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

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

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

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

♣

命题 0.1

设 A 是 n 阶矩阵, 证明: $\lim\limits_{k\to\infty} A^k$ 存在的充要条件是 A 的特征值的模长小于 1, 或者特征值等于 1 并且 A 关于特征值 1 的 Jordan 块都是一阶的. 此时, 极限矩阵为

$$\lim_{k\to\infty} A^k = P\mathrm{diag}\{1, \cdots, 1, 0, \cdots, 0\}P^{-1},$$

其中 1 的个数等于 A 的特征值 1 的代数重数.

♠

证明 设 P 为非异阵, 使得 $P^{-1}AP = J = \mathrm{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \cdots, J_{r_s}(\lambda_s)\}$ 为 A 的 Jordan 标准型, 则

$$A^k = PJ^kP^{-1} = P\mathrm{diag}\{J_{r_1}(\lambda_1)^k, J_{r_2}(\lambda_2)^k, \cdots, J_{r_s}(\lambda_s)^k\}P^{-1},$$

因此 $\lim\limits_{k\to\infty} A^k$ 存在当且仅当 $\lim\limits_{k\to\infty} J_{r_i}(\lambda_i)^k (1 \leqslant i \leqslant s)$ 都存在. 不妨取 $k > n$, 由引理 0.1计算可得 Jordan 块 $J_{r_i}(\lambda_i)$ 的 k 次幂为

$$J_{r_i}(\lambda_i)^k = \begin{pmatrix} \lambda_i^k & \mathrm{C}_k^1\lambda_i^{k-1} & \mathrm{C}_k^2\lambda_i^{k-2} & \cdots & \mathrm{C}_k^{r_i-1}\lambda_i^{k-r_i+1} \\ & \lambda_i^k & \mathrm{C}_k^1\lambda_i^{k-1} & \cdots & \mathrm{C}_k^{r_i-2}\lambda_i^{k-r_i+2} \\ & & \lambda_i^k & \cdots & \mathrm{C}_k^{r_i-3}\lambda_i^{k-r_i+3} \\ & & & \ddots & \vdots \\ & & & & \lambda_i^k \end{pmatrix},$$

故当 $|\lambda_i| \geqslant 1$ 且 $\lambda_i \neq 1$ 时, $\lim\limits_{k\to\infty} \lambda_i^k$ 发散; 当 $\lambda_i = 1$ 且 $r_i \geqslant 2$ 时, $\lim\limits_{k\to\infty} \mathrm{C}_k^1\lambda_i^{k-1}$ 发散; 当 $\lambda_i = 1$ 且 $r_1 = 1$ 时, $\lim\limits_{k\to\infty} J_{r_i}(\lambda_i)^k = J_1(1)$; 当 $|\lambda_i| < 1$ 时, $\lim\limits_{k\to\infty} J_{r_i}(\lambda_i)^k = O$. 因此, $\lim\limits_{k\to\infty} A^k$ 存在的充要条件是 A 的特征值的模长小于 1, 或者特征值等于 1 并且 A 关于特征值 1 的 Jordan 块都是一阶的. 此时, 极限矩阵 $\lim\limits_{k\to\infty} A^k = P\,\mathrm{diag}\,\{1, \cdots, 1, 0, \cdots, 0\}\,P^{-1}$, 其中 1 的个数等于 A 的特征值 1 的代数重数. □

定理 0.1

设 $f(z)=\sum_{i=0}^{\infty}a_iz^i$ 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 $f_p(z)=a_0+a_1z+a_2z^2+\cdots+a_pz^p$ 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.

定理 0.3

证明:$\cos A=\dfrac{e^{\mathrm{i}A}+e^{-\mathrm{i}A}}{2},\quad \sin A=\dfrac{e^{\mathrm{i}A}-e^{-\mathrm{i}A}}{2\mathrm{i}}$.

