矩阵论

第 07 讲

矩阵级数与矩阵函数

上讲回顾

第六讲　Jordan 标准形应用

- Jordan 标准形变换矩阵求解
- Jordan 标准形的幂及
- 矩阵序列

同一特征值可能出现在不同的 Jordan 块中

各 Jordan 块阶数高低一次进行处理，高阶先处理，低阶后处理，同阶同时处理。

最高阶（没有属于同一特征值的 Jordan 块同阶）

$(A-\lambda_i I)^{m_i}x=0$ 但 $(A-\lambda_i I)^{m_i-1}x\neq 0$ 的 x 作为 P_{im_i}

$P_{i(j-1)}=(A-\lambda_i I)P_j$ $P_{im_{i-1}},P_{im_{i-2}},\cdots$，直至 P_{i1}

对于上述新 Jordan 块，它的 P_{im_i} 不仅要考虑到满足

$$(A - \lambda_i I)^{m_i} x = 0$$

但 $$(A - \lambda_i I)^{m_i - 1} x \neq 0$$

而且还应与前述 P_{ij} 线性无关。

其它属于同一特征值的 Jordan 块处理时，按照较低阶的原则处理即可。

出现多个属于同一特征值的 Jordan 块同阶时，还应考虑线性无关问题。

Jordan 标准形的幂及多项式

$$J^k=\begin{bmatrix} J_1^k & & & \\ & J_2^k & & \\ & & \mathbf{O} & \\ & & & J_s^k \end{bmatrix}$$

$$J_i=\begin{bmatrix} \lambda_i & 1 & & 0 \\ & \lambda_i & \mathbf{O} & \\ & & \mathbf{O} & 1 \\ 0 & & & \lambda_i \end{bmatrix}$$

J_i^k 亦为类似的上三角形条带矩阵，在与主对角线平行的斜线上各元素相等 .

$$J_i^2=\begin{bmatrix} \lambda_i^2 & 2\lambda_i & 1 & & & 0 \\ & \lambda_i^2 & 2\lambda_i & 1 & & \\ & & \lambda_i^2 & \mathbf{O} & & \\ & & & \mathbf{O} & \mathbf{O} & 1 \\ & 0 & & & \lambda_i^2 & 2\lambda_i \\ & & & & & \lambda_i^2 \end{bmatrix}$$

$$f(x)=\sum_{k=0}^{m} a_k x^k = a_0 + a_1 x + \ldots + a_m x^m$$

$$f(J_i)=a_0\begin{bmatrix}\begin{bmatrix}1 & & & & 0\\ & 1 & & & \\ & & 0 & 0 & \\ & & 0 & 0 & \\ 0 & & & 1 & \\ & & & & 1\end{bmatrix}\end{bmatrix}+a_1\begin{bmatrix}\begin{bmatrix}\lambda & 1 & & & & 0\\ & \lambda & 1 & & & \\ & & & 0 & 0 & \\ & & & 0 & 0 & \\ & 0 & & & \lambda & 1\\ & & & & & \lambda\end{bmatrix}\end{bmatrix}+a_2\begin{bmatrix}\begin{bmatrix}\lambda_i^2 & 2\lambda_i & 1 & & & 0\\ & \lambda_i^2 & 2\lambda_i & 1 & & \\ & & \lambda_i^2 & 0 & 0 & \\ & & & 0 & 0 & 1\\ & 0 & & & \lambda_i^2 & 2\lambda_i\\ & & & & & \lambda_i^2\end{bmatrix}\end{bmatrix}+L$$

