矩阵分析与应用

第十讲 矩阵分析及其应用之一

信息与通信工程学院

吕旌阳

本讲主要内容

- 矩阵序列
- 矩阵级数
- 矩阵函数

引言：

- 一元多项式 $f(t)=c_0+c_1t+\cdots+c_mt^m$

- 矩阵多项式 $f(A)=c_0I+c_1A+\cdots+c_mA^m$，$(\forall A\in C^{n\times n})$

- $f(A)$以矩阵为自变量且取值为矩阵的一类函数

本章研究一般的以矩阵为自变量且取值为矩阵的函数

——矩阵函数

一、敛散性

定义：将矩阵序列 $A^{(k)}=\left(a_{ij}^{(k)}\right)_{m\times n}$,记作 $\left\{A^{(k)}\right\}$

当 $\lim\limits_{k\to\infty} a_{ij}^{(k)}=a_{ij}(\forall i,j)$ 时，称矩阵序列$\{A^{(k)}\}$收敛于

矩阵$A=(a_{ij})$。记作

$$\lim_{k\to\infty} A^{(k)}=A \quad ，或者 \quad A^{(k)}\to A\left(k\to\infty\right)$$

若数列 $\left(a_{ij}^{(k)}\right)$ 之一发散，称 $\left\{A^{(k)}\right\}$ 发散

性质：

(1) 若 $\lim\limits_{k\to\infty} A^{(k)} = A_{m\times n}, \lim\limits_{k\to\infty} B^{(k)} = B_{m\times n}$ 则

$$\lim_{k\to\infty}(aA^{(k)} + bB^{(k)}) = aA + bB, \quad \forall a, b$$

(2) 若 $\lim\limits_{k\to\infty} A^{(k)} = A_{m\times n}, \lim\limits_{k\to\infty} B^{(k)} = B_{n\times l}$ 则

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

(3) 若 $A^{(k)}$ 与 A 是可逆矩阵，且 $\lim\limits_{k\to\infty} A^{(k)} = A$ ，则

$$\lim_{k\to\infty}(A^{(k)})^{-1} = A^{-1}$$

定理1：设 $A^{(k)}, A \in C^{m\times n}$，则

（1）$\lim\limits_{k\to\infty} A^{(k)} = 0 \Leftrightarrow \forall \|\bullet\|, \lim\limits_{k\to\infty} \left\|A^{(k)}\right\| = 0$

（2）$\lim\limits_{k\to\infty} A^{(k)} = A \Leftrightarrow \forall \|\bullet\|, \lim\limits_{k\to\infty} \left\|A^{(k)} - A\right\| = 0$

证明：(1)考虑F-矩阵范数

$$\lim_{k\to\infty} A^{(k)} = 0 \Leftrightarrow \lim_{k\to\infty} a_{ij}^{(k)} = 0 \quad (all\ i, j)$$

$$\Leftrightarrow \lim_{k\to\infty} \sum_{i=1}^{m}\sum_{j=1}^{n} \left|a_{ij}^{(k)}\right|^2 = 0$$

$$\Leftrightarrow \lim_{k\to\infty} \left\|A^{(k)}\right\|_F = 0$$

(2)由 $\lim\limits_{k\to\infty} A^{(k)} = A \Leftrightarrow \lim\limits_{k\to\infty}(A^{(k)} - A) = 0$ 可直接的得到

敛散性的另一定义：矩阵序列 $\left\{A^{(k)}\right\}$收敛的充要条件

为对任给$\varepsilon>0$ 存在$N(\varepsilon)$, 当 $k, l \geq N(\varepsilon)$ 时有

$$\left\|A^{(k)}-A^{(l)}\right\|<\varepsilon$$

其中 $\|\bullet\|$ 为任意的广义矩阵范数。

例1： $\mathbf{A}^{(n)}=\begin{pmatrix} 1-\frac{1}{n} & \sin(\frac{1}{n}) \\ e^{-n} & \sum_{k=1}^{n}\frac{\sin(k)}{k^2} \end{pmatrix}$ ，证明其敛散性

因为求不出 $\mathbf{A}^{(n)}$ 的极限从而很难应用定义证明收敛。

相反，由于 $\left|\sum_{k=m+1}^{n}\frac{\sin(k)}{k^2}\right|<\left|\sum_{k=m+1}^{n}\frac{1}{k^2}\right|<\left|\sum_{k=m+1}^{n}\frac{1}{k(k-1)}\right|<\frac{1}{m}$

