

0.1 矩阵函数

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

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

则称矩阵序列 $\{A_p\}$ 收敛于 B , 记为

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

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 (复矩阵极限的性质)

设 n 阶复方阵序列 $\{A_p\}$ 和 $\{B_p\}$.

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

f 是任意一个复系数多项式, 即 $f \in \mathbb{C}[x]$. 我们有

- (1) $\lim_{p \rightarrow +\infty} (A_p \pm B_p) = \lim_{p \rightarrow +\infty} A_p \pm \lim_{p \rightarrow +\infty} B_p$;
- (2) $\lim_{p \rightarrow +\infty} (A_p B_p) = \lim_{p \rightarrow +\infty} A_p \cdot \lim_{p \rightarrow +\infty} B_p$;
- (3) $\lim_{p \rightarrow +\infty} f(A_p) = f\left(\lim_{p \rightarrow +\infty} A_p\right)$, 特别地, $\left| \lambda I - \lim_{p \rightarrow +\infty} A_p \right| = \lim_{p \rightarrow +\infty} |\lambda I - A_p|$.

证明

(1) 由极限的四则运算性质可知

$$\lim_{p \rightarrow +\infty} (a_{ij}^{(p)} \pm b_{ij}^{(p)}) = \lim_{p \rightarrow +\infty} a_{ij}^{(p)} \pm \lim_{p \rightarrow +\infty} b_{ij}^{(p)}, \quad \forall i, j \in \{1, 2, \cdots, n\}.$$

因此 $\lim_{p \rightarrow +\infty} (A_p \pm B_p) = \lim_{p \rightarrow +\infty} A_p \pm \lim_{p \rightarrow +\infty} B_p$.

(2) 由极限的四则运算性质可知

$$\lim_{p \rightarrow +\infty} \sum_{i,j=1}^n \sum_{k=1}^n a_{ik}^{(p)} b_{kj}^{(p)} = \sum_{i,j=1}^n \sum_{k=1}^n \left(\lim_{p \rightarrow +\infty} a_{ik}^{(p)} \cdot \lim_{p \rightarrow +\infty} b_{kj}^{(p)} \right), \quad \forall i, j \in \{1, 2, \cdots, n\}.$$

因此 $\lim_{p \rightarrow +\infty} (A_p B_p) = \lim_{p \rightarrow +\infty} A_p \cdot \lim_{p \rightarrow +\infty} B_p$.

(3) 由 (1)(2) 的结论立得. 特别地, 记 $\lambda I - A_p = (c_{ij}^{(p)})$, 则由行列式的组合定义可知

$$\begin{aligned} \left| \lambda I - \lim_{p \rightarrow +\infty} A_p \right| &= \left| \lim_{p \rightarrow +\infty} (\lambda I - A_p) \right| = \sum_{1 \leq i_1, i_2, \dots, i_n \leq n} \left[(-1)^{\tau(i_1 i_2 \dots i_n)} \lim_{p \rightarrow +\infty} c_{1i_1}^{(p)} \lim_{p \rightarrow +\infty} c_{2i_2}^{(p)} \cdots \lim_{p \rightarrow +\infty} c_{ni_n}^{(p)} \right] \\ &= \lim_{p \rightarrow +\infty} \sum_{1 \leq i_1, i_2, \dots, i_n \leq n} \left[(-1)^{\tau(i_1 i_2 \dots i_n)} c_{1i_1}^{(p)} c_{2i_2}^{(p)} \cdots c_{ni_n}^{(p)} \right] = \lim_{p \rightarrow +\infty} |\lambda I - A_p|. \end{aligned}$$

□

命题 0.1

设 A 是 n 阶矩阵, 证明: $\lim_{k \rightarrow \infty} A^k$ 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

$$\lim_{k \rightarrow \infty} A^k = P \operatorname{diag}\{1, \dots, 1, 0, \dots, 0\} P^{-1},$$

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 $P^{-1}AP = J = \operatorname{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \dots, J_{r_s}(\lambda_s)\}$ 为 A 的 Jordan 标准型, 则

$$A^k = P J^k P^{-1} = P \operatorname{diag}\{J_{r_1}(\lambda_1)^k, J_{r_2}(\lambda_2)^k, \dots, J_{r_s}(\lambda_s)^k\} P^{-1},$$

因此 $\lim_{k \rightarrow \infty} A^k$ 存在当且仅当 $\lim_{k \rightarrow \infty} J_{r_i}(\lambda_i)^k (1 \leq i \leq s)$ 都存在. 不妨取 $k > n$, 由命题 ??(5) 计算可得 Jordan 块 $J_{r_i}(\lambda_i)$ 的 k 次幂为