证明 由定义 0.2可知

$$\begin{aligned}
e^{\mathrm{i}A} &= I+\frac{1}{1!}\mathrm{i}A-\frac{1}{2!}A^2-\frac{1}{3!}\mathrm{i}A^3-\cdots+\frac{(-1)^k}{2k!}A^{2k}+\frac{(-1)^k}{(2k+1)!}\mathrm{i}A^{2k+1}+\cdots,\\
e^{-\mathrm{i}A} &= I-\frac{1}{1!}\mathrm{i}A-\frac{1}{2!}A^2+\frac{1}{3!}\mathrm{i}A^3-\cdots+\frac{(-1)^k}{2k!}A^{2k}+\frac{(-1)^{k+1}}{(2k+1)!}\mathrm{i}A^{2k+1}+\cdots.
\end{aligned}$$

从而

$$\begin{aligned}
\frac{e^{\mathrm{i}A}+e^{-\mathrm{i}A}}{2} &= I-\frac{1}{2!}A^2+\cdots+\frac{(-1)^k}{2k!}A^{2k}+\cdots=\cos A,\\
\frac{e^{\mathrm{i}A}-e^{-\mathrm{i}A}}{2\mathrm{i}} &= \frac{1}{1!}A-\frac{1}{3!}A^3+\cdots+\frac{(-1)^k}{(2k+1)!}A^{2k+1}\cdots=\sin A.
\end{aligned}$$

□

命题 0.2

求证: 若 n 阶矩阵 A,B 乘法可交换, 则 $\mathrm{e}^A\cdot\mathrm{e}^B=\mathrm{e}^B\cdot\mathrm{e}^A$.

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

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

如对

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

通过计算不难验证 $AB=A$, $BA=B$, 并且

$$\mathrm{e}^A=\begin{pmatrix}\mathrm{e}&\mathrm{e}-1\\0&1\end{pmatrix},\quad \mathrm{e}^B=\begin{pmatrix}1&0\\\mathrm{e}-1&\mathrm{e}\end{pmatrix},\quad \mathrm{e}^{A+B}=\begin{pmatrix}\frac{\mathrm{e}^2+1}{2}&\frac{\mathrm{e}^2-1}{2}\\\frac{\mathrm{e}^2-1}{2}&\frac{\mathrm{e}^2+1}{2}\end{pmatrix}$$

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

证明 设 $f(z)=\mathrm{e}^z$, 并且 $f_p(z)=1+\frac{1}{1!}z+\frac{1}{2!}z^2+\cdots+\frac{1}{p!}z^p$ 为 $f(z)$ 的部分和, 因为 $f(z)$ 的收敛半径为 $+\infty$, 所以对任一矩阵 A, $\lim\limits_{p\to\infty}f_p(A)=f(A)=\mathrm{e}^A$. 由于 $AB=BA$, 故对任意的正整数 p,q, 成立 $f_p(A)f_q(B)=f_q(B)f_p(A)$. 先固定 p, 令 $q\to\infty$, 则可得

$$\begin{aligned}
f_p(A)f(B) &= f_p(A)\left(\lim_{q\to\infty}f_q(B)\right)=\lim_{q\to\infty}\left(f_p(A)f_q(B)\right)=\lim_{q\to\infty}\left(f_q(B)f_p(A)\right)\\
&= \left(\lim_{q\to\infty}f_q(B)\right)f_p(A)=f(B)f_p(A)
\end{aligned}$$

同理, 再对上式令 $p\to\infty$, 则可得 $f(A)f(B)=f(B)f(A)$, 即结论成立. □

推论 0.1

若 $f(z), g(z)$ 是两个收敛半径都是 $+\infty$ 的复幂级数, 则对任意乘法可交换的 A, B, 均有 $f(A)g(B) = g(B)f(A)$. ♡

证明 由命题 0.2类似的讨论可证明. □

定义 0.3 (矩阵的范数)

设 $A = (a_{ij})$ 是 n 阶复矩阵, 定义 A 的**范数**为其所有元素模长的平方和的算术平方根, 即 $\|A\| = \sqrt{\sum\limits_{i,j=1}^{n} |a_{ij}|^2}$. ♣

命题 0.3 (矩阵的范数的基本性质)

设 $A = (a_{ij})$ 是 n 阶复矩阵,$B = (b_{ij})$ 也是 n 阶复矩阵, 求证:

(1) $\|A\| \geq 0$, 等号成立当且仅当 $A = O$;

(2) $\|A + B\| \leq \|A\| + \|B\|$;

(3) $\|AB\| \leq \|A\| \cdot \|B\|$. ♠

证明 (1) 显然成立.

