0.1 矩阵函数

引理 0.1

设 $A \in M_n(\mathbb{C})$, 则存在 P 是非异阵, 使

$$P^{-1}AP = J = \mathrm{diag}\{J_1, J_2, \cdots, J_k\}$$

是 A 的 Jordan 标准型, 其中 J_i 是 A 的特征值 λ_i 的 r 阶 Jordan 块. 若 $f(x) = a_0 + a_1x + \cdots + a_px^p$, 则

$$f(A) = P\mathrm{diag}\left\{f\left(J_1\right), f\left(J_2\right), \cdots, f\left(J_k\right)\right\} P^{-1},$$

其中

$$f(J_i) = \begin{pmatrix} f(\lambda_i) & \frac{1}{1!}f'(\lambda_i) & \frac{1}{2!}f^{(2)}(\lambda_i) & \cdots & \frac{1}{(r-1)!}f^{(r-1)}(\lambda_i) \\ & f(\lambda_i) & \frac{1}{1!}f'(\lambda_i) & \cdots & \frac{1}{(r-2)!}f^{(r-2)}(\lambda_i) \\ & & f(\lambda_i) & \cdots & \frac{1}{(r-3)!}f^{(r-3)}(\lambda_i) \\ & & & \ddots & \vdots \\ & & & & f(\lambda_i) \end{pmatrix}.$$

证明 注意到

$$J^m = \mathrm{diag}\{J_1^m, J_2^m, \cdots, J_k^m\}.$$

又

$$A^m = (PJP^{-1})^m = PJ^mP^{-1},$$

因此要计算 $f(A)$, 只需计算出 J_i^m 即可. 利用二项式定理和数学归纳法不难证明

$$J_i^m = \left[\lambda_i I_r + \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & \ddots & \\ & & & \ddots & 1 \\ & & & & 0 \end{pmatrix}\right]^m = \begin{pmatrix} \lambda_i^m & \mathrm{C}_m^1\lambda_i^{m-1} & \mathrm{C}_m^2\lambda_i^{m-2} & \cdots & \cdots \\ & \lambda_i^m & \mathrm{C}_m^1\lambda_i^{m-1} & \cdots & \cdots \\ & & \lambda_i^m & \cdots & \cdots \\ & & & \ddots & \vdots \\ & & & & \lambda_i^m \end{pmatrix}.$$

则不难算出

$$f(J_i) = \begin{pmatrix} f(\lambda_i) & \frac{1}{1!}f'(\lambda_i) & \frac{1}{2!}f^{(2)}(\lambda_i) & \cdots & \frac{1}{(r-1)!}f^{(r-1)}(\lambda_i) \\ & f(\lambda_i) & \frac{1}{1!}f'(\lambda_i) & \cdots & \frac{1}{(r-2)!}f^{(r-2)}(\lambda_i) \\ & & f(\lambda_i) & \cdots & \frac{1}{(r-3)!}f^{(r-3)}(\lambda_i) \\ & & & \ddots & \vdots \\ & & & & f(\lambda_i) \end{pmatrix}.$$

再由

$$\begin{aligned} f(A) &= f(PJP^{-1}) = Pf(J)P^{-1} \\ &= Pf(\mathrm{diag}\{J_1, J_2, \cdots, J_k\})P^{-1} \\ &= P\mathrm{diag}\{f(J_1), f(J_2), \cdots, f(J_k)\}P^{-1}, \end{aligned}$$

即可计算出 $f(A)$. □

定义 0.1 (复方阵幂级数)

设有 n 阶复方阵序列 $\{A_p\}$:

$$A_p = \begin{pmatrix} a_{11}^{(p)} & \cdots & a_{1n}^{(p)} \\ \vdots & & \vdots \\ a_{n1}^{(p)} & \cdots & a_{nn}^{(p)} \end{pmatrix},$$

$B = (b_{ij})$ 是一个同阶方阵, 若对每个 (i, j), 序列 $\{a_{ij}^{(p)}\}$ 均收敛于 b_{ij}, 即

$$\lim_{p\to\infty} a_{ij}^{(p)} = b_{ij},$$

则称矩阵序列 $\{A_p\}$ 收敛于 B, 记为

$$\lim_{p\to\infty} A_p = B.$$

否则称 $\{A_p\}$ 发散.