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

故当 $|\lambda_i| \geq 1$ 且 $\lambda_i \neq 1$ 时, $\lim_{k \rightarrow \infty} \lambda_i^k$ 发散; 当 $\lambda_i = 1$ 且 $r_i \geq 2$ 时, $\lim_{k \rightarrow \infty} C_k^1 \lambda_i^{k-1}$ 发散; 当 $\lambda_i = 1$ 且 $r_i = 1$ 时, $\lim_{k \rightarrow \infty} J_{r_i}(\lambda_i)^k = J_1(1)$; 当 $|\lambda_i| < 1$ 时, $\lim_{k \rightarrow \infty} J_{r_i}(\lambda_i)^k = O$. 因此, $\lim_{k \rightarrow \infty} A^k$ 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 $\lim_{k \rightarrow \infty} A^k = P \operatorname{diag}\{1, \dots, 1, 0, \dots, 0\} P^{-1}$, 其中 1 的个数等于 A 的特征值 1 的代数重数. □

定理 0.2

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

$$f(X) = \operatorname{diag}\{f(X_1), \dots, f(X_k)\};$$

(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

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



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

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

(2) 注意到 $f_p(z)$ 是多项式, 从而有

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

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

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

(3) 由命题??(5) 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 $p \rightarrow \infty$, 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 0.3

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

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

- (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

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

都收敛;

(4) 若 $f(A)$ 收敛, 则 $f(A)$ 的特征值为

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



证明

- (1) 设 A 的 Jordan 标准型为 $J = \text{diag}\{J_1, J_2, \dots, J_k\}$. 显然 $f(A)$ 的收敛性等价于所有 $f(J_i) (i = 1, \dots, k)$ 的收敛性. 由定理 0.2 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 $f(\lambda_1), f(\lambda_2), \dots, f(\lambda_n)$.

□

定义 0.2

于是对一切方阵, 定义

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

都有意义. 若 A 所有特征值的模长都小于 1, 则

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

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:

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

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 0.3 可知 $e^A, \sin A, \cos A, \ln A$ 都收敛, 从而都有意义. 故上述定义是良定义的.

定理 0.4

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



证明 由定义 0.2 可知

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

从而

$$\frac{e^{iA} + e^{-iA}}{2} = I - \frac{1}{2!}A^2 + \dots + \frac{(-1)^k}{2k!}A^{2k} + \dots = \cos A,$$

$$\frac{e^{iA} - e^{-iA}}{2i} = \frac{1}{1!}A - \frac{1}{3!}A^3 + \cdots + \frac{(-1)^k}{(2k+1)!}A^{2k+1} \cdots = \sin A.$$

□

命题 0.2

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

◆

注 对一般来说对矩阵 A, B , 下面的等式并不一定成立:

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

如对

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

通过计算不难验证 $AB = A, BA = B$, 并且

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

故 $e^A \cdot e^B \neq e^{A+B}$.

证明 设 $f(z) = e^z$, 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 $A, \lim_{p \rightarrow \infty} f_p(A) = f(A) = e^A$. 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 $q \rightarrow \infty$, 则可得

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

同理, 再对上式令 $p \rightarrow \infty$, 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 $\|A\| = \sqrt{\sum_{i,j=1}^n |a_{ij}|^2}$.

♣

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

$$\|A + B\|^2 = \sum_{i,j=1}^n |a_{ij} + b_{ij}|^2 = \sum_{i,j=1}^n (|a_{ij}|^2 + |b_{ij}|^2 + 2|a_{ij}b_{ij}|)$$

$$= \sum_{i,j=1}^n |a_{ij}|^2 + \sum_{i,j=1}^n |b_{ij}|^2 + 2 \sum_{i,j=1}^n |a_{ij}b_{ij}|$$

于是

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

故结论得证.

$$(3) \text{ 注意到 } \|AB\|^2 = \sum_{i,j=1}^n \left| \sum_{k=1}^n a_{ik}b_{kj} \right|^2, \|A\|^2 \cdot \|B\|^2 = \left(\sum_{i,k=1}^n |a_{ik}|^2 \right) \cdot \left(\sum_{k,j=1}^n |b_{kj}|^2 \right), \text{ 任取 } i, j \in \{1, 2, \dots, n\},$$

