

第一章 相似标准型

1.1 多项式矩阵

定义 1.1 (λ -矩阵)

一般地, 下列形式的矩阵:

$$A(\lambda) = \begin{pmatrix} a_{11}(\lambda) & a_{12}(\lambda) & \cdots & a_{1n}(\lambda) \\ a_{21}(\lambda) & a_{22}(\lambda) & \cdots & a_{2n}(\lambda) \\ \vdots & \vdots & & \vdots \\ a_{m1}(\lambda) & a_{m2}(\lambda) & \cdots & a_{mn}(\lambda) \end{pmatrix},$$

其中 $a_{ij}(\lambda)$ 是以 λ 为未定元的数域 \mathbb{K} 上的多项式, 称为**多项式矩阵**, 或 **λ -矩阵**. λ -矩阵的加法、数乘及乘法与数域上的矩阵运算一样, 只需在运算过程中将数的运算代之以多项式即可.

定义 1.2 (λ -矩阵的初等变换)

对 λ -矩阵 $A(\lambda)$ 施行的下列 3 种变换称为 λ -矩阵的初等行变换:


- (1) 将 $A(\lambda)$ 的两行对换;
- (2) 将 $A(\lambda)$ 的第 i 行乘以 \mathbb{K} 中的非零常数 c ;
- (3) 将 $A(\lambda)$ 的第 i 行乘以 \mathbb{K} 上的多项式 $f(\lambda)$ 后加到第 j 行上去.

同理我们可以定义 3 种 λ -矩阵的初等列变换.

定义 1.3 (λ -矩阵的相抵)

若 $A(\lambda), B(\lambda)$ 是同阶 λ -矩阵且 $A(\lambda)$ 经过 λ -矩阵的初等变换后可变为 $B(\lambda)$, 则称 $A(\lambda)$ 与 $B(\lambda)$ **相抵**. 与数字矩阵一样, λ -矩阵的相抵关系也是一种等价关系, 即

- (1) $A(\lambda)$ 与自身相抵;
- (2) 若 $A(\lambda)$ 与 $B(\lambda)$ 相抵, 则 $B(\lambda)$ 与 $A(\lambda)$ 相抵;
- (3) 若 $A(\lambda)$ 与 $B(\lambda)$ 相抵, $B(\lambda)$ 与 $C(\lambda)$ 相抵, 则 $A(\lambda)$ 与 $C(\lambda)$ 相抵.

 **笔记** λ -矩阵的相抵关系也是一种等价关系的证明与数域上相同, 类似易证.

定义 1.4 (初等 λ -矩阵)

下列 3 种矩阵称为初等 λ -矩阵:

- (1) 将 n 阶单位阵的第 i 行与第 j 行对换, 记为 P_{ij} ;
- (2) 将 n 阶单位阵的第 i 行乘以非零常数 c , 记为 $P_i(c)$;
- (3) 将 n 阶单位阵的第 i 行乘以多项式 $f(\lambda)$ 后加到第 j 行上去得到的矩阵, 记为 $T_{ij}(f(\lambda))$.

注 第一类与第二类初等 λ -矩阵与数域上的第一类与第二类初等矩阵没有什么区别. 第三类初等 λ -矩阵的形状如

下:

$$T_{ij}(f(\lambda)) = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & \vdots & \ddots & \\ & & f(\lambda) & \cdots & 1 \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}.$$

定理 1.1

对 λ -矩阵 $A(\lambda)$ 施行第 k ($k = 1, 2, 3$) 类初等行 (列) 变换等于用第 k 类初等 λ -矩阵左 (右) 乘以 $A(\lambda)$.



注 下列 λ -矩阵的变换不是 λ -矩阵的初等变换:

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} \lambda & \lambda \\ 0 & 1 \end{pmatrix}.$$

这是因为前面一个矩阵的第一行乘以 λ 不是 λ -矩阵的初等变换. 同理下面的变换需第一行乘以 λ^{-1} , 因此也不是 λ -矩阵的初等变换:

$$\begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

证明 证明是显然的.



定义 1.5 (可逆 λ -矩阵)

若 $A(\lambda), B(\lambda)$ 都是 n 阶 λ -矩阵, 且

$$A(\lambda)B(\lambda) = B(\lambda)A(\lambda) = I_n,$$

则称 $B(\lambda)$ 是 $A(\lambda)$ 的逆 λ -矩阵. 这时称 $A(\lambda)$ 为可逆 λ -矩阵, 在不引起混淆的情形下, 有时简称为可逆阵.



笔记 容易证明, 有限个可逆 λ -矩阵之积仍是可逆 λ -矩阵, 而初等 λ -矩阵都是可逆 λ -矩阵, 因此有限个初等 λ -矩阵之积也是可逆的 λ -矩阵.

注 注意不要将数字矩阵中的一些结论随意搬到 λ -矩阵上. 比如下面的 λ -矩阵的行列式不为零, 但它不是可逆 λ -矩阵:

$$\begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}.$$

这是因为矩阵

$$\begin{pmatrix} \lambda^{-1} & 0 \\ 0 & 1 \end{pmatrix}$$

不是 λ -矩阵之故.

引理 1.1

设 $M(\lambda)$ 是一个 n 阶 λ -矩阵, 则 $M(\lambda)$ 可以化为如下形状:

$$M(\lambda) = M_m \lambda^m + M_{m-1} \lambda^{m-1} + \cdots + M_0,$$

其中 M_i 为数域 \mathbb{F} 上的 n 阶数字矩阵. 因此, 一个多项式矩阵可以化为系数为矩阵的多项式, 反之亦然.



证明 证明是显然的.



引理 1.2

设 $M(\lambda)$ 与 $N(\lambda)$ 是两个 n 阶 λ -矩阵且都不等于零. 又设 B 为 n 阶数字矩阵, 则必存在 λ -矩阵 $Q(\lambda)$ 及 $S(\lambda)$ 和数字矩阵 R 及 T , 使下式成立:

$$M(\lambda) = (\lambda I - B)Q(\lambda) + R, \quad (1.1)$$

$$N(\lambda) = S(\lambda)(\lambda I - B) + T. \quad (1.2)$$

♡

证明 将 $M(\lambda)$ 写为

$$M(\lambda) = M_m \lambda^m + M_{m-1} \lambda^{m-1} + \cdots + M_0,$$

其中 $M_m \neq O$. 可对 m 用归纳法, 若 $m = 0$, 则已符合要求 (取 $Q(\lambda) = O$). 现设对小于 m 次的矩阵多项式, (1.1) 式成立. 令

$$Q_1(\lambda) = M_m \lambda^{m-1},$$

则

$$M(\lambda) - (\lambda I - B)Q_1(\lambda) = (BM_m + M_{m-1})\lambda^{m-1} + \cdots + M_0. \quad (1.3)$$

上式是一个次数小于 m 的矩阵多项式, 由归纳假设得

$$M(\lambda) - (\lambda I - B)Q_1(\lambda) = (\lambda I - B)Q_2(\lambda) + R.$$

于是

$$M(\lambda) = (\lambda I - B)[Q_1(\lambda) + Q_2(\lambda)] + R.$$

令 $Q(\lambda) = Q_1(\lambda) + Q_2(\lambda)$ 即得 (1.1) 式. 同理可证 (1.2) 式. □

定理 1.2

设 A, B 是数域 \mathbb{F} 上的矩阵, 则 A 与 B 相似的充分必要条件是 $\lambda I - A$ 与 $\lambda I - B$ 相抵.

♡

证明 若 A 与 B 相似, 则存在 \mathbb{F} 上的非异阵 P , 使 $B = P^{-1}AP$, 于是

$$P^{-1}(\lambda I - A)P = \lambda I - P^{-1}AP = \lambda I - B.$$

把 P 看成是常数 λ -矩阵, 上式表明 $\lambda I - A$ 与 $\lambda I - B$ 相抵.

反过来, 若 $\lambda I - A$ 与 $\lambda I - B$ 相抵, 即存在 $M(\lambda)$ 及 $N(\lambda)$, 使

$$M(\lambda)(\lambda I - A)N(\lambda) = \lambda I - B, \quad (1.4)$$

其中 $M(\lambda)$ 与 $N(\lambda)$ 都是有限个初等矩阵之积, 因而都是可逆阵. 因此可将 (1.4) 式写为

$$M(\lambda)(\lambda I - A) = (\lambda I - B)N(\lambda)^{-1}, \quad (1.5)$$

由引理 1.2 可设

$$M(\lambda) = (\lambda I - B)Q(\lambda) + R,$$

代入 (1.5) 式经整理得

$$R(\lambda I - A) = (\lambda I - B)[N(\lambda)^{-1} - Q(\lambda)(\lambda I - A)].$$

上式左边是次数小于等于 1 的矩阵多项式, 因此上式右边中括号内的矩阵多项式的次数必须小于等于零, 也即必是一个常数矩阵, 设为 P . 于是

$$R(\lambda I - A) = (\lambda I - B)P. \quad (1.6)$$

(1.6) 式又可整理为

$$(R - P)\lambda = RA - BP.$$

再次比较次数得 $R = P, RA = BP$. 现只需证明 P 是一个非异阵即可. 由假设

$$P = N(\lambda)^{-1} - Q(\lambda)(\lambda I - A),$$

将上式两边右乘 $N(\lambda)$ 并移项得

$$PN(\lambda) + Q(\lambda)(\lambda I - A)N(\lambda) = I.$$

但由(1.4)式可得

$$(\lambda I - A)N(\lambda) = M(\lambda)^{-1}(\lambda I - B),$$

因此

$$PN(\lambda) + Q(\lambda)M(\lambda)^{-1}(\lambda I - B) = I. \quad (1.7)$$

再由引理 1.2 可设

$$N(\lambda) = S(\lambda)(\lambda I - B) + T,$$

代入(1.7)式并整理得

$$[PS(\lambda) + Q(\lambda)M(\lambda)^{-1}](\lambda I - B) = I - PT.$$

上式右边是次数小于等于零的矩阵多项式, 因此上式左边中括号内的矩阵多项式必须为零, 从而 $PT = I$, 即 P 是非异阵. \square

1.2 矩阵的法式

引理 1.3

设 $A(\lambda) = (a_{ij}(\lambda))_{m \times n}$ 是任一非零 λ -矩阵, 则 $A(\lambda)$ 必相抵于这样的一个 λ -矩阵 $B(\lambda) = (b_{ij}(\lambda))_{m \times n}$, 其中 $b_{11}(\lambda) \neq 0$ 且 $b_{11}(\lambda)$ 可整除 $B(\lambda)$ 中的任一元素 $b_{ij}(\lambda)$.



证明 设 $k = \min\{\deg a_{ij}(\lambda) \mid a_{ij}(\lambda) \neq 0, 1 \leq i \leq m; 1 \leq j \leq n\}$, 我们对 k 用数学归纳法. 首先, 经行对换及列对换可将 $A(\lambda)$ 的第 $(1, 1)$ 元素变成次数最低的非零多项式, 因此不妨设 $a_{11}(\lambda) \neq 0$ 且 $\deg a_{11}(\lambda) = k$. 若 $k = 0$, 则 $a_{11}(\lambda)$ 是一个非零常数, 结论显然成立. 假设对非零元素次数的最小值小于 k 的任一 λ -矩阵, 引理的结论成立, 现考虑非零元素次数的最小值等于 k 的 λ -矩阵 $A(\lambda)$. 若 $a_{11}(\lambda)$ 可整除所有的 $a_{ij}(\lambda)$, 则结论已成立. 若否, 设在第一列中有元素 $a_{i1}(\lambda)$ 不能被 $a_{11}(\lambda)$ 整除, 作带余除法:

$$a_{i1}(\lambda) = a_{11}(\lambda)q(\lambda) + r(\lambda).$$

用 $-q(\lambda)$ 乘以第一行加到第 i 行上, 第 $(i, 1)$ 元素就变为 $r(\lambda)$. 注意到 $r(\lambda) \neq 0$ 且 $\deg r(\lambda) < \deg a_{11}(\lambda) = k$, 由归纳假设即知结论成立.

同样的方法可施于第一行. 因此我们不妨设 $a_{11}(\lambda)$ 可整除第一行及第一列. 这时, 设 $a_{21}(\lambda) = a_{11}(\lambda)g(\lambda)$. 将第一行乘以 $-g(\lambda)$ 加到第二行上, 则第 $(2, 1)$ 元素变为零. 用同样的方法可消去第一行、第一列除 $a_{11}(\lambda)$ 以外的所有元素, 于是 $A(\lambda)$ 经初等变换后变成下列形状:

$$\begin{pmatrix} a_{11}(\lambda) & 0 & \cdots & 0 \\ 0 & a'_{22}(\lambda) & \cdots & a'_{2n}(\lambda) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a'_{m2}(\lambda) & \cdots & a'_{mn}(\lambda) \end{pmatrix}.$$

这时, 若 $a_{11}(\lambda)$ 可整除所有其他元素, 则结论已成立. 若否, 比如 $a_{11}(\lambda)$ 不能整除 $a'_{ij}(\lambda)$, 则将第 i 行加到第一行上去, 这时在第一行又出现了一元素 $a'_{ij}(\lambda)$, 它不能被 $a_{11}(\lambda)$ 整除. 重复上面的做法, 通过归纳假设即可得到结论. \square

定理 1.3

设 $A(\lambda)$ 是一个 n 阶 λ -矩阵, 则 $A(\lambda)$ 相抵于对角阵

$$\text{diag}\{d_1(\lambda), d_2(\lambda), \dots, d_r(\lambda), 0, \dots, 0\}, \quad (1.8)$$

其中 $d_i(\lambda)$ 是非零首一多项式且 $d_i(\lambda) \mid d_{i+1}(\lambda)$ ($i = 1, 2, \dots, r-1$). 我们称上式中的对角 λ -矩阵为 $A(\lambda)$ 的**法式或相抵标准型或 Smith 标准型**.