(2) 注意到

$$\begin{aligned}
\|A+B\|^2 &= \sum_{i,j=1}^{n} \left|a_{ij}+b_{ij}\right|^2 = \sum_{i,j=1}^{n} \left(\left|a_{ij}\right|^2 + \left|b_{ij}\right|^2 + 2\left|a_{ij}b_{ij}\right|\right) \\
&= \sum_{i,j=1}^{n} \left|a_{ij}\right|^2 + \sum_{i,j=1}^{n} \left|b_{ij}\right|^2 + 2\sum_{i,j=1}^{n} \left|a_{ij}b_{ij}\right|
\end{aligned}$$

于是

$$\begin{aligned}
(\|A\|+\|B\|)^2 &= \|A\|^2 + \|B\|^2 + 2\|A\|\cdot\|B\| \\
&= \sum_{i,j=1}^{n} \left|a_{ij}\right|^2 + \sum_{i,j=1}^{n} \left|b_{ij}\right|^2 + 2\sqrt{\left(\sum_{i,j=1}^{n} \left|a_{ij}\right|^2\right)\left(\sum_{i,j=1}^{n} \left|b_{ij}\right|^2\right)} \\
&\overset{\text{Cauchy-Schwarz 不等式}}{\geqslant} \sum_{i,j=1}^{n} \left|a_{ij}\right|^2 + \sum_{i,j=1}^{n} \left|b_{ij}\right|^2 + 2\sqrt{\left(\sum_{i,j=1}^{n} \left|a_{ij}b_{ij}\right|\right)^2} \\
&= \sum_{i,j=1}^{n} \left|a_{ij}\right|^2 + \sum_{i,j=1}^{n} \left|b_{ij}\right|^2 + 2\sum_{i,j=1}^{n} \left|a_{ij}b_{ij}\right| \\
&= \|A+B\|^2 .
\end{aligned}$$

故结论得证.

(3) 注意到 $\|AB\|^2 = \sum\limits_{i,j=1}^{n} \left|\sum\limits_{k=1}^{n} a_{ik}b_{kj}\right|^2$,$\|A\|^2 \cdot \|B\|^2 = \left(\sum\limits_{i,k=1}^{n} |a_{ik}|^2\right) \cdot \left(\sum\limits_{k,j=1}^{n} |b_{kj}|^2\right)$, 任取 $i, j \in \{1, 2, \cdots, n\}$, 固定 i 和 j, 由 Cauchy-Schwarz 不等式可知

$$\left|\sum_{k=1}^{n} a_{ik}b_{kj}\right|^2 \leqslant \left(\sum_{k=1}^{n} a_{ik}^2\right) \cdot \left(\sum_{k=1}^{n} b_{kj}^2\right).$$

从而先对 i 求和可得

$$\sum_{i=1}^{n} \left|\sum_{k=1}^{n} a_{ik}b_{kj}\right|^2 \leqslant \sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_{ik}^2\right) \cdot \left(\sum_{k=1}^{n} b_{kj}^2\right) = \left(\sum_{k=1}^{n} b_{kj}^2\right) \left(\sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_{ik}^2\right)\right).$$

再对 j 求和可得

$$\begin{aligned}\sum_{i,j=1}^{n}\left|\sum_{k=1}^{n}a_{ik}b_{kj}\right|^2 &\leqslant \sum_{j=1}^{n}\left[\left(\sum_{k=1}^{n}b_{kj}^2\right)\left(\sum_{i=1}^{n}\left(\sum_{k=1}^{n}a_{ik}^2\right)\right)\right]\\ &=\left(\sum_{i=1}^{n}\left(\sum_{k=1}^{n}a_{ik}^2\right)\right)\left(\sum_{j=1}^{n}\left(\sum_{k=1}^{n}b_{kj}^2\right)\right)\\ &=\left(\sum_{i,k=1}^{n}|a_{ik}|^2\right)\cdot\left(\sum_{k,j=1}^{n}|b_{kj}|^2\right).\end{aligned}$$

由此即得结论. □

命题 0.4

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 利用这个命题也可给出命题 0.2的另一证明.