固定 i 和 j , 由 Cauchy-Schwarz 不等式可知

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

从而先对 i 求和可得

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

再对 j 求和可得

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

由此即得结论. □

命题 0.4

如果 A 与 B 乘法可交换, 即 $AB = BA$, 则 $e^A \cdot e^B = e^{A+B}$ 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 $f(z) = e^z$, 并且 $f_p(z) = 1 + \frac{1}{1!}z + \frac{1}{2!}z^2 + \dots + \frac{1}{p!}z^p$ 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 $\frac{A^i B^j}{i! j!}$ 的形式, 其中 $i+j > p$, 故由矩阵的范数的基本性质 (3) 可得

$$\|f_p(A)f_p(B) - f_p(A+B)\| \leq \sum_{k>p} \left(\sum_{i+j=k} \frac{\|A\|^i}{i!} \frac{\|B\|^j}{j!} \right) = \sum_{k>p} \left(\sum_{i=0}^k \frac{\|A\|^i}{i!} \frac{\|B\|^{k-i}}{(k-i)!} \right)$$

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

由于数项级数 $\sum_{k=0}^{\infty} \frac{1}{k!} (\|A\| + \|B\|)^k$ 收敛到 $e^{\|A\| + \|B\|}$, 故当 p 充分大时, 上式右边趋于零. 令 $p \rightarrow \infty$, 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 $e^A \cdot e^B = e^{A+B}$. \square

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

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

收敛到

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



注 注意矩阵级数的 Cauchy 积有

$$\begin{aligned}
e^{A+B} &= \sum_{p=0}^{\infty} \frac{(A+B)^p}{p!} = \sum_{p=0}^{\infty} \left(\sum_{i=0}^p C_p^i \frac{A^i B^{p-i}}{p!} \right) \\
&= \sum_{p=0}^{\infty} \left(\sum_{i=0}^p \frac{p!}{i!(p-i)!} \frac{A^i B^{p-i}}{p!} \right) = \sum_{p=0}^{\infty} \left(\sum_{i=0}^p \frac{A^i B^{p-i}}{i!(p-i)!} \right) \\
&= \sum_{p=0}^{\infty} \left(\sum_{i+j=p} \frac{A^i B^j}{i! j!} \right).
\end{aligned}$$

证明 由命题 0.4 立得. \square

命题 0.5

设 t 是一个数值变量, A 是一个 n 阶复方阵. $P^{-1}AP = J = \text{diag}\{J_1, \dots, J_k\}$ 是 A 的 Jordan 标准型, J_i 是特征值为 λ_i 的 r 阶 Jordan 块, 则

$$e^{tA} = P e^{tJ} P^{-1},$$

其中

$$e^{tJ} = \begin{pmatrix} e^{tJ_1} & & \\ & \ddots & \\ & & e^{tJ_k} \end{pmatrix}, \quad e^{tJ_i} = e^{t\lambda_i} \begin{pmatrix} 1 & t & \frac{1}{2!}t^2 & \frac{1}{3!}t^3 & \cdots & \frac{1}{(r-1)!}t^{r-1} \\ & 1 & t & \frac{1}{2!}t^2 & \cdots & \frac{1}{(r-2)!}t^{r-2} \\ & & 1 & t & \cdots & \frac{1}{(r-3)!}t^{r-3} \\ & & & \ddots & \ddots & \vdots \\ & & & & \ddots & t \\ & & & & & 1 \end{pmatrix}.$$



证明 证法一: 若令 $f(z) = e^{tz}$, 则由定理 0.2 即得 $f(A) = e^{tA}$ 的计算结果.

证法二: 注意到

$$J_i = \lambda_i I + N,$$

其中 N 是 r 阶基础幂零阵, 即

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

于是

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

因为 $(\lambda_i I)N = N(\lambda_i I)$, 故由命题 0.4 可知

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

同理

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

又注意到

$$\begin{aligned} e^{tA} &= e^{P(tJ)P^{-1}} = Pe^{tJ}P^{-1}, \\ e^{tJ} &= \begin{pmatrix} e^{tJ_1} & & \\ & \ddots & \\ & & e^{tJ_k} \end{pmatrix}. \end{aligned}$$

于是将 e^{tJ_i} 的式子代入上面的式子即可求出 e^{tA} . □

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 0.4 可知

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

由命题 0.2 和命题 0.4 可知

$$e^{iA}e^{-iA} = e^{-iA}e^{iA} = e^{iA-iA} = I_n, (e^{iA})^2 = e^{2iA}, (e^{-iA})^2 = e^{-2iA}. \quad (4)$$

(1) 由(3)和(4)式可得

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

(2) 由(3)和(4)式可得

$$2 \sin A \cos A = \frac{1}{2} (e^{iA} - e^{-iA}) (e^{iA} + e^{-iA}) = \frac{1}{2} (e^{i2A} - e^{-i2A}) = \sin 2A.$$

□

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

解 由指数矩阵函数的定义可得

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

因此

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

□

命题 0.7

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

▲

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

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

□

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 $|e^A| = e^{\text{tr}(A)} \neq 0$, 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 $(e^A)^{-1} = e^{-A}$.

□