Matrix Analysis

矩阵分析

定义：对于任何一个矩阵 $A\in C^{m\times n}$，用 $\|A\|$ 表示按照某一确定法则与矩阵 A 相对应的一个实数，且满足

（1）非负性：当 $A\neq 0,\ \|A\|>0$，当且仅当 $A=0$ 时，$\|A\|=0.$

（2）齐次性：$\|kA\|=|k|\|A\|$，k 为任意复数

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

（4）矩阵乘法的相容性：对于任意两个可以相乘的矩阵 A,B，都有

$$\|AB\| \le \|A\|\|B\|$$

那么我们称 $\|A\|$ 是矩阵 A 的范数.

例1：对于任意 $A \in C^{m\times n}$，定义

$$\|A\| = \sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|$$

可以证明如此定义的 $\|A\|$ 的确为矩阵 A 的范数.

证明：只需要验证此定义满足矩阵范数的四条性质即可。非负性，齐次性与三角不等式容易证明。现在我们验证乘法的相容性。设 $A\in C^{m\times p}, B\in C^{p\times n}$，则

$$\|AB\| = \sum_{i=1}^{m}\sum_{j=1}^{n}\left|\sum_{k=1}^{p} a_{ik}b_{kj}\right| \le \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p}|a_{ik}||b_{kj}|$$

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

$$= (\sum_{i=1}^{m}\sum_{k=1}^{p}|a_{ik}|)(\sum_{j=1}^{n}\sum_{k=1}^{p}|b_{kj}|) = \|A\|\|B\|$$

例2：设矩阵 $A\in C^{n\times n}$，证明：

$$\|A\|=n\max_{i,j}|a_{ij}|$$

是矩阵范数.

证明：非负性，齐次性和三角不等式容易证得。现在我们考虑乘法的相容性。设 $A\in C^{n\times n},B\in C^{n\times n}$，那么

$$\|AB\| = n\max_{i,j}\left|\sum_{k=1}^{n} a_{ik}b_{kj}\right| \le n\max_{i,j}\sum_{k=1}^{n}|a_{ik}||b_{kj}|$$

$$\le n \cdot n\max_{i,k}|a_{ik}|\max_{k,j}|b_{kj}|$$

$$= n\max_{i,k}|a_{ik}| \cdot n\max_{k,j}|b_{kj}|$$

$$= \|A\|\|B\|$$

因此 $\|A\|$ 为矩阵 A 的范数.

例3： 对于任意$A\in C^{m\times n}$，定义

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

可以证明$\|A\|$也是矩阵A的范数. 我们称此范数为矩阵A的**Frobenious范数**.

证明： 此定义的非负性，齐次性是显然的. 利用Minkowski不等式容易证明三角不等式. 现在我们验证乘法的相容性. 设$A\in C^{m\times l},B\in C^{l\times n}$，则

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

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

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

于是有 $\|AB\|_F \le \|A\|_F\,\|B\|_F$

Frobenious范数的酉不变性：

（1）如果$A=\begin{bmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_n\end{bmatrix}$，那么 $\|A\|_F^2=\sum_{i=1}^{n}\|\alpha_i\|_2^2$

（2）$\|A\|_F^2=\mathrm{Tr}(A^H A)=\sum_{i=1}^{n}\lambda_i(A^H A)$

（3）对于任何m阶酉矩阵U与n阶酉矩阵V都有等式

$$\|A\|_F=\|UA\|_F=\|A^H\|_F$$

$$=\|AV\|_F=\|UAV\|_F$$

关于矩阵范数的等价性定理

定理：设 $\|A\|_\alpha$，$\|A\|_\beta$ 是矩阵 A 的任意两种范数，则

总存在正数 d_1，d_2 使得

$$d_1\|A\|_\beta \le \|A\|_\alpha \le d_2\|A\|_\beta, \quad \forall A \in C^{m\times n}$$

诱导范数

定义：设 $\|X\|_\alpha$ 是向量范数，$\|A\|_\beta$ 是矩阵范数，如果对于任何矩阵 A 与向量 X 都有

$$\|AX\|_\alpha \le \|A\|_\beta \|X\|_\alpha$$

则称矩阵范数 $\|A\|_\beta$ 与向量范数 $\|X\|_\alpha$ 是相容的.