从而只要取l充分大，则当$m, n>l$ 时就有

$$\left|\sum_{k=m+1}^{n}\frac{\sin(k)}{k^2}\right|\le\varepsilon$$

$$\because \frac{1}{k(k-1)}=\frac{1}{k-1}-\frac{1}{k}$$

$$\sum_{k=m+1}^{n}\frac{1}{k(k-1)}=\frac{1}{m}-\frac{1}{n}<\frac{1}{m}$$

这样 $\mathbf{A}^{(n)}$ 收敛

定义：若 $A_{n\times n}$ 满足 $\lim\limits_{k\to\infty} A^k = 0_{n\times n}$，称A为收敛矩阵

定理2： A为收敛矩阵 ⟺ $\rho(A)<1$

证明：充分性。已知$\rho(A)<1$，对 $\varepsilon=\frac{1}{2}[1-\rho(A)]>0$

存在矩阵范数 $\|\bullet\|_M$，使得

$$\|A\|_M\le\rho(A)+\varepsilon=\frac{1}{2}[1+\rho(A)]<1$$

于是有 $\|A^k\|_M\le\|A\|_M^k\to 0$，故由定理1可得 $A^k\to 0$

必要性：已知 $A^k\to 0$，设$Ax=\lambda x(x\neq 0)$，则有

$$\lambda^k x=A^k x\to 0 \quad\Rightarrow\lambda^k\to 0\quad\Rightarrow|\lambda|<1$$

故 $\rho(A)<1$

定理3：若矩阵范数 $\|\bullet\|_M$ 使 $\|A\|_M < 1$ ，则 $A^k \to 0$

证明： $\rho(A) \le \|A\|_M < 1 \quad \Rightarrow A^k \to 0$

例： $A = \begin{bmatrix} 0.1 & 0.3 \\ 0.7 & 0.6 \end{bmatrix}$

$$\|A\|_1 = 0.9 < 1 \Rightarrow A^k \to 0$$

定义：设有矩阵序列 $\{A^{(k)}\}$，其中 $A^{(k)}=(a_{ij}^{(k)})\in C^{m\times n}$

称 $A^{(0)}+A^{(1)}+\cdots+A^{(k)}+\cdots$ 为**矩阵级数**。记为 $\sum_{k=0}^{\infty}A^{(k)}$

部分和 $S^{(N)}=\sum_{k=0}^{N}A^{(k)}$ 构成矩阵序列 $\{S^{(N)}\}$

敛散性：若 $\lim_{N\to\infty}S^{(N)}=S$，称 $\sum A^{(k)}$ 收敛于 S，记做

$$\sum A^{(k)}=S$$

若 $\{S^{(N)}\}$ 发散，称 $\sum A^{(k)}$ 发散

性质1： $\sum A^{(k)} = S \Longleftrightarrow \sum_{k=0}^{\infty} a_{ij}^{(k)} = s_{ij} \ (all \quad i,j)$

证明： 左 $\Longleftrightarrow \lim_{N\to\infty} S^{(N)} = S$

$\Longleftrightarrow \lim_{N\to\infty} s_{ij}^{(N)} = s_{ij} \ (all \quad i,j)$

$\Longleftrightarrow \lim_{N\to\infty} \sum_{k=0}^{N} a_{ij}^{(k)} = s_{ij} \ (all \quad i,j)$

$\Longleftrightarrow$ 右

性质2： 若 $\sum\left|a_{ij}^{(k)}\right|$ 收敛 $(all\quad i,j)$，称 $\sum A^{(k)}$ 绝对收敛。

（1） $\sum A^{(k)}$ 绝对收敛 $\Rightarrow \sum A^{(k)}$ 收敛

（2） 若 $\sum A^{(k)}$ 绝对收敛于S，对 $\sum A^{(k)}$ 任意重组重排得 $\sum B^{(k)}$，则 $\sum B^{(k)}$ 绝对收敛于S。

性质3: $\sum A^{(k)}=S$ 绝对收敛 $\Leftrightarrow \forall\|\bullet\|,\sum\|A^{(k)}\|$ 收敛

证明： 只需考虑矩阵范数 $\|\bullet\|_{m1}$

证明： （必要性）由于矩阵级数 $\sum_{k=0}^{\infty}A^{(k)}$ 绝对收敛，

因此mn个数项级数 $\sum_{k=0}^{\infty}a_{ij}^{(k)}$ 绝对收敛。因此存在

$M>0$ ，使得对任意N，都有 $\sum_{k=0}^{N}|a_{ij}^{(k)}|<M,(\forall i,j)$

故 $$\sum_{k=0}^{N}\|A^{(k)}\|_{m1}=\sum_{k=0}^{N}(\sum_{i,j}|a_{ij}^{(k)}|)\le mnM$$

因此 $\sum_{k=0}^{\infty}\|A^{(k)}\|_{m1}$ 收敛

（充分性）$\sum_{k=0}^{\infty}\|A^{(k)}\|_{m1}$ 收敛。因此$\sum_{k=0}^{\infty}|a_{ij}^{(k)}|\le\sum_{k=0}^{\infty}\|A^{(k)}\|_{m1}$