设

$$f(z) = a_0 + a_1 z + a_2 z^2 + \cdots$$

是一个幂级数, 记

$$f_p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_p z^p$$

是其部分和. 若矩阵序列 $\{f_p(A)\}$ 收敛于 B, 则称矩阵级数

$$f(A) = a_0 I + a_1 A + a_2 A^2 + \cdots .$$

收敛, 极限为 B, 记为 $f(A) = B$. 否则称 $f(A)$ 发散. 用变量矩阵 X 代替 A, 便可定义矩阵幂级数

$$f(X) = a_0 I + a_1 X + a_2 X^2 + \cdots .$$

♣

定理 0.1

设 $f(z) = \sum_{i=0}^{\infty} a_i z^i$ 是复幂级数, 则

(1) 方阵幂级数 $f(X)$ 收敛的充分必要条件是对任一非异阵 P,$f(P^{-1}XP)$ 都收敛, 这时

$$f(P^{-1}XP) = P^{-1} f(X) P;$$

(2) 若 $X = \mathrm{diag}\{X_1, \cdots, X_k\}$, 则 $f(X)$ 收敛的充分必要条件是 $f(X_1), \cdots, f(X_k)$ 都收敛, 这时

$$f(X) = \mathrm{diag}\{f(X_1), \cdots, f(X_k)\};$$

(3) 若 $f(z)$ 的收敛半径为 r,J_0 是特征值为 λ_0 的 n 阶 Jordan 块

$$J_0 = \begin{pmatrix} \lambda_0 & 1 & & & \\ & \lambda_0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_0 \end{pmatrix},$$

则当 $|\lambda_0| < r$ 时 $f(J_0)$ 收敛, 且

$$f(J_0) = \begin{pmatrix} f(\lambda_0) & \frac{1}{1!}f'(\lambda_0) & \frac{1}{2!}f^{(2)}(\lambda_0) & \cdots & \frac{1}{(n-1)!}f^{(n-1)}(\lambda_0) \\ & f(\lambda_0) & \frac{1}{1!}f'(\lambda_0) & \cdots & \frac{1}{(n-2)!}f^{(n-2)}(\lambda_0) \\ & & f(\lambda_0) & \cdots & \frac{1}{(n-3)!}f^{(n-3)}(\lambda_0) \\ & & & \ddots & \vdots \\ & & & & f(\lambda_0) \end{pmatrix}. \tag{1}$$

♡

证明 设 $f_p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_p z^p$ 是 $f(z)$ 前 $p+1$ 项的部分和.

(1) 注意到 $f_p(z)$ 是多项式, 从而有

$$f_p(P^{-1}XP) = P^{-1}f_p(X)P.$$

由于 n 阶矩阵序列的收敛等价于 n^2 个数值序列的收敛, 故

$$\begin{aligned} f(P^{-1}XP) &= \lim_{p\to\infty} f_p(P^{-1}XP) = \lim_{p\to\infty} P^{-1}f_p(X)P \\ &= P^{-1}(\lim_{p\to\infty} f_p(X))P = P^{-1}f(X)P. \end{aligned}$$

(2) 注意到 $f_p(z)$ 是多项式, 从而有

$$f_p(X) = f_p(\mathrm{diag}\{X_1, \cdots, X_k\}) = \mathrm{diag}\{f_p(X_1), \cdots, f_p(X_k)\}.$$

由于分块矩阵序列的收敛等价于每个分块的矩阵序列的收敛, 故

$$\begin{aligned} f(X) &= \lim_{p\to\infty} f_p(X) = \lim_{p\to\infty} \mathrm{diag}\{f_p(X_1), \cdots, f_p(X_k)\} \\ &= \mathrm{diag}\{\lim_{p\to\infty} f_p(X_1), \cdots, \lim_{p\to\infty} f_p(X_k)\} = \mathrm{diag}\{f(X_1), \cdots, f(X_k)\}. \end{aligned}$$

