbol{B}\|$

(3) 连续性. $\left\|\boldsymbol{A}^{(k)} - \boldsymbol{A}\right\| \to 0 \Rightarrow \left\|\boldsymbol{A}^{(k)}\right\| \to \|\boldsymbol{A}\|$

例2.7 已知 $A=\left(a_{ij}\right)\in C^{n\times n}$， 证明以下两函数

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

都是 $\mathbf{C}^{n\times n}$上的矩阵范数.

证：仅就三角不等式与相容性加以验证.

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

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

$$\|AB\|_{m_1} = \sum_{i,j=1} \left| a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} \right|$$

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

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

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

因此 $||A||_{m_1}$ 是 A 的矩阵范数.

注意：此处改为 $A = \left(a_{ij}\right) \in C^{m\times n}$ 依然成立 .

矩阵与向量混用？

$$Ax$$

矩阵与向量混用

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

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

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

例2.8 设 $A\in \mathbb{C}^{m\times n}$，证明函数

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

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

证：设A 的第 j 列为 $a_j\,(j=1,2,\cdots,n)$

$B\in \mathbb{C}^{m\times n}$的第 j 列为 $b_j\,(j=1,2,\cdots,n)$，则有

$$\|A+B\|_F^2=\|a_1+b_1\|_2^2+\|a_2+b_2\|_2^2+\cdots+\|a_n+b_n\|_2^2$$

$$\leq\left(\|a_1\|_2+\|b_1\|_2\right)^2+\left(\|a_2\|_2+\|b_2\|_2\right)^2+\cdots+\left(\|a_n\|_2+\|b_n\|_2\right)^2$$

$$=\|a_1\|_2^2+\|a_2\|_2^2+\cdots+\|a_n\|_2^2$$

$$+2(\|a_1\|_2\|b_1\|_2+\|a_2\|_2\|b_2\|_2+\cdots+\|a_n\|_2\|b_n\|_2)$$

$$+(\|b_1\|_2^2+\|b_2\|_2^{2\,2}+\cdots+\|b_n\|_2^2)$$

所以

$$\|A+B\|_F^2\le\|A\|_F^2+2\|A\|_F\|B\|_F+\|B\|_F^2=\left(\|A\|_F+\|B\|_F\right)^2$$

即三角不等式成立.

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

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

$$\le \sum_{i=1}^{m}\sum_{j=1}^{l}\left[(\sum_{k=1}^{n}|a_{ik}|^2)(\sum_{k=1}^{n}|b_{kj}|^2)\right]$$

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

即 $\|A\|_F$ 是 A 的矩阵范数.

取 $B = x \in C^{n\times 1}$, 则有

$$\| Ax\|_2 = \| AB\|_F \le \| A\|_F \| B\|_F = \| A\|_F \| x\|_2$$

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

范数 $\|\cdot\|_F$ 称为**Frobennius**范数，或简称为**F-范数**，又记为 $\|\cdot\|_{m_2}$ 即 m_2 -范数.

$\|A\|_F$ 的特点：

定理2.3 设 $A\in C^{m\times n}$，且 $P\in C^{m\times m}$ 与 $Q\in C^{n\times n}$ 都是酉矩阵，则 $\|PA\|_F=\|A\|_F=\|AQ\|_F$.

证：因为 $\|PA\|_F^2=\left\|P\left(a_1,a_2,\cdots,a_n\right)\right\|_F^2=\left\|Pa_1,Pa_2,\cdots,Pa_n\right\|_F^2$

$$=\sum_{i=1}^{n}\left\|Pa_i\right\|_2^2=\sum_{i=1}^{n}\left\|a_i\right\|_2^2=\|A\|_F^2$$

即 $\|PA\|_F=\|A\|_F$，于是

$$\|AQ\|_F=\left\|(AQ)^H\right\|_F=\|(Q^HA^H)\|_F=\|A^H\|_F=\|A\|_F$$

推论 与A酉相似的矩阵的F-范数是相同的.

