§3.2 矩阵序列和矩阵级数

(Matrix Sequences，Matrix Series)

§ 3.2 矩阵序列

定义 设有矩阵序列 $\left\{A^{(k)}=\left(a_{ij}^{(k)}\right)\right\}_{k\in\mathbb{N}}\subseteq C^{n\times n}$. 若存在数列 $\{a_{ij}^{(k)}\}$, 使得

$$\lim_{k\to\infty} a_{ij}^{(k)}=a_{ij},\qquad \forall i,j=1,2,\cdots,n.$$

则称矩阵序列 $\left\{A^{(k)}\right\}$ 依元素收敛于矩阵 $A=(a_{ij})$.

记为: $\lim\limits_{k\to\infty} A^{(k)}=\boldsymbol{A}$ 或 $A^{(k)}\to\boldsymbol{A}$.

若有一个元素数列是发散的, 则称该矩阵序列发散.

说明: 矩阵序列收敛等价于序列中所有的元素同时收敛.

Example

$$\begin{bmatrix}\left(1+\dfrac{1}{k}\right)^k & 1+\dfrac{1}{k}\\ -1 & \dfrac{(-1)^k}{k}\end{bmatrix}\to\begin{bmatrix}e & 1\\ -1 & 0\end{bmatrix}.$$

§ 3.2 矩阵序列

一、矩阵序列收敛的性质

(1) 若 $A^{(k)} \to A,\ B^{(k)} \to B$, 则

$$hA^{(k)} + lB^{(k)} \to hA + lB, \qquad \forall h, l \in C;$$

(2) 若 $A^{(k)} \to A,\ B^{(k)} \to B$, 则

$$A^{(k)}B^{(k)} \to AB;$$

(3) 若 $A^{(k)} \to A$ 且 $A^{(k)}$ 与 A 均可逆, 则

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

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

§ 3.2 矩阵序列

证明
(宏观)

(2) $\| A^{(k)}B^{(k)} - AB \|$

$$= \| A^{(k)}B^{(k)} - AB^{(k)} + AB^{(k)} - AB \|$$

$A^{(k)}B^{(k)} \to AB$

$$\leq \| A^{(k)}B^{(k)} - AB^{(k)} \| + \| AB^{(k)} - AB \|$$

$$\leq \| A^{(k)} - A \| \| B^{(k)} \| + \| A \| \| B^{(k)} - B \|.$$

知 $\| A^{(k)}B^{(k)} - AB \| \to 0.$

(3) **利用零化多项式可将** A^{-1} **表示为** A **的多项式:**

$\left(A^{(k)}\right)^{-1} \to A^{-1}$

$$f(A) = A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I = 0$$

$$\Rightarrow \ A^{-1} = -\frac{1}{a_0}(A^{n-1} + \cdots + a_2 A + a_1 I) = g(A).$$

由(1)(2), **易见** $g\left(A^{(k)}\right) \to g(A).$

证明
(宏观)

(2) $\| A^{(k)}B^{(k)} - AB \|$

$$= \| A^{(k)}B^{(k)} - AB^{(k)} + AB^{(k)} - AB \|$$

$A^{(k)}B^{(k)} \to AB$

$$\leq \| A^{(k)}B^{(k)} - AB^{(k)} \| + \| AB^{(k)} - AB \|$$

$$\leq \| A^{(k)} - A \| \| B^{(k)} \| + \| A \| \| B^{(k)} - B \|.$$

知 $\| A^{(k)}B^{(k)} - AB \| \to 0.$

(3) 注意$A^{(k)}$和A均要可逆。

$\left(A^{(k)}\right)^{-1} \to A^{-1}$

可逆 $\begin{bmatrix} \frac{1}{k} & 0 \\ 0 & \frac{-1}{k} \end{bmatrix} \to \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. 不可逆

证明 (微观) (2) 令 $C^{(k)}=A^{(k)}B^{(k)}=\left(c_{ij}^{k}\right)$, 则

$$c_{ij}^{(k)}=\sum_{k} a_{ik}^{(k)} b_{kj}^{(k)} \to \sum_{k} a_{ik} b_{ik}\,,$$