证明 对 n 用数学归纳法, 当 $n = 1$ 时结论显然, 现设 $A(\lambda)$ 是 n 阶 λ -矩阵. 由引理 1.3 可知 $A(\lambda)$ 相抵于 n 阶 λ -矩阵 $B(\lambda) = (b_{ij}(\lambda))$, 其中 $b_{11}(\lambda) \mid b_{ij}(\lambda)$ 对一切 i, j 成立. 因此, 将 $B(\lambda)$ 的第一行乘以 λ 的某个多项式加到第二行上去便可消去 $b_{21}(\lambda)$. 同理可依次消去第一列除 $b_{11}(\lambda)$ 以外的所有元素. 再用类似方法消去第一行其余元素. 这样便得到了一个矩阵:

$$\begin{pmatrix} b_{11}(\lambda) & 0 & \cdots & 0 \\ 0 & b'_{22}(\lambda) & \cdots & b'_{2n}(\lambda) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & b'_{n2}(\lambda) & \cdots & b'_{nn}(\lambda) \end{pmatrix}.$$

不难看出, 这时 $b_{11}(\lambda)$ 仍可整除所有的 $b'_{ij}(\lambda)$. 设 c 为 $b_{11}(\lambda)$ 的首项系数, 记 $d_1(\lambda) = c^{-1}b_{11}(\lambda)$, 设 $\bar{B}(\lambda)$ 为上面的矩阵中右下方的 $n-1$ 阶 λ -矩阵, 则由归纳假设可知存在 $P(\lambda)$ 及 $Q(\lambda)$, 使

$$P(\lambda)\bar{B}(\lambda)Q(\lambda) = \text{diag}\{d_2(\lambda), \dots, d_r(\lambda); 0, \dots, 0\},$$

且 $d_i(\lambda) \mid d_{i+1}(\lambda)$ ($i = 2, \dots, r-1$), 其中 $P(\lambda)$ 与 $Q(\lambda)$ 可写成为有限个 $n-1$ 阶初等 λ -矩阵之积. 因此

$$\begin{pmatrix} 1 & O \\ O & P(\lambda) \end{pmatrix} \begin{pmatrix} d_1(\lambda) & O \\ O & \bar{B}(\lambda) \end{pmatrix} \begin{pmatrix} 1 & O \\ O & Q(\lambda) \end{pmatrix} = \text{diag}\{d_1(\lambda), d_2(\lambda), \dots, d_r(\lambda); 0, \dots, 0\},$$

且

$$\begin{pmatrix} 1 & O \\ O & P(\lambda) \end{pmatrix}, \begin{pmatrix} 1 & O \\ O & Q(\lambda) \end{pmatrix}$$

可写成有限个 n 阶初等 λ -矩阵之积. 于是只需证明 $d_1(\lambda) \mid d_2(\lambda)$ 即可. 但这点很容易看出, 事实上由于 $\bar{B}(\lambda)$ 中的任一元素均可被 $d_1(\lambda)$ 整除, 因此 $P(\lambda)\bar{B}(\lambda)Q(\lambda)$ 中的任一元素也可被 $d_1(\lambda)$ 整除, 这就证明了定理. \square

注 我们上面对 n 阶 λ -矩阵证明了它必相抵于一个对角阵. 事实上, 对长方 λ -矩阵, 结论也同样成立, 证明也类似. (1.8) 式中的 r 通常称为 $A(\lambda)$ 的秩. 但要注意即使某个 n 阶 λ -矩阵的秩等于 n , 它也未必是可逆 λ -矩阵.

推论 1.1

任一 n 阶可逆 λ -矩阵都可表示为有限个初等 λ -矩阵之积.



证明 由定理 1.3, 存在 $P(\lambda), Q(\lambda)$, 使可逆阵 $A(\lambda)$ 适合

$$P(\lambda)A(\lambda)Q(\lambda) = \text{diag}\{d_1(\lambda), d_2(\lambda), \dots, d_r(\lambda); 0, \dots, 0\},$$

其中 $P(\lambda), Q(\lambda)$ 为有限个初等 λ -矩阵之积. 因为上式左边是个可逆阵, 故右边的矩阵也可逆, 从而 $r = n$. 注意一个对角 λ -矩阵要可逆必须 $d_1(\lambda), d_2(\lambda), \dots, d_n(\lambda)$ 皆为非零常数, 又它们都是首一多项式, 故只能是 $d_1(\lambda) = d_2(\lambda) = \dots = d_n(\lambda) = 1$, 于是

$$A(\lambda) = P(\lambda)^{-1}Q(\lambda)^{-1}.$$

因为初等 λ -矩阵的逆仍是初等 λ -矩阵, 故 $P(\lambda)^{-1}$ 与 $Q(\lambda)^{-1}$ 都是有限个初等 λ -矩阵之积, 从而 $A(\lambda)$ 也是有限个初等 λ -矩阵之积. \square

推论 1.2

设 A 是数域 \mathbb{K} 上的 n 阶矩阵, 则 A 的特征矩阵 $\lambda I_n - A$ 必相抵于

$$\text{diag}\{1, \cdots, 1, d_1(\lambda), \cdots, d_m(\lambda)\},$$

其中 $d_i(\lambda) \mid d_{i+1}(\lambda)$ ($i = 1, 2, \cdots, m-1$). 我们称上式中的对角 λ -矩阵为 $A(\lambda)$ 的特征矩阵 $\lambda I_n - A$ 的**法式**或**相抵标准型**.



证明 由定理 1.3, 存在 $P(\lambda), Q(\lambda)$, 使

$$P(\lambda)(\lambda I_n - A)Q(\lambda) = \text{diag}\{d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda); 0, \cdots, 0\},$$

其中 $P(\lambda), Q(\lambda)$ 为有限个初等 λ -矩阵之积. 根据 λ -矩阵初等变换的定义以及行列式的性质可得, 上式左边的行列式等于 $c|\lambda I_n - A|$, 其中 c 是一个非零常数, 从而上式右边的行列式不为零, 故 $r = n$. 把 $d_i(\lambda)$ 中的常数多项式写出来 (因是首一多项式, 故为常数 1), 即得结论. \square

例题 1.1 求 $\lambda I - A$ 的**法式**, 其中

$$A = \begin{pmatrix} 0 & 1 & -1 \\ 3 & -2 & 0 \\ -1 & 1 & -1 \end{pmatrix}.$$

解

$$\begin{aligned} \lambda I - A &= \begin{pmatrix} \lambda & -1 & 1 \\ -3 & \lambda+2 & 0 \\ 1 & -1 & \lambda+1 \end{pmatrix} \xrightarrow{r_1 \leftrightarrow r_3} \begin{pmatrix} 1 & -1 & \lambda+1 \\ -3 & \lambda+2 & 0 \\ \lambda & -1 & 1 \end{pmatrix} \\ &\xrightarrow{3r_1+r_2, -\lambda r_1+r_3} \begin{pmatrix} 1 & -1 & \lambda+1 \\ 0 & \lambda-1 & 3\lambda+3 \\ 0 & \lambda-1 & -\lambda^2-\lambda+1 \end{pmatrix} \xrightarrow{j_1+j_2, -(\lambda+1)j_1+j_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda-1 & 3\lambda+3 \\ 0 & \lambda-1 & -\lambda^2-\lambda+1 \end{pmatrix} \\ &\xrightarrow{-3j_2+j_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda-1 & 6 \\ 0 & \lambda-1 & -\lambda^2-4\lambda+4 \end{pmatrix} \xrightarrow{j_2 \leftrightarrow j_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 6 & \lambda-1 \\ 0 & -\lambda^2-4\lambda+4 & \lambda-1 \end{pmatrix} \\ &\xrightarrow{6j_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 6 & 6(\lambda-1) \\ 0 & -\lambda^2-4\lambda+4 & 6(\lambda-1) \end{pmatrix} \xrightarrow{-(\lambda-1)j_2+j_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & -\lambda^2-4\lambda+4 & (\lambda-1)(\lambda^2+4\lambda+2) \end{pmatrix} \\ &\xrightarrow{\frac{1}{6}j_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\lambda^2-4\lambda+4 & (\lambda-1)(\lambda^2+4\lambda+2) \end{pmatrix} \xrightarrow{-(-\lambda^2-4\lambda+4)r_2+r_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (\lambda-1)(\lambda^2+4\lambda+2) \end{pmatrix}. \end{aligned}$$

\square

1.3 不变因子

定义 1.6 (k 阶行列式因子)

设 $A(\lambda)$ 是 n 阶 λ -矩阵, k 是小于等于 n 的正整数. 如果 $A(\lambda)$ 有一个 k 阶子式不为零, 则定义 $A(\lambda)$ 的 k 阶**行列式因子** $D_k(\lambda)$ 为 $A(\lambda)$ 的所有 k 阶子式的最大公因式 (首一多项式). 如果 $A(\lambda)$ 的所有 k 阶子式都等于零, 则定义 $A(\lambda)$ 的 k 阶行列式因子 $D_k(\lambda)$ 为零.



引理 1.4

设 $D_1(\lambda), D_2(\lambda), \dots, D_r(\lambda)$ 是 $A(\lambda)$ 的非零行列式因子, 则

$$D_i(\lambda) \mid D_{i+1}(\lambda), \quad i = 1, 2, \dots, r-1.$$



证明 设 A_{i+1} 是 $A(\lambda)$ 的任一 $i+1$ 阶子式, 即在 $A(\lambda)$ 中任意取出 $i+1$ 行及 $i+1$ 列组成的行列式. 将这个行列式按某一行展开, 则它的每一个展开项都是一个多项式与一个 i 阶子式的乘积. 由于 $D_i(\lambda)$ 是所有 i 阶子式的公因子, 因此 $D_i(\lambda) \mid A_{i+1}$. 而 $D_{i+1}(\lambda)$ 是所有 $i+1$ 阶子式的最大公因子, 因此 $D_i(\lambda) \mid D_{i+1}(\lambda)$ 对一切 $i = 1, 2, \dots, r-1$ 成立. \square

定义 1.7 (不变因子)

设 $D_1(\lambda), D_2(\lambda), \dots, D_r(\lambda)$ 是 λ -矩阵 $A(\lambda)$ 的非零行列式因子, 则

$$\begin{aligned} g_1(\lambda) &= D_1(\lambda), \\ g_2(\lambda) &= D_2(\lambda)/D_1(\lambda), \\ &\dots \\ g_r(\lambda) &= D_r(\lambda)/D_{r-1}(\lambda) \end{aligned}$$

称为 $A(\lambda)$ 的不变因子.



笔记 由不变因子和行列式因子的定义可知, 不变因子和行列式因子相互唯一确定.

注 以后特征矩阵 $\lambda I - A$ 的行列式因子和不变因子均简称为 A 的行列式因子和不变因子.

命题 1.1

求下列矩阵的行列式因子和不变因子:

$$A(\lambda) = \begin{pmatrix} d_1(\lambda) & & & & \\ & \ddots & & & \\ & & d_r(\lambda) & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix},$$

其中 $d_i(\lambda)$ 为非零首一多项式且 $d_i(\lambda) \mid d_{i+1}(\lambda)$ ($i = 1, 2, \dots, r-1$).



解 $A(\lambda)$ 的非零行列式因子为

$$\begin{aligned} D_1(\lambda) &= d_1(\lambda), \\ D_2(\lambda) &= d_1(\lambda)d_2(\lambda), \\ &\dots \\ D_r(\lambda) &= d_1(\lambda)d_2(\lambda)\cdots d_r(\lambda). \end{aligned}$$

根据不变因子的定义可知 $A(\lambda)$ 的不变因子分别为: $d_1(\lambda), d_2(\lambda), \dots, d_n(\lambda)$. \square

定理 1.4

相抵的 λ -矩阵有相同的行列式因子, 从而有相同的不变因子.



证明 我们只需证明行列式因子在三类初等变换下不改变就可以了. 对第一类初等变换, 交换 λ -矩阵 $A(\lambda)$ 的任意两行 (列), 显然 $A(\lambda)$ 的 i 阶子式最多改变一个符号, 因此行列式因子不改变.

对第二类初等变换, $A(\lambda)$ 的 i 阶子式与变换后矩阵的 i 阶子式最多差一个非零常数, 因此行列式因子也不改

变.