例2.9 设$\|\cdot\|_M$是 $C^{m\times n}$上的矩阵范数，任取C^n中的非零列向量y，则函数$\|x\|_V=\|xy^H\|_M, \forall x\in C^n$ 是 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$x\neq O$时，$xy^H\neq O$, 从而$\|x\|_V>0$

当$x=O$时，$xy^H=O$, 从而$\|x\|_V=0$

齐次性. 对$\forall k\in C$，有

$$\|kx\|_V=\|kxy^H\|_M=|k|\|xy^H\|_M=|k|\|x\|_V$$

三角不等式. 对 $\forall x_1, x_2 \in C^n$，有

$$\|x_1+x_2\|_V = \left\|\left(x_1+x_2\right)y^H\right\|_M = \left\|x_1 y^H + x_2 y^H\right\|_M$$

$$\leq \left\|x_1 y^H\right\|_M + \left\|x_2 y^H\right\|_M = \|x_1\|_V + \|x_2\|_V$$

因此，$\| x \|_V$是**C^n**上的向量范数.

相容性. 当 $A \in C^{m\times n}$，$x \in C^n$ 时，

$$\| Ax\|_V = \left\|(Ax)y^H\right\|_M = \left\|A\left(xy^H\right)\right\|_M \leq \| A\|_M \left\|xy^H\right\|_M = \| A\|_M \| x\|_V$$

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

2.2.2 几种常用的矩阵范数

定理2.4 已知 C^n 与 C^m 上的同类向量范数 $\|\cdot\|$，设 $A\in C^{m\times n}$，则函数 $\|A\|=\max\limits_{\|x\|=1}\|Ax\|$ 是 $C^{m\times n}$ 上的矩阵范数，且与已知向量范数相容.

证 因为 $\|Ax\|$ 连续，所以 $\max\limits_{\|x\|=1}\|Ax\|$ 存在.

非负性. 当 $A\neq O$ 时，$\exists x_0\in C^n$ 满足 $\|x_0\|=1$，

使得 $Ax_0\neq O$，从而 $\|A\|\geq\|Ax_0\|>0$

当 $A=O$ 时，$\|A\|=\max\limits_{\|x\|=1}\|Ox\|=0.$

齐次性. 设 $a\in C$，则有

$$\|aA\|=\max_{\|x\|=1}\|aAx\|=|a|\max_{\|x\|=1}\|Ax\|=|a|\|A\|$$

三角不等式. 设 $B \in C^{m\times n}$，对于**A+B**，$\exists x_1 \in C^n$

满足$\left\| x_1 \right\| = 1$，使得$\| A+B \| = \| (A+B)x_1 \|$，于是

$$\| A+B \| = \| Ax_1 + Bx_1 \| \le \| Ax_1 \| + \| Bx_1 \| \le \| A \| + \| B \|$$

与向量范数的相容性. 对于 $\forall y \in C^n$及 $A \in C^{m\times n}$

当$y = O$，结论显然成立，当 $y \ne O$ 时，令

$\boxed{y_0 = \dfrac{1}{\|y\|} y}$，　则 $\|y_0\| = 1$，且有 $\|Ay_0\| \le \|A\|$，

于是，$\| Ay \| = \left\| A\left(\| y \| y_0 \right) \right\| = \| y \| \| Ay_0 \| \le \| y \| \| A \|$

相容性. 对 $\forall A \in C^{m\times n}, B \in C^{n\times l}$，$\exists x_2 \in C^l$

满足 $\| x_2 \| = 1$，使得 $\| AB \| = \|(AB)x_2\|$.

因为 $\|(AB)x_2\| \le \| A \| \| Bx_2 \| \le \| A \| \| B \| \| x_2 \| = \| A \| \| B \|$

所以 $\| AB \| \le \| A \| \| B \|$

因而$\|A\|$是矩阵范数，且与已知向量范数相容.

证毕

从属范数： 由向量范数导出的矩阵范数.

定理2.5 设 $A=\left(a_{ij}\right)\in C^{m\times n}, x=\left(\xi_1\xi_2\ldots\xi_n\right)^T\in C^n$ 则从属于向量 x 的三种范数 $\|x\|_1, \|x\|_2, \|x\|_\infty$ 的矩阵范数依次是:

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

(2) $\|A\|_2=\sqrt{\lambda_1}$, λ_1是A^HA的最大特征值

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

通常称 $\|A\|_1, \|A\|_2, \|A\|_\infty$ 依次为列和范数、谱范数、行和范数.