即 $C^{(k)}=A^{(k)}B^{(k)} \to AB.$

(3) **3阶矩阵为例**

$$\left|A^{(k)}\right|=(-1)^{\tau(j_1,j_2,j_3)}\sum a_{1j_1}^{(k)}a_{2j_2}^{(k)}a_{3j_3}^{(k)} \quad (j_i\text{互异})$$

$$\to (-1)^{\tau(j_1 j_2 j_3)}\sum a_{1j_1}a_{2j_2}a_{3j_3}=|A|.$$

同理 $\left(A^{(k)}\right)^{*}=\begin{bmatrix} |A_{11}^{(k)}| & |A_{12}^{(k)}| & |A_{13}^{(k)}| \\ |A_{21}^{(k)}| & |A_{22}^{(k)}| & |A_{23}^{(k)}| \\ |A_{31}^{(k)}| & |A_{32}^{(k)}| & |A_{33}^{(k)}| \end{bmatrix} \to A^{*}.$

从而 $\left(A^{(k)}\right)^{-1}=\left|A^{(k)}\right|^{-1}\left(A^{(k)}\right)^{*} \to A^{-1}.$ ■

二、矩阵序列的有界性

定义 若矩阵序列 $\left\{A^{(k)}=\left(a_{ij}^{(k)}\right)\right\}_{k\in\mathbb{N}}\subseteq C^{n\times n}$ 每个位置的元素均构成有界数列，即存在数 $M>0$

s.t.
$$\sup_{k\in\mathbb{N}}\left|a_{ij}^{(k)}\right|\le M,\qquad \forall i,j=1,2,\cdots,n.$$

则称矩阵序列 $\left\{A^{(k)}\right\}$ 有界.

注 $\left\{A^{(k)}\right\}$ 有界 $\Leftrightarrow$ $\|A^{(k)}\|_1=\max\limits_j\left\{\sum_{i=1}^{n}\left|a_{ij}^{(k)}\right|\right\}\le M'$

$\Leftrightarrow$ $\left\{A^{(k)}\right\}$ 在任意矩阵范数下有界

有界矩阵序列必有收敛子列

定义 称方阵 A 为**收敛矩阵**, 如果 $\lim\limits_{k\to\infty} A^k = \mathbf{0}$.

定理 设A是方阵，则当 $\| A \| < 1$, 有 $\lim\limits_{k\to\infty} A^k = \mathbf{0}$

证明 由矩阵范数的相容性,

$$\| A^k - \mathbf{0} \| = \| A^k \| \le \| A \|^k \to 0,$$

故 $\lim\limits_{k\to\infty} A^k = \mathbf{0}$, 即 A 为收敛矩阵. ■

注意：反之不一定成立。如$A = \begin{bmatrix} \mathbf{0.8} & \mathbf{1} \\ \mathbf{0} & \mathbf{0.8} \end{bmatrix}$

$$\boldsymbol{A}^{(k)} = \begin{bmatrix} \mathbf{0.8}^k & k\mathbf{0.8}^{k-1} \\ \mathbf{0} & \mathbf{0.8}^k \end{bmatrix} \longrightarrow \mathbf{0}.$$ 而$\| A \|_{\mathbf{1}} = 1.8 > 1.$

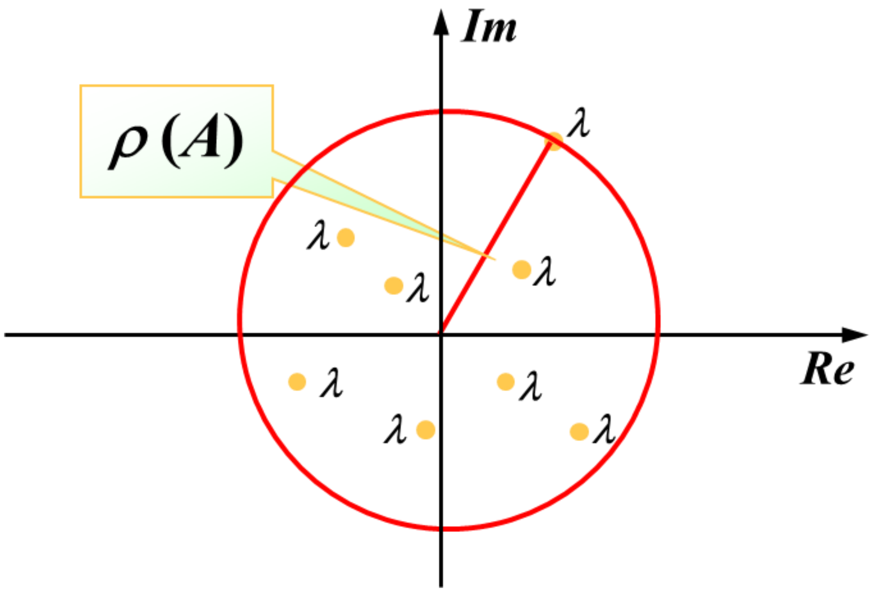
三、方阵的谱半径

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

$$\rho(A) := \max_{i=1,2,\cdots,n} |\lambda_i|$$

为 A 的**谱半径**.