故矩阵级数 $\sum_{k=0}^{\infty}A^{(k)}$ 绝对收敛

性质4：$\sum A^{(k)}$ 收敛于$S \Rightarrow \sum_{k=0}^{\infty} PA^{(k)}Q$ 收敛于 PSQ

$\sum A^{(k)}$ 绝对收敛$\Rightarrow \sum_{k=0}^{\infty} PA^{(k)}Q$ 绝对收敛

证明： 只需考虑矩阵范数 $\|\bullet\|_{m1}$

证明：(1) $S^{(N)} = \sum_{k=0}^{N} A^{(k)} \to S \Rightarrow \sum_{k=0}^{N} PA^{(k)}Q = PS^{(N)}Q \to PSQ$

$$\Rightarrow \sum_{k=0}^{\infty} PA^{(k)}Q \to PSQ$$

性质4：

$$\sum A^{(k)} \text{ 绝对收敛} \Rightarrow \sum_{k=0}^{\infty} PA^{(k)}Q \text{ 绝对收敛}$$

(2)矩阵范数 $\|\bullet\|$，由性质3知 $\sum \left\|A^{(k)}\right\|$ 收敛

因为 $\left\|PA^{(k)}Q\right\| \le \|P\| \left\|A^{(k)}\right\| \|Q\| = M\left\|A^{(k)}\right\| \qquad \left(M = \|P\|\|Q\|\right)$

所以 $\sum_{k=0}^{N}\left\|PA^{(k)}Q\right\| \le \sum_{k=0}^{N}\left(M\left\|A^{(k)}\right\|\right) = M\sum_{k=0}^{N}\left\|A^{(k)}\right\|$ 有界

故 $\sum A^{(k)}$ 绝对收敛 $\Rightarrow \sum_{k=0}^{\infty} PA^{(k)}Q$ 绝对收敛

性质5： $\sum\limits_{k=1}^{\infty} \boldsymbol{A}^{(k)}$ 绝对收敛于 $\boldsymbol{S}_{m\times n}$，$\sum\limits_{k=1}^{\infty} \boldsymbol{B}^{(k)}$ 绝对收敛于 $\boldsymbol{T}_{n\times l}$

则Cauchy积

$$\boldsymbol{A}^{(1)}\boldsymbol{B}^{(1)} + \left[\boldsymbol{A}^{(1)}\boldsymbol{B}^{(2)} + \boldsymbol{A}^{(2)}\boldsymbol{B}^{(1)}\right] + \left[\boldsymbol{A}^{(1)}\boldsymbol{B}^{(3)} + \boldsymbol{A}^{(2)}\boldsymbol{B}^{(2)} + \boldsymbol{A}^{(3)}\boldsymbol{B}^{(1)}\right]$$

$$+\cdots+\left[\boldsymbol{A}^{(1)}\boldsymbol{B}^{(k)} + \boldsymbol{A}^{(2)}\boldsymbol{B}^{(k-1)} + \cdots + \boldsymbol{A}^{(k)}\boldsymbol{B}^{(1)}\right] + \cdots$$

绝对收敛于$\boldsymbol{ST}$，记作 $\sum \boldsymbol{A}^{(k)} \bullet \sum \boldsymbol{B}^{(k)} = \boldsymbol{ST}$

证明：只需利用性质3即可

Neumann级数： $A_{n\times n},\sum_{k=0}^{\infty}A^k,\left(A^0=I\right)$

定理4： $A_{n\times n},\sum A^k$ 收敛 $\Leftrightarrow A^k\to 0$

$\sum A^k$ 收敛时，其和为 $(I-A)^{-1}$

证明： 必要性。$\sum A^k$ 收敛时 $\sum\left(A^k\right)_{ij},\forall i,j$ 收敛

即 $\left(A^k\right)_{ij}\to 0$ ，也就是 $A^k\to 0$

充分性。$A^k\to 0$ 由定理2可知 $\rho(A)<1\Leftrightarrow(I-A)$ 可逆

$$\left(I+A+A^2+\cdots+A^N\right)(I-A)=I-A^{N+1}$$

$$\left(I+A+A^2+\cdots+A^N\right)=(I-A)^{-1}-A^{N+1}(I-A)^{-1}\to(I-A)^{-1},N\to\infty$$

即 $\sum A^k=(I-A)^{-1}$

定理5：$A_{n\times n}, \|A\|<1 \Rightarrow \left\|(I-A)^{-1}-\sum_{k=0}^{N}A^k\right\| \le \frac{\|A\|^{N+1}}{1-\|A\|}, N=0,1,2$