证明 设 $f(z)=\mathrm{e}^z$, 并且 $f_p(z)=1+\frac{1}{1!}z+\frac{1}{2!}z^2+\cdots+\frac{1}{p!}z^p$ 为 $f(z)$ 的部分和. 注意到 $AB=BA$, 经简单的计算可知,$f_p(A)f_p(B)$ 展开后的单项包含 $f_p(A+B)$ 展开后的所有单项, 且剩余单项可表示为 $\frac{A^i}{i!}\frac{B^j}{j!}$ 的形式, 其中 $i+j>p$, 故由矩阵的范数的基本性质 (3)可得

$$\begin{aligned}\|f_p(A)f_p(B)-f_p(A+B)\| &\le \sum_{k>p}\left(\sum_{i+j=k}\frac{\|A\|^i}{i!}\frac{\|B\|^j}{j!}\right)=\sum_{k>p}\left(\sum_{i=0}^{k}\frac{\|A\|^i}{i!}\frac{\|B\|^{k-i}}{(k-i)!}\right)\\ &=\sum_{k>p}\left(\sum_{i=0}^{k}\frac{k!}{i!(k-i)!}\frac{\|A\|^i\|B\|^{k-i}}{k!}\right)=\sum_{k>p}\left(\sum_{i=0}^{k}\mathrm{C}_k^i\frac{\|A\|^i\|B\|^{k-i}}{k!}\right)\\ &=\sum_{k>p}\frac{(\|A\|+\|B\|)^k}{k!}.\end{aligned}$$

由于数项级数 $\sum_{k=0}^{\infty}\frac{1}{k!}(\|A\|+\|B\|)^k$ 收敛到 $\mathrm{e}^{\|A\|+\|B\|}$, 故当 p 充分大时, 上式右边趋于零. 令 $p\to\infty$, 则由上式即得 $\|f(A)f(B)-f(A+B)\|=0$, 再次由矩阵的范数的基本性质 (1)可得 $\mathrm{e}^A\cdot\mathrm{e}^B=\mathrm{e}^{A+B}$. □

推论 0.2

若矩阵幂级数 e^A 绝对收敛, 则矩阵级数的 Cauchy 乘积

$$\mathrm{e}^{A+B}=\sum_{p=0}^{\infty}\left(\sum_{i+j=p}\frac{A^i}{i!}\frac{B^j}{j!}\right)$$

收敛到

$$\left(\sum_{i=0}^{\infty}\frac{A^i}{i!}\right)\cdot\left(\sum_{j=0}^{\infty}\frac{B^j}{j!}\right)=\mathrm{e}^A\cdot\mathrm{e}^B.$$

注 注意矩阵级数的 Cauchy 积有

$$\begin{aligned}\mathrm{e}^{A+B} &=\sum_{p=0}^{\infty}\frac{(A+B)^p}{p!}=\sum_{p=0}^{\infty}\left(\sum_{i=0}^{p}\mathrm{C}_p^i\frac{A^iB^{p-i}}{p!}\right)\\ &=\sum_{p=0}^{\infty}\left(\sum_{i=0}^{p}\frac{p!}{i!(p-i)!}\frac{A^iB^{p-i}}{p!}\right)=\sum_{p=0}^{\infty}\left(\sum_{i=0}^{p}\frac{A^i}{i!}\frac{B^{p-i}}{(p-i)}\right)\end{aligned}$$

$$= \sum_{p=0}^{\infty}\left(\sum_{i+j=p}\frac{A^i}{i!}\frac{B^j}{j!}\right).$$

证明 由命题 0.4立得. □

命题 0.5

设 t 是一个数值变量,$\boldsymbol{A}$ 是一个 n 阶复方阵. $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{J}=\mathrm{diag}\{\boldsymbol{J}_1,\cdots,\boldsymbol{J}_k\}$ 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 0.4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

$$\mathrm{e}^{t\boldsymbol{J}_i}=\mathrm{e}^{t\lambda_i}\cdot\mathrm{e}^{t\boldsymbol{N}}=\mathrm{e}^{t\lambda_i}\left[t\boldsymbol{I}+t\boldsymbol{N}+\frac{t}{2!}\boldsymbol{N}^2+\frac{t}{3!}\boldsymbol{N}^3+\cdots+\frac{t}{(r-1)!}\boldsymbol{N}^{r-1}\right]$$

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

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 $\mathrm{e}^{t\boldsymbol{A}}$. □

命题 0.6 (矩阵三角函数的性质)

设 A 是 n 阶矩阵, 求证:

(1) $\sin^2 A + \cos^2 A = I_n$.