(3) 由引理 0.1可知

$$f_p(J_0) = \begin{pmatrix} f_p(\lambda_0) & \frac{1}{1!}f_p'(\lambda_0) & \frac{1}{2!}f_p^{(2)}(\lambda_0) & \cdots & \frac{1}{(n-1)!}f_p^{(n-1)}(\lambda_0) \\ & f_p(\lambda_0) & \frac{1}{1!}f_p'(\lambda_0) & \cdots & \frac{1}{(n-2)!}f_p^{(n-2)}(\lambda_0) \\ & & f_p(\lambda_0) & \cdots & \frac{1}{(n-3)!}f_p^{(n-3)}(\lambda_0) \\ & & & \ddots & \vdots \\ & & & & f_p(\lambda_0) \end{pmatrix}.$$

令 $p \to \infty$, 由矩阵序列收敛与 n^2 个数值序列收敛的等价性和幂级数的相关性质即得结论.

□

定理 0.2

设 $f(z)$ 是复幂级数, 收敛半径为 r. 设 A 是 n 阶复方阵, 特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 定义 A 的**谱半径**

$$\rho(A) = \max_{1\le i\le n} |\lambda_i|.$$

(1) 若 $\rho(A) < r$, 则 $f(A)$ 收敛;

(2) 若 $\rho(A) > r$, 则 $f(A)$ 发散;

(3) 若 $\rho(A) = r$, 则 $f(A)$ 收敛的充分必要条件是: 对每一模长等于 r 的特征值 λ_j, 若 A 的属于 λ_j 的初等因子中最高幂为 n_j 次, 则 n_j 个数值级数

$$f(\lambda_j), f'(\lambda_j), \cdots, f^{(n_j-1)}(\lambda_j). \tag{2}$$

都收敛;

(4) 若 $f(A)$ 收敛, 则 $f(A)$ 的特征值为

$$f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_n).$$

证明

(1) 设 A 的 Jordan 标准型为 $J = \mathrm{diag}\{J_1, J_2, \cdots, J_k\}$. 显然 $f(A)$ 的收敛性等价于所有 $f(J_i)$（$i = 1, \cdots, k$）的收敛性. 由定理 0.1即知 (1) 成立.

(2) 若某一个 $|\lambda_j| > r$, 则 $f(\lambda_j)$ 发散, 因此 $f(J_j)$ 发散, 故 $f(A)$ 发散, 这就证明了 (2).

(3) 当 $\rho(A) = r$ 时, 对 $|\lambda_i| < r$ 的 J_i,$f(J_i)$ 收敛. 对 $|\lambda_j| = r$ 的特征值 λ_j, 注意到 $f(z)$ 的任意次导数的收敛半径仍为 r, 又初等因子 $(\lambda - \lambda_j)^{n_j}$ 对应的 Jordan 块为 n_j 阶, 从(1)式即可知道 $f(J_j)$ 的收敛性等价于(2)式中 n_j 个级数的收敛性.

(4) 最后若 $f(A)$ 收敛, 则 $f(A)$ 与 $f(J)$ 有相同的特征值, 即为 $f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_n)$.

□

定义 0.2

于是对一切方阵, 定义

$$\begin{aligned} \mathrm{e}^{\boldsymbol{A}} &= \boldsymbol{I} + \frac{1}{1!}\boldsymbol{A} + \frac{1}{2!}\boldsymbol{A}^2 + \frac{1}{3!}\boldsymbol{A}^3 + \cdots, \\ \sin \boldsymbol{A} &= \boldsymbol{A} - \frac{1}{3!}\boldsymbol{A}^3 + \frac{1}{5!}\boldsymbol{A}^5 - \frac{1}{7!}\boldsymbol{A}^7 + \cdots, \\ \cos \boldsymbol{A} &= \boldsymbol{I} - \frac{1}{2!}\boldsymbol{A}^2 + \frac{1}{4!}\boldsymbol{A}^4 - \frac{1}{6!}\boldsymbol{A}^6 + \cdots \end{aligned}$$

都有意义. 若 $\boldsymbol{A}$ 所有特征值的模长都小于 1, 则

$$\ln(\boldsymbol{I} + \boldsymbol{A}) = \boldsymbol{A} - \frac{1}{2}\boldsymbol{A}^2 + \frac{1}{3}\boldsymbol{A}^3 - \frac{1}{4}\boldsymbol{A}^4 + \cdots$$

也有意义. 同理还可以定义幂函数、双曲函数等.