证明：$\|A\|<1 \Rightarrow \rho(A)<1 \quad \Rightarrow (I-A)$ 可逆

$$\left(I+A+A^2+\cdots+A^N\right)(I-A)=I-A^{N+1}$$

右乘 $(I-A)^{-1}$ ，移项可得

$$(I-A)^{-1}-\left(I+A+A^2+\cdots+A^N\right)=A^{N+1}(I-A)^{-1}$$

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

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

$$\left\|A^{N+1}(I-A)^{-1}\right\| \le \left\|A^{N+1}\right\|+\left\|A^{N+1}(I-A)^{-1}\right\|\bullet\|A\|$$

故 $\left\|A^{N+1}(I-A)^{-1}\right\| \le \frac{\left\|A^{N+1}\right\|}{1-\|A\|} \Rightarrow \left\|(I-A)^{-1}-\sum_{k=0}^{N}A^k\right\| \le \frac{\|A\|^{N+1}}{1-\|A\|}$

幂级数： 对函数 $f(z)=\sum c_k z^k,\left(|z|<r\right)$ 方阵 $A_{n\times n}$，

构造矩阵幂级数 $f(A)=\sum c_k A^k$

定理6： （1） $\rho\left(A\right)<r \quad \Rightarrow \sum c_k A^k$ 绝对收敛

（2） $\rho\left(A\right)>r \quad \Rightarrow \sum c_k A^k$ 发散

证明:（1）对A，取 $\varepsilon=\frac{1}{2}[r-\rho(A)]>0,\exists\|\bullet\|_{\varepsilon}$，使得

$$\|A\|_{\varepsilon}\le\rho(A)+\varepsilon=\frac{1}{2}[r+\rho(A)]<r$$

$$\left\|c_k A^k\right\|_{\varepsilon}\le|c_k|\|A\|_{\varepsilon}^{k}\le|c_k|[\rho(A)+\varepsilon]^k$$

当 $|z|<r$ 时，$\sum|c_k||z|^k$ 收敛，于是

$\sum|c_k|[\rho(A)+\varepsilon]^k$ 收敛 $\Rightarrow\sum\left\|c_k A^k\right\|_{\varepsilon}$ 收敛 $\Rightarrow\sum c_k A^k$ 收敛

(2)设A的特征值λ满足$|\lambda|=\rho(A)$，x为λ相应的特征向量

$$\sum_{k=0}^{n}c_k(A^k x)=\sum_{k=0}^{n}c_k(\lambda^k x)=\left(\sum_{k=0}^{n}c_k\lambda^k\right)x$$

由于$\rho(A)>r$,那么 $\left(\sum_{k=0}^{n}c_k\lambda^k\right)x$ 发散(注意x为非零向量)

从而 $\sum_{k=0}^{n}c_k(A^k x)$ 发散，这样 $\sum_{k=0}^{\infty}c_k A^k$ 发散

定义：设一元函数 $f(z)$ 能展开为z的幂级数

$$f(z)=\sum_{k=0}^{\infty}c_k z^k \quad \left(|z|<r, r>0\right)$$

其中r>0表示该幂级数的收敛半径。当n阶矩阵A的谱半径$\rho(A)<r$ 时，把收敛的矩阵幂级数$\sum_{k=0}^{\infty}c_k A^k$ 的和为f(A), 即 $f(A)=\sum_{k=0}^{\infty}c_k A^k$

性质(代入规则)： 若$f(z)=g(z)$, 则$f(A)=g(A)$.

例1：

$$e^{z}=1+\frac{1}{1!}z+\cdots+\frac{1}{k!}z^{k}+\cdots\quad(r=+\infty)$$

$$e^{A}=I+\frac{1}{1!}A+\cdots+\frac{1}{k!}A^{k}+\cdots\quad(\forall A_{n\times n})$$

$$\sin z=z-\frac{1}{3!}z^{3}+\cdots+(-1)^{k}\frac{1}{(2k+1)!}z^{(2k+1)}+\cdots\quad(r=+\infty)$$

$$\sin A=A-\frac{1}{3!}A^{3}+\cdots+(-1)^{k}\frac{1}{(2k+1)!}A^{(2k+1)}+\cdots\quad(\forall A_{n\times n})$$

例2：$f(z)=\dfrac{1}{1-z}=\sum_{k=0}^{\infty} z^k \quad \left(|z|<1\right)$， $f(A)=\dfrac{1}{I-A}=\sum_{k=0}^{\infty} A^k \quad \left(\rho(A)<1\right)$

例3： $\forall A_{n\times n}, e^{jA}=\cos A+j\sin A \quad \left(j=\sqrt{-1}\right)$