$$=\begin{bmatrix}
f(\lambda_i) & f'(\lambda_i) & \frac{1}{2!}f^{(2)}(\lambda_i) & \ldots & \frac{1}{m!}f^{(m)}(\lambda_i) & \frac{1}{(m+1)!}f^{(m+1)}(\lambda_i) & \ldots & \frac{1}{m_i!}f^{(m_i)}(\lambda_i)\\
 & f(\lambda_i) & f'(\lambda_i) & \frac{1}{2!}f^{(2)}(\lambda_i) & 0 & & & 0\\
 & & f(\lambda_i) & f'(\lambda_i) & 0 & 0 & & \frac{1}{(m+1)!}f^{(m+1)}(\lambda_i)\\
 & & & f(\lambda_i) & 0 & 0 & 0 & \frac{1}{m!}f^{(m)}(\lambda_i)\\
 & & & & 0 & f'(\lambda_i) & \frac{1}{2!}f^{(2)}(\lambda_i) & 0\\
 & & & & & f(\lambda_i) & f'(\lambda_i) & \frac{1}{2!}f^{(2)}(\lambda_i)\\
 & & & & & & f(\lambda_i) & f'(\lambda_i)\\
 & & & & & & & f(\lambda_i)
\end{bmatrix}$$

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

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

不收敛的级数则称为发散的，其中又分为有界和无界的情况.

2. 收敛矩阵序列的性质：

设 $A^{(k)}$，$B^{(k)}$ 分别收敛于 A,B 则

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

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

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

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

3 收敛矩阵：设 A 为方阵，且当 $k\to\infty$ 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.

上讲作业：习题 1.3 ， 11

11. 对于下列矩阵A，求正交（酉）矩阵P，使 $P^{-1}AP$ 为对角矩阵。

(1) $A=\begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \end{pmatrix}$ (2) $A=\begin{pmatrix} 0 & i & 1 \\ -i & 0 & 0 \end{pmatrix}$

第一步：求特征值。$\det(\lambda I-A)=(\lambda-1)^2(\lambda-10)$

$$\lambda_1=\lambda_2=1,\quad \lambda_3=10$$

第二步：求特征向量。对应 $\lambda_1=\lambda_2=1$：

$$x_1=(-2,1,0)^T,\quad x_2=(2,0,1)^T$$

第三步：施密特正交化 + 单位化。正交化得：

$$y_1=(-2,1,0)^T,\quad y_2=\left(\frac{2}{5},\frac{4}{5},1\right)^T$$

（重根对应的特征向量需先正交化，再单位化）

上讲作业：习题 1.3 ， 11

单位化 + 正交矩阵 P　单位化 y_1, y_2，并求 $\lambda_3 = 10$ 的特征向量：

$$p_1 = \left(-\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}, 0\right)^T$$

$$p_2 = \left(\frac{2}{3\sqrt{5}}, \frac{4}{3\sqrt{5}}, \frac{5}{3\sqrt{5}}\right)^T$$

$$p_3 = \left(-\frac{1}{3}, -\frac{2}{3}, \frac{2}{3}\right)^T$$

以 p_1, p_2, p_3 为列，得正交矩阵：

$$\mathrm{P} = \begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{2}{3\sqrt{5}} & -\frac{1}{3} \\ \frac{1}{\sqrt{5}} & \frac{4}{3\sqrt{5}} & -\frac{2}{3} \\ 0 & \frac{5}{3\sqrt{5}} & \frac{2}{3} \end{pmatrix}$$

使 $P^T AP$ 为对角矩阵：

$$P^T \mathrm{AP} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 10 \end{pmatrix}$$

上讲作业：习题 1.3 ， 11

(2) $A=\begin{pmatrix} 0 & i & 1 \\ -i & 0 & 0 \end{pmatrix}$

特征值：$\det(\lambda I-A)=\lambda(\lambda-\sqrt{2})(\lambda+\sqrt{2})$　　$\lambda_1=0,\quad \lambda_2=\sqrt{2},\quad \lambda_3=-\sqrt{2}$

特征向量（已单位化）：

$p_1=\left(0,\frac{i}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)^T$

$p_2=\left(\frac{1}{\sqrt{2}},-\frac{i}{2},\frac{1}{2}\right)^T$

$p_3=\left(-\frac{1}{\sqrt{2}},-\frac{i}{2},\frac{1}{2}\right)^T$

得酉矩阵：

$$\mathrm{P}=\begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & -\frac{i}{2} & -\frac{i}{2} \\ \frac{1}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

使 $P^H\mathrm{AP}=\begin{pmatrix} 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix}$

矩阵级数与矩阵函数

- 矩阵级数
- 方阵的幂级数
- 矩阵函数
- 矩阵函数的计算
- 利用 Jordan 标准形求矩阵函数
- 利用零化多项式求解矩阵函数

矩阵级数

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

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

不收敛的级数必为发散的 .