证 (1) 设$\|x\|_1=1$，则

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

$$\le \left(\max_j \sum_{i=1}^{m}|a_{ij}|\right)\sum_{j=1}^{n}|\xi_j| = \max_j \sum_{i=1}^{m}|a_{ij}|$$

因此有 $$\|A\|_1 = \max_{\|x\|_1=1}\|Ax\|_1 \le \max_j \sum_{i=1}^{m}|a_{ij}|$$

选取k,使得 $$\sum_{i=1}^{m}|a_{ik}| = \max_j \sum_{i=1}^{m}|a_{ij}|$$

因为 $$Ae_k = (a_{1k}, a_{2k}, \cdots, a_{mk})^T$$

所以 $$\|Ae_k\|_1 = \sum_{i=1}^{m}|a_{ik}| = \max_j \sum_{i=1}^{m}|a_{ij}|$$

因此 $$\|A\|_1 = \max_j \sum_{i=1}^{m}|a_{ij}|$$

(2) 因为 $A^H A$ 是Hermite 正定矩阵，所以其特征值都是非负实数，设为 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$

又 $x_1, x_2, \cdots, x_n$为对应的单位特征向量且相互正交.

于是，$\forall x \in C^n$，$\| x \|_2 = 1$ 可由特征向量线性表示，

即有 $x = k_1 x_1 + k_2 x_2 + \cdots + k_n x_n$

由于 $A^H Ax = A^H A\left(\sum_{i=1}^{n} k_i x_i\right) = \sum_{i=1}^{n} k_i \lambda_i x_i$

因此 $\|Ax\|_2^2 = (x, A^H Ax) = (\sum_{i=1}^{n} k_i x_i, \sum_{i=1}^{n} k_i \lambda_i x_i)$

$$= \lambda_1 |k_1|^2 + \lambda_2 |k_2|^2 + \cdots + \lambda_n |k_n|^2$$

$$\le \lambda_1 (|k_1|^2 + |k_2|^2 + \cdots + |k_n|^2) = \lambda_1$$

从而有 $\|A\|_2 = \max_{\|x\|_2=1} \|Ax\| \le \sqrt{\lambda_1}$

又 $\|x_1\|_2 = 1$ $\|Ax_1\|_2^2 = \left(x_1, A^H Ax_1\right) = (x_1, \lambda_1 x_1) = \lambda_1$

所以 $\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2 = \|Ax\|_1 = \sqrt{\lambda_1}$

(3) 设$\|x\|_\infty=1$，则

$$\|Ax\|_\infty=\max_i\left|\sum_{j=1}^n a_{ij}\xi_j\right|\le\max_i\sum_{j=1}^n|a_{ij}||\xi_j|\le\max_i\sum_{j=1}^n|a_{ij}|$$

选取k，使得$\sum_{j=1}^n|a_{kj}|=\max_i\sum_{j=1}^n|a_{ij}|$

令 $y=(\eta_1,\eta_2,\cdots,\eta_n)^T,\eta_i=\begin{cases}1, & a_{kj}=0\\ \dfrac{|a_{kj}|}{a_{kj}}, & a_{kj}\neq 0\end{cases}$

则有 $\|y\|_\infty=1$，且 $Ay=(*,\cdots,*,\sum_{j=1}^n|a_{kj}|,*,\cdots,*)^T$

从而 $\|Ay\|_\infty\ge\sum_{j=1}^n|a_{kj}|=\max_i\sum_{j=1}^n|a_{ij}|$

所以 $\|A\|_\infty=\max_i\sum_{j=1}^n|a_{ij}|$

例 已知矩阵 $A=\begin{pmatrix}1 & 2\\ 1 & -3\end{pmatrix}$ ，求 $\|A\|_1,\|A\|_2,\|A\|_\infty$

解 $\|A\|_1=5$

$A^HA=\begin{pmatrix}2 & -1\\ -1 & 13\end{pmatrix}$, 由于 $\lambda_{1,2}=\frac{5}{2}\left(3\pm\sqrt{5}\right)$