(2) $\sin 2A = 2\sin A\cos A$.

证明 由定理 0.3可知

$$\cos A = \frac{1}{2}(e^{\mathrm{i}A} + e^{-\mathrm{i}A}), \quad \sin A = \frac{1}{2\mathrm{i}}(e^{\mathrm{i}A} - e^{-\mathrm{i}A}). \tag{3}$$

由命题 0.2和命题 0.4可知

$$e^{\mathrm{i}A}e^{-\mathrm{i}A} = e^{-\mathrm{i}A}e^{\mathrm{i}A} = e^{\mathrm{i}A-\mathrm{i}A} = I_n, (e^{\mathrm{i}A})^2 = e^{2\mathrm{i}A}, (e^{-\mathrm{i}A})^2 = e^{-2\mathrm{i}A}. \tag{4}$$

(1) 由(3)和(4)式可得

$$\sin^2 A + \cos^2 A = \frac{1}{4}(e^{2\mathrm{i}A} + 2I_n + e^{-2\mathrm{i}A}) - \frac{1}{4}(e^{2\mathrm{i}A} - 2I_n + e^{-2\mathrm{i}A}) = I_n.$$

(2) 由(3)和(4)式可得

$$2\sin A\cos A = \frac{1}{2}\left(e^{\mathrm{i}A} - e^{-\mathrm{i}A}\right)\left(e^{\mathrm{i}A} + e^{-\mathrm{i}A}\right) = \frac{1}{2}\left(e^{\mathrm{i}2A} - e^{-\mathrm{i}2A}\right) = \sin 2A.$$

□

例题 0.1 计算 $\sin(e^{cI})$ 及 $\cos(e^{cI})$, 其中 c 是非零常数.

解 由指数矩阵函数的定义可得

$$\begin{aligned} e^{cI} &= I + \frac{1}{1!}(cI) + \frac{1}{2!}(cI)^2 + \frac{1}{3!}(cI)^3 + \cdots \\ &= \left(1 + \frac{1}{1!}c + \frac{1}{2!}c^2 + \frac{1}{3!}c^3 + \cdots\right)I = e^cI. \end{aligned}$$

因此

$$\begin{aligned} \sin(e^{cI}) = \sin(e^cI) &= e^cI - \frac{1}{3!}(e^cI)^3 + \frac{1}{5!}(e^cI)^5 - \frac{1}{7!}(e^cI)^7 + \cdots \\ &= \left(e^c - \frac{1}{3!}(e^c)^3 + \frac{1}{5!}(e^c)^5 - \frac{1}{7!}(e^c)^7 + \cdots\right)I = (\sin e^c)I, \\ \cos(e^{cI}) = \cos(e^cI) &= I - \frac{1}{2!}(e^cI)^2 + \frac{1}{4!}(e^cI)^4 - \frac{1}{6!}(e^cI)^6 + \cdots \\ &= \left(1 - \frac{1}{2!}(e^c)^2 + \frac{1}{4!}(e^c)^4 - \frac{1}{6!}(e^c)^6 + \cdots\right)I = (\cos e^c)I. \end{aligned}$$

□

命题 0.7

设 A 是 n 阶方阵, 证明:e^A 的行列式为 $e^{\operatorname{tr}(A)}$.

证明 设 A 的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 则由命题**??**可知 e^A 的特征值为 $e^{\lambda_1}, e^{\lambda_2}, \cdots, e^{\lambda_n}$, 因此

$$|e^A| = e^{\lambda_1} e^{\lambda_2} \cdots e^{\lambda_n} = e^{\lambda_1+\lambda_2+\cdots+\lambda_n} = e^{\operatorname{tr}(A)}.$$

□

推论 0.3

求证: 对任一 n 阶方阵 A,e^A 总是非异阵.

证明 由命题 0.7可知 $|e^A| = e^{\operatorname{tr}(A)} \neq 0$, 从而 e^A 非异. 也可由命题 0.4得到

$$e^A e^{-A} = e^{A-A} = I_n,$$

于是 e^A 非异且 $(e^A)^{-1} = e^{-A}$. □ □