注 由复分析知道:

$$\begin{aligned} \mathrm{e}^z &= 1 + \frac{1}{1!}z + \frac{1}{2!}z^2 + \frac{1}{3!}z^3 + \cdots, \\ \sin z &= z - \frac{1}{3!}z^3 + \frac{1}{5!}z^5 - \frac{1}{7!}z^7 + \cdots, \\ \cos z &= 1 - \frac{1}{2!}z^2 + \frac{1}{4!}z^4 - \frac{1}{6!}z^6 + \cdots, \\ \ln(1+z) &= z - \frac{1}{2}z^2 + \frac{1}{3}z^3 - \frac{1}{4}z^4 + \cdots. \end{aligned}$$

前 3 个级数在整个复平面上收敛, 而 $\ln(1+z)$ 的收敛半径为 1. 于是由定理 0.2可知 $\mathrm{e}^{\boldsymbol{A}}, \sin \boldsymbol{A}, \cos \boldsymbol{A}, \ln \boldsymbol{A}$ 都收敛, 从而都有意义. 故上述定义是良定义的.

命题 0.1

如果 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 乘法可交换, 即 $\boldsymbol{AB} = \boldsymbol{BA}$, 则 $\mathrm{e}^{\boldsymbol{A}} \cdot \mathrm{e}^{\boldsymbol{B}} = \mathrm{e}^{\boldsymbol{A}+\boldsymbol{B}}$ 必成立.

注 对一般来说对矩阵 $\boldsymbol{A}, \boldsymbol{B}$, 下面的等式并不一定成立:

$$\mathrm{e}^{\boldsymbol{A}} \cdot \mathrm{e}^{\boldsymbol{B}} = \mathrm{e}^{\boldsymbol{A}+\boldsymbol{B}}.$$

如对

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \boldsymbol{B} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix},$$

不难验证 $\mathrm{e}^{\boldsymbol{A}} \cdot \mathrm{e}^{\boldsymbol{B}} \neq \mathrm{e}^{\boldsymbol{A}+\boldsymbol{B}}$.

证明

□

命题 0.2

设 t 是一个数值变量,$\boldsymbol{A}$ 是一个 n 阶复方阵. $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{J}=\operatorname{diag}\{\boldsymbol{J}_1,\cdots,\boldsymbol{J}_k\}$ 是 $\boldsymbol{A}$ 的 Jordan 标准型,$\boldsymbol{J}_i$ 是特征值为 λ_i 的 r 阶 Jordan 块, 则

$$\mathrm{e}^{t\boldsymbol{A}}=\boldsymbol{P}\mathrm{e}^{t\boldsymbol{J}}\boldsymbol{P}^{-1},$$

其中

$$\mathrm{e}^{t\boldsymbol{J}}=\begin{pmatrix}\mathrm{e}^{t\boldsymbol{J}_1} & & \\ & \ddots & \\ & & \mathrm{e}^{t\boldsymbol{J}_k}\end{pmatrix},\quad \mathrm{e}^{t\boldsymbol{J}_i}=\mathrm{e}^{t\lambda_i}\begin{pmatrix}1 & t & \frac{1}{2!}t^2 & \frac{1}{3!}t^3 & \cdots & \frac{1}{(r-1)!}t^{r-1}\\ & 1 & t & \frac{1}{2!}t^2 & \cdots & \frac{1}{(r-2)!}t^{r-2}\\ & & 1 & t & \cdots & \frac{1}{(r-3)!}t^{r-3}\\ & & & \ddots & \ddots & \vdots\\ & & & & \ddots & t\\ & & & & & 1\end{pmatrix}.$$

证明 证法一: 若令 $f(z)=\mathrm{e}^{tz}$, 则由定理 0.1即得 $f(\boldsymbol{A})=\mathrm{e}^{t\boldsymbol{A}}$ 的计算结果.