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. 绝对收敛矩阵的性质

(1) 绝对收敛级数一定收敛，且任意调换它的项所得的级数仍收敛，并具有相同的和 .

(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

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

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

I. 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 $\sum_{k=0}^{\infty} A^k$ 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 $(A^k)_{ij} \underset{k\to\infty}{\to} 0$, 即 $A^k \underset{k\to\infty}{\to} 0$

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 λ，$(I - A)$ 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： $1 - \mu = \lambda \rightarrow \mu = 1 - \lambda$

故 $0 < |\mu| < 2 \rightarrow \mu \neq 0$ ，$(I - A)$ 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 $(I - A)^{-1}$

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

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 $\varphi(z)=\sum_{k=0}^{\infty} c_k z^k$ 的收敛圆内，则矩阵幂级数

$$\varphi(A)=\sum_{k=0}^{\infty} c_k A^k,\ (A^0=I)$$

是绝对收敛的．反之，若 A 存在落在 $\varphi(z)$ 的收敛圆外的特征值，则 $\varphi(A)$ 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，$\varphi(A)$ 均收敛．

矩阵函数

如： e^{A} , sinA, cosA

以矩阵为自变量的" 函数" (实际上是"函矩阵")

我们知道 ,

$$e^{z} = 1 + z + \frac{1}{2!}z^{2} + \ldots = \sum_{n=0}^{\infty} \frac{1}{n!}z^{n}$$

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

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

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 $e^{A}e^{B}$ ，$e^{B}e^{A}$， e^{A+B} 三者互不相等．例如

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

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

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

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

可见

$$e^{A}e^{B}\neq e^{B}e^{A}$$

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

$$(A+B)^3=2^2(A+B) \qquad e^{A+B}=I+(\sum_{n=1}^{\infty}\frac{1}{n!}2^{n-1})(A+B)=I+\frac{1}{2}(e^2-1)(A+B)=\begin{bmatrix}e^2 & 0\\ 0 & 1\end{bmatrix}$$

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：

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

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

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

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

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

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} 总存在逆矩阵

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

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

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

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

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

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

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

即 $\varphi(A)=0$

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：、 π 、-π 0、0, 求 sinA、cosA、 e^{A}

解

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

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

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

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

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

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

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

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

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

$$= I + A + \frac{\cos\pi - 1}{\pi^{2}} A^{2} + \frac{\sin\pi - \pi}{\pi^{3}} A^{3}$$

利用 Jordan 标准形求矩阵函数。

对于矩阵的多项式，我们曾导出 $f(A)=Pf(J)P^{-1}$， f：多项式

$$f(J)=\begin{bmatrix} f(J_1) & & & \\ & f(J_2) & & \\ & & \ddots & \\ & & & f(J_s) \end{bmatrix}$$

$$f(J_i)=\begin{bmatrix} f(\lambda_i) & f'(\lambda_i) & \frac{1}{2!}f''(\lambda_i) & \cdots & \frac{1}{(m_i-1)!}f^{(m_i-1)}(\lambda_i) \\ & \ddots & \ddots & \ddots & \\ & & \ddots & & \end{bmatrix}$$