(A 的所有谱分布在复平面以原点为圆心, $\rho(A)$ 为半径的圆盘上.)



三、方阵的谱半径

定理 $A^k \to 0 \Leftrightarrow \rho(A) < 1.$

证明(若当化方法) 设 $A = PJP^{-1}$, $A^k = PJ^kP^{-1}$.

$$\text{故 } A^k \to 0 \Leftrightarrow J^k = \operatorname{diag}\left(J_1^k(\lambda_1), J_2^k(\lambda_2), \cdots, J_s^k(\lambda_s)\right) \to 0$$

$$\Leftrightarrow J_i^k(\lambda_i) \to 0, \ \forall i = 1,2,\cdots,s.$$

而

$$J_i^k(\lambda_i) = \begin{bmatrix} \lambda_i^k & C_k^1\lambda_i^{k-1} & \cdots & C_k^{n_i-1}\lambda_i^{k-n_i+1} \\ & \ddots & \ddots & \vdots \\ & & \lambda_i^k & C_k^1\lambda_i^{k-1} \\ & & & \lambda_i^k \end{bmatrix} \qquad J_i^k(\lambda_i) = \begin{bmatrix} \lambda_i^k & (\lambda_i^k)' & \cdots & \frac{1}{(n_i-1)!}(\lambda_i^k)^{(n_i-1)} \\ & \ddots & \ddots & \vdots \\ & & \lambda_i^k & (\lambda_i^k)' \\ & & & \lambda_i^k \end{bmatrix}$$

$$J_i^k(\lambda_i) \to 0 \Leftrightarrow |\lambda_i| < 1, \ \forall 1 \le i \le s. \ \Leftrightarrow \rho(A) < 1. \ \blacksquare$$

四、方阵的谱半径的界

定理 设 $A \in C^{n\times n}$, 则对 $C^{n\times n}$ 上的任意范数 $\|\cdot\|$,

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

证明 设 λ 是 A 的任意特征值: $AX = \lambda X,\ X \ne 0.$

令 $B = (X, 0, \cdots, 0) \in C^{n\times n}$, 则

$$AB = (AX, 0, \cdots, 0) = \lambda B,$$

因此, $|\lambda| \| B \| = \| \lambda B \| = \| AB \| \le \| A \| \| B \|.$

因为 B 非零, $\| B \| > 0$, 故 $|\lambda| \le \| A \|.$

由 λ 的任意性, $\rho(A) \le \| A \|.$ ■

说明: 方阵的谱半径是矩阵范数的下确界.

定理 设 $A \in C^{n\times n}$, 则对任意 $\varepsilon > 0$ 存在范数 $\|\cdot\|_*$, 使得

$$\| A \|_* \le \rho(A) + \varepsilon.$$

若满足 $|\lambda| = \rho(A)$ 的特征值 λ 均对应 1 阶Jordan块,

$$\| A \|_* = \rho(A).$$

证明 设 A 的Jordan标准型 $P^{-1}AP = J = \mathrm{diag}(J_1, J_2, \cdots, J_m)$.

略，跳过

令 $D = \mathrm{diag}(1, \varepsilon, \cdots, \varepsilon^{n-1})$. 考虑线性变换 $J \to D^{-1}JD$.

该变换将 J 的Jordan块 J_i 变为 $\widehat{J_i} = \begin{bmatrix} \lambda_i & \varepsilon & & \\ & \lambda_i & \ddots & \\ & & \ddots & \varepsilon \\ & & & \lambda_i \end{bmatrix}$.

于是 $\| (PD)^{-1}APD \|_\infty = \| D^{-1}JD \|_\infty = \max \| \widehat{J_i} \|_\infty$

$$\le \max |\lambda_i| + \varepsilon = \rho(A) + \varepsilon.$$

令 $\| A \|_* = \| (PD)^{-1}APD \|_\infty$, 此即满足条件的范数.