例： 矩阵的Frobenius范数与向量的2-范数是相容的.

证明： $\|AX\|_2 = \|AX\|_F \le \|A\|_F \|X\|_F = \|A\|_F \|X\|_2$

例　设 $\|X\|_\alpha$ 是向量的范数，则

$$\|A\|_i = \max_{X \neq 0} \frac{\|AX\|_\alpha}{\|X\|_\alpha}$$

满足矩阵范数的定义，且 $\|A\|_i$ 是与向量范数 $\|X\|_\alpha$ 相容的矩阵范数.

证明：首先我们验证此定义满足范数的四条性质。非负性，齐次性与三角不等式易证. 现在考虑矩阵范数的相容性.

设 $B \neq 0$，那么

$$\|AB\|_i = \max_{X\neq 0}\frac{\|ABX\|_\alpha}{\|X\|_\alpha} = \max_{BX\neq 0}\left(\frac{\|A(BX)\|_\alpha}{\|BX\|_\alpha}\frac{\|BX\|_\alpha}{\|X\|_\alpha}\right)$$

$$\leq \max_{BX\neq 0}\frac{\|A(BX)\|_\alpha}{\|BX\|_\alpha}\max_{X\neq 0}\frac{\|BX\|_\alpha}{\|X\|_\alpha}$$

$$\leq \max_{Y\neq 0}\frac{\|AY\|_\alpha}{\|Y\|_\alpha}\max_{X\neq 0}\frac{\|BX\|_\alpha}{\|X\|_\alpha}$$

$$= \|A\|_i\|B\|_i$$

因此$\|A\|_i$的确满足矩阵范数的定义.

最后证明 $\|A\|_i$ 与 $\|X\|_\alpha$ 是相容的.

由 $\|A\|_i$ 的定义可知, 当 $X \neq 0$ 时,

$$\|A\|_i \geq \frac{\|AX\|_\alpha}{\|X\|_\alpha}, \rightarrow \|AX\|_\alpha \leq \|A\|_i \|X\|_\alpha$$

当 $X = 0$ 时, $\|AX\|_\alpha = \|A\|_i \|X\|_\alpha = 0$,

这说明 $\|A\|_i$ 与 $\|X\|_\alpha$ 相容的.

定义：上面所定义的矩阵范数称为由向量范数 $\|X\|_\alpha$ 所诱导的诱导范数或算子范数. 由向量 p--范数 $\|X\|_p$ 所诱导的矩阵范数称为矩阵 p--范数. 即

$$\|A\|_p = \max_{X\neq 0}\frac{\|AX\|_p}{\|X\|_p}$$

常用的矩阵 p--范数为 $\|A\|_1$，$\|A\|_2$ 和 $\|A\|_\infty$.

定理：设 $A \in C^{m\times n}$，则

（1）列和范数 $\|A\|_1 = \max\limits_j (\sum_{i=1}^{m} |a_{ij}|)$, $j = 1, 2, \cdots, n$

（2）谱范数 $\|A\|_2 = \max\limits_j (\lambda_j (A^H A))^{1/2}$, $j = 1, 2, \cdots, n$

其中 $\lambda_j (A^H A)$ 表示矩阵 $A^H A$ 的第 j 个特征值.