实际上，以上结果不仅对矩阵的多项式成立，对矩阵的幂级数也成立。由此引出矩阵函数的另一种定义及计算方法。

1. 定义：设 n 阶矩阵 A 的 Jordan 标准形为 J

$$J=\begin{bmatrix} J_1 & & & & \\ & J_2 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & J_s \end{bmatrix} \qquad J(\lambda_i)=\begin{bmatrix} \lambda_i & 1 & & & 0 \\ & \lambda_i & 1 & & \\ & & \lambda_i & \ddots & \\ & & & \ddots & 1 \\ 0 & & & & \lambda_i \end{bmatrix}$$

且有非奇异矩阵 P 使得：$P^{-1}AP=J$

对于函数 f(z)，若下列函数

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

均有意义，则称矩阵函数 f(A) 有意义，且

$$f(A) = Pf(J)P^{-1} = P\begin{bmatrix} f(J_1) & & & & \\ & f(J_2) & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & f(J_s) \end{bmatrix} P^{-1}$$

$$f(J_i) = \begin{bmatrix} f(\lambda_i) & f'(\lambda_i) & \frac{1}{2!}f''(\lambda_i) & \cdots & \frac{1}{(m_i-1)!}f^{(m_i-1)}(\lambda_i) \\ & \ddots & \ddots & \ddots & \\ & & & & \end{bmatrix}_{m_i \times m_i}$$

2. 矩阵函数的求法（步骤）：

(1) 求出 A 的 Jordan 标准形及变换矩阵 P，$P^{-1}AP = J$

(2) 对于 J 的各 Jordan 块 J_i 求出，即计算 $f(J_i)$ 出

$$f(\lambda_i), f'(\lambda_i), \ldots\ldots, f^{(m_i-1)}(\lambda_i)$$

并按照顺序构成 $f(J_i)$

$$f(J_i) = \begin{bmatrix} f(\lambda_i) & f'(\lambda_i) & \frac{1}{2!}f''(\lambda_i) & \cdots & \frac{1}{(m_i-1)!}f^{(m_i-1)}(\lambda_i) \\ & \ddots & \ddots & \ddots & \\ & & & & \end{bmatrix}_{m_i \times m_i}$$

(3) 合成

$$f(J)=\begin{bmatrix} f(J_1) & & & & \\ & f(J_2) & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & f(J_s) \end{bmatrix}$$

(4) 矩阵乘积给出

$$f(A)=Pf(J)P^{-1}$$

需要说明的是，计算结果与 Jordan 标准形中 Jordan 块的顺序无关。

例 1 （教材 P176 例 3-8 ）

$$A=\begin{bmatrix}1 & 2 & 3 & 4\\ & 1 & 2 & 3\\ & & 1 & 2\\ & & & 1\end{bmatrix} \qquad \sqrt{A}$$

[解] (1) 求出 J 及 P

$$J=\begin{bmatrix}1 & 1 & 0 & 0\\ & 1 & 1 & 0\\ & & 1 & 1\\ & & & 1\end{bmatrix},P=\begin{bmatrix}8 & 4 & 0 & 0\\ & 4 & -1 & 1\\ & & 2 & -2\\ & & & 1\end{bmatrix},P^{-1}=\frac{1}{16}\begin{bmatrix}2 & -2 & -1 & 0\\ & 4 & 2 & 0\\ & & 8 & 16\\ & & & 16\end{bmatrix}$$

(3) 求出$f(\lambda_i), f'(\lambda_i), \ldots\ldots, f^{(m_i-1)}(\lambda_i)$ 并构成(J_i) ：

$$\lambda_1 = 1, m_1 = 4, f(z) = \sqrt{z}$$

$$f(1)=1$$

$$f'(1) = \frac{1}{2} z^{-\frac{1}{2}} |_{z=1} = \frac{1}{2}, f''(1) = -\frac{1}{4} z^{-\frac{3}{2}} |_{z=1} = -\frac{1}{4}, f'''(1) = \frac{3}{8} z^{-\frac{5}{2}} |_{z=1} = \frac{3}{8}$$

$$f(J_1) = \begin{bmatrix} 16 & 8 & -2 & 1 \\ & 16 & 8 & -2 \\ & & 16 & 8 \\ & & & 16 \end{bmatrix} \frac{1}{16}$$