对第三类初等变换, 记变换后的矩阵为 $B(\lambda)$, 则 $B(\lambda)$ 与 $A(\lambda)$ 的 i 阶子式可能出现以下 3 种情形: 子式完全相同; $B(\lambda)$ 子式中的某一行 (列) 等于 $A(\lambda)$ 中相应子式的同一行 (列) 加上该子式中某一行 (列) 与某个多项式之积; $B(\lambda)$ 子式中的某一行 (列) 等于 $A(\lambda)$ 中相应子式的同一行 (列) 加上不在该子式中的某一行 (列) 与某个多项式之积. 在前面两种情形, 行列式的值不改变, 因此不影响行列式因子. 现在来讨论第三种情形. 设 B_i 为 $B(\lambda)$ 的 i 阶子式, 相应的 $A(\lambda)$ 的 i 阶子式记为 A_i , 则由行列式的性质得


$$B_i = A_i + f(\lambda)\tilde{A}_i,$$

其中 \tilde{A}_i 由 $A(\lambda)$ 中的 i 行与 i 列组成, 因此它与 $A(\lambda)$ 的某个 i 阶子式最多差一个符号. $f(\lambda)$ 是乘以某一行 (列) 的那个多项式, 于是 $A(\lambda)$ 的行列式因子 $D_i(\lambda) \mid A_i, D_i(\lambda) \mid \tilde{A}_i$, 故 $D_i(\lambda) \mid B_i$. 这说明, $D_i(\lambda)$ 可整除 $B(\lambda)$ 的所有 i 阶子式, 因此 $D_i(\lambda)$ 可整除 $B(\lambda)$ 的 i 阶行列式因子 $\tilde{D}_i(\lambda)$. 但 $B(\lambda)$ 也可用第三类初等变换变成 $A(\lambda)$, 于是 $\tilde{D}_i(\lambda) \mid D_i(\lambda)$. 由于 $D_i(\lambda)$ 及 $\tilde{D}_i(\lambda)$ 都是首一多项式, 因此必有 $D_i(\lambda) = \tilde{D}_i(\lambda)$. \square

推论 1.3

设 n 阶 λ -矩阵 $A(\lambda)$ 的法式为

$$\Lambda = \text{diag}\{d_1(\lambda), d_2(\lambda), \dots, d_r(\lambda); 0, \dots, 0\},$$

其中 $d_i(\lambda)$ 是非零首一多项式且 $d_i(\lambda) \mid d_{i+1}(\lambda)$ ($i = 1, 2, \dots, r-1$), 则 $A(\lambda)$ 的不变因子为 $d_1(\lambda), d_2(\lambda), \dots, d_r(\lambda)$ 特别地, **法式和不变因子之间相互唯一确定.** 

证明 首先, 由定理 1.4 可知, $A(\lambda)$ 与 Λ 有相同的不变因子. 再由命题 1.1 可知, Λ 的不变因子为 $d_1(\lambda), d_2(\lambda), \dots, d_r(\lambda)$, 从而它们也是 $A(\lambda)$ 的不变因子. 故 $A(\lambda)$ 的法式可以唯一确定其不变因子.

接着, 设 $A(\lambda)$ 的不变因子为 $d_1(\lambda), d_2(\lambda), \dots, d_r(\lambda)$, 由定理 1.3, 可设 $A(\lambda)$ 相抵于对角阵

$$B(\lambda) = \text{diag}\{d'_1(\lambda), d'_2(\lambda), \dots, d'_r(\lambda); 0, \dots, 0\}, \quad (1.9)$$

其中 $d'_i(\lambda)$ 是非零首一多项式且 $d'_i(\lambda) \mid d'_{i+1}(\lambda)$ ($i = 1, 2, \dots, r-1$). 再由命题 1.1 可知, $B(\lambda)$ 的不变因子为 $d'_1(\lambda), d'_2(\lambda), \dots, d'_r(\lambda)$. 由定理 1.4 可知, $A(\lambda)$ 和 $B(\lambda)$ 的不变因子相同, 故根据 $d_i(\lambda), d'_i(\lambda)$ 的整除关系, 我们就有


$$\begin{aligned} d_1(\lambda) &= d'_1(\lambda), \\ d_2(\lambda) &= d'_2(\lambda), \\ &\dots\dots\dots \\ d_r(\lambda) &= d'_r(\lambda). \end{aligned}$$

因此 $A(\lambda)$ 的相抵于对角阵

$$\Lambda = \text{diag}\{d_1(\lambda), d_2(\lambda), \dots, d_r(\lambda); 0, \dots, 0\},$$


其中 $d_i(\lambda)$ 是非零首一多项式且 $d_i(\lambda) \mid d_{i+1}(\lambda)$ ($i = 1, 2, \dots, r-1$). 上式也就是 $A(\lambda)$ 的法式. 故 $A(\lambda)$ 的不变因子可以唯一确定其法式. \square

推论 1.4

设 $A(\lambda), B(\lambda)$ 为 n 阶 λ -矩阵, 则 $A(\lambda)$ 与 $B(\lambda)$ 相抵当且仅当它们有相同的法式. 

证明 若 $A(\lambda)$ 与 $B(\lambda)$ 有相同的法式, 显然它们相抵. 若 $A(\lambda)$ 与 $B(\lambda)$ 相抵, 由定理 1.4 知 $A(\lambda)$ 与 $B(\lambda)$ 有相同的不变因子, 从而由推论 1.3 可知, $A(\lambda)$ 与 $B(\lambda)$ 有相同的法式. \square

推论 1.5

n 阶 λ -矩阵 $A(\lambda)$ 的法式与初等变换的选取无关. 

证明 设 Λ_1, Λ_2 是 $A(\lambda)$ 通过不同的初等变换得到的两个法式, 则 Λ_1 与 Λ_2 相抵, 由推论 1.4 可得 $\Lambda_1 = \Lambda_2$. \square

定理 1.5

数域 \mathbb{K} 上 n 阶矩阵 A 与 B 相似的充分必要条件是它们的特征矩阵 $\lambda I - A$ 与 $\lambda I - B$ 具有相同的行列式因子或不变因子.



证明 显然不变因子与行列式因子之间相互唯一确定. 再由定理 1.2、推论 1.4 及推论 1.3 即得结论. \square

推论 1.6

设 $\mathbb{F} \subseteq \mathbb{K}$ 是两个数域, A, B 是 \mathbb{F} 上的两个矩阵, 则 A 与 B 在 \mathbb{F} 上相似的充分必要条件是它们在 \mathbb{K} 上相似.



笔记 这个推论告诉我们: 矩阵的相似关系在基域扩张下不变. 事实上, 这个推论的证明过程也说明: 矩阵的不变因子在基域扩张下也不变.

证明 若 A 与 B 在 \mathbb{F} 上相似, 由于 $\mathbb{F} \subseteq \mathbb{K}$, 它们当然在 \mathbb{K} 上也相似. 反之, 若 A 与 B 在 \mathbb{K} 上相似, 则 $\lambda I - A$ 与 $\lambda I - B$ 在 \mathbb{K} 上有相同的不变因子, 也就是说它们有相同的法式. 由推论 1.5 可知, 求法式与初等变换的选取无关. 注意到 $\lambda I - A$ 与 $\lambda I - B$ 是数域 \mathbb{F} 上的 λ -矩阵, 故可用 \mathbb{F} 上 λ -矩阵的初等变换就能将它们变成法式, 其中只涉及 \mathbb{F} 中数的加、减、乘、除运算以及 \mathbb{F} 上的多项式的加、减、乘、数乘运算, 最后得到法式中的不变因子 $d_i(\lambda)$ 仍是 \mathbb{F} 上的多项式. 这就是说存在 \mathbb{F} 上的可逆 λ -矩阵 $P(\lambda), Q(\lambda), M(\lambda), N(\lambda)$, 使

$$P(\lambda)(\lambda I - A)Q(\lambda) = M(\lambda)(\lambda I - B)N(\lambda) = \text{diag}\{d_1(\lambda), \dots, d_n(\lambda)\},$$

从而

$$M(\lambda)^{-1}P(\lambda)(\lambda I - A)Q(\lambda)N(\lambda)^{-1} = \lambda I - B,$$

即 $\lambda I - A$ 与 $\lambda I - B$ 在 \mathbb{F} 上相抵, 由定理 1.2 可得 A 与 B 在 \mathbb{F} 上相似. \square

1.4 有理标准型

命题 1.2

设矩阵 A 的特征矩阵 $\lambda I - A$ 的法式为

$$\text{diag}\{1, \dots, 1, d_1(\lambda), \dots, d_k(\lambda)\},$$

其中 $d_i(\lambda)$ 为非常数首一多项式且 $d_i(\lambda) \mid d_{i+1}(\lambda)$ ($i = 1, 2, \dots, k-1$), 则 A 的不变因子就是

$$1, \dots, 1, d_1(\lambda), \dots, d_k(\lambda).$$



证明 由推论 1.2 可知, 矩阵 A 的特征矩阵 $\lambda I - A$ 的法式为

$$\text{diag}\{1, \dots, 1, d_1(\lambda), \dots, d_k(\lambda)\},$$

其中 $d_i(\lambda)$ 为非常数首一多项式且 $d_i(\lambda) \mid d_{i+1}(\lambda)$ ($i = 1, 2, \dots, k-1$), 则根据不变因子的定义可知, A 的不变因子就是

$$1, \dots, 1, d_1(\lambda), \dots, d_k(\lambda).$$



引理 1.5 (Frobenius 块的基本性质)

设 r 阶矩阵

$$F = F(f(\lambda)) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_r & -a_{r-1} & -a_{r-2} & \cdots & -a_1 \end{pmatrix},$$

其中 $f(\lambda) = \lambda^r + a_1\lambda^{r-1} + \cdots + a_r$, 则

(1) $|F(f(\lambda))| = (-1)^{r+2}a_r = (-1)^{r+2}f(0)$.

(2) $F = F(f(\lambda))$ 的特征多项式 $|\lambda I - F|$ 为 $f(\lambda)$.

(3) F 的行列式因子为

$$1, \cdots, 1, f(\lambda), \quad (1.10)$$

其中共有 $r-1$ 个 1, $f(\lambda) = \lambda^r + a_1\lambda^{r-1} + \cdots + a_r$, F 的不变因子也由(1.10)式给出, F 的不变因子分别为:

$$1, \cdots, 1, f(\lambda).$$

进而, $\lambda I - F$ 相抵于 $\text{diag}\{1, \cdots, 1, f(\lambda)\}$.

(4) F 的极小多项式等于 $f(\lambda)$.



证明

(1) 注意到 $f(0) = a_r$, 于是就有

$$|F(f(\lambda))| = \begin{vmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_r & -a_{r-1} & -a_{r-2} & \cdots & -a_1 \end{vmatrix} \xrightarrow{\text{按第一列展开}} (-1)^{r+2}a_r = (-1)^{r+2}f(0).$$

(2) 由命题??(1) 同理可得

$$|\lambda I - F| = \lambda^r + a_1\lambda^{r-1} + \cdots + a_r = f(\lambda).$$

(3) F 的 r 阶行列式因子就是它的特征多项式, 由命题??(1) 同理可得

$$|\lambda I - F| = \lambda^r + a_1\lambda^{r-1} + \cdots + a_r = f(\lambda).$$

对任一 $k < r$, $\lambda I - F$ 总有一个 k 阶子式其值等于 $(-1)^k$, 故 $D_k(\lambda) = 1$. 又由推论 1.3 可知, $\lambda I - F$ 的法式为 $\text{diag}\{1, \cdots, 1, f(\lambda)\}$. 故 $\lambda I - F$ 相抵于 $\text{diag}\{1, \cdots, 1, f(\lambda)\}$.

(4) 因为 F 的特征多项式为 $f(\lambda)$, 所以 F 适合多项式 $f(\lambda)$. 设 e_i ($i = 1, 2, \cdots, r$) 是 r 维标准单位行向量, 则不难算出:

$$e_1 F = e_2, \quad e_1 F^2 = e_2 F = e_3, \quad \cdots, \quad e_1 F^{r-1} = e_{r-1} F = e_r.$$

显然, $e_1, e_1 F, \cdots, e_1 F^{r-1}$ 是一组线性无关的向量, 从而任取 $g(x) \in P_{r-1}[x]$ 且 $g(x)$ 非零, 则存在一组不全为零的数 a_1, a_2, \cdots, a_r , 使得

$$g(x) = a_1 x^{r-1} + a_2 x^{r-2} + \cdots + a_r.$$

于是将 F 代入上式, 再在等式两边同乘 e_1 得到

$$e_1 g(F) = a_1 e_1 F^{r-1} + a_2 e_1 F^{r-2} + \cdots + a_r e_1 F.$$

又因为 $e_1, e_1 F, \cdots, e_1 F^{r-1}$ 是一组线性无关的向量, 且 a_1, a_2, \cdots, a_r 不全为零, 所以 $e_1 g(F) \neq 0$. 即 $g(F)$ 的第

一行不为零, 故 $g(F) \neq 0$. 因此 F 不可能适合一个次数不超过 $r-1$ 的非零多项式, 从而 F 的极小多项式就是 $f(\lambda)$. □

引理 1.6

设 λ -矩阵 $A(\lambda)$ 相抵于对角 λ -矩阵

$$\text{diag}\{d_1(\lambda), d_2(\lambda), \dots, d_n(\lambda)\}, \quad (1.11)$$

λ -矩阵 $B(\lambda)$ 相抵于对角 λ -矩阵

$$\text{diag}\{d'_1(\lambda), d'_2(\lambda), \dots, d'_n(\lambda)\}, \quad (1.12)$$

且 $d'_1(\lambda), d'_2(\lambda), \dots, d'_n(\lambda)$ 是 $d_1(\lambda), d_2(\lambda), \dots, d_n(\lambda)$ 的一个置换 (即若不计次序, 这两组多项式完全相同), 则 $A(\lambda)$ 相抵于 $B(\lambda)$. ♡

证明 利用初等行对换及初等列对换即可将 (1.11) 式变成 (1.12) 式, 因此 (1.11) 式所示的矩阵与 (1.12) 式所示的矩阵相抵, 从而 $A(\lambda)$ 与 $B(\lambda)$ 相抵. □

定理 1.6 (有理标准型/Frobenius 标准型)

设 A 是数域 \mathbb{F} 上的 n 阶方阵, A 的不变因子组为

$$1, \dots, 1, d_1(\lambda), \dots, d_k(\lambda),$$

其中 $\deg d_i(\lambda) = m_i \geq 1$, 则 A 相似于下列分块对角阵:

$$F = \begin{pmatrix} F_1 & & & \\ & F_2 & & \\ & & \ddots & \\ & & & F_k \end{pmatrix}, \quad (1.13)$$

