

Linear Algebra II

Supplementaries - 7

ζ

April 20, 2023

1 λ -矩阵初步

Problem 1.1. 设 $f(\lambda), g(\lambda)$ 是数域 \mathbb{F} 上的首一多项式, $d(\lambda) = (f(\lambda), g(\lambda)), m(\lambda) = [f(\lambda), g(\lambda)]$ 分别是其最大公因式与最小公倍式, 证明下列 λ -矩阵相抵:

$$\begin{pmatrix} f(\lambda) & 0 \\ 0 & g(\lambda) \end{pmatrix}, \begin{pmatrix} g(\lambda) & 0 \\ 0 & f(\lambda) \end{pmatrix}, \begin{pmatrix} d(\lambda) & 0 \\ 0 & m(\lambda) \end{pmatrix}.$$

Problem 1.2. 求 Frobenius 矩阵的特征矩阵的不变因子, 初等因子 (组), 其定义如下:

$$\mathbf{F} = \begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \dots & 0 & -a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -a_{n-1} \end{pmatrix}.$$

Problem 1.3. 求如下 Jordan 块 $J(\lambda, n)$ 的特征矩阵的不变因子, 初等因子 (组).

注意到 λ -矩阵为 $(\mathbb{F}[\lambda])^{m \times n}$, 我们补充一拓展知识.

Problem 1.4. 我们仅考虑含么环. 对于环 R , 定义矩阵环 $R^{m \times n}$ 为 $m \times n$ 矩阵, 其分量均为环 R 中元素. 则 $R^{m \times m}$ 也为环, $R^{m \times n}$ 是 $R^{m \times m}$ 左模, 也是 $R^{n \times n}$ 右模.

- (1) \mathbf{A} (或 φ) 可对角化;
- (2) \mathbf{A} (或 φ) 的极小多项式无重根;
- (3) \mathbf{A} (或 φ) 的初等因子全是一次因子.

7.1.4 矩阵函数

1. 矩阵序列的收敛

设有 n 阶矩阵序列 $\{\mathbf{A}_k\}$:

$$\mathbf{A}_k = \begin{pmatrix} a_{11}^{(k)} & \cdots & a_{1n}^{(k)} \\ \vdots & & \vdots \\ a_{n1}^{(k)} & \cdots & a_{nn}^{(k)} \end{pmatrix},$$

$\mathbf{B} = (b_{ij})$ 也是一个 n 阶矩阵. 若对每个 (i, j) , 都有 $\lim_{k \rightarrow \infty} a_{ij}^{(k)} = b_{ij}$, 则称矩阵序列 $\{\mathbf{A}_k\}$ 收敛于 \mathbf{B} , 记为 $\lim_{k \rightarrow \infty} \mathbf{A}_k = \mathbf{B}$.

2. 矩阵幂级数

设 $f(z) = a_0 + a_1z + a_2z^2 + \cdots + a_nz^n + \cdots$ 是一个复幂级数, $f_k(z)$ 是其部分和. 若矩阵序列 $\{f_k(\mathbf{A})\}$ 收敛于 \mathbf{B} , 则称矩阵幂级数 $f(\mathbf{A})$ 收敛于 \mathbf{B} .

3. 定理

设 $f(z) = a_0 + a_1z + a_2z^2 + \cdots + a_nz^n + \cdots$ 是一个复幂级数, 则

(1) 矩阵幂级数 $f(\mathbf{X})$ 收敛的充要条件是对任一可逆矩阵 \mathbf{P} , $f(\mathbf{P}^{-1}\mathbf{X}\mathbf{P})$ 收敛, 这时

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

(2) 若 $\mathbf{X} = \text{diag}\{\mathbf{X}_1, \cdots, \mathbf{X}_m\}$ 是分块对角矩阵, 则矩阵幂级数 $f(\mathbf{X})$ 收敛的充要条件是 $f(\mathbf{X}_i)$ ($i = 1, \cdots, m$) 收敛, 这时

$$f(\mathbf{X}) = \text{diag}\{f(\mathbf{X}_1), \cdots, f(\mathbf{X}_m)\};$$

(3) 若 $f(z)$ 的收敛半径为 r , \mathbf{J}_0 为 n 阶 Jordan 块, 且主对角元素为 λ_0 , 则当 $|\lambda_0| < r$ 时, $f(\mathbf{J}_0)$ 收敛.