(3) 合成 $f(J)=f(J_1)$

(4) 求 $f(A)=Pf(J)P^{-1}$

$$f(A)=\begin{bmatrix} 1 & 1 & 1 & 1 \\ & 1 & 1 & 1 \\ & & 1 & 1 \\ & & & 1 \end{bmatrix}$$

说明：

（1）$f(z)=\sqrt{z}$ ，在 z = 0 不存在泰勒展开 (而存在洛朗展开), 如按原先的幂级数定义，则根本无从谈 $\sqrt{A}$ 的计算，可见新的定义延拓了原来的定义；

(2)

$$[f(A)]^2=\begin{bmatrix}1&1&1&1\\&1&1&1\\&&1&1\\&&&1\end{bmatrix}^2=\begin{bmatrix}1&2&3&4\\&1&2&3\\&&1&2\\&&&1\end{bmatrix}=A$$

可见这样的 $\sqrt{A}$ 确与 A^2 构成反函数；

(3) 矩阵函数的种类不仅是我们介绍的这种，如辛矩阵。以

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

为例，以我们这里的 $\sqrt{A}=\begin{bmatrix}\pm i & 0\\ 0 & \pm i\end{bmatrix}$ ，但 $B=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}$

亦满足 $B^2=A$ ，即 B 也可以看作某种 $\sqrt{A}$

利用零化多项式求解矩阵函数

利用 Jordan 标准型求解矩阵函数的方法比较复杂，它需要求 J 和 P 。下面我们介绍根据零化多项式求解矩阵函数的一种方法。

定律： n 阶方阵 A 的最小多项式等于它的特征矩阵的第 n 个（也就是最后一个）不变因子 $d_n(\lambda)$ （可参见张远达《线性代数原理》 P215 ）

设 n 阶方阵 A 的不变因子反向依次为 $d_n(\lambda)$, $d_{n-1}(\lambda),\cdots,d_1(\lambda)$
由它们给出的初等因子分别为

$$(\lambda-\lambda_1)^{m_1},(\lambda-\lambda_2)^{m_2},\cdots,(\lambda-\lambda_r)^{m_r}$$

$$(\lambda-\lambda_{r+1})^{m_{r+1}},\cdots,(\lambda-\lambda_s)^{m_s}$$

$$\sum_{i=1}^{s} m_i = n$$

由于 $d_1(\lambda)\,|\,d_2(\lambda), d_2(\lambda)\,|\,d_3(\lambda),\cdots,d_{n-1}(\lambda)\,|\,d_n(\lambda)$

故 (1) $\lambda_{r+1}\sim\lambda_s$ 必定出现在 $\lambda_1\sim\lambda_r$ 中;

(2) 若 $\lambda_i(i>r)=\lambda_j(j\le r)$ 则 $m_i\le m_j$

根据上述定理， A 的最小多项式

$$\varphi_0(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2}\cdots(\lambda - \lambda_r)^{m_r}$$

$$(\lambda_1 I - A)^{m_1}(\lambda_2 I - A)^{m_2}\cdots(\lambda_r I - A)^{m_r} = O$$

令 $m = \sum_{i=1}^{r} m_i$ ，则 可见 A^m 可以由 $A^0 = I, A, A^2, \cdots, A^{m-1}$

线性表示，从而 $A^{m+i}(\lambda > 0)$ 亦可由 $A^0 = I, A, A^2, \cdots, A^{m-1}$ 线性表

所以，矩阵函数 f(A) 若存在，也必定可由 $A^0 \sim A^{m-1}$ 线性表示。

因此，我们定义一个系数待定的 (m-1) 次多项式 $g(\lambda)=\sum_{i=0}^{m-1} c_i\lambda^i$，根据以上论述，适当选择系数 $c_0 \sim c_{m-1}$，就可以使 $f(A)=g(A)$