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

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

证明：在 e^{jA} 中,视“ jA”为整体,并按奇偶次幂分开

$$e^{jA}=\left[I+\frac{1}{2!}(jA)^2+\frac{1}{4!}(jA)^4+\cdots\right]+\left[\frac{1}{1!}(jA)+\frac{1}{3!}(jA)^3+\cdots\right]$$

$$=\cos A+j\sin A \quad \left(j=\sqrt{-1}\right)$$

例4： $A=\begin{bmatrix}1&1\\0&0\end{bmatrix}, B=\begin{bmatrix}1&-1\\0&0\end{bmatrix}$， 求 e^A, e^B 及 e^{A+B}

解：$A^2=A$： $e^A=I+\left(\frac{1}{1!}+\cdots+\frac{1}{k!}+\cdots\right)A=I+(e-1)A$

$$=\begin{bmatrix}e&e-1\\0&1\end{bmatrix}$$

$B^2=B$： $e^B=I+\left(\frac{1}{1!}+\cdots+\frac{1}{k!}+\cdots\right)B=I+(e-1)B$

$$=\begin{bmatrix}e&1-e\\0&1\end{bmatrix}$$

$A+B=\begin{bmatrix}2&0\\0&0\end{bmatrix}=2\begin{bmatrix}1&0\\0&0\end{bmatrix}$ $\quad (A+B)^k=2^k\begin{bmatrix}1&0\\0&0\end{bmatrix}$

$$e^{A+B}=I+\left(\frac{2}{1!}+\frac{2^2}{2!}+\cdots\right)\begin{bmatrix}1&0\\0&0\end{bmatrix}=I+(e^2-1)\begin{bmatrix}1&0\\0&0\end{bmatrix}=\begin{bmatrix}e^2&0\\0&1\end{bmatrix}$$

注意:

$$e^{A+B}=\begin{bmatrix} e^2 & 0 \\ 0 & 1 \end{bmatrix}$$

$$e^Ae^B=\begin{bmatrix} e^2 & -(e-1)^2 \\ 0 & 1 \end{bmatrix},\quad e^Be^A=\begin{bmatrix} e^2 & (e-1)^2 \\ 0 & 1 \end{bmatrix}$$

$$e^{A+B}\neq e^Ae^B\neq e^Be^A$$

定理7： $A_{n\times n}, B_{n\times n}, AB=BA \quad \Rightarrow e^{A+B}=e^Ae^B=e^Be^A$

证明：$e^Ae^B=\left[I+\frac{1}{1!}A+\frac{1}{2!}A^2+\cdots\right]\left[I+\frac{1}{1!}B+\frac{1}{2!}B^2+\cdots\right]$

$$=I+(A+B)+\frac{1}{2!}\left(A^2+2AB+B^2\right)+\frac{1}{3!}\left(A^3+3AB^2+3A^2B+B^3\right)+\cdots$$

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

$$=e^{A+B}$$

同理： $e^Be^A=e^{B+A}=e^{A+B}$

注： （1） $e^Ae^{-A}=e^o=I\Rightarrow\left(e^A\right)^{-1}=e^{-A} \quad \forall A$

（2） $\left(e^A\right)^m=e^{mA} \quad m=2,3,\cdots$

例5： $A_{n\times n}, B_{n\times n}, AB = BA$

$$\cos(A+B) = \cos A\cos B - \sin A\sin B$$

$$\sin(A+B) = \sin A\cos B + \cos A\sin B$$

证明： $\cos A\cos B - \sin A\sin B =$

$$= \frac{1}{2}\left[e^{jA} + e^{-jA}\right]\bullet\frac{1}{2}\left[e^{jB} + e^{-jB}\right] - \frac{1}{2j}\left[e^{jA} - e^{-jA}\right]\bullet\frac{1}{2j}\left[e^{jB} - e^{-jB}\right]$$

$$= \frac{1}{4}\left[e^{j(A+B)} + e^{j(A-B)} + e^{j(-A+B)} + e^{-j(A+B)}\right] + \frac{1}{4}\left[e^{j(A+B)} - e^{j(A-B)} - e^{j(-A+B)} + e^{-j(A+B)}\right]$$

$$= \frac{1}{2}\left[e^{j(A+B)} + e^{-j(A+B)}\right]$$

$$= \cos(A+B)$$

矩阵函数值的求法

1.待定系数法:设n阶矩阵A 的特征多项式$\varphi(\lambda)=\det(\lambda I-A)$

如果首1多项式 $\psi(\lambda)=\lambda^m+b_1\lambda^{m-1}+\cdots+b_{m-1}\lambda+b_m\ (1\le m\le n)$

以 A 为根且满足$\psi(\lambda)\mid\varphi(\lambda)$，分解

$$\psi(\lambda)=(\lambda-\lambda_1)^{m_1}\cdots(\lambda-\lambda_s)^{m_s}\quad \lambda_i\neq\lambda_j,\sum m_i=m$$