4. 定理

设 $f(z) = a_0 + a_1z + a_2z^2 + \cdots + a_nz^n + \cdots$ 是一个复幂级数且其收敛半径为 r .

A 是 n 阶矩阵, 特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$, 记 $\lambda = \max_{1 \leq j \leq n} |\lambda_j|$.

- (1) 若 $\lambda < r$, 则 $f(A)$ 收敛;
- (2) 若 $\lambda > r$, 则 $f(A)$ 发散;
- (3) 若 $f(A)$ 收敛, 则 $f(A)$ 的特征值为 $f(\lambda_1), f(\lambda_2), \dots, f(\lambda_n)$.

§ 7.2 例题分析

7.2.1 λ -矩阵和初等因子

下面 3 个例题是 λ -矩阵和初等因子的基本性质, 我们在后面将会用到.

例 7.1 设 $f(\lambda), g(\lambda)$ 是数域 \mathbb{K} 上的首一多项式, $d(\lambda) = (f(\lambda), g(\lambda))$, $m(\lambda) = [f(\lambda), g(\lambda)]$ 分别是 $f(\lambda)$ 和 $g(\lambda)$ 的最大公因式和最小公倍式, 证明下列 λ -矩阵相抵:

$$\begin{pmatrix} f(\lambda) & 0 \\ 0 & g(\lambda) \end{pmatrix}, \begin{pmatrix} g(\lambda) & 0 \\ 0 & f(\lambda) \end{pmatrix}, \begin{pmatrix} d(\lambda) & 0 \\ 0 & m(\lambda) \end{pmatrix}.$$

证明 由已知, 存在多项式 $u(\lambda), v(\lambda)$, 使 $f(\lambda)u(\lambda) + g(\lambda)v(\lambda) = d(\lambda)$. 设 $f(\lambda) = d(\lambda)h(\lambda)$, 则 $m(\lambda) = g(\lambda)h(\lambda)$. 作下列 λ -矩阵的初等变换:

$$\begin{aligned} \begin{pmatrix} f(\lambda) & 0 \\ 0 & g(\lambda) \end{pmatrix} &\rightarrow \begin{pmatrix} f(\lambda) & 0 \\ f(\lambda)u(\lambda) & g(\lambda) \end{pmatrix} \rightarrow \begin{pmatrix} f(\lambda) & 0 \\ f(\lambda)u(\lambda) + g(\lambda)v(\lambda) & g(\lambda) \end{pmatrix} = \\ &\begin{pmatrix} f(\lambda) & 0 \\ d(\lambda) & g(\lambda) \end{pmatrix} \rightarrow \begin{pmatrix} 0 & -g(\lambda)h(\lambda) \\ d(\lambda) & g(\lambda) \end{pmatrix} \rightarrow \begin{pmatrix} 0 & g(\lambda)h(\lambda) \\ d(\lambda) & 0 \end{pmatrix} \rightarrow \begin{pmatrix} d(\lambda) & 0 \\ 0 & m(\lambda) \end{pmatrix}. \end{aligned}$$

另一结论同理可得. \square

设 $f(\lambda)$ 为数域 \mathbb{K} 上的多项式, $p(\lambda)$ 是 \mathbb{K} 上的首一不可约多项式, 若存在正整数 k , 使得 $p(\lambda)^k \mid f(\lambda)$, 但 $p(\lambda)^{k+1} \nmid f(\lambda)$, 则称 $p(\lambda)^k$ 为 $f(\lambda)$ 的一个准素因子. 事实上, 若设 $f(\lambda)$ 在 \mathbb{K} 上的标准因式分解为

$$f(\lambda) = cP_1(\lambda)^{e_1} P_2(\lambda)^{e_2} \cdots P_t(\lambda)^{e_t},$$

其中 c 为非零常数, $P_i(\lambda)$ 为互异的首一不可约多项式, $e_i > 0 (i = 1, \dots, t)$, 则 $f(\lambda)$ 的所有准素因子为 $P_1(\lambda)^{e_1}, P_2(\lambda)^{e_2}, \dots, P_t(\lambda)^{e_t}$. 因此等价地, 矩阵 A 的初等因子组就是 A 的所有不变因子的准素因子组. 下面的例题将初等因子组的这一等价定义进行了推广.