定理 设 $A \in C^{n\times n}$, 则对任意 $\varepsilon > 0$ 存在范数 $\|\cdot\|_*$, 使得

$$\| A \|_* \le \rho(A) + \varepsilon.$$

若满足 $|\lambda| = \rho(A)$ 的特征值 λ 均对应 1 阶Jordan块,

$$\| A \|_* = \rho(A).$$

证明 设 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 为互异特征值并按 $|\lambda_i|$ 大小排列, 则存在 $1 \le k \le s$ 和 $\delta > 0$ 使得

$$|\lambda_1| = \cdots = |\lambda_k| = \rho(A), \quad |\lambda_i| \le \rho(A) - \delta, \quad i > k.$$

因此对所有的 $\varepsilon < \delta$, 由于 $J_1, \cdots, J_k$ 为 1 阶若当块, 有

$$\begin{cases} \| \widehat{J}_i \|_\infty = |\lambda_1| = \rho(A), & i \le k, \\ \| \widehat{J}_i \|_\infty \le |\lambda_i| + \varepsilon < \rho(A), & i > k. \end{cases}$$

从而, $\| A \|_* = \| (PD)^{-1} APD \|_\infty = \max \| \widehat{J}_i \|_\infty$

$$= \rho(A). \quad \blacksquare$$

§3.2　矩阵幂级数

(Matrix Power Series)

$$f(A)=\sum_{k=0}^{\infty} a_k A^k$$

§ 3.2 矩阵幂级数

一、矩阵幂级数的定义

定义 设 $A \in C^{n\times n}$, $a_k \in C$. A 的以a_k为系数的**幂级数**为

$$\sum_{k=0}^{\infty} a_k A^k = a_0 I + a_1 A + a_2 A^2 + \cdots .$$

定义 称矩阵幂级数 $\sum_{k=0}^{\infty} a_k A^k$ 是收敛的, 如果其**部分和序列** $S_N = \sum_{k=0}^{N} a_k A^k$ 是收敛的. 若 $\lim_{N\to\infty} S_N = S$, 则称 S 为幂级数 $\sum_{k=0}^{\infty} a_k A^k$ 的**和矩阵**.

§ 3.2 矩阵幂级数

下面讨论矩阵幂级数收敛性的判别.

对复变量 z 的幂级数 $f(z)=\sum_{k=0}^{\infty} a_k z^k$, 其收敛半径定义为

$$R=\sup\left\{|x|: f(|x|)=\sum_{k=0}^{\infty} a_k |x|^k \text{ 收敛}\right\}.$$

引理(补) 对 $f(z)=\sum_{k=0}^{\infty} a_k z^k$, 如果

$$\lim_{n\to\infty}\frac{|a_{n+1}|}{|a_n|}=\rho, \text{ 或 } \lim_{n\to\infty}\sqrt[n]{|a_n|}=\rho,$$

其中, $0\le\rho\le\infty$, 则收敛半径

$$R=\frac{1}{\rho}.$$

定理 设复变量 z 的幂级数 $f(z)=\sum_{k=0}^{\infty} a_k z^k$ 的收敛半径为 R，方阵 A 的谱半径为 $\rho(A)$. 则

(1) $\rho(A)<R$ 时，矩阵幂级数 $f(A)=\sum_{k=0}^{\infty} a_k A^k$ 收敛；

(2) $\rho(A)>R$ 时，矩阵幂级数 $f(A)=\sum_{k=0}^{\infty} a_k A^k$ 发散.

证明 设 A 的Jordan标准型 $A=PJP^{-1}=P\mathrm{diag}(J_1,J_2,\cdots,J_m)P^{-1}$.

则 $S_N(A)=\sum_{i=0}^{N} a_k A^k = P\mathrm{diag}(S_N(J_1),\cdots,S_N(J_m))P^{-1}$.

于是 $S_N(A)$ 收敛当且仅当每个 $S_N(J_i)$ 收敛. 又 $S_N(J_i)$ 决定于它的第一行元素:

$$\left(S_N(\lambda_i), S_N'(\lambda_i), \frac{S_N''(\lambda_i)}{2!}, \cdots, \frac{S_N^{(n_i-1)}(\lambda_i)}{(n_i-1)!}\right),$$

其中, n_i 为特征值 λ_i 的重数.