（3）行和范数 $\|A\|_\infty = \max\limits_i (\sum_{j=1}^{n} |a_{ij}|)$, $i = 1, 2, \cdots, m$

证：

$$\|A\|_1 = \max_{X\neq 0}\frac{\|AX\|_1}{\|X\|_1} = \max_{\|X\|_1=1}\|AX\|_1$$

$$\|Ax\|_1 = \|A(x_1e_1+\cdots+x_ie_i+\cdots x_ne_n)\|_1$$

$$\leq |x_1|\|Ae_1\|_1+\cdots+|x_i|\|Ae_i\|_1+\cdots|x_n|\|Ae_n\|_1 \leq \max_j\|Ae_j\|_1\sum_{i=1}^{n}|x_i|$$

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

$$\therefore\quad \|A\|_1 = \max_{X\neq 0}\frac{\|Ax\|_1}{\|x\|_1} \leq \max_j\sum_{i=1}^{m}|a_{ij}|$$

又 $\|A\|_1 = \max_{X \neq 0} \frac{\|Ax\|_1}{\|x\|_1} \geq \max_j \frac{\|Ae_j\|_1}{\|e_j\|_1} = \max_j \sum_{i=1}^{m} |a_{ij}|$

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

A 的最大奇异值

（2） $\|A\|_2 = \max\limits_j(\lambda_j(A^H A))^{1/2}$,

$\lambda_j(A^H A)$ 表示矩阵 $A^H A$ 的第 j 个特征值。我们称此范数为矩阵 A 的谱范数。

证：

$$\|A\|_2 = \max_{X\neq 0}\frac{\|Ax\|_2}{\|x\|_2} = \max_{X\neq 0}\frac{\left(x^H A^H A x\right)^{1/2}}{\left(x^H x\right)^{1/2}}$$

$$= \max_j\left(\lambda_j(A^H A)\right)^{1/2}$$

（3）$\|A\|_\infty = \max_i \left(\sum_{j=1}^{n} |a_{ij}|\right), \quad i = 1, 2, \cdots, m$

我们称此范数为矩阵 A 的行和范数。 (了解)

证明: (不要求)

$$\|Ax\|_\infty = \max_{1\le i\le m}\left|\sum_{j=1}^{n} a_{ij}x_j\right| \le \max_{1\le i\le m}\sum_{j=1}^{n}|a_{ij}x_j|$$

$$\le \|x\|_\infty \cdot \max_{1\le i\le m}\sum_{j=1}^{n}|a_{ij}| \Longrightarrow \|A\|_\infty \le \max_{1\le i\le m}\sum_{j=1}^{n}|a_{ij}|$$

如果 $A=0$, 结论显然. 假设 $A\neq 0$, 定义

$Z_j=(z_k^{(j)})\in C^{n\times 1}$, 其中

$$\begin{cases} z_k^{(j)}=\dfrac{\bar{a}_{jk}}{|a_{jk}|}, & \text{if } a_{jk}\neq 0 \\ z_k^{(j)}=1, & \text{if } a_{jk}=0 \end{cases} \Rightarrow \quad \text{显然 } \|Z_j\|_\infty=1, \quad a_{jk}z_k^{(j)}=|a_{jk}|$$

$$\max_{x\neq 0}\frac{\|Ax\|_\infty}{\|x\|_\infty}\geq \|AZ_j\|_\infty=\max_{1\leq i\leq m}\left|\sum_{k=1}^n a_{ik}z_k^{(j)}\right|$$

$$\geq \sum_{k=1}^n a_{jk}z_k^{(j)}=\sum_{k=1}^n |a_{jk}| \qquad \therefore \quad \|A\|_\infty\geq \max_{1\leq j\leq m}\sum_{k=1}^n |a_{jk}|$$

例：已知 $A=\begin{bmatrix}2&-1&0\\0&2&3\\1&2&0\end{bmatrix}$，计算 $\|A\|_1$，$\|A\|_2$，$\|A\|_\infty$ 和 $\|A\|_F$.