证明 相似的矩阵具有相同的迹. 若 $P^{-1}CP$ 是主对角线上元素全为零的矩阵 (根据上题, 它是存在的), 则

$$(P^{-1}AP)(P^{-1}BP) - (P^{-1}BP)(P^{-1}AP) = P^{-1}CP.$$

故不失一般性, 可假定 C 是一个主对角线上元素全为零的矩阵. 设 $C = (c_{ij})$, 令 B 是对角矩阵 $B = \text{diag}\{b_1, b_2, \dots, b_n\}$, 且假定 b_i 互不相同. 又设 $A = (a_{ij})$, 则

$$\begin{aligned} AB - BA &= \begin{pmatrix} a_{11}b_1 & a_{12}b_2 & \cdots & a_{1n}b_n \\ a_{21}b_1 & a_{22}b_2 & \cdots & a_{2n}b_n \\ \vdots & \vdots & & \vdots \\ a_{n1}b_1 & a_{n2}b_2 & \cdots & a_{nn}b_n \end{pmatrix} - \begin{pmatrix} a_{11}b_1 & a_{12}b_1 & \cdots & a_{1n}b_1 \\ a_{21}b_2 & a_{22}b_2 & \cdots & a_{2n}b_2 \\ \vdots & \vdots & & \vdots \\ a_{n1}b_n & a_{n2}b_n & \cdots & a_{nn}b_n \end{pmatrix} \\ &= \begin{pmatrix} 0 & c_{12} & \cdots & c_{1n} \\ c_{21} & 0 & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & 0 \end{pmatrix}. \end{aligned}$$

当 $i \neq j$ 时, $a_{ij}b_j - a_{ij}b_i = c_{ij}$, 因此 $a_{ij} = \frac{c_{ij}}{b_j - b_i}$, 矩阵 A 存在. \square

7.2.9 矩阵函数

矩阵函数在微分方程理论中有重要应用. 下面介绍几道例题供读者参考. 首先必须注意到, 在具体计算矩阵函数时不能随便套用数值函数的性质. 比如在数值函数中, 成立 $e^x \cdot e^y = e^{x+y} = e^y \cdot e^x$, 但对一般的矩阵 A, B 来说, $e^A \cdot e^B = e^{A+B} = e^B \cdot e^A$ 并不一定成立. 例如,

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

通过计算不难验证 $AB \neq A, BA \neq B$, 并且

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

因此 $AB \neq BA$, 并且在 $e^A \cdot e^B, e^B \cdot e^A$ 以及 e^{A+B} 这 3 个矩阵中, 任意两个都不相等. 但若 A, B 乘法可交换, 则在上述 3 个矩阵中, 任意两个都相等, 这就是下面的例 7.55 和例 7.57.

例 7.55 求证: 若 n 阶矩阵 A 和 B 乘法可交换, 则 $e^A \cdot e^B = e^B \cdot e^A$.

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

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

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. \square

注 由例 7.55 类似的讨论可证明: 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 $f(A)g(B) = g(B)f(A)$.

要证明 $e^A \cdot e^B = e^{A+B}$, 却没有例 7.55 那么简单, 因为这里面涉及到级数的乘积. 比如在数学分析中考虑数项级数乘积 $\sum_{i,j=1}^{\infty} a_i b_j$ 的收敛性, 一般来说, 只有当 $\sum_{i=1}^{\infty} a_i$ 和 $\sum_{j=1}^{\infty} b_j$ 都是绝对收敛时, 上述级数乘积的收敛性才能得到保证, 特别地, 还可以得到 Cauchy 乘积的收敛性, 即 $\sum_{n=1}^{\infty} \left(\sum_{i+j=n} a_i b_j\right)$ 收敛到 $\left(\sum_{i=1}^{\infty} a_i\right) \cdot \left(\sum_{j=1}^{\infty} b_j\right)$. 因此, 为了类似地讨论矩阵幂级数乘积的收敛性, 我们必须引入矩阵范数的概念 (类似于复数的模长). 由于更一般的范数概念及其性质将在第 9 章内积空间中详细定义和研究, 故这里只讨论矩阵范数的一些简单性质. 另外, 矩阵范数可以有多种, 这里我们选取由 Frobenius 内积诱导的范数 (具体可参考第 9 章).