证法二: 注意到

$$\boldsymbol{J}_i=\lambda_i\boldsymbol{I}+\boldsymbol{N},$$

其中 $\boldsymbol{N}$ 是 r 阶基础幂零阵, 即

$$\boldsymbol{N}^r=\boldsymbol{O},\quad \boldsymbol{N}=\begin{pmatrix}0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1\\ & & & & 0\end{pmatrix}.$$

于是

$$\mathrm{e}^{\boldsymbol{N}}=\boldsymbol{I}+\boldsymbol{N}+\frac{1}{2!}\boldsymbol{N}^2+\frac{1}{3!}\boldsymbol{N}^3+\cdots+\frac{1}{(r-1)!}\boldsymbol{N}^{r-1}.$$

因为 $(\lambda_i\boldsymbol{I})\boldsymbol{N}=\boldsymbol{N}(\lambda_i\boldsymbol{I})$, 故由命题 0.1可知

$$\begin{aligned}\mathrm{e}^{\boldsymbol{J}_i}&=\mathrm{e}^{\lambda_i\boldsymbol{I}+\boldsymbol{N}}=\mathrm{e}^{\lambda_i\boldsymbol{I}}\cdot\mathrm{e}^{\boldsymbol{N}}=\mathrm{e}^{\lambda_i}\cdot\mathrm{e}^{\boldsymbol{N}}\\ &=\mathrm{e}^{\lambda_i}\boldsymbol{I}+\mathrm{e}^{\lambda_i}\boldsymbol{N}+\frac{1}{2!}\mathrm{e}^{\lambda_i}\boldsymbol{N}^2+\cdots+\frac{1}{(r-1)!}\mathrm{e}^{\lambda_i}\boldsymbol{N}^{r-1}.\end{aligned}$$

同理

$$\begin{aligned}\mathrm{e}^{t\boldsymbol{J}_i}&=\mathrm{e}^{t\lambda_i}\cdot\mathrm{e}^{t\boldsymbol{N}}=\mathrm{e}^{t\lambda_i}\left[t\boldsymbol{I}+t\boldsymbol{N}+\frac{t}{2!}\boldsymbol{N}^2+\frac{t}{3!}\boldsymbol{N}^3+\cdots+\frac{t}{(r-1)!}\boldsymbol{N}^{r-1}\right]\\ &=\mathrm{e}^{t\lambda_i}\begin{pmatrix}1 & t & \frac{1}{2!}t^2 & \frac{1}{3!}t^3 & \cdots & \frac{1}{(r-1)!}t^{r-1}\\ & 1 & t & \frac{1}{2!}t^2 & \cdots & \frac{1}{(r-2)!}t^{r-2}\\ & & 1 & t & \cdots & \frac{1}{(r-3)!}t^{r-3}\\ & & & \ddots & \ddots & \vdots\\ & & & & \ddots & t\\ & & & & & 1\end{pmatrix}.\end{aligned}$$

又注意到

$$
\mathrm{e}^{t\boldsymbol{A}} = \mathrm{e}^{\boldsymbol{P}(t\boldsymbol{J})\boldsymbol{P}^{-1}} = \boldsymbol{P}\mathrm{e}^{t\boldsymbol{J}}\boldsymbol{P}^{-1},
$$

$$
\mathrm{e}^{t\boldsymbol{J}} = \begin{pmatrix} \mathrm{e}^{t\boldsymbol{J}_1} & & \\ & \ddots & \\ & & \mathrm{e}^{t\boldsymbol{J}_k} \end{pmatrix}.
$$

于是将 $\mathrm{e}^{t\boldsymbol{J}_i}$ 的式子代入上面的式子即可求出 $\mathrm{e}^{t\boldsymbol{A}}$. □