其中 F_i 的阶等于 m_i , 且 F_i 是形如引理 1.5 中的矩阵, F_i 的最后一行由 $d_i(\lambda)$ 的系数 (除首项系数之外) 的负值组成. 此即, 设 $d_i = \lambda^{m_i} + a_{1i}\lambda^{m_i-1} + \dots + a_{m_i,i}$, 则

$$F = F(d_i(\lambda)) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_{m_i,i} & -a_{m_i-1,i} & -a_{m_i-2,i} & \cdots & -a_{1i} \end{pmatrix}.$$

(1.13) 式称为矩阵 A 的**有理标准型**或**Frobenius 标准型**, 每个 F_i 称为**Frobenius 块**. ♡

证明 注意到 $\lambda I - A$ 的第 n 个行列式因子就是 A 的特征多项式 $|\lambda I - A|$, 再由不变因子的定义可知:

$$|\lambda I - A| = d_1(\lambda)d_2(\lambda)\cdots d_k(\lambda).$$

而 $|\lambda I - A|$ 是一个 n 次多项式, 因此 $m_1 + m_2 + \dots + m_k = n$. 一方面, $\lambda I - A$ 的法式为

$$\text{diag}\{1, \dots, 1, d_1(\lambda), d_2(\lambda), \dots, d_k(\lambda)\},$$

其中有 $n-k$ 个 1. 另一方面, 对 $\lambda I - F$ 的每个分块都施以 λ -矩阵的初等变换, 由引理 1.5 可知, $\lambda I - F$ 相抵于如下对角阵:

$$\text{diag}\{1, \dots, 1, d_1(\lambda); 1, \dots, 1, d_2(\lambda); \dots; 1, \dots, 1, d_k(\lambda)\}, \quad (1.14)$$

其中每个 $d_i(\lambda)$ 前各有 $m_i - 1$ 个 1, 从而共有 $\sum_{i=1}^k (m_i - 1) = n - k$ 个 1. 因此 (1.14) 式所示的矩阵与 $\lambda I - A$ 的法式

只相差主对角线上元素的置换, 由引理 1.6 可得 $\lambda I - A$ 与 $\lambda I - F$ 相抵, 从而 A 与 F 相似. \square

例题 1.2 设 6 阶矩阵 A 的不变因子为

$$1, 1, 1, \lambda - 1, (\lambda - 1)^2, (\lambda - 1)^2(\lambda + 1),$$

则 A 的有理标准型为

$$\begin{pmatrix} 1 & & & & & \\ & 0 & 1 & & & \\ & -1 & 2 & & & \\ & & & 0 & 1 & 0 \\ & & & 0 & 0 & 1 \\ & & & -1 & 1 & 1 \end{pmatrix}.$$

定理 1.7

设数域 \mathbb{K} 上的 n 阶矩阵 A 的不变因子为

$$1, \dots, 1, d_1(\lambda), \dots, d_k(\lambda),$$

其中 $d_i(\lambda) \mid d_{i+1}(\lambda)$ ($i = 1, \dots, k-1$), 则 A 的特征多项式为 $d_1(\lambda)d_2(\lambda)\cdots d_k(\lambda)$, 极小多项式为 $m(\lambda) = d_k(\lambda)$. \heartsuit

证明 首先证明特征多项式, 根据不变因子和行列式因子的定义可知

$$1 \cdot 1 \cdot d_1(\lambda)d_2(\lambda)\cdots d_k(\lambda) = D_n(\lambda).$$

其中 $D_n(\lambda)$ 为 $\lambda I_n - A$ 的 n 阶行列式因子, 即 A 的特征多项式 $|\lambda I_n - A|$. 因此

$$|\lambda I_n - A| = d_1(\lambda)d_2(\lambda)\cdots d_k(\lambda).$$

然后证明极小多项式, 设 A 的有理标准型为

$$F = \begin{pmatrix} F_1 & & & \\ & F_2 & & \\ & & \ddots & \\ & & & F_k \end{pmatrix}.$$

因为相似矩阵有相同的极小多项式, 故只需证明 F 的极小多项式是 $d_k(\lambda)$ 即可. 但 F 是分块对角阵, 由极小多项式的性质 (6) 知 F 的极小多项式是诸 F_i 极小多项式的最小公倍式. 又由引理 1.5 知 F_i 的极小多项式为 $d_i(\lambda)$. 因为 $d_i(\lambda) \mid d_{i+1}(\lambda)$, 故诸 $d_i(\lambda)$ 的最小公倍式等于 $d_k(\lambda)$. \square

例题 1.3 下面两个 4 阶矩阵

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

的不变因子分别为 $A: 1, \lambda, \lambda, \lambda^2$ 和 $B: 1, 1, \lambda^2, \lambda^2$. 它们的特征多项式和极小多项式分别相等, 但它们不相似.

定义 1.8 (循环子空间)

设 V 是数域 \mathbb{K} 上的 n 维线性空间, φ 是 V 上的线性变换. 设 $0 \neq \alpha \in V$, 则 $U = L(\alpha, \varphi(\alpha), \varphi^2(\alpha), \dots)$ 称为 V 的循环子空间, 记为 $U = C(\varphi, \alpha)$, α 称为 U 的循环向量. 若 $U = V$, 则称 V 为循环空间. \clubsuit

定理 1.8 (循环子空间的基本性质)

设 V 是数域 \mathbb{K} 上的 n 维线性空间, φ 是 V 上的线性变换, $0 \neq \alpha \in V$, $U = C(\varphi, \alpha)$ 为循环子空间, 则循环子空间 U 是 V 的 φ -不变子空间, 并且是包含 α 的最小 φ -不变子空间. \heartsuit

证明 U 是 V 的 φ -不变子空间是显然的. 下证 U 是包含 α 的最小 φ -不变子空间.

设 $\alpha \in W$, 且 W 为 φ -不变子空间, 则由数学归纳法易知

$$\alpha, \varphi^k(\alpha) \in W, \forall k \in \mathbb{N}_1.$$

于是

$$U = L(\alpha, \varphi(\alpha), \dots) \in W.$$

□

定理 1.9

设 V 是数域 \mathbb{F} 上的 n 维线性空间, φ 是 V 上的线性变换, $0 \neq \alpha \in V$, $U = C(\varphi, \alpha)$ 为循环子空间, 若 $\dim U = r$, 求证: $\{\alpha, \varphi(\alpha), \dots, \varphi^{r-1}(\alpha)\}$ 是 U 的一组基.

♡

证明 设 $m = \max\{k \in \mathbb{Z}^+ \mid \alpha, \varphi(\alpha), \dots, \varphi^{k-1}(\alpha) \text{ 线性无关}\}$, 则显然 $1 \leq m \leq r$, 故 m 是良定义的. 于是由命题??和数学归纳法容易验证: 对任意的 $k \geq m$, $\varphi^k(\alpha)$ 都是 $\alpha, \varphi(\alpha), \dots, \varphi^{m-1}(\alpha)$ 的线性组合, 于是 $\{\alpha, \varphi(\alpha), \dots, \varphi^{m-1}(\alpha)\}$ 是 U 的一组基, 从而 $m = \dim U = r$. □

定理 1.10

设 U 是 V 的 φ -不变子空间, 求证: U 为循环子空间的充要条件是 $\varphi|_U$ 在 U 的某组基下的表示矩阵为某个首一多项式的友阵.

♡

证明 充分性: 设 $\varphi|_U$ 在 U 的一组基 $\{e_1, e_2, \dots, e_r\}$ 下的表示矩阵是友阵 $C(d(\lambda))$, 其中 $d(\lambda) = \lambda^r + a_1\lambda^{r-1} + \dots + a_{r-1}\lambda + a_r$, 则由友阵的定义可知 $\varphi(e_i) = e_{i+1} (1 \leq i \leq r-1)$, $\varphi(e_r) = -\sum_{i=1}^r a_{r-i+1}e_i$. 因此 $e_i = \varphi^{i-1}(e_1) (2 \leq i \leq r)$, $U = L(e_1, e_2, \dots, e_r) = C(\varphi, e_1)$ 为循环子空间.

必要性: 设 $U = C(\varphi, \alpha)$ 是 r 维循环子空间, 则由定理 1.9 可知, $\{\alpha, \varphi(\alpha), \dots, \varphi^{r-1}(\alpha)\}$ 是 U 的一组基. 设

$$\varphi^r(\alpha) = -a_r\alpha - a_{r-1}\varphi(\alpha) - \dots - a_1\varphi^{r-1}(\alpha)$$


令 $d(\lambda) = \lambda^r + a_1\lambda^{r-1} + \dots + a_{r-1}\lambda + a_r$, 容易验证: $\varphi|_U$ 在基 $\{\alpha, \varphi(\alpha), \dots, \varphi^{r-1}(\alpha)\}$ 下的表示矩阵就是友阵 $C(d(\lambda))$. □

定理 1.11 (有理标准型的几何意义)

设 V 是数域 \mathbb{F} 上的 n 维线性空间, φ 是 V 上的线性变换, 且 φ 的不变因子组是 $1, \dots, 1, d_1(\lambda), \dots, d_k(\lambda)$, 其中 $d_i(\lambda)$ 是非常数首一多项式, $d_i(\lambda) \mid d_{i+1}(\lambda) (1 \leq i \leq k-1)$, 则 V 存在一个循环子空间的直和分解:

$$V = C(\varphi, \alpha_1) \oplus C(\varphi, \alpha_2) \oplus \dots \oplus C(\varphi, \alpha_k) \quad (1.15)$$

使得 $\varphi|_{C(\varphi, \alpha_i)}$ 在基 $\{\alpha_i, \varphi(\alpha_i), \dots, \varphi^{r_i-1}(\alpha_i)\}$ 下的表示矩阵就是友阵 $C(d_i(\lambda))$, 其中 $r_i = \dim C(\varphi, \alpha_i)$. ♡

 **笔记** 线性变换 φ 的有理标准型诱导的 V 的上述循环子空间直和分解 (1.15) 就是有理标准型的几何意义.

证明 由定理 1.6 可知, 存在 V 的一组基 $\{e_1, e_2, \dots, e_n\}$, 使得 φ 在这组基下的表示矩阵为

$$C = \text{diag}\{C(d_1(\lambda)), C(d_2(\lambda)), \dots, C(d_k(\lambda))\}$$

其中 $\varphi|_{L(e_{i1}, \dots, e_{ir_i})}$ 的表示矩阵就是友阵 $C(d_i(\lambda))$, $i = 1, 2, \dots, k$. 再结合定理 1.10 的讨论可知, $L(e_{i1}, \dots, e_{ir_i})$ 就是一个循环子空间. 于是任取 $\alpha_i \in L(e_{i1}, \dots, e_{ir_i})$ 作为循环向量, 则

$$C(\varphi, \alpha_i) = L(e_{i1}, \dots, e_{ir_i}) = L(\alpha_i, \varphi(\alpha_i), \dots, \varphi^{r_i-1}(\alpha_i))$$

其中 $\dim C(\varphi, \alpha_i) = r_i$.

综上所述, 此时 V 存在一个循环子空间的直和分解:

$$V = C(\varphi, \alpha_1) \oplus C(\varphi, \alpha_2) \oplus \dots \oplus C(\varphi, \alpha_k)$$

使得 $\varphi|_{C(\varphi, \alpha_i)}$ 在基 $\{\alpha_i, \varphi(\alpha_i), \dots, \varphi^{r_i-1}(\alpha_i)\}$ 下的表示矩阵就是友阵 $C(d_i(\lambda))$, 其中 $r_i = \dim C(\varphi, \alpha_i)$. □

定理 1.12 (循环子空间的刻画)

设 φ 是数域 \mathbb{K} 上 n 维线性空间 V 上的线性变换, φ 的特征多项式和极小多项式分别为 $f(\lambda)$ 和 $m(\lambda)$, 证明以下 4 个结论等价:

- (1) φ 的行列式因子组或不变因子组为 $1, \dots, 1, f(\lambda)$;
- (2) φ 的初等因子组为 $P_1(\lambda)^{r_1}, P_2(\lambda)^{r_2}, \dots, P_k(\lambda)^{r_k}$, 其中 $P_i(\lambda)$ 是 \mathbb{K} 上互异的首一不可约多项式, $r_i \geq 1, 1 \leq i \leq k$;
- (3) φ 的极小多项式 $m(\lambda)$ 等于特征多项式 $f(\lambda)$;
- (4) V 是关于线性变换 φ 的循环空间.



证明 (1) \Leftrightarrow (2): 由不变因子和初等因子之间的相互转换即得.

(1) \Leftrightarrow (3): 由极小多项式等于最大的不变因子, 以及所有不变因子的乘积等于特征多项式即得.

(1) \Leftrightarrow (4): 若 V 是循环空间, 则由 **定理 1.10** 可知, φ 在某组基下的表示矩阵是友阵 $C(g(\lambda))$, 再由友阵的性质 (**引理 1.5**) 可知, φ 的行列式因子组和不变因子组均为 $1, \dots, 1, g(\lambda) = f(\lambda)$. 若 φ 的不变因子组为 $1, \dots, 1, f(\lambda)$, 则由有理标准型的几何意义 (**定理 1.11**) 可知, V 是循环空间. \square

1.5 初等因子

定义 1.9 (初等因子)

设 $d_1(\lambda), d_2(\lambda), \dots, d_k(\lambda)$ 是数域 \mathbb{K} 上矩阵 A 的非常数不变因子, 由于 $d_i(\lambda) \mid d_{i+1}(\lambda)$, 因此可以在 \mathbb{K} 上把 $d_i(\lambda)$ 分解成不可约因式之积:

$$\begin{aligned} d_1(\lambda) &= p_1(\lambda)^{e_{11}} p_2(\lambda)^{e_{12}} \cdots p_t(\lambda)^{e_{1t}}, \\ d_2(\lambda) &= p_1(\lambda)^{e_{21}} p_2(\lambda)^{e_{22}} \cdots p_t(\lambda)^{e_{2t}}, \\ &\dots\dots\dots \\ d_k(\lambda) &= p_1(\lambda)^{e_{k1}} p_2(\lambda)^{e_{k2}} \cdots p_t(\lambda)^{e_{kt}}, \end{aligned} \tag{1.16}$$

其中 e_{ij} 是非负整数 (注意 e_{ij} 可以为零!), 并且

$$e_{1j} \leq e_{2j} \leq \cdots \leq e_{kj}, \quad j = 1, 2, \dots, t.$$

若 (1.16) 式中的 $e_{ij} > 0$, 则称 $p_j(\lambda)^{e_{ij}}$ 为 A 的一个**初等因子**, A 的全体初等因子称为 A 的**初等因子组**.