例 7.56 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 $\|A\| = \sqrt{\sum_{i,j=1}^n |a_{ij}|^2}$. 设 $B = (b_{ij})$ 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) 就是一般范数的三角不等式, 读者可参考第 9 章的相关内容. (3) 注意到 $\|AB\|^2 = \sum_{i,j=1}^n \left| \sum_{k=1}^n a_{ik} b_{kj} \right|^2$, $\|A\|^2 \cdot \|B\|^2 = \left(\sum_{i,k=1}^n |a_{ik}|^2 \right) \cdot \left(\sum_{k,j=1}^n |b_{kj}|^2 \right)$, 由 Cauchy-Schwarz 不等式 (参考例 2.68 的复数形式) 即得结论. \square

例 7.57 求证: 若 n 阶矩阵 A 和 B 乘法可交换, 则 $e^A \cdot e^B = e^{A+B}$.

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. 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 $i+j > p$, 故由例 7.56 可得

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

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 $\|f(A)f(B) - f(A+B)\| = 0$, 再次由例 7.56 可得 $e^A \cdot e^B = e^{A+B}$. \square

注 从例 7.57 的证明可以看出, 矩阵幂级数 e^A 的绝对收敛性保证了矩阵级数的 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 $\left(\sum_{i=0}^{\infty} \frac{A^i}{i!} \right) \cdot \left(\sum_{j=0}^{\infty} \frac{B^j}{j!} \right) = e^A \cdot e^B$. 另外, 利用例 7.57 也可给出例 7.55 的另一证明.

例 7.58 设 A 是 n 阶矩阵, 证明: $\sin^2 A + \cos^2 A = I_n$.

证明 由定义可知 $\cos A = \frac{1}{2}(e^{iA} + e^{-iA})$, $\sin A = \frac{1}{2i}(e^{iA} - e^{-iA})$. 求平方和 (注意交换性) 即可得到结论. \square

例 7.59 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

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

例 7.60 设 A 是 n 阶矩阵, 求 e^A 的行列式.

解 设 A 的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$, 则 e^A 的特征值为 $e^{\lambda_1}, e^{\lambda_2}, \dots, e^{\lambda_n}$, 因此

$$|e^A| = e^{\lambda_1} e^{\lambda_2} \dots e^{\lambda_n} = e^{\lambda_1 + \lambda_2 + \dots + \lambda_n} = e^{\text{tr} A}. \quad \square$$

例 7.61 设 A 是 n 阶矩阵, 求 $\lim_{k \rightarrow \infty} A^k$ 存在的充要条件以及极限矩阵.

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

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

因此 $\lim_{k \rightarrow \infty} A^k$ 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 $k > n$, 经计算可得 Jordan 块 $J_{r_i}(\lambda_i)$ 的 k 次幂为

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

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 $\lim_{k \rightarrow \infty} A^k = P \text{diag}\{1, \dots, 1, 0, \dots, 0\}P^{-1}$, 其中 1 的个数等于 A 中特征值 1 的代数重数. \square

7.2.10 Jordan 标准型的几何方法

矩阵的标准型理论通常可采用代数方法或几何方法来做. 我们在教材 [1] 中采用的是代数方法, 即用 λ -矩阵的方法求有理标准型和 Jordan 标准型. 另外一种常用的方法是几何方法, 我们在这一节中将对这一方法作一简要的介绍. 这两种方法各有长处, 代数方法不仅证明了有理标准型和 Jordan 标准型的存在性, 而且同时给出了计算这两类标准型的方法, 比较适合初学者掌握; 而几何方法只能证明标准型的存在性, 并不能精确地计算出标准型, 不过几何方法比较直观, 利于读者从几何的层面上