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$\{A^{(k)}\}$，其中$A^{(k)}=\left(a_{ij}^{(k)}\right)$，且当$k\to\infty$时$a_{ij}^{(k)}\to a_{ij}$，则称$\{A^{(k)}\}$收敛，并把$A=\left(a_{ij}\right)$叫做$\{A^{(k)}\}$的极限，或称$\{A^{(k)}\}$收敛于$A$，记为

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

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$\{A^{(k)}\}$，若存在常数$M>0$，使得对一切k都有

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

则称$\{A^{(k)}\}$为有界的。

2. 收敛矩阵序列的性质:

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

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

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

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

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

3. 收敛矩阵的定义:

设 A 为方阵,若当 $k\to\infty$ 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。

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

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

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

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

$A^k \to 0$就等价于$J_i^k \to 0 (i = 1,2,...,s)$，等价于$\lambda_i^k \to 0 (i = 1,2,...,s)$，而这只有$|\lambda_i| < 1$才可能也必能。 [证毕]

二、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，记为$\sum_{k=1}^{\infty}A^{(k)}$。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$S$，则称该矩阵级数收敛，且有和$S$。 记为

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

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

$$\sum_{k=1}^{\infty}\Big(\sum_{i=1}^{k}A^{(i)}B^{(k+1-i)}\Big)=S_1S_2$$

三、方阵的幂级数

1. 定义： 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。

2. 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

$$(A^k)_{ij} \xrightarrow[k\to\infty]{} 0\,,\quad \text{即 } A^k \xrightarrow[k\to\infty]{} 0$$

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$1-\mu=\lambda\to\mu=1-\lambda$，故$0<|\mu|<2\to\mu\neq0$，$(I-A)$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，$A^{k+1}\to0$，故$A^{k+1}(I-A)^{-1}\to0$。 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$(I-A)^{-1}$。

[证毕]

3. 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。

[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。

四、矩阵函数

以 n 阶矩阵为自变量和函数值（因变量）的一种函数。

1. 定义： 设一元函数 $f(z)$ 能展开为 z 的幂级数：

$$f(z)=\sum_{k=0}^{\infty}c_k z^k \qquad (|z|<r)$$

其中 $r>0$ 表示该幂级数的收敛半径。当 n 阶矩阵 A 的特征值全部落在该幂级数的收敛圆内时，把收敛的矩阵幂级数 $\sum_{k=0}^{\infty}c_k A^k$ 的和称为矩阵函数，即为 $f(A)$，即：

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

例如：已知$e^{z}=1+z+\frac{1}{2!}z^{2}+\cdots=\sum_{n=0}^{\infty}\frac{1}{n!}z^{n}$

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

均为整个复平面上收敛的级数，故对任意方阵 A

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

均绝对收敛。三者分别称为矩阵指数函数、矩阵正弦函数和矩阵余弦函数。

2. 性质

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

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

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

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

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

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

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

可见$e^Ae^B\neq e^Be^A$

$$A+B=\begin{bmatrix}2 & 0\\ 0 & 0\end{bmatrix}, \quad (A+B)^2=2\begin{bmatrix}2 & 0\\ 0 & 0\end{bmatrix}=2(A+B), \quad (A+B)^3=2^2(A+B), \cdots$$

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

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

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

$$\cos(A\pm B)=\cos A\cos B\mp\sin A\sin B$$

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

证明:

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

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

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

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

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

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
$$\varphi(A)=A^n+c_1A^{n-1}+\cdots+c_{n-1}A+c_nI=0$$

[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 $\varphi(A)=0$ [证毕]

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。

3. 最小多项式:

首项系数为 1 （简称首 1 ），次数最小，且以矩阵 A 为零点的 λ 的多项式，称为 A 的最小多项式，常用 $m(\lambda)$ 表示。

例 1: 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 $\varphi(A)=A^4-\pi^2A^2=0 \to A^4=\pi^2A^2, A^5=\pi^2A^3, A^6=\pi^2A^4=\pi^4A^2,\cdots$