解： $\|A\|_1=5$, $\|A\|_F=\sqrt{23}$, $\|A\|_\infty=5$,

$A^HA=\begin{bmatrix}5&0&0\\0&9&6\\0&6&9\end{bmatrix}$，所以 A^HA 的特征值为 5, 15, 3. $\|A\|_2=\sqrt{15}$.

矩阵的谱半径及其性质

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

$$\rho(A)=\max\{|\lambda_1|,|\lambda_2|,\cdots,|\lambda_n|\}$$

为矩阵 A 的谱半径.

例：设 $A\in C^{n\times n}$，那么 $\rho(A)\le\|A\|$.

其中 $\|A\|$ 是矩阵 A 的任何一种范数.

$$AX=\lambda X,\ X\neq 0$$

$$|\lambda|\|X\|=\|\lambda X\|=\|AX\|$$

$$\le\|A\|\|X\|,\ \longrightarrow\ |\lambda|\le\|A\|.$$

例：证明：对于任何矩阵$A \in C^{m\times n}$都有

$$(a)\quad \|A^H\|_1 = \|A^T\|_1 = \|A\|_\infty$$

$$(b)\quad \|A^H\|_2 = \|A^T\|_2 = \|A\|_2$$

$$(c)\quad \|A^H A\|_2 = \|A\|_2^2$$

$$(d)\quad \|A\|_2^2 \le \|A\|_1 \|A\|_\infty$$

$$(c)\ \left\|A^H A\right\|_2^2 = \max_j \lambda_j[(A^H A)^H (A^H A)]$$

$$= \max_j \lambda_j[(A^H A)^2] = [\max_j \lambda_j (A^H A)]^2 = \left\|A\right\|_2^4$$

$$\Rightarrow \left\|A^H A\right\|_2 = \left\|A\right\|_2^2$$

$$(d)\ \left\|A\right\|_2^2 = \max_j \lambda_j (A^H A) = \rho(A^H A) \le \left\|A^H A\right\|_1 \le \left\|A^H\right\|_1 \left\|A\right\|_1 = \left\|A\right\|_\infty \left\|A\right\|_1$$

如何由矩阵范数构造与之相容的向量范数?

定理：设 $\|A\|_*$ 是矩阵范数，则存在向量范数 $\|X\|$ 使得

$$\|AX\| \le \|A\|_* \|X\|$$

证明：对于任意的非零向量 α，定义向量范数

$$\|X\| = \|X\alpha^H\|_*$$

容易验证此定义满足向量范数的三个性质，且

$$\|AX\| = \|AX\alpha^H\|_* \le \|A\|_* \|X\alpha^H\|_* \le \|A\|_* \|X\|.$$

例： 已知矩阵范数

$$\|A\|_* = \|A\| = \sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|,$$

求与之相容的一个向量范数.

解： 取 $\alpha = [0 \quad 1 \quad \cdots \quad 0]^T$，设 $X = [x_1 \quad x_2 \quad \cdots \quad x_n]^T$

那么

$$\|X\| = \|X\alpha^H\|_* = \sum_{i=1}^{n}|x_i| = \|X\|_1$$

例：设 A 是一个 n 阶正规矩阵，则 $\rho(A)=\|A\|_2$

证明：设 A 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，$A$ 是正规矩阵，所以存在酉矩阵 U 使得 $A=U\,\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)U^H$，从而

$$A^H A=U\,\mathrm{diag}(|\lambda_1|^2,|\lambda_2|^2,\cdots,|\lambda_n|^2)U^H$$

所以 $\|A\|_2=\max\limits_j(\lambda_j(A^H A))^{1/2}=\max\limits_j|\lambda_j|=\rho(A).$

例：设 $\|\cdot\|$ 是 $C^{n\times n}$ 上的矩阵范数. 证明：

(1) $\|I\| \geq 1$

(2) A为可逆矩阵，λ 为 A 的特征值，则有

$$\|A^{-1}\|^{-1} \leq |\lambda| \leq \|A\|$$