所以 $\|A\|_2=\left(\frac{5}{2}\left(3+\sqrt{5}\right)\right)^{\frac{1}{2}}$

$\|A\|_\infty=4$

例 已知矩阵$A=\begin{pmatrix} 1 & -2 \\ -3 & 4 \end{pmatrix}$，求 $\|A\|_1, \|A\|_2, \|A\|_\infty, \|A\|_F$

解 $\|A\|_1 = 6$

$$\|A\|_2 = \sqrt{15+\sqrt{221}}$$

$$\|A\|_\infty = 7$$

$$\|A\|_F = \sqrt{30}$$

2.3 范数的一些应用

一. 矩阵的非奇异条件

定理2.6 设 $A \in C^{n\times n}$，且对 $C^{n\times n}$ 上的某种矩阵范数 $\|\bullet\|$，有 $\|A\| < 1$，则矩阵$I-A$ 非奇异，且有

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

证 设矩阵范数 $\|\bullet\|$ 与向量范数 $\|\bullet\|_V$ 相容, λ是A 的任一特征值，x 是对应的特征向量;

因为 $Ax = \lambda x$， 所以 $\|Ax\|_V = |\lambda| \|x\|_V$;

又因 $\|Ax\|_V \le \|A\| \|x\|_V$， 所以 $|\lambda| \le \|A\| < 1$.

因此, $I - A$ 非奇异.

由 $(I - A)^{-1}(I - A) = I$, 可得

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

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

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

定理2.7 设 $A \in C^{n\times n}$，且对 $C^{n\times n}$ 上的某种矩阵范数 $\|\bullet\|$，有 $\|A\| < 1$，则

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

证 因为 $\|A\| < 1$，所以 $(I-A)^{-1}$ 存在，又

$$(I-A)\left(I-(I-A)^{-1}\right) = -A$$

设 $S = I-(I-A)^{-1}$ 得

$$S - AS = -A \quad 即 \quad S = AS - A$$

两端取范数，$\|S\| = \|AS - A\| \le \|AS\| + \|A\| \le \|A\|\|S\| + \|A\|$

所以 $\|S\| \le \dfrac{\|A\|}{1-\|A\|}$ 故 $\left\|I-(I-A)^{-1}\right\| \le \dfrac{\|A\|}{1-\|A\|}$

二. 近似逆矩阵的误差---逆矩阵的摄动

定理2.8 设 $A \in \mathbb{C}^{n\times n}$ 非奇异，$B \in \mathbb{C}^{n\times n}$，且对 $\mathbb{C}^{n\times n}$ 上的某种矩阵范数 $\|\bullet\|$，有 $\|A^{-1}B\| < 1$，则

（1）$A+B$ 非奇异；

（2）记 $F = I - (I - A^{-1}B)^{-1}$, $\|F\| \le \dfrac{\|A^{-1}B\|}{1-\|A^{-1}B\|}$

（3）$$\frac{\|A^{-1} - (A+B)^{-1}\|}{\|A^{-1}\|} \le \frac{\|A^{-1}B\|}{1-\|A^{-1}B\|}$$

证　因为 $A+B=A(I+A^{-1}B)$，又因 $\|A^{-1}B\|<1$

所以 $A+B$ 非奇异.

由定理2.7，可得（2）

再由　$A^{-1}-(A+B)^{-1}=[I-(I+A^{-1}B)^{-1}]A^{-1}$

利用（2）的结论得（3）.

若令 $\mathrm{cond}(A)=\|A\|\|A^{-1}\|$

则当 $\|A^{-1}\|\|\delta A\|<1$时，有结论（2）与（3）

可得 $$\|I-(I+A^{-1}\delta A)^{-1}\|\le\frac{cond(A)\frac{\|\delta A\|}{\|A\|}}{1-cond(A)\frac{\|\delta A\|}{\|A\|}}$$

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

称cond(A) 为矩阵A 的条件数.

三. 矩阵的谱半径及其性质

定义2.5 设 $A \in C^{n\times n}$ 的 n 个特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$

称 $\rho(A) = \max_i |\lambda_i|$ 为 A 的谱半径.