又，假设 J、P 分别为 A 的 Jordan 标准形及相应变换矩阵，$A=PJP^{-1}$

$$\begin{matrix} f(A)=Pf(J)P^{-1} \\ g(A)=Pg(J)P^{-1} \end{matrix} \rightarrow \quad f(J)=g(J) \quad \rightarrow \quad f(J_i)=g(J_i)$$

$$f(\lambda_i)=g(\lambda_i), f'(\lambda_i)=g'(\lambda_i), \cdots, f^{(m_i-1)}(\lambda_i)=g^{(m_i-1)}(\lambda_i) \quad (i=1,2,\cdots,r)$$

由于 g(A) 为待定系数的多项式，上面就成为关于 $c_0 \sim c_{m-1}$ 的线性方程组。且方程 的个数为 $n=\sum_{i=1}^{r} m_i$ 等于未知数个数，正好可以确定

由此给出根据最小多项式求矩阵函数的一般方法。

(1) 求出最小多项式 $\varphi_0(\lambda)=d_n(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_r)^{m_r},\sum_{i=1}^{r}m_i=m$

(2) 形式上写出待定多项式 $g(\lambda)=\sum_{i=0}^{m-1}c_i\lambda^i=c_0+c_1\lambda+c_2\lambda^2+\cdots+c_{m-1}\lambda^{m-1}$

(3) 求解关于 $c_0\sim c_{m-1}$ 的线性方程组

$$g^{(k)}(\lambda_i)=f^{(k)}(\lambda_i)\quad(k=1,2,\cdots,m_i;i=1,2,\cdots,r)$$

(4) 求出 g(A)，即可得 f(A)=g(A).

例 2 ，采用新方法计算 A 的函数 。

$$A=\begin{bmatrix}1 & 2 & 3 & 4\\ & 1 & 2 & 3\\ & & 1 & 2\\ & & & 1\end{bmatrix} \qquad \sqrt{A}$$

(1) $\varphi(\lambda)=\varphi_0(\lambda)=(\lambda-1)^4.\quad m_1=4=m=n,\lambda_1=1$

(2) $g(\lambda)=c_0+c_1\lambda+c_2\lambda^2+c_3\lambda^3$

(3) 方程组为

$$g(1)=f(1)=1=c_0+c_1+c_2+c_3$$

$$g'(1)=f'(1)=\frac{1}{2}=c_1+2c_2+3c_3$$

$$g''(1)=f''(1)=-\frac{1}{4}=2c_2+6c_3$$

$$g'''(1)=f'''(1)=\frac{3}{8}=6c_3$$

$$c_3=\frac{1}{16},c_2=-\frac{5}{16},c_1=\frac{15}{16},c_0=\frac{5}{16}$$

(4)

$$g(A)=\frac{1}{16}(5I+15A-5A^2+A^3)$$

$$A^2=\begin{bmatrix}1&4&10&20\\&1&4&10\\&&1&4\\&&&1\end{bmatrix}$$

$$A^3=\begin{bmatrix}1&6&21&56\\&1&6&21\\&&1&6\\&&&1\end{bmatrix}$$

$$f(A)=\frac{1}{16}\left\{\begin{bmatrix}5&0&0&0\\&5&0&0\\&&5&0\\&&&5\end{bmatrix}+\begin{bmatrix}15&30&45&60\\&15&30&45\\&&15&30\\&&&15\end{bmatrix}-\begin{bmatrix}5&20&50&100\\&5&20&50\\&&5&20\\&&&5\end{bmatrix}+\begin{bmatrix}1&6&21&56\\&1&6&21\\&&1&6\\&&&1\end{bmatrix}\right\}=\begin{bmatrix}1&1&1&1\\&1&1&1\\&&1&1\\&&&1\end{bmatrix}$$

与 Jordan 标准形方法完全一致。

作业： 习题 3.3 ， 3,4,5,6