**命题 1.3**

矩阵 A 的初等因子组与不变因子组相互唯一确定.



证明 由因式分解的唯一性可知 A 的初等因子被 A 的不变因子唯一确定.

反过来, 若给定一组初等因子 $p_j(\lambda)^{e_{ij}}$, 适当增加一些 1 (表示为 $p_j(\lambda)^{e_{ij}}$, 其中 $e_{ij} = 0$), 则可将这组初等因子按不可约因式的降幂排列如下:

$$\begin{aligned} &p_1(\lambda)^{e_{k1}}, p_1(\lambda)^{e_{k-1,1}}, \dots, p_1(\lambda)^{e_{11}}, \\ &p_2(\lambda)^{e_{k2}}, p_2(\lambda)^{e_{k-1,2}}, \dots, p_2(\lambda)^{e_{12}}, \\ &\dots\dots\dots \\ &p_t(\lambda)^{e_{kt}}, p_t(\lambda)^{e_{k-1,t}}, \dots, p_t(\lambda)^{e_{1t}}, \end{aligned} \tag{1.17}$$

令

$$\begin{aligned} d_k(\lambda) &= p_1(\lambda)^{e_{k1}} p_2(\lambda)^{e_{k2}} \cdots p_t(\lambda)^{e_{kt}}, \\ d_{k-1}(\lambda) &= p_1(\lambda)^{e_{k-1,1}} p_2(\lambda)^{e_{k-1,2}} \cdots p_t(\lambda)^{e_{k-1,t}}, \end{aligned}$$

.....

$$d_1(\lambda) = p_1(\lambda)^{e_{11}} p_2(\lambda)^{e_{12}} \cdots p_t(\lambda)^{e_{1t}},$$

则 $d_i(\lambda) \mid d_{i+1}(\lambda)$ ($i = 1, \cdots, k-1$), 且 $d_1(\lambda), \cdots, d_k(\lambda)$ 的初等因子组就如 (1.17) 所示. 因此, 给定 A 的初等因子组, 我们可唯一地确定 A 的不变因子组. 这一事实表明, A 的不变因子组与初等因子组在讨论矩阵相似关系中的作用是相同的. \square

定理 1.13

数域 \mathbb{K} 上的两个矩阵 A 与 B 相似的充分必要条件是它们有相同的初等因子组, 即矩阵的初等因子组是矩阵相似关系的全系不变量.



证明 由定理 1.5 可知, 矩阵 A 和 B 相似等价于 A 和 B 有相同的不变因子. 又由命题 1.3 可知, A 和 B 有相同的不变因子等价于它们有相同的初等因子组. 故矩阵 A 与 B 相似的充分必要条件是它们有相同的初等因子组. \square

例题 1.4 设 9 阶矩阵 A 的不变因子组为

$$1, \cdots, 1, (\lambda-1)(\lambda^2+1), (\lambda-1)^2(\lambda^2+1)(\lambda^2-2),$$

试分别在有理数域、实数域和复数域上求 A 的初等因子组.

解 A 在有理数域上的初等因子组为

$$\lambda-1, (\lambda-1)^2, \lambda^2+1, \lambda^2+1, \lambda^2-2.$$

A 在实数域上的初等因子组为

$$\lambda-1, (\lambda-1)^2, \lambda^2+1, \lambda^2+1, \lambda+\sqrt{2}, \lambda-\sqrt{2}.$$

A 在复数域上的初等因子组为

$$\lambda-1, (\lambda-1)^2, \lambda+i, \lambda+i, \lambda-i, \lambda-i, \lambda+\sqrt{2}, \lambda-\sqrt{2}.$$

 \square

例题 1.5 设 A 是一个 10 阶矩阵, 它的初等因子组为

$$\lambda-1, \lambda-1, (\lambda-1)^2, (\lambda+1)^2, (\lambda+1)^3, \lambda-2.$$

求 A 的不变因子组.

解 将上述多项式按不可约因式的降幂排列:

$$\begin{array}{lll} (\lambda-1)^2, & \lambda-1, & \lambda-1; \\ (\lambda+1)^3, & (\lambda+1)^2, & 1; \\ \lambda-2, & 1, & 1. \end{array}$$

于是

$$d_3(\lambda) = (\lambda-1)^2(\lambda+1)^3(\lambda-2), \quad d_2(\lambda) = (\lambda-1)(\lambda+1)^2, \quad d_1(\lambda) = \lambda-1.$$

从而 A 的不变因子组为

$$1, \cdots, 1, \lambda-1, (\lambda-1)(\lambda+1)^2, (\lambda-1)^2(\lambda+1)^3(\lambda-2),$$

其中有 7 个 1. \square

1.6 Jordan 标准型

引理 1.7

r 阶矩阵

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

的初等因子组为 $(\lambda - \lambda_0)^r$.



证明 显然 J 的特征多项式为 $(\lambda - \lambda_0)^r$. 对任一小于 r 的正整数 k , $\lambda I - J$ 总有一个 k 阶子式, 其值等于 $(-1)^k$, 因此 J 的行列式因子为

$$1, \dots, 1, (\lambda - \lambda_0)^r. \quad (1.18)$$

(1.18) 式也是 J 的不变因子组, 故 J 的初等因子组只有一个多项式 $(\lambda - \lambda_0)^r$. \square

引理 1.8

设特征矩阵 $\lambda I - A$ 经过初等变换化为下列对角阵:

$$\begin{pmatrix} f_1(\lambda) & & & \\ & f_2(\lambda) & & \\ & & \ddots & \\ & & & f_n(\lambda) \end{pmatrix}, \quad (1.19)$$

其中 $f_i(\lambda)$ ($i = 1, \dots, n$) 为非零首一多项式. 将 $f_i(\lambda)$ 作不可约分解, 若 $(\lambda - \lambda_0)^k$ 能整除 $f_i(\lambda)$, 但 $(\lambda - \lambda_0)^{k+1}$ 不能整除 $f_i(\lambda)$, 就称 $(\lambda - \lambda_0)^k$ 是 $f_i(\lambda)$ 的一个**准素因子**, 所有 $f_i(\lambda)$ 的准素因子称为 A 的**准素因子组**, 则矩阵 A 的初等因子组等于所有 $f_i(\lambda)$ 的准素因子组.



注 这个引理给出了求矩阵初等因子组的另外一个方法, 它可以不必先求不变因子组而直接用初等变换把特征矩阵化为对角阵, 再分解主对角线上的多项式即可. 另外, 这个引理的结论及其证明在一般的数域 \mathbb{K} 上也成立.

证明 第一步, 先证明下列事实:

若 $f_i(\lambda), f_j(\lambda)$ ($i \neq j$) 的最大公因式和最小公倍式分别为 $g(\lambda), h(\lambda)$, 则

$$\text{diag}\{f_1(\lambda), \dots, f_i(\lambda), \dots, f_j(\lambda), \dots, f_n(\lambda)\}$$

经过初等变换可以变为

$$\text{diag}\{f_1(\lambda), \dots, g(\lambda), \dots, h(\lambda), \dots, f_n(\lambda)\},$$

且这两个对角阵具有相同的准素因子组.

不失一般性, 令 $i = 1, j = 2$. 因为 $(f_1(\lambda), f_2(\lambda)) = g(\lambda)$, 所以存在 $u(\lambda), v(\lambda)$, 使

$$f_1(\lambda)u(\lambda) + f_2(\lambda)v(\lambda) = g(\lambda).$$

又令 $f_1(\lambda) = g(\lambda)q(\lambda), f_2(\lambda) = g(\lambda)q'(\lambda)$. 则 $h(\lambda) = g(\lambda)q(\lambda)q'(\lambda) = f_2(\lambda)q(\lambda)$. 对 (1.19) 式作下列初等变换:

$$\begin{pmatrix} f_1(\lambda) & & & \\ & f_2(\lambda) & & \\ & & \ddots & \\ & & & f_n(\lambda) \end{pmatrix} \xrightarrow{u(\lambda) \cdot r_1 + r_2} \begin{pmatrix} f_1(\lambda) & 0 & & \\ f_1(\lambda)u(\lambda) & f_2(\lambda) & & \\ & & \ddots & \\ & & & f_n(\lambda) \end{pmatrix} \xrightarrow{v(\lambda)j_2 + j_1}$$

$$\begin{pmatrix} f_1(\lambda) & & & \\ g(\lambda) & f_2(\lambda) & & \\ & & \ddots & \\ & & & f_n(\lambda) \end{pmatrix} \xrightarrow{-q(\lambda)r_2+r_1} \begin{pmatrix} 0 & -h(\lambda) & & \\ g(\lambda) & f_2(\lambda) & & \\ & & \ddots & \\ & & & f_n(\lambda) \end{pmatrix} \xrightarrow[r_2 \leftrightarrow r_1]{(-1) \cdot r_2} \begin{pmatrix} g(\lambda) & f_2(\lambda) & & \\ & h(\lambda) & & \\ & & \ddots & \\ & & & f_n(\lambda) \end{pmatrix}$$

$$\begin{pmatrix} g(\lambda) & f_2(\lambda) & & \\ & h(\lambda) & & \\ & & \ddots & \\ & & & f_n(\lambda) \end{pmatrix} \xrightarrow{-q'(\lambda)j_1+j_2} \begin{pmatrix} g(\lambda) & & & \\ & h(\lambda) & & \\ & & \ddots & \\ & & & f_n(\lambda) \end{pmatrix}.$$

现来考察 $g(\lambda)$ 与 $h(\lambda)$ 的准素因子. 将 $f_1(\lambda), f_2(\lambda)$ 作标准因式分解, 其分解式不妨设为

$$\begin{aligned} f_1(\lambda) &= (\lambda - \lambda_1)^{c_1} (\lambda - \lambda_2)^{c_2} \cdots (\lambda - \lambda_t)^{c_t}, \\ f_2(\lambda) &= (\lambda - \lambda_1)^{d_1} (\lambda - \lambda_2)^{d_2} \cdots (\lambda - \lambda_t)^{d_t}, \end{aligned}$$

其中 c_i, d_i 为非负整数. 令

$$e_i = \max\{c_i, d_i\}, \quad k_i = \min\{c_i, d_i\},$$

则

$$\begin{aligned} g(\lambda) &= (\lambda - \lambda_1)^{k_1} (\lambda - \lambda_2)^{k_2} \cdots (\lambda - \lambda_t)^{k_t}, \\ h(\lambda) &= (\lambda - \lambda_1)^{e_1} (\lambda - \lambda_2)^{e_2} \cdots (\lambda - \lambda_t)^{e_t}. \end{aligned}$$

不难看出 $g(\lambda), h(\lambda)$ 的准素因子组与 $f_1(\lambda), f_2(\lambda)$ 的准素因子组相同.

第二步证明 (1.19) 式所示矩阵的法式可通过上述变换得到.

先将第 (1, 1) 位置的元素依次和第 (2, 2) 位置, \cdots , 第 (n, n) 位置的元素进行上述变换, 此时第 (1, 1) 元素的所有一次因式的幂都是最小的; 再将第 (2, 2) 位置的元素依次和第 (3, 3) 位置, \cdots , 第 (n, n) 位置的元素进行上述变换; \cdots ; 最后将第 $(n-1, n-1)$ 位置的元素和第 (n, n) 位置的元素进行上述变换. 可以看出, 最后得到的对角阵就是 (1.19) 式所示矩阵的法式. 注意到在每一次变换的过程中, 准素因子组都保持不变, 这就证明了结论. \square

例题 1.6 设 $\lambda I - A$ 经过初等变换后化为下列对角阵:

$$\begin{pmatrix} 1 & & & \\ & (\lambda - 1)^2(\lambda + 2) & & \\ & & \lambda + 2 & \\ & & & 1 \\ & & & & \lambda - 1 \end{pmatrix},$$

求 A 的初等因子组.

