第七讲 矩阵级数与矩阵函数

一、 矩阵序列

1. 定义：设有矩阵序列 $\left\{A^{(k)}\right\}$，其中 $A^{(k)}=\left(a_{ij}^{(k)}\right)$，且当 $k\to\infty$ 时 $a_{ij}^{(k)}\to a_{ij}$，则称 $\left\{A^{(k)}\right\}$ 收敛，并把 $A=\left(a_{ij}\right)$ 叫做 $\left\{A^{(k)}\right\}$ 的极限，或称 $\left\{A^{(k)}\right\}$ 收敛于 A，记为

$$\lim_{k\to\infty} A^{(k)} = A \quad \text{或} \quad A^{(k)} \underset{k\to\infty}{\longrightarrow} A$$

不收敛的矩阵序列则称为发散的，其中又分为有界和无界的情况。

对于矩阵序列 $\left\{A^{(k)}\right\}$，若存在常数 $M>0$，使得对一切 k 都有

$$\left|a_{ij}^{(k)}\right| < M$$

则称 $\left\{A^{(k)}\right\}$ 为有界的。

2. 收敛矩阵序列的性质:

设 $\left\{A^{(k)}\right\},\left\{B^{(k)}\right\}$分别收敛于$A, B$则

(1) $\alpha A^{(k)}+\beta B^{(k)} \underset{k\to\propto}{\longrightarrow} \alpha A+\beta B$

(2) $A^{(k)}B^{(k)} \underset{k\to\propto}{\longrightarrow} AB$

(3) 若$(A^{(k)})^{-1}, A^{-1}$存在，则 $(A^{(k)})^{-1} \underset{k\to\propto}{\longrightarrow} A^{-1}$

(4) $PA^{(k)}Q \underset{k\to\propto}{\longrightarrow} PAQ$

3 收敛矩阵的定义: 设A为方阵,若当$k\to\propto$时 $A^k \to 0$，则称A为收敛矩阵。

[定理] 方阵A为收敛矩阵的充要条件是A的所有特征值的模值均小于1.

证明: 对任何方阵A,均存在可逆矩阵P, 使得

$$A = PJP^{-1}$$

其中J为A的 Jordan 标准形

$$J = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{bmatrix}, \quad J_i = \begin{bmatrix} \lambda_i & 1 & & 0 \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{bmatrix}$$

$$A^k = PJ^kP^{-1} = P\begin{bmatrix} J_1^k & & & \\ & J_2^k & & \\ & & \ddots & \\ & & & J_s^k \end{bmatrix}P^{-1}$$

$$J_i^k = \begin{bmatrix} \lambda_i^k & k\lambda_i^{k-1} & \cdots & \dfrac{k!}{(m_i-1)!(k-m_i+1)!}\lambda_i^{k-m_i+1} \\ & \ddots & \ddots & \ddots \\ & & \ddots & \ddots \\ & & & \end{bmatrix}，当 k > m_i$$

$A^k \to 0$ 就等价于 $J_i^k \to 0 (i=1,2,\ldots,s)$，等价于 $\lambda_i^k \to 0 (i=1,2,\ldots,s)$，

而这只有$\left|\lambda_i\right|<1$才可能也必能。

［得证］

二、 矩阵级数

1.定义：矩阵序列$\left\{A^{(k)}\right\}$的无穷和$A^{(1)}+A^{(2)}+\cdots+A^{(k)}+\cdots$叫做矩阵级数，记为$\sum_{k=1}^{\infty}\mathbf{A}^{(k)}$。而$S^{(N)}=\sum_{k=1}^{N}A^{(k)}$称为其部分和，若矩阵序列$\left\{S^{(N)}\right\}$收敛,且有极限$S$，则称该矩阵级数收敛,且有和$S$．记为

$$\mathbf{S}=\sum_{k=1}^{\infty}\mathbf{A}^{(k)}$$

不收敛的矩阵级数称为是发散的。

若矩阵级数$\sum_{k=1}^{\infty}A^{(k)}$的所有元素$\sum_{k=1}^{\infty}a_{ij}^{(k)}$均绝对收敛,则称该级数为绝对收敛。

2. 绝对收敛矩阵级数的性质

(1) 绝对收敛矩阵级数一定收敛，且任意调换它的项所得的级数仍收敛,且其和不变。

(2) $\sum_{k=1}^{\infty}A^{(k)}$绝对收敛，则$\sum_{k=1}^{\infty}PA^{(k)}Q$也绝对收敛且等于$P\sum_{k=1}^{\infty}A^{(k)}Q$。