因为λ_i 是 A 的特征值，所以 $|\lambda_i|\le\rho(A)<r$ ，从而

$f(\lambda_i)=\sum c_k\lambda_i^k$ 绝对收敛。

设 $$f(z)=\sum c_k z^k=\psi(\lambda)g(z)+r(z)$$

$$r(z)=b_0+b_1z+\cdots+b_{m-1}z^{m-1}$$

由 $\psi(\lambda_i)=0,\psi^{(1)}(\lambda_i)=0,\cdots,\psi^{(m_i-1)}(\lambda_i)=0$ 可得

$$r(\lambda_i)=f(\lambda_i)$$

$$\psi(\lambda)=(\lambda-\lambda_i)^{m_i}g(\lambda)$$

$$r'(\lambda_i)=f'(\lambda_i)$$

$$\psi'(\lambda)=(\lambda-\lambda_i)^{m_i-1}g(\lambda)+(\lambda-\lambda_i)^{m_i}g'(\lambda)$$

$$=(\lambda-\lambda_i)^{m_i-1}h(\lambda)$$

$$\cdots$$

$$r^{(m_i-1)}(\lambda_i)=f^{(m_i-1)}(\lambda_i) \qquad i=1,2,\cdots,s$$

解此方程组得出 $b_0,b_1,\cdots,b_{m-1}$。因为 $\psi(A)=O$ 所以

$$f(A)=\sum c_k A^k=\psi(A)g(A)+r(A)\quad =r(A)$$

即 $$f(A)=b_0I+b_1A+\cdots+b_{m-1}A^{m-1}$$

例6：$A=\begin{bmatrix}2&0&0\\1&1&1\\1&-1&3\end{bmatrix}$，求 $e^A, e^{tA}\quad (t\in R)$

解：$\varphi(\lambda)=\det(\lambda I-A)=(\lambda-2)^3$

$$(A-2I)=\begin{bmatrix}0&0&0\\1&-1&1\\1&-1&1\end{bmatrix},(A-2I)^2=O$$

取 $\psi(\lambda)=m(\lambda)=(\lambda-2)^2$

（1）$f(\lambda)=e^{\lambda}=\psi(\lambda)g(\lambda)+(a+b\lambda)$

$$f'(\lambda)=e^{\lambda}=[\psi(\lambda)g(\lambda)]'+b$$

$$\left.\begin{aligned}f(2)=e^2:&\quad (a+2b)=e^2\\ f'(2)=e^2:&\quad b=e^2\end{aligned}\right\}\Rightarrow\left\{\begin{aligned}a&=-e^2\\ b&=e^2\end{aligned}\right.$$

$$e^A=e^2(A-I)=e^2\begin{bmatrix}1&0&0\\1&0&1\\1&-1&2\end{bmatrix}$$

（2） $f(\lambda)=e^{t\lambda}=\psi(\lambda)g(\lambda)+(a+b\lambda)$

$$f'(\lambda)=te^{t\lambda}=\left[\psi(\lambda)g(\lambda)\right]'+b$$

$$\left.\begin{aligned} f(2)=e^{2t}:&\quad (a+2b)=e^{2t} \\ f'(2)=te^{2t}:&\quad b=te^{2t} \end{aligned}\right\} \Rightarrow \left\{\begin{aligned} a&=(1-2t)e^{2t} \\ b&=te^{2t} \end{aligned}\right.$$

$$e^{tA}=e^{2t}\left[(1-2t)I+tA\right]=e^{2t}\begin{bmatrix} 1 & 0 & 0 \\ t & 1-t & t \\ t & -t & 1+t \end{bmatrix}$$

2. 数项级数求和法。

利用首一多项式 $\psi(\lambda)$ ，且满足 $\psi(A)=0$，即

$$A^m + b_1 A^{m-1} + \cdots + b_{m-1} A + b_m I = 0$$

或者 $A^m = k_0^{(0)} I + k_1^{(0)} A + \cdots + k_{m-1}^{(0)} A^{m-1} \quad \left(k_i^{(0)} = -b_{m-i}\right)$

可以求出 $A^{m+1} = A^m A = k_0^{(1)} I + k_1^{(1)} A + \cdots + k_{m-1}^{(1)} A^{m-1}$

$$\vdots$$

$$A^{m+l} = k_0^{(l)} I + k_1^{(l)} A + \cdots + k_{m-1}^{(l)} A^{m-1}$$

$$\vdots$$

于是 $f(A) = \sum_{k=0}^{\infty} c_k A^k = \left(c_0 I + c_1 A + \cdots + c_{m-1} A^{m-1}\right) + c_m \left(k_0^{(0)} I + k_1^{(0)} A + \cdots + k_{m-1}^{(0)} A^{m-1}\right) + \cdots$

$$= \left(c_0 + \sum_{l=0}^{\infty} c_{m+l} k_0^{(l)}\right) I + \left(c_1 + \sum_{l=0}^{\infty} c_{m+l} k_1^{(l)}\right) A + \cdots + \left(c_{m-1} + \sum_{l=0}^{\infty} c_{m+l} k_{m-1}^{(l)}\right) A^{m-1}$$