§ 3.2 矩阵幂级数

定理 设复变量 z 的幂级数 $f(z)=\sum_{k=0}^{\infty} a_k z^k$ 的收敛半径为 R, 方阵 A 的谱半径为 $\rho(A)$. 则

(1) $\rho(A)<R$ 时, 矩阵幂级数 $f(A)=\sum_{k=0}^{\infty} a_k A^k$ 收敛;

(2) $\rho(A)>R$ 时, 矩阵幂级数 $f(A)=\sum_{k=0}^{\infty} a_k A^k$ 发散.

证明

$$\left(S_N(\lambda_i), S_N'(\lambda_i), \frac{S_N''(\lambda_i)}{2!}, \cdots, \frac{S_N^{(n_i-1)}(\lambda_i)}{(n_i-1)!}\right)$$

故当 $\rho(A)<R$ 时, $|\lambda_i|<R$,

$S_N^{(j)}(\lambda_i)$ 均收敛, 故 $S_N(A)$ 收敛;

而当 $\rho(A)>R$ 时, 存在 $|\lambda_{i_0}|>R$,

此时 $S_N(\lambda_{i_0})$ 发散, 故 $S_N(A)$ 发散. ■

§ 3.2 矩阵幂级数

定理 设复变量 z 的幂级数 $f(z)=\sum_{k=0}^{\infty} a_k z^k$ 的收敛半径为 R, 方阵 A 的谱半径为 $\rho(A)$. 则

(1) $\rho(A)<R$ 时, 矩阵幂级数 $f(A)=\sum_{k=0}^{\infty} a_k A^k$ 收敛;

(2) $\rho(A)>R$ 时, 矩阵幂级数 $f(A)=\sum_{k=0}^{\infty} a_k A^k$ 发散.

推论 若存在某矩阵范数使得 $\| A \| < R$, 则矩阵幂级数 $f(A)$ 收敛.

(Tips: $\rho(A) \le \| A \|$)

例1 讨论矩阵幂级数 $f(A)=\sum_{k=0}^{\infty}A^k$ 的敛散性. 当收敛时, 求其和矩阵.

解 易见 $f(z)$ 的收敛半径 $R=1$.

故当 $\rho(A)<1$时, $f(A)$ 收敛.

此时, A 的特征值的模均小于 1,

故 1 不是 A 的特征值, $I-A$ 可逆.

由 $S_N(A)=\sum_{k=0}^{N}A^k$, $(I-A)S_N(A)=I-A^{N+1}$,

于是 $(I-A)S=\lim\limits_{N\to\infty}(I-A)S_N(A)$

$$=\lim_{N\to\infty}(I-A^{N+1})=I,$$

故 $S=\lim\limits_{N\to\infty}S_N(A)=(I-A)^{-1}$. ■

公式: $(I-A)^{-1}=\sum_{k=0}^{\infty}A^k$, $(\rho(A)<1)$.

例2 讨论矩阵幂级数 $f(A)=\sum_{k=0}^{\infty}A^k$ 的敛散性，其中

$$A=\begin{bmatrix}0.2 & 0.1 & 0.2\\ 0.5 & 0.5 & 0.4\\ 0.1 & 0.3 & 0.2\end{bmatrix}.$$

解 易见 $f(z)$ 的收敛半径 $R=1$.

故当 $\rho(A)<1$ 时，$f(A)$ 收敛.

由 $\|A\|_1=0.9<1$,

$\rho(A)\le\|A\|_1<1$，$f(A)$ 收敛. ■

例3 讨论矩阵幂级数 $f(A)=\sum_{k=1}^{\infty}\frac{1}{k^2}A^k$ 的敛散性, 其中

$$A=\begin{bmatrix}1 & 4\\ -1 & -3\end{bmatrix}.$$

解 A 的特征值 $\lambda_1=\lambda_2=-1$, $\rho(A)=1=R$, 定理失效.

由其 Jordan 标准型 $A=P\begin{bmatrix}-1 & 1\\ 0 & -1\end{bmatrix}P^{-1}$, 知

$$S_N(A)=\sum_{k=1}^{N}\frac{1}{k^2}A^k=P\begin{bmatrix}\sum_{k=1}^{N}\frac{(-1)^k}{k^2} & \sum_{k=1}^{N}\frac{(-1)^{k-1}}{k}\\ 0 & \sum_{k=1}^{N}\frac{(-1)^k}{k^2}\end{bmatrix}P^{-1}$$

收敛. ■