(3) $\sum_{k=1}^{\infty}A^{(k)}$，$\sum_{k=1}^{\infty}B^{(k)}$均绝对收敛,且和分别为$S_1,S_2$则

$$\sum_{k=1}^{\infty}\left(\sum_{i=1}^{k} A^{(i)} B^{(k+1-i)}\right)=S_1 S_2$$

三、 方阵的幂级数

A为方阵，$\sum_{k=0}^{\infty} c_k A^k, (A^0 = I)$ 称为 A 的幂级数. $\sum_{k=0}^{\infty} A^k$ 称为 A 的 Neumann 级数。

1. Neumann 级数收敛的充要条件

[定理] Neumann 级数收敛的充要条件是A为收敛矩阵,且在收敛时其和为$(I-A)^{-1}$。

证明: [必要性]

级数$\sum_{k=0}^{\propto} A^k$收敛，其元素为

$$\delta_{ij} + (A)_{ij} + (A^2)_{ij} + (A^3)_{ij} + \cdots$$

显然也是收敛的．作为数项级数，其通项趋于零是级数收敛的必要条件。 故

$$(A^k)_{ij} \underset{k \to \propto}{\longrightarrow} 0\text{，即} A^k \underset{k \to \propto}{\longrightarrow} 0$$

也就是说A为收敛矩阵。

[充分性]:

A为收敛矩阵，则其特征值的模值均小于 1。 设A的特征值为$\boldsymbol{\lambda}$，

$(I-A)$的特征值为μ. 则由

$$\det(\mu I-(I-A))=\det((\mu-1)I+A)=(-1)^n\det((1-\mu)I-A)$$

可见$1-\mu=\lambda\to\mu=1-\lambda$

故$0<|\mu|<2\to\mu\neq 0$, $(I-A)$的行列式不为零,$(I-A)^{-1}$存在.

而$(I+A+A^2+\ldots+A^k)(I-A)=I-A^{k+1}$

右乘$(I-A)^{-1}$得

$$I+A+A^2+\ldots+A^k=(I-A^{k+1})(I-A)^{-1}$$

当$k\to\infty$时, $A^{k+1}\to 0$, 故$A^{k+1}(I-A)^{-1}\to 0$. 所以

$$\sum_{i=0}^{\infty}A^i=\lim_{k\to\infty}\sum_{i=0}^{k}A^i=(I-A)^{-1}$$

即 Neumann 级数收敛于$(I-A)^{-1}$。

2. 收敛圆

[定理] 若矩阵A的特征值全部落在幂级数$\boldsymbol{\varphi}(z)=\sum_{k=0}^{\infty}c_k z^k$的收敛圆内，则矩阵幂级数$\boldsymbol{\varphi}(A)=\sum_{k=0}^{\infty}c_k A^k,(A^0=I)$是绝对收敛的。 反之，若$A$存在落在$\boldsymbol{\varphi}(z)$的收敛圆外的特征值，则$\boldsymbol{\varphi}(A)$是发散的。

证明略.

[推论] 若幂级数在整个复平面上收敛，则对任何的方阵A，$\boldsymbol{\varphi}(A)$均收敛。

四、 矩阵函数

如：e^A，$\sin A$，$\cos A$

以矩阵为自变量的“函数”（实际上是“函矩阵”）

我们知道，$e^z = 1 + z + \frac{1}{2!}z^2 + \cdots = \sum_{n=0}^{\infty} \frac{1}{n!} z^n$

$$\sin(z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1}$$

$$\cos(z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n}$$

均为整个复平面上收敛的级数，故对任何的方阵 A

$$e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n$$

$$\sin(A) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} A^{2n+1}$$

$$\cos(A) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} A^{2n}$$

均绝对收敛. 三者分别称为矩阵指数函数、矩阵正弦函数、矩阵余弦函数。

[性质]

$$e^{jA} = \cos A + j\sin A$$

$$\cos A = \frac{1}{2}(e^{jA} + e^{-jA})$$

$$\sin A = \frac{1}{2j}(e^{jA} - e^{-jA})$$

$$\cos(-A) = \cos A$$

$$\sin(-A) = -\sin A$$

$$\left.\begin{aligned}\cos(A \pm B) &= \cos A \cos B \mp \sin A \sin B \\ \sin(A \pm B) &= \sin A \cos B \pm \cos A \sin B\end{aligned}\right\} \leftarrow AB = BA$$