例7： $A=\begin{bmatrix}\pi & 0 & 0 & 0\\ & -\pi & 0 & 0\\ & & 0 & 1\\ & & & 0\end{bmatrix}$ ，求 $\sin A$

解： $\varphi(\lambda)=|\lambda I-A|=\lambda^4-\pi^2\lambda^2$，取 $\psi(\lambda)=\varphi(\lambda)$

$$\psi(A)=0\Rightarrow A^4=\pi^2A^2,\ A^5=\pi^2A^3,\ A^7=\pi^4A^3,\cdots$$

$$\sin A=A-\frac{1}{3!}A^3+\frac{1}{5!}A^5-\frac{1}{7!}A^7+\cdots$$

$$=A+\left[-\frac{1}{3!}+\frac{\pi^2}{5!}-\frac{\pi^4}{7!}+\cdots\right]A^3$$

$$=A+\frac{1}{\pi^3}[\sin\pi-\pi]A^3=A-\frac{1}{\pi^2}A^3=\begin{bmatrix}0 & 0 & 0 & 0\\ & 0 & 0 & 0\\ & & 0 & 1\\ & & & 0\end{bmatrix}$$

$$\because A^3=\mathrm{diag}\left(\pi^3,-\pi^3,0,0\right)$$

3. 对角阵法

设 $P^{-1}AP=\mathrm{diag}\left(\lambda_1,\cdots,\lambda_n\right)=\Lambda$ ，则 $A^k=P\Lambda^k P^{-1}$,

且有 $$\sum_{k=0}^{N} c_k A^k = P\sum_{k=0}^{N} c_k \Lambda^k P^{-1}$$

$$= P\,\mathrm{diag}\left(\sum_{k=0}^{N} c_k\lambda_1^k,\cdots,\sum_{k=0}^{N} c_k\lambda_n^k\right)P^{-1}$$

于是

$$f(A)=\sum_{k=0}^{\infty} c_k A^k = P\bullet\mathrm{diag}\left(f(\lambda_1),\cdots,f(\lambda_n)\right)\bullet P^{-1}$$

$$f(A)=\sum_{k=0}^{\infty}c_k A^k = P\bullet\mathrm{diag}\left(f(\lambda_1),\cdots,f(\lambda_n)\right)\bullet P^{-1}$$

例8： $P^{-1}AP=\Lambda$ ：

$$e^A = P\bullet\mathrm{diag}\left(e^{\lambda_1},\cdots,e^{\lambda_n}\right)\bullet P^{-1}$$

$$\because e^A=\sum_{k=0}^{\infty}\frac{1}{k!}A^k\left(\forall A_{n\times n}\right) \qquad e^z=\sum_{k=0}^{\infty}\frac{1}{k!}z^k$$

$$e^{tA} = P\bullet\mathrm{diag}\left(e^{\lambda_1 t},\cdots,e^{\lambda_n t}\right)\bullet P^{-1}$$

$$\sin A = P\bullet\mathrm{diag}\left(\sin\lambda_1,\cdots,\sin\lambda_n\right)\bullet P^{-1}$$

移位矩阵

$$I^{(1)}=\begin{bmatrix}0&1&0\\&0&1\\&&0\end{bmatrix}\qquad I^{(2)}=\begin{bmatrix}0&0&1\\&0&0\\&&0\end{bmatrix}=I^{(1)}I^{(1)}$$

$$I^{(1)}A=\begin{bmatrix}0&1&0\\&0&1\\&&0\end{bmatrix}\begin{bmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{bmatrix}=\begin{bmatrix}a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\\0&0&0\end{bmatrix}$$

$$AI^{(1)}=\begin{bmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{bmatrix}\begin{bmatrix}0&1&0\\&0&1\\&&0\end{bmatrix}=\begin{bmatrix}0&a_{11}&a_{12}\\0&a_{21}&a_{22}\\0&a_{31}&a_{32}\end{bmatrix}$$

$$I^{(k)}I^{(1)}=I^{(1)}I^{(k)}=I^{(k+1)},I_n^{(n)}=O$$

4.Jordan标准型法

设 $P^{-1}AP = J = \text{diag}\left(J_1,\cdots,J_s\right),\ J_i = \lambda_i I + I^{(1)}$