解 由引理 1.8 知, A 的初等因子组为 $\lambda - 1, (\lambda - 1)^2, \lambda + 2, \lambda + 2$. \square

引理 1.9

设 J 是分块对角阵:

$$\begin{pmatrix} J_1 & & \\ & J_2 & \\ & & \ddots \\ & & & J_k \end{pmatrix},$$

其中每个 J_i 都是形如引理 1.7 中的矩阵, J_i 的初等因子组为 $(\lambda - \lambda_i)^{r_i}$, 则 J 的初等因子组为

$$(\lambda - \lambda_1)^{r_1}, (\lambda - \lambda_2)^{r_2}, \cdots, (\lambda - \lambda_k)^{r_k}.$$



证明 $\lambda I - J$ 是一个分块对角 λ -矩阵. 由于对分块对角阵中某一块施行初等变换时其余各块保持不变, 故由引理

1.7及命题 1.2知, $\lambda I - J$ 相抵于下列分块对角阵:

$$H = \begin{pmatrix} H_1 & & \\ & H_2 & \\ & & \ddots \\ & & & H_k \end{pmatrix},$$

其中 $H_i = \text{diag}\{1, \dots, 1, (\lambda - \lambda_i)^{r_i}\}$. 再由引理 1.8 即得结论. □

定理 1.14 (Jordan 标准型)

设 A 是复数域上的矩阵且 A 的初等因子组为

$$(\lambda - \lambda_1)^{r_1}, (\lambda - \lambda_2)^{r_2}, \dots, (\lambda - \lambda_k)^{r_k},$$

则 A 相似于分块对角阵:

$$J = \begin{pmatrix} J_1(\lambda_1) & & \\ & J_2(\lambda_2) & \\ & & \ddots \\ & & & J_k(\lambda_k) \end{pmatrix}, \quad (1.20)$$

其中 J_i 为 r_i 阶矩阵, 且

$$J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & 1 & \\ & & \ddots & \ddots \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix}.$$

(1.20) 式中的矩阵 J 称为 A 的 **Jordan 标准型**, 每个 $J_i(\lambda_i)$ 称为 A 的一个 **Jordan 块**. ♥

注 由引理 1.7 可以看出, 若交换任意两个 Jordan 块的位置, 得到的矩阵与原来的矩阵仍有相同的初等因子组, 它们仍相似. 因此矩阵 A 的 Jordan 标准型中 Jordan 块的排列可以是任意的. 但是, 由于每个初等因子唯一确定了一个 Jordan 块, 故若不计 Jordan 块的排列次序, 则矩阵的 Jordan 标准型是唯一确定的.

证明 由定理 1.13 知, A 与 J 有相同的初等因子组, 因此 A 与 J 相似. □

命题 1.4 (Jordan 块的性质)

- (1) $J_n(\lambda_0)$ 的初等因子组为 $(\lambda - \lambda_0)^n$.
- (2) 设 $J_0 = J_n(0)$, 则 J_0 是基础幂零阵, 且 $J_0^n = 0$.
- (3) 设

$$J = \begin{pmatrix} J_1(\lambda_1) & & \\ & J_2(\lambda_2) & \\ & & \ddots \\ & & & J_r(\lambda_r) \end{pmatrix},$$

则 $J_i^k(\lambda_i)$ 与 $J_i(\lambda_i)$ 相似 ($i = 1, 2, \dots, r$), 进而, J^k 与 J 也相似.

- (4) 设

$$J = \begin{pmatrix} J_1(\lambda_1) & & \\ & J_2(\lambda_2) & \\ & & \ddots \\ & & & J_r(\lambda_r) \end{pmatrix},$$

其中 $\lambda_i \neq 0, i = 1, 2, \dots, r$, 则 $J_i^{-1}(\lambda_i)$ 与 $J_i(\lambda_i)$ 相似 ($i = 1, 2, \dots, r$), 进而, J^{-1} 与 J 也相似.

- (5) 设 $J = J_n(\lambda_0)$ 是特征值为 λ_0 的 n 阶 Jordan 块, 则和 J 乘法可交换的 n 阶矩阵必可表示为 J 的次数不超过 $n-1$ 的多项式.

证明

- (1) 由引理 1.7 即得.
 (2) 由 Jordan 块的定义可直接得到

$$J_0 = J_n(0) = \begin{pmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & \ddots & \ddots \\ & & & \ddots & 1 \\ & & & & 0 \end{pmatrix}.$$

从而结论显然成立.

- (3) 不妨设 $J = J_n(1)$ 只有一个 Jordan 块, 则 $J = I_n + J_0$, 其中 $J_0 = J_n(0)$ 是特征值为 0 的 n 阶 Jordan 块. 注意到

$$J^k = (I_n + J_0)^k = I_n + C_k^1 J_0 + C_k^2 J_0^2 + \dots + J_0^k.$$

故 J^k 是一个上三角矩阵, 其主对角线上的元素全为 1, 上次对角线上的元素全为 k , 从而它的特征多项式为 $(\lambda - 1)^n$. 为了确定它的极小多项式, 我们可进行如下计算:

$$(J^k - I_n)^{n-1} = (C_k^1 J_0 + C_k^2 J_0^2 + \dots + J_0^k)^{n-1} = k^{n-1} J_0^{n-1} \neq O.$$

于是 J^k 的极小多项式为 $(\lambda - 1)^n$, 其不变因子组为 $1, \dots, 1, (\lambda - 1)^n$. 因此 J^k 与 J 有相同的不变因子, 从而 J^k 与 J 相似.

- (4) 不妨设 $J = J_n(\lambda_0)$ 只有一个 Jordan 块, 则 $J = \lambda_0 I_n + J_0$, 其中 $\lambda_0 \neq 0, J_0 = J_n(0)$ 是特征值为 0 的 n 阶 Jordan 块. 注意到

$$\lambda_0^n I_n = (\lambda_0 I_n)^n - (-J_0)^n = (\lambda_0 I_n + J_0)(\lambda_0^{n-1} I_n - \lambda_0^{n-2} J_0 + \dots + (-1)^{n-1} J_0^{n-1}).$$

以及 $\lambda_0^{-1} = \lambda_0$, 故可得

$$J^{-1} = (\lambda_0 I_n + J_0)^{-1} = \lambda_0 I_n - \lambda_0^2 J_0 + \dots + (-1)^{n-1} \lambda_0^n J_0^{n-1}.$$

因此 J^{-1} 是一个上三角矩阵, 其主对角线上的元素全为 λ_0 , 上次对角线上的元素全为 $-\lambda_0^2$, 从而它的特征多项式为 $(\lambda - \lambda_0)^n$. 为了确定它的极小多项式, 我们可进行如下计算:

$$(J^{-1} - \lambda_0 I)^{n-1} = (-\lambda_0^2 J_0 + \dots + (-1)^{n-1} \lambda_0^n J_0^{n-1})^{n-1} = \lambda_0 J_0^{n-1} \neq O.$$

于是 J^{-1} 的极小多项式为 $(\lambda - \lambda_0)^n$, 其不变因子组为 $1, \dots, 1, (\lambda - \lambda_0)^n$. 因此 J^{-1} 与 J 有相同的不变因子组, 从而 J^{-1} 与 J 相似.

- (5) 根据 Jordan 标准型的几何意义, $\mathbb{C}^n = C(J - \lambda_0 I_n, e_n)$ 是关于线性变换 $J - \lambda_0 I_n$ 的循环空间, 循环向量是标准单位列向量中的最后一个 $e_n = (0, \dots, 0, 1)'$, 再由定理 1.24 即得结论. 当然也可以通过代数方法直接进行证明. 设 A 和 J 可交换, 注意到 $J = \lambda_0 I_n + J_0$, 其中 $J_0 = J_n(0)$ 是特征值为零的 Jordan 块, 故 A, J 乘法可交换当且仅当 A, J_0 乘法可交换. 经计算得到 A 必为下列形状的上三角矩阵:

$$A = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ & a_1 & \ddots & \vdots \\ & & \ddots & a_2 \\ & & & a_1 \end{pmatrix}$$

于是

$$A = a_1 I_n + a_2 J_0 + \dots + a_n J_0^{n-1} = a_1 I_n + a_2 (J - \lambda_0 I_n) + \dots + a_n (J - \lambda_0 I_n)^{n-1}$$

可表示为 J 的次数不超过 $n-1$ 的多项式.

□

定理 1.15

设 φ 是复数域上线性空间 V 上的线性变换, 则必存在 V 的一组基, 使得 φ 在这组基下的表示矩阵为 (1.20) 式所示的 Jordan 标准型.

♡

证明

□

推论 1.7

设 A 是 n 阶复矩阵, 则下列结论等价:

- (1) A 可对角化;
- (2) A 的极小多项式无重根;
- (3) A 的初等因子都是一次多项式.

♡

证明 (1) \Rightarrow (2): 由可对角化的判定条件 (5) 的结论即得.

(2) \Rightarrow (3): 设 A 的极小多项式 $m(\lambda)$ 无重根. 由于 $m(\lambda)$ 是 A 的最后一个不变因子, 故 A 的所有不变因子都无重根, 从而 A 的初等因子都是一次多项式.

(3) \Rightarrow (1): 设 A 的初等因子组为 $\lambda - \lambda_1, \lambda - \lambda_2, \dots, \lambda - \lambda_n$, 则由定理 1.14 知, A 相似于对角阵 $\text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$, 即 A 可对角化.

□

推论 1.8

设 φ 是复线性空间 V 上的线性变换, 则 φ 可对角化当且仅当 φ 的极小多项式无重根, 当且仅当 φ 的初等因子都是一次多项式.

♡

证明

□

推论 1.9

设 φ 是复线性空间 V 上的线性变换, V_0 是 φ 的不变子空间. 若 φ 可对角化, 则 φ 在 V_0 上的限制也可对角化.

♡

证明 设 $\varphi, \varphi|_{V_0}$ 的极小多项式分别为 $g(\lambda), h(\lambda)$, 则由推论 1.8 知, $g(\lambda)$ 无重根. 又 $g(\varphi|_{V_0}) = g(\varphi)|_{V_0} = \mathbf{0}$, 故 $h(\lambda) \mid g(\lambda)$, 于是 $h(\lambda)$ 也无重根, 再次由推论 1.8 知, $\varphi|_{V_0}$ 可对角化.

□

推论 1.10

设 φ 是复线性空间 V 上的线性变换, 且 $V = V_1 \oplus V_2 \oplus \dots \oplus V_k$, 其中每个 V_i 都是 φ 的不变子空间, 则 φ 可对角化的充分必要条件是 φ 在每个 V_i 上的限制都可对角化.

♡

证明 必要性由推论 1.9 即得, 下证充分性. 若 φ 在每个 V_i 上的限制都可对角化, 则由定义存在 V_i 的一组基, 使得 $\varphi|_{V_i}$ 在这组基下的表示矩阵是对角阵. 再由定理 ?? 知 V_i 的一组基可以拼成 V 的一组基, 因此 φ 在这组基下的表示矩阵是对角阵, 即 φ 可对角化.

□

推论 1.11

设 A 是数域 \mathbb{K} 上的矩阵, 如果 A 的特征值全在 \mathbb{K} 中, 则 A 在 \mathbb{K} 上相似于其 Jordan 标准型.

♡

证明 由于 A 的特征值全在 \mathbb{K} 中, 故 A 的 Jordan 标准型 J 实际上是 \mathbb{K} 上的矩阵. 因为 A 在复数域上相似于 J , 由推

论 1.10 知, A 在 \mathbb{R} 上也相似于 J . □

例题 1.7 设 A 是 7 阶矩阵, 其初等因子组为

$$\lambda - 1, (\lambda - 1)^3, (\lambda + 1)^2, \lambda - 2,$$

求其 Jordan 标准型.

解 A 的 Jordan 标准型为

$$J = \begin{pmatrix} 1 & & & & & & \\ & 1 & 1 & 0 & & & \\ & 0 & 1 & 1 & & & \\ & 0 & 0 & 1 & & & \\ & & & & -1 & 1 & \\ & & & & 0 & -1 & \\ & & & & & & 2 \end{pmatrix},$$

J 含有 4 个 Jordan 块. □

例题 1.8 设复数域上的四维线性空间 V 上的线性变换 φ 在一组基 $\{e_1, e_2, e_3, e_4\}$ 下的表示矩阵为

$$A = \begin{pmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 6 & 1 & 2 & 1 \\ -14 & -5 & -1 & 0 \end{pmatrix},$$

求 V 的一组基, 使 φ 在这组基下的表示矩阵为 Jordan 标准型, 并求出从原来的基到新基的过渡矩阵.

解 用初等变换把 $\lambda I - A$ 化为对角 λ -矩阵并求出它的初等因子组为

$$(\lambda - 1)^2, (\lambda - 1)^2.$$

因此, A 的 Jordan 标准型为

$$J = \begin{pmatrix} 1 & 1 & & \\ 0 & 1 & & \\ & & 1 & 1 \\ & & 0 & 1 \end{pmatrix}.$$

设矩阵 P 是从 $\{e_1, e_2, e_3, e_4\}$ 到新基的过渡矩阵, 则

$$P^{-1}AP = J,$$

此即

$$AP = PJ. \tag{1.21}$$

设 $P = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$, 其中 α_i 是四维列向量, 代入(1.21)式得

$$(A\alpha_1, A\alpha_2, A\alpha_3, A\alpha_4) = (\alpha_1, \alpha_2, \alpha_3, \alpha_4) \begin{pmatrix} 1 & 1 & & \\ 0 & 1 & & \\ & & 1 & 1 \\ & & 0 & 1 \end{pmatrix},$$

化成方程组为

$$(A - I)\alpha_1 = \mathbf{0},$$

$$(A - I)\alpha_2 = \alpha_1,$$

$$(A - I)\alpha_3 = \mathbf{0},$$

$$(A - I)\alpha_4 = \alpha_3.$$

由于 α_1, α_3 都是 A 的属于特征值 1 的特征向量, 故 α_2, α_4 称为属于特征值 1 的广义特征向量. 我们可取方程组

$(A - I)x = \mathbf{0}$ 的两个线性无关的解分别作为 α_1, α_3 (注意不能取线性相关的两个解, 因为 P 是非异阵), 然后再分别求出 α_2, α_4 (注意诸 α_i 的解可能不唯一, 只需取比较简单的一组解) 即可. 经计算可得

$$\alpha_1 = \begin{pmatrix} 1 \\ -2 \\ 1 \\ -5 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \alpha_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \quad \alpha_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

于是

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ -5 & 0 & -1 & 1 \end{pmatrix}.$$

因此新基为

$$\{e_1 - 2e_2 + e_3 - 5e_4, e_2, e_3 - e_4, e_4\}.$$

□

1.7 Jordan 标准型的进一步讨论和应用

定理 1.16

线性变换 φ 的特征值 λ_1 的度数等于 φ 的 Jordan 标准型中属于特征值 λ_1 的 Jordan 块的个数, λ_1 的重数等于所有属于特征值 λ_1 的 Jordan 块的阶数之和.