但是一般来说e^Ae^B，e^Be^A，e^{A+B}三者互不相等. 例如

$$A=\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix},\ B=\begin{bmatrix}1 & -1\\ 0 & 0\end{bmatrix},$$ 则

$$A^2=\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}=A^3=A^4=\cdots$$

$$B^2=\begin{bmatrix}1 & -1\\ 0 & 0\end{bmatrix}=B^3=B^4=\cdots$$

$$e^A=I+(\sum_{n=1}^{\infty}\frac{1}{n!})A=I+(e-1)A=\begin{bmatrix}e & e-1\\ 0 & 1\end{bmatrix}$$

$$e^{B}=I+(\sum_{n=1}^{\infty}\frac{1}{n!})B=I+(e-1)B=\begin{bmatrix} e & 1-e \\ 0 & 1 \end{bmatrix}$$

可见$e^{A}e^{B}\neq e^{B}e^{A}$

$$A+B=\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix},\ (A+B)^{2}=2\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}=2(A+B),$$

$$(A+B)^{3}=2^{2}(A+B),\cdots$$

$$e^{A+B}=I+(\sum_{n=1}^{\infty}\frac{1}{n!}2^{n-1})(A+B)=I+\frac{1}{2}(e^{2}-1)(A+B)=\begin{bmatrix} e^{2} & 0 \\ 0 & 1 \end{bmatrix}$$

所以，$e^{A+B}\neq e^{A}e^{B}$ ，$e^{A+B}\neq e^{B}e^{A}$

[定理] 若 $AB=BA$, 则 $e^{A+B}=e^{A}e^{B}=e^{B}e^{A}$

证明:

$$e^{A}e^{B}=(I+A+\frac{1}{2!}A^{2}+\ldots)(I+B+\frac{1}{2!}B^{2}+\ldots)$$
$$=I+(A+B)+\frac{1}{2!}(A^{2}+2AB+B^{2})+\frac{1}{3!}(A^{3}+3A^{2}B+3AB^{2}+B^{3})+\ldots$$
$$=I+(A+B)+\frac{1}{2!}(A+B)^{2}+\frac{1}{3!}(A+B)^{3}+\ldots=e^{A+B}$$

$$(A+B)^{2}=(A+B)(A+B)=A^{2}+AB+BA+B^{2}=A^{2}+2AB+B^{2}$$

$$(A+B)^{3}=\cdots=A^{3}+3A^{2}B+3AB^{2}+B^{3}$$

同理, 有 $e^{B}e^{A}=e^{A+B}$

[推论] $e^A e^{-A} = e^{-A} e^A = e^0 = I$，$(e^A)^{-1} = e^{-A}, (e^A)^m = e^{mA}$，$e^A$总存在逆阵。

五、 矩阵函数的初步计算

1. Hamilton-Cayley 定理

n阶矩阵A是其特征多项式的零点，即令

$$\varphi(\lambda) = \det(\lambda I - A) = \lambda^n + c_1 \lambda^{n-1} + \cdots + c_{n-1}\lambda + c_n$$

则$\varphi(A) = A^n + c_1 A^{n-1} + \cdots + c_{n-1} A + c_n I = 0$

[证明]: 设A的特征值为$\lambda_1, \lambda_2, \cdots, \lambda_n$，则$\varphi(\lambda)$又可写成

$$\varphi(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n)$$

由 Schur 引理知，存在酉矩阵U，使得

$$U^{-1}AU=\begin{bmatrix}\lambda_1 & & & *\\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n\end{bmatrix}$$

而

$$U^{-1}\varphi(A)U=\varphi(U^{-1}AU)=(U^{-1}AU-\lambda_1 I)(U^{-1}AU-\lambda_2 I)\cdots(U^{-1}AU-\lambda_n I)$$

$$=\begin{bmatrix}0 & & & & *\\ & \lambda_2-\lambda_1 & & & \\ & & \lambda_3-\lambda_1 & & \\ & & & \ddots & \\ 0 & & & & \lambda_n-\lambda_1\end{bmatrix}\begin{bmatrix}\lambda_1-\lambda_2 & & & & *\\ & 0 & & & \\ & & \lambda_3-\lambda_2 & & \\ & & & \ddots & \\ 0 & & & & \lambda_n-\lambda_2\end{bmatrix}\cdots\begin{bmatrix}\lambda_1-\lambda_n & & & & *\\ & \lambda_2-\lambda_n & & & \\ & & \ddots & & \\ & & & \lambda_{n-1}-\lambda_n & \\ 0 & & & & 0\end{bmatrix}$$