定理2.9 设 $A \in C^{n\times n}$，则对 $C^{n\times n}$ 上的任何一种矩阵范数 $\|\bullet\|$，都有 $\rho(A) \le \|A\|$.

证 设 A 的属于特征值 λ 的特征向量为 x，取与矩阵范数 $\|\bullet\|$ 相容的向量范数 $\|\bullet\|_V$，

则由 $Ax = \lambda x$，可得 $|\lambda|\|x\|_V = \|\lambda x\|_V = \|Ax\|_V \le \|A\|\|x\|_V$

因为 $x \ne O$，所以 $|\lambda| \le \|A\|$，从而 $\rho(A) \le \|A\|$.

习题1 设 $\boldsymbol{A} \in \mathrm{C}^{n\times n}$，则 $\rho(\boldsymbol{A}^k) = \rho^k(\boldsymbol{A})$.

习题2 对任意非奇异矩阵 $\boldsymbol{A} \in \mathrm{C}^{n\times n}$，则 A 的的谱范数为 $\|\boldsymbol{A}\|_2 = \rho^{1/2}(\boldsymbol{A}^H\boldsymbol{A}) = \rho^{1/2}(\boldsymbol{A}\boldsymbol{A}^H)$，当 A 是Hermite 矩阵时，则 $\|\boldsymbol{A}\|_2 = \rho(\boldsymbol{A})$.

定理2.10 设 $\boldsymbol{A} \in \mathrm{C}^{n\times n}$，对 $\forall \varepsilon > 0$，存在某种矩阵范数 $\|\bullet\|_M$，使得 $\|\boldsymbol{A}\|_M \le \rho(\boldsymbol{A}) + \varepsilon$.

证 $\exists$可逆矩阵$P \in \mathbb{C}^{n\times n}$,使得$P^{-1}AP = J$.

设 $\Lambda = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$,

$$\tilde{I} = \begin{pmatrix} 0 & \delta_1 & & & \\ & 0 & \delta_2 & & \\ & & \ddots & \ddots & \\ & & & 0 & \delta_{n-1} \\ & & & & 0 \end{pmatrix}, \delta_i = 0或1$$

则有 $J = \Lambda + \tilde{I}$. 其中$\lambda_1, \lambda_2, \cdots, \lambda_n$是 A 的n 个特征值.

令 $\boldsymbol{D} = diag(1, \varepsilon, \cdots, \varepsilon^{n-1})$，则有

$$(PD)^{-1}\boldsymbol{A}(\boldsymbol{PD}) = \boldsymbol{D}^{-1}\boldsymbol{JD} = \boldsymbol{\varLambda} + \varepsilon\tilde{\boldsymbol{I}}$$

设 ***S*=*PD*,** 那么***S*** 可逆，且有

$$\left\|\boldsymbol{S}^{-1}\boldsymbol{AS}\right\|_1 = \left\|\boldsymbol{\varLambda} + \varepsilon\tilde{\boldsymbol{I}}\right\|_1 \le \rho(\boldsymbol{A}) + \varepsilon$$

令 $\left\|\boldsymbol{A}\right\|_M = \left\|\boldsymbol{S}^{-1}\boldsymbol{AS}\right\|_1$，容易验证它是 $C^{n\times n}$ 上的矩阵范数. 所以 $\left\|\boldsymbol{A}\right\|_M \le \rho(\boldsymbol{A}) + \varepsilon$

证毕

第二章　　总结

一. 向量范数

1. 定义（非负性，齐次性，三角不等式）

2. 等价性 $c_1\|x\|_\beta \le \|x\|_\alpha \le c_2\|x\|_\beta$

3. 应用（$x^{(k)} \to x \Leftrightarrow \left\|x^{(k)} - x\right\| \to 0$）

二. 矩阵范数

1. 定义（非负性，齐次性，三角不等式，相容性）

2. 等价性

三. 应用

1. $A^{(k)} \to A \Leftrightarrow \left\| A^{(k)} - A \right\| \to 0$

2. $\rho(A) \le \|A\|$

3. $\forall \varepsilon > 0, \exists \|\bullet\|$使$\|A\| \le \rho(A) + \varepsilon$

4. ***A*****的逆的相对误差与*****A*****的条件数Cond(A)有关.**