易证 $I^{(k)}I^{(1)} = I^{(1)}I^{(k)} = I^{(k+1)}, I^{(m_i)} = O$

$$k \le m_i - 1: J_i^k = \lambda_i^k I + C_k^1 \lambda_i^{k-1} I^{(1)} + \cdots + C_k^{k-1} \lambda_i I^{(k-1)} + I^{(k)}$$

$$k \ge m_i : J_i^k = \lambda_i^k I + C_k^1 \lambda_i^{k-1} I^{(1)} + \cdots + C_k^{m_i-1} \lambda_i^{k-m_i+1} I^{(m_i-1)}$$

$$f\left(J_i\right) = \sum_{k=0}^{\infty} c_k J_i^k = f\left(\lambda_i\right) I + \frac{f'\left(\lambda_i\right)}{1!} I^{(1)} + \cdots + \frac{f^{(m_i-1)}\left(\lambda_i\right)}{(m_i-1)!} I^{(m_i-1)}$$

$$f\left(A\right) = \sum_{k=0}^{\infty} c_k A^k = P \bullet \sum_{k=0}^{\infty} c_k J^k \bullet P^{-1} = P \bullet diag\left(f(J_1),\cdots, f(J_s)\right) \bullet P^{-1}$$

三、矩阵函数的一般定义

展开式 $f(z)=\sum c_k z^k,\quad (|z|<r, r>0)$，要求

（1） $f^{(k)}(0)$ 存在 $\quad (k=0,1,2,\cdots)$

（2） $\lim\limits_{k\to\infty}\dfrac{f^{(k+1)}(\xi)}{(k+1)!}\xi^{k+1}=0 \quad (|z|<r)$

对于一元函数 $f(z)=\dfrac{1}{z}$ 等，还不能定义矩阵函数。

基于矩阵函数值的Jordan标准形算法，拓宽定义

矩阵函数的一般定义

设 $P^{-1}AP = J = \mathrm{diag}\left(J_1,\cdots,J_s\right),\ J_i = \lambda_i I + I^{(1)}$

如果 $f\left(z\right)$ 在 λ_i 处有 $m_i - 1$ 阶导数，令

$$f\left(J_i\right) = \sum_{k=0}^{\infty} c_k J_i^k = f\left(\lambda_i\right)I + \frac{f'\left(\lambda_i\right)}{1!} I^{(1)} + \cdots + \frac{f^{(m_i-1)}\left(\lambda_i\right)}{(m_i - 1)!} I^{(m_i-1)}$$

$$f\left(A\right) = \sum_{k=0}^{\infty} c_k A^k = P\bullet\sum_{k=0}^{\infty} c_k J^k \bullet P^{-1} = P\bullet diag\left(f(J_1),\cdots, f(J_s)\right)\bullet P^{-1}$$

称 $f\left(A\right)$ 为对应于 $f\left(z\right)$ 的矩阵函数

[注] ① 拓宽定义不要求$f(z)$能展为"z"的幂级数，但要求在A的特征值 λ_i （重数为m_i）处有 m_i-1 阶导数,后者较前者弱！

② 当能够展为"z"的幂级数时, 矩阵函数的拓宽定义与级数原始定义是一致的.

例9： $A=\begin{bmatrix} 2 & 1 & & \\ & 2 & 1 & \\ & & 2 & 1 \\ & & & 2 \end{bmatrix}, f(z)=\frac{1}{z}$ ，求 $f(A)$

解： $f(z)=\frac{1}{z}, f'(z)=-z^{-2}, f''(z)=2z^{-3}, f'''(z)=-6z^{-4}$

$$f(A)=f(J)$$

$$=f(2)\bullet I+f'(2)\bullet I^{(1)}+\frac{f''(2)}{2!}\bullet I^{(2)}+\frac{f'''(2)}{3!}\bullet I^{(3)}$$

$$=\begin{bmatrix} 0.5 & -0.25 & 0.125 & -0.0625 \\ & 0.5 & -0.25 & 0.125 \\ & & 0.5 & -0.25 \\ & & & 0.5 \end{bmatrix}$$

例10： $A=\begin{bmatrix}1&1&0\\0&1&0\\0&0&2\end{bmatrix}, f(z)=\sqrt{z}$ ，求 $f(A)$

解： $f(z)=\sqrt{z}, f'(z)=\dfrac{1}{2\sqrt{z}}$

$$J_1=\begin{bmatrix}1&1\\&1\end{bmatrix}:\quad f(J_1)=f(1)\bullet I+f'(1)\bullet I^{(1)}\ =\begin{bmatrix}1&1/2\\&1\end{bmatrix}$$

$$J_2=[2]:\quad f(J_2)=f(2)\bullet I=\left[\sqrt{2}\right]$$

$$f(A)=f(J)\ =\begin{bmatrix}f(J_1)&\\&f(J_2)\end{bmatrix}=\begin{bmatrix}1&1/2&0\\&1&0\\&&\sqrt{2}\end{bmatrix}$$