$$
=\begin{bmatrix} 0 & 0 & & & \\ \vdots & \vdots & & & \\ & & & * & \\ \vdots & \vdots & & & \\ 0 & 0 & & & \end{bmatrix}
\begin{bmatrix} \lambda_1-\lambda_3 & & & & \\ & \lambda_2-\lambda_3 & & * & \\ & & 0 & & \\ & & & \lambda_4-\lambda_3 & \\ 0 & & & & \ddots \end{bmatrix}
\cdots
\begin{bmatrix} \lambda_1-\lambda_n & & & & \\ & \lambda_2-\lambda_n & & * & \\ & & \ddots & & \\ & & & \lambda_{n-1}-\lambda_n & \\ 0 & & & & 0 \end{bmatrix}
$$

$$
=\begin{bmatrix} 0 & 0 & 0 & & \\ 0 & 0 & 0 & & \\ \vdots & \vdots & \vdots & & * \\ & & & & \\ 0 & 0 & 0 & & \end{bmatrix}
\begin{bmatrix} \lambda_1-\lambda_4 & & & & * \\ & \lambda_2-\lambda_4 & & & \\ & & \lambda_3-\lambda_4 & & \\ & & & 0 & \\ 0 & & & & \ddots \end{bmatrix}
\cdots
\begin{bmatrix} \lambda_1-\lambda_n & & & & * \\ & \lambda_2-\lambda_n & & & \\ & & \ddots & & \\ & & & \lambda_{n-1}-\lambda_n & \\ 0 & & & & 0 \end{bmatrix}
$$

$$
=\begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \\ 0 & 0 & \cdots & \cdots & 0 \end{bmatrix}=0
$$

即 $\boldsymbol{\varphi}(A)=0$

2.零化多项式

多项式 $f(z)$,若 $f(A)=0$,则称其为 A 的零化多项式。

由以上定理可知，方阵 A 的特征多项式为 A 的零化多项式。

3. 矩阵指数函数、正弦函数、余弦函数的计算

例: 已知四阶矩阵的特征值是π、$-\pi$、 0、 0, 求$\sin A$、 $\cos A$、e^A

解: $\varphi(\lambda)=(\lambda-\pi)(\lambda+\pi)(\lambda-0)(\lambda-0)=\lambda^4-\pi^2\lambda^2$

故

$$\varphi(A)=A^4-\pi^2A^2=0\to A^4=\pi^2A^2, A^5=\pi^2A^3, A^6=\pi^2A^4=\pi^4A^2,\cdots$$

$$\sin(A) = A + \sum_{n=1}^{\infty} \frac{(-1)^n}{(2n+1)!} A^{2n+1} = A + \sum_{n=1}^{\infty} \frac{(-1)^n}{(2n+1)!} \boldsymbol{\pi}^{2(n-1)} A^3$$

$$= A + \frac{1}{\pi^3} \left(\sum_{n=1}^{\infty} \frac{(-1)^n}{(2n+1)!} \boldsymbol{\pi}^{2n+1} \right) A^3$$

$$= A + \frac{1}{\boldsymbol{\pi}^3} (\sin \boldsymbol{\pi} - \boldsymbol{\pi}) A^3 = A - \boldsymbol{\pi}^{-2} A^3$$

$$\cos(A) = I + \sum_{n=1}^{\infty} \frac{(-1)^n}{(2n)!} A^{2n} = I + \sum_{n=1}^{\infty} \frac{(-1)^n}{(2n)!} \boldsymbol{\pi}^{2(n-1)} A^2$$

$$= I + \frac{1}{\boldsymbol{\pi}^2} (\cos \boldsymbol{\pi} - 1) A^2 = I - 2\boldsymbol{\pi}^{-2} A^2$$

$$e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n = I + A + \sum_{n=1}^{\infty} \frac{1}{(2n)!} A^{2n} + \sum_{n=1}^{\infty} \frac{1}{(2n+1)!} A^{2n+1}$$

$$= I + A + \sum_{n=1}^{\infty} \frac{1}{(2n)!} \boldsymbol{\pi}^{2(n-1)} A^2 + \sum_{n=1}^{\infty} \frac{1}{(2n+1)!} \boldsymbol{\pi}^{2(n-1)} A^3$$

$$= I + A + \frac{\cosh \boldsymbol{\pi} - 1}{\boldsymbol{\pi}^2} A^2 + \frac{\sinh \boldsymbol{\pi} - \boldsymbol{\pi}}{\boldsymbol{\pi}^3} A^3$$

作业　P163　 3, 4, 5