♡

证明 设 V 是 n 维复线性空间, φ 是 V 上的线性变换. 设 φ 的初等因子组为

$$(\lambda - \lambda_1)^{r_1}, (\lambda - \lambda_2)^{r_2}, \dots, (\lambda - \lambda_k)^{r_k}, \quad (1.22)$$

定理 1.15 告诉我们, 存在 V 的一组基 $\{e_{11}, e_{12}, \dots, e_{1r_1}; e_{21}, e_{22}, \dots, e_{2r_2}; \dots; e_{k1}, e_{k2}, \dots, e_{kr_k}\}$, 使得 φ 在这组基下的表示矩阵为

$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{pmatrix}.$$

上式中每个 J_i 是相应于初等因子 $(\lambda - \lambda_i)^{r_i}$ 的 Jordan 块, 其阶正好为 r_i . 令 V_i 是由基向量 $e_{i1}, e_{i2}, \dots, e_{ir_i}$ 生成的子空间, 则

$$\begin{aligned} \varphi(e_{i1}) &= \lambda_i e_{i1}, \\ \varphi(e_{i2}) &= e_{i1} + \lambda_i e_{i2}, \\ &\dots\dots\dots \\ \varphi(e_{ir_i}) &= e_{i,r_i-1} + \lambda_i e_{ir_i}. \end{aligned} \quad (1.23)$$

这表明 $\varphi(V_i) \subseteq V_i$, 即 $V_i (i = 1, 2, \dots, k)$ 是 φ 的不变子空间. 显然我们有

$$V = V_1 \oplus V_2 \oplus \dots \oplus V_k.$$

线性变换 φ 限制在 V_1 上 (仍记为 φ) 便成为 V_1 上的线性变换. 这个线性变换在基 $\{e_{11}, e_{12}, \dots, e_{1r_1}\}$ 下的表

示矩阵为

$$J_1 = \begin{pmatrix} \lambda_1 & 1 & & & \\ & \lambda_1 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_1 \end{pmatrix}.$$

注意到 J_1 的特征值全为 λ_1 , 并且 $\lambda_1 I - J_1$ 的秩等于 $r_1 - 1$, 故 J_1 只有一个线性无关的特征向量, 不妨选为 e_{11} . 显然 e_{11} 也是 φ 作为 V 上线性变换关于特征值 λ_1 的特征向量. 不失一般性, 不妨设在 φ 的初等因子组即(1.22)式中

$$\lambda_1 = \lambda_2 = \cdots = \lambda_s, \quad \lambda_i \neq \lambda_1 (i = s+1, \cdots, k),$$

则 J_1, \cdots, J_s 都以 λ_1 为特征值, 且相应于每一块有且只有一个线性无关的特征向量. 相应的特征向量可取为

$$e_{11}, e_{21}, \cdots, e_{s1}, \quad (1.24)$$

显然这是 s 个线性无关的特征向量. 如果 $\lambda_i \neq \lambda_1$, 则容易看出 $r(\lambda_1 I - J_i) = r_i$, 于是

$$r(\lambda_1 I - J) = \sum_{i=1}^k r(\lambda_1 I - J_i) = (r_1 - 1) + \cdots + (r_s - 1) + r_{s+1} + \cdots + r_k = n - s.$$

因此 φ 关于特征值 λ_1 的特征子空间 V_{λ_1} 的维数等于 $n - r(\lambda_1 I - J) = s$, 从而特征子空间 V_{λ_1} 以(1.24)式中的向量为一组基. 又 λ_1 是 φ 的 $r_1 + r_2 + \cdots + r_s$ 重特征值, 因此 λ_1 的重数与度数之差等于

$$(r_1 + r_2 + \cdots + r_s) - s.$$

□

定义 1.10 (根子空间)

设 λ_0 是 n 维复线性空间 V 上线性变换 φ 的特征值, 则

$$R(\lambda_0) = \{v \in V \mid (\varphi - \lambda_0 I)^n(v) = 0\}$$

构成了 V 的一个子空间, 称为属于特征值 λ_0 的根子空间.

♣

定理 1.17

设 φ 是 n 维复线性空间 V 上的线性变换.

(1) 若 φ 的初等因子组为

$$(\lambda - \lambda_1)^{r_1}, (\lambda - \lambda_2)^{r_2}, \cdots, (\lambda - \lambda_k)^{r_k},$$

则 V 可分解为 k 个不变子空间的直和:

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k, \quad (1.25)$$

其中 V_i 是维数等于 r_i 的关于 $\varphi - \lambda_i I$ 的循环子空间;

(2) 若 $\lambda_1, \cdots, \lambda_s$ 是 φ 的全体不同特征值, 则 V 可分解为 s 个不变子空间的直和:

$$V = R(\lambda_1) \oplus R(\lambda_2) \oplus \cdots \oplus R(\lambda_s),$$

其中 $R(\lambda_i)$ 是 λ_i 的根子空间, $R(\lambda_i)$ 的维数等于 λ_i 的重数, 且每个 $R(\lambda_i)$ 又可分解为(1.25)式中若干个 V_j 的直和.

♡

证明 在定理 1.16 的证明的基础上, 现在再来看 J_1 所对应的子空间 V_1 , 由(1.23)式中诸等式可知

$$(\varphi - \lambda_1 I)(e_{1r_1}) = e_{1,r_1-1}, \cdots, (\varphi - \lambda_1 I)(e_{12}) = e_{11}, (\varphi - \lambda_1 I)(e_{11}) = 0,$$

因此, 若记 $\alpha = e_{1r_1}, \psi = \varphi - \lambda_1 I$, 则

$$\psi(\alpha) = e_{1,r_1-1}, \psi^2(\alpha) = e_{1,r_1-2}, \cdots, \psi^{r_1-1}(\alpha) = e_{11}, \psi^{r_1}(\alpha) = 0.$$

也就是说

$$\{\alpha, \psi(\alpha), \psi^2(\alpha), \dots, \psi^{r_1-1}(\alpha)\}$$

构成了 V_1 的一组基.

上面的事实说明, 每个 Jordan 块 J_i 对应的子空间 V_i 是一个循环子空间. 把属于同一个特征值, 比如属于 λ_1 的所有循环子空间加起来构成 V 的一个子空间:

$$R(\lambda_1) = V_1 \oplus \dots \oplus V_s.$$

若 $v \in R(\lambda_1)$, 则不难算出 $(\varphi - \lambda_1 I)^s(v) = 0$, 其中

$$s = \dim R(\lambda_1) = r_1 + \dots + r_s.$$

事实上, 我们可以证明

$$R(\lambda_1) = \{v \in V \mid (\varphi - \lambda_1 I)^n(v) = 0\}. \quad (1.26)$$

为证明 (1.26) 式成立, 设 $U = \{v \in V \mid (\varphi - \lambda_1 I)^n(v) = 0\}$, 则由上面的分析知道, $R(\lambda_1) \subseteq U$. 另一方面, 任取 $v \in U$, 设 $v = v_1 + v_2$, 其中 $v_1 \in R(\lambda_1), v_2 \in V_{s+1} \oplus \dots \oplus V_k$. 因为 $(\lambda - \lambda_1)^n$ 与 $(\lambda - \lambda_{s+1})^n \dots (\lambda - \lambda_k)^n$ 互素, 故存在多项式 $p(\lambda), q(\lambda)$, 使

$$(\lambda - \lambda_1)^n p(\lambda) + (\lambda - \lambda_{s+1})^n \dots (\lambda - \lambda_k)^n q(\lambda) = 1.$$

将 $\lambda = \varphi$ 代入上式并作用在 v 上可得

$$\begin{aligned} v &= p(\varphi)(\varphi - \lambda_1 I)^n(v) + q(\varphi)(\varphi - \lambda_{s+1} I)^n \dots (\varphi - \lambda_k I)^n(v) \\ &= q(\varphi)(\varphi - \lambda_{s+1} I)^n \dots (\varphi - \lambda_k I)^n(v_1) + q(\varphi)(\varphi - \lambda_{s+1} I)^n \dots (\varphi - \lambda_k I)^n(v_2) \\ &= q(\varphi)(\varphi - \lambda_{s+1} I)^n \dots (\varphi - \lambda_k I)^n(v_1) \in R(\lambda_1). \end{aligned}$$

这就证明了 (1.26) 式.

上面的结果表明: 特征值 λ_0 的根子空间可表示为若干个循环子空间的直和, 每个循环子空间对应于一个 Jordan 块. 虽然我们前面的讨论是对特征值 λ_1 进行的, 其实对任一特征值 λ_i 均适用. \square

命题 1.5

证明: 复数域上的方阵 A 必可分解为两个对称阵的乘积.

证明 设 P 是非异阵且使 $P^{-1}AP = J$ 为 A 的 Jordan 标准型, 于是 $A = PJP^{-1}$. 设 J_i 是 J 的第 i 个 Jordan 块, 则

$$J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & 1 & \\ & & \ddots & \ddots \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix} = \begin{pmatrix} & & & 1 & \lambda_i \\ & & & & \lambda_i \\ & & \ddots & \ddots & \\ & & & \ddots & \\ 1 & & & & \\ \lambda_i & & & & \end{pmatrix} \begin{pmatrix} & & & & 1 \\ & & & 1 & \\ & & \ddots & & \\ & & & \ddots & \\ 1 & & & & \end{pmatrix},$$

即 J_i 可分解为两个对称阵之积. 因此 J 也可以分解为两个对称阵之积, 记为 S_1, S_2 , 于是

$$A = PJP^{-1} = PS_1S_2P^{-1} = (PS_1P')(P^{-1})'S_2P^{-1}.$$

显然 PS_1P' 和 $(P^{-1})'S_2P^{-1}$ 都是对称矩阵, 故 A 必可分解为两个对称阵的乘积. \square

例题 1.9 已知

$$A = \begin{pmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 6 & 1 & 2 & 1 \\ -14 & -5 & -1 & 0 \end{pmatrix},$$

计算 A^k .

解 用初等变换把 $\lambda I - A$ 化为对角 λ -矩阵并求出它的初等因子组为

$$(\lambda - 1)^2, (\lambda - 1)^2.$$

因此, A 的 Jordan 标准型为

$$J = \begin{pmatrix} 1 & 1 & & \\ 0 & 1 & & \\ & & 1 & 1 \\ & & 0 & 1 \end{pmatrix}.$$

因为

$$P^{-1}A^kP = (P^{-1}AP)^k = J^k,$$

故先计算 J^k . 注意 J 是分块对角阵, 它的 k 次方等于将各对角块 k 次方, 因此

$$\begin{aligned} J^k &= \begin{pmatrix} 1 & 1 & & \\ 0 & 1 & & \\ & & 1 & 1 \\ & & 0 & 1 \end{pmatrix}^k = \begin{pmatrix} 1 & k & & \\ 0 & 1 & & \\ & & 1 & k \\ & & 0 & 1 \end{pmatrix}, \\ A^k &= PJ^kP^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ -5 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & k & & \\ 0 & 1 & & \\ & & 1 & k \\ & & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ -5 & 0 & -1 & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 2k+1 & k & 0 & 0 \\ -4k & -2k+1 & 0 & 0 \\ 6k & k & k+1 & k \\ -14k & -5k & -k & -k+1 \end{pmatrix}. \end{aligned}$$

□

定理 1.18 (Jordan-Chevalley 分解)

设 A 是 n 阶复矩阵, 则 A 可分解为 $A = B + C$, 其中 B, C 适合下面条件:

- (1) B 是一个可对角化矩阵;
- (2) C 是一个幂零阵;
- (3) $BC = CB$;
- (4) B, C 均可表示为 A 的多项式.

不仅如此, 上述满足条件 (1)(3) 的分解是唯一的 (即只要满足条件 (1)(3) 的分解就是唯一的). 进而, 上述满足条件 (1)(2)(3)(4) 的分解也是唯一的.

♡

证明 先对 A 的 Jordan 标准型 J 证明结论. 设 A 的全体不同特征值为 $\lambda_1, \lambda_2, \dots, \lambda_s$ 且

$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{pmatrix},$$

其中 J_i 是属于特征值 λ_i 的根子空间对应的块, 其阶设为 m_i . 显然对每个 i 均有 $J_i = M_i + N_i$, 其中 $M_i = \lambda_i I$ 是对

角阵, N_i 是幂零阵且 $M_i N_i = N_i M_i$. 令

$$M = \begin{pmatrix} M_1 & & \\ & M_2 & \\ & & \ddots \\ & & & M_s \end{pmatrix}, \quad N = \begin{pmatrix} N_1 & & \\ & N_2 & \\ & & \ddots \\ & & & N_s \end{pmatrix},$$

则 $J = M + N$, $MN = NM$, M 是对角阵, N 是幂零阵.

因为 $(J_i - \lambda_i I)^{m_i} = O$, 所以 J_i 适合多项式 $(\lambda - \lambda_i)^{m_i}$. 而 λ_i 互不相同, 因此多项式 $(\lambda - \lambda_1)^{m_1}, (\lambda - \lambda_2)^{m_2}, \dots, (\lambda - \lambda_s)^{m_s}$ 两两互素. 由中国剩余定理, 存在多项式 $g(\lambda)$ 满足条件

$$g(\lambda) = h_i(\lambda)(\lambda - \lambda_i)^{m_i} + \lambda_i,$$

对所有 $i = 1, 2, \dots, s$ 成立 (这里 $h_i(\lambda)$ 也是多项式). 代入 J_i 得到

$$g(J_i) = h_i(J_i)(J_i - \lambda_i I)^{m_i} + \lambda_i I = \lambda_i I = M_i.$$

于是

$$g(J) = \begin{pmatrix} g(J_1) & & \\ & g(J_2) & \\ & & \ddots \\ & & & g(J_s) \end{pmatrix} = \begin{pmatrix} M_1 & & \\ & M_2 & \\ & & \ddots \\ & & & M_s \end{pmatrix} = M.$$

又因为 $N = J - M = J - g(J)$, 所以 N 也是 J 的多项式.