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

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

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

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

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

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

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

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

六、利用Jordan标准形求矩阵函数

对于矩阵的多项式，我们曾导出 $f(A) = Pf(J)P^{-1}$

$$f(J) = \begin{bmatrix} f(J_1) & & & & \\ & f(J_2) & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & f(J_s) \end{bmatrix}, \quad f(J_i) = \begin{bmatrix} f(\lambda_i) & f'(\lambda_i) & \frac{1}{2!}f''(\lambda_i) & \cdots & \frac{1}{(m_i-1)!}f^{(m_i-1)}(\lambda_i) \\ & \ddots & \ddots & \ddots & \end{bmatrix}$$

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

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

且有非奇异矩阵 P 使得： $P^{-1}AP=J$

若函数 $f(z)$ 在 λ_i 处具有直到 m_i-1 阶导数（$i=1,2,\cdots,s$），则称

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

为对应于 $f(z)$ 的矩阵函数。这拓展了矩阵函数的定义，使得无法展开成收敛的幂级数的函数也有了意义。如： $f(z)=\dfrac{1}{z}$。

2. 利用 Jordan 标准形求矩阵函数的步骤

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 $f(J_i)$，即计算出

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

并按照顺序构成 $f(J_i)$，

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

计算结果与 Jordan 标准形中 Jordan 块的顺序无关。

例 2. $A=\begin{bmatrix}1&2&3&4\\ &1&2&3\\ &&1&2\\ &&&1\end{bmatrix}$，求$\sqrt{A}$

解：1° 求出 J 及 P

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

2° 求出 $f(\lambda_i), f'(\lambda_i),\ldots\ldots, f^{(m_i-1)}(\lambda_i)$ 并构成 $f(J_i)$：

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

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

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

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 $f(A)$ 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；

七、用零化多项式求解矩阵函数

定理：n 阶方阵 A 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 A 的最小多项式为：

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

即

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

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$A^{m+i}(i>0)$亦可由$A^0=I,A,A^2,\cdots,A^{m-1}$线性表示。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，可以定义一个系数待定的$(m\text{-}1)$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$f(A)=g(A)$

假设J、P分别为A的 Jordan 标准形及相应变换矩阵：

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

则 $f(A)=Pf(J)P^{-1}$，$g(A)=Pg(J)P^{-1} \to f(J)=g(J) \to f(J_i)=g(J_i)$

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

由于$g(\lambda)$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$m=\sum_{i=1}^{r} m_i$等于未知数个数，正好可以确定$c_0 \sim c_{m-1}$。

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

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

4° 求出 $g(A)$，即可得 $f(A)=g(A)$。

从推导的过程看，似乎不仅最小多项式可用于矩阵函数的计算，一般零化多项式也可以，其中以特征多项式最为方便。

例 3. 计算 $A=\begin{bmatrix}1&2&3&4\\ &1&2&3\\ & &1&2\\ & & &1\end{bmatrix}$ 的函数 $\sqrt{A}$ 。（ $f(\lambda)=\sqrt{\lambda}$ ）

解： 1° $\varphi(\lambda)=\varphi_0(\lambda)=(\lambda-1)^4$. $m_1=4=n$ ， $\lambda_1=1$;

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

3° 方程组为

$$\left.\begin{aligned}
g(1)&=f(1)=1=c_0+c_1+c_2+c_3\\
g'(1)&=f'(1)=\frac{1}{2}=c_1+2c_2+3c_3\\
g''(1)&=f''(1)=-\frac{1}{4}=2c_2+6c_3\\
g'''(1)&=f'''(1)=\frac{3}{8}=6c_3
\end{aligned}\right\} \Rightarrow c_3=\frac{1}{16},c_2=-\frac{5}{16},c_1=\frac{15}{16},c_0=\frac{5}{16}$$

4° $g(A)=\dfrac{1}{16}(5I+15A-5A^2+A^3)$

$$A^2=\begin{bmatrix}1&4&10&20\\&1&4&10\\&&1&4\\&&&1\end{bmatrix},\quad A^3=\begin{bmatrix}1&6&21&56\\&1&6&21\\&&1&6\\&&&1\end{bmatrix}$$

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

$$=\begin{bmatrix}1&1&1&1\\&1&1&1\\&&1&1\\&&&1\end{bmatrix}$$

作业：P163　3, 4, 5,6