现考虑一般情形, 设 $P^{-1}AP = J$, 则 $A = PJP^{-1} = P(M + N)P^{-1}$. 令 $B = PMP^{-1}$, $C = PNP^{-1}$, 则 B 是可对角化矩阵, C 是幂零阵, $BC = CB$ 并且

$$g(A) = g(PJP^{-1}) = Pg(J)P^{-1} = PMP^{-1} = B,$$

从而 $C = A - g(A)$.

最后证明唯一性. 假设 A 有另一满足条件 (1) (3) 的分解 $A = B_1 + C_1$, 则 $B - B_1 = C_1 - C$. 由 $B_1 C_1 = C_1 B_1$ 不难验证 $AB_1 = B_1 A$, $AC_1 = C_1 A$. 因为 $B = g(A)$, 故 $BB_1 = B_1 B$. 同理 $CC_1 = C_1 C$. 设 $C^r = O$, $C_1^t = O$, 用二项式定理即知 $(C_1 - C)^{r+t} = O$. 于是

$$(B - B_1)^{r+t} = (C_1 - C)^{r+t} = O.$$

因为 $BB_1 = B_1 B$, 它们都是可对角化矩阵, 由命题??知它们可同时对角化, 即存在可逆阵 Q , 使 $Q^{-1}BQ$ 和 $Q^{-1}B_1Q$ 都是对角阵. 注意到

$$(Q^{-1}BQ - Q^{-1}B_1Q)^{r+t} = (Q^{-1}(B - B_1)Q)^{r+t} = Q^{-1}(B - B_1)^{r+t}Q = O,$$

两个对角阵之差仍是一个对角阵, 这个差的幂要等于零矩阵, 则这两个矩阵必相等, 由此即得 $B = B_1$, 从而 $C = C_1$. \square

1.8 矩阵函数

引理 1.10

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

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

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 $f(x) = a_0 + a_1x + \dots + a_px^p$, 则

$$f(A) = P \text{diag}\{f(J_1), f(J_2), \dots, f(J_k)\} P^{-1},$$

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

$$J^m = \text{diag}\{J_1^m, J_2^m, \dots, J_k^m\}.$$

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

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

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

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

即可计算出 $f(A)$. □

定义 1.11 (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 1.19

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 $X = \text{diag}\{X_1, \cdots, X_k\}$, 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

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

(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

$$f(J_0) = \begin{pmatrix} f(\lambda_0) & \frac{1}{1!}f'(\lambda_0) & \frac{1}{2!}f^{(2)}(\lambda_0) & \cdots & \frac{1}{(n-1)!}f^{(n-1)}(\lambda_0) \\ & f(\lambda_0) & \frac{1}{1!}f'(\lambda_0) & \cdots & \frac{1}{(n-2)!}f^{(n-2)}(\lambda_0) \\ & & f(\lambda_0) & \cdots & \frac{1}{(n-3)!}f^{(n-3)}(\lambda_0) \\ & & & \ddots & \vdots \\ & & & & f(\lambda_0) \end{pmatrix}. \quad (1.27)$$

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

$$f(P^{-1}XP) = \lim_{p \rightarrow \infty} f_p(P^{-1}XP) = \lim_{p \rightarrow \infty} P^{-1}f_p(X)P$$

$$= P^{-1}(\lim_{p \rightarrow \infty} f_p(X))P = P^{-1}f(X)P.$$

(2) 注意到 $f_p(z)$ 是多项式, 从而有

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

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

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

(3) 由引理 1.10 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 1.20

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 $\lambda_1, \lambda_2, \dots, \lambda_n$, 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

$$f(\lambda_j), f'(\lambda_j), \dots, f^{(n_j-1)}(\lambda_j). \quad (1.28)$$

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 1.19 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 (1.27) 式即可知道 $f(J_j)$ 的收敛性等价于 (1.28) 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 1.12

于是对一切方阵, 定义

$$e^A = I + \frac{1}{1!}A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots,$$

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

都有意义. 若 A 所有特征值的模长都小于 1, 则

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

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:

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

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 1.20 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.

命题 1.6

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 对一般来说对矩阵 A, B , 下面的等式并不一定成立:

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

如对

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

不难验证 $e^A \cdot e^B \neq e^{A+B}$.

证明

□

命题 1.7

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

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

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 1.19 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 1.6 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} . □

1.9 矩阵相似的全系不变量

1.9.1 矩阵相似的判定准则之一: 特征矩阵相抵

回顾定理 1.2 中矩阵相似的充要条件.

命题 1.8

设 A, B 是数域 \mathbb{F} 上的 n 阶矩阵, $\lambda I_n - A$ 相抵于 $\text{diag}\{f_1(\lambda), f_2(\lambda), \cdots, f_n(\lambda)\}$, $\lambda I_n - B$ 相抵于 $\text{diag}\{f_{i_1}(\lambda), f_{i_2}(\lambda), \cdots, f_{i_n}(\lambda)\}$, 其中 $f_{i_1}(\lambda), f_{i_2}(\lambda), \cdots, f_{i_n}(\lambda)$ 是 $f_1(\lambda), f_2(\lambda), \cdots, f_n(\lambda)$ 的一个排列. 求证: A 与 B 相似. ◆

证明 对换 λ -矩阵 $\text{diag}\{f_1(\lambda), f_2(\lambda), \dots, f_n(\lambda)\}$ 的第 i, j 行, 再对换第 i, j 列, 可将 $f_i(\lambda)$ 与 $f_j(\lambda)$ 互换位置. 由于任一排列都可由若干次对换实现, 故 $\text{diag}\{f_1(\lambda), f_2(\lambda), \dots, f_n(\lambda)\}$ 相抵于 $\text{diag}\{f_{i_1}(\lambda), f_{i_2}(\lambda), \dots, f_{i_n}(\lambda)\}$, 于是 $\lambda I_n - A$ 相抵于 $\lambda I_n - B$, 从而 A 与 B 相似. \square

例题 1.10 设 n 阶方阵 A, B, C, D 中 A, C 可逆, 求证: 存在可逆矩阵 P, Q , 使得 $A = PCQ, B = PDQ$ 的充要条件是 $\lambda A - B$ 与 $\lambda C - D$ 相抵.

证明 必要性由 $\lambda A - B = P(\lambda C - D)Q$ 即得. 下证充分性.

设 $\lambda A - B$ 与 $\lambda C - D$ 相抵, 则由 A, C 可逆知, $\lambda I_n - A^{-1}B$ 与 $\lambda I_n - C^{-1}D$ 相抵, 于是 $A^{-1}B$ 与 $C^{-1}D$ 相似. 设 Q 为可逆矩阵, 使得 $A^{-1}B = Q^{-1}(C^{-1}D)Q$, 令 $P = AQ^{-1}C^{-1}$, 则 P 可逆且 $A = PCQ, B = PDQ$. \square

1.9.2 矩阵相似的判定准则二: 有相同的行列式因子组

回顾定理 1.5 中矩阵相似的充要条件和 λ -矩阵的行列式因子相关定义和性质.

命题 1.9 (矩阵必与其转置相似)

求证: 任一 n 阶矩阵 A 都与它的转置 A' 相似.

证明 注意到 $(\lambda I_n - A)' = \lambda I_n - A'$, 并且行列式的值在转置下不改变, 故 $\lambda I_n - A$ 和 $\lambda I_n - A'$ 有相同的行列式因子组, 从而 A 和 A' 相似. \square

命题 1.10

求证: 对任意的 $b \neq 0, n$ 阶方阵 $A(a, b)$ 均相互相似:

$$A(a, b) = \begin{pmatrix} a & b & \cdots & b & b \\ & a & \ddots & \ddots & b \\ & & \ddots & \ddots & \vdots \\ & & & a & b \\ & & & & a \end{pmatrix}$$

证明 只要证明对任意的 $b \neq 0, A(a, b)$ 的行列式因子组都一样即可. 显然 $D_n(\lambda) = (\lambda - a)^n \lambda I_n - A(a, b)$ 的前 $n-1$ 行、前 $n-1$ 列构成的子式, 其值为 $(\lambda - a)^{n-1}$; $\lambda I_n - A(a, b)$ 的前 $n-1$ 行、后 $n-1$ 列构成的子式, 其值设为 $g(\lambda)$. 注意到 $g(a)$ 是 $n-1$ 阶上三角行列式, 主对角元素全为 $-b$, 从而 $g(a) = (-b)^{n-1} \neq 0$. 因此 $(\lambda - a)^{n-1}$ 与 $g(\lambda)$ 没有公共根, 故 $((\lambda - a)^{n-1}, g(\lambda)) = 1$, 于是 $D_{n-1}(\lambda) = 1$, 从而 $A(a, b)$ 的行列式因子组为 $1, \dots, 1, (\lambda - a)^n$, 结论得证. \square

注

- (1) 在上(下)三角矩阵(如 Jordan 块)或类上(下)三角矩阵(如友阵或 Frobenius 块)中, 若上(下)次对角线上的元素全部非零, 可以尝试计算行列式因子组. 对一般的矩阵(如数字矩阵), 不建议计算行列式因子组, 推荐使用 λ -矩阵的初等变换算法, 得到不变因子组.
- (2) 注意到 $A(a, 0) = aI_n$ 的行列式因子组为 $D_i(\lambda) = (\lambda - a)^i (1 \leq i \leq n)$. 因此, 在求相似标准型的过程中, 注意千万不能使用摄动法!

1.9.3 矩阵相似的判定准则三: 有相同的不变因子组

回顾定理 1.7 可知, 所有不变因子的乘积等于特征多项式, 整除关系下最大的那个不变因子等于极小多项式. 因此, 确定特征多项式和极小多项式可帮助确定不变因子组.

命题 1.11 (同阶幂零阵必相似)

设 A 是 n 阶 n 次幂零矩阵, 即 $A^n = O$ 但 $A^{n-1} \neq O$. 若 B 也是 n 阶 n 次幂零矩阵, 求证: A 相似于 B .

证明 显然 A 的极小多项式为 λ^n , 故 A 的不变因子组是 $1, \dots, 1, \lambda^n$. 同理 B 的不变因子组也是 $1, \dots, 1, \lambda^n$, 因此 A 和 B 相似. \square

命题 1.12

设 A 为 n 阶矩阵, 证明以下 3 个结论等价:

- (1) $A = cI_n$, 其中 c 为常数;
- (2) A 的 $n-1$ 阶行列式因子是一个 $n-1$ 次多项式;
- (3) A 的不变因子组中无常数.

证明 (1) \Rightarrow (2): 显然成立.

(2) \Rightarrow (3): 由于 A 的 n 阶行列式因子 $D_n(\lambda)$ 是一个 n 次多项式, 故 A 的最后一个不变因子 $d_n(\lambda) = D_n(\lambda)/D_{n-1}(\lambda)$ 是一个一次多项式, 设为 $\lambda - c$. 因为其他不变因子都要整除 $d_n(\lambda)$, 并且所有不变因子的乘积等于 n 阶行列式因子 $D_n(\lambda)$, 故 A 的不变因子组只能是 $\lambda - c, \lambda - c, \dots, \lambda - c$.

(3) \Rightarrow (1): 设 A 的不变因子组为 $d_1(\lambda), d_2(\lambda), \dots, d_n(\lambda)$, 则 $\deg d_i(\lambda) \geq 1$. 注意到 $d_1(\lambda)d_2(\lambda)\cdots d_n(\lambda) = D_n(\lambda)$ 的次数为 n , 因此 $\deg d_i(\lambda) = 1$. 又 $d_i(\lambda) \mid d_n(\lambda)$, 故只能是 $d_1(\lambda) = d_2(\lambda) = \dots = d_n(\lambda) = \lambda - c$. 因此 A 与 cI_n 有相同不变因子组, 从而它们相似, 即存在可逆矩阵 P , 使得 $A = P^{-1}(cI_n)P = cI_n$. 另外, 也可以利用 A 的极小多项式等于 $\lambda - c$ 或 A 的 Jordan 标准型来证明. \square

例题 1.11 设 n 阶矩阵 A 的特征值全为 1, 求证: 对任意的正整数 k, A^k 与 A 相似.

注 证法一 是用“三段论法”和极小多项式来证明的 (当然用行列式因子和几何重数替代也可以); 后面利用命题 1.38 给出了第二种证法; 而命题 1.39 (当 $a = \pm 1$ 时) 给出了第三种证法.

证明 证法一: 由 A 的特征值全为 1 可知 A^k 的特征值也全为 1. 设 P 为可逆矩阵, 使得 $P^{-1}AP = J = \text{diag}\{J_{r_1}(1), \dots, J_{r_s}(1)\}$ 为 Jordan 标准型. 由于 $P^{-1}A^kP = (P^{-1}AP)^k = J^k$, 故只要证明 J^k 与 J 相似即可 (见 Jordan 块的性质 (3)). 又因为 $J^k = \text{diag}\{J_{r_1}(1)^k, \dots, J_{r_s}(1)^k\}$, 故问题可进一步归结到每个 Jordan 块, 即只要证明 $J_{r_i}(1)^k$ 与 $J_{r_i}(1)$ 相似即可. 因此不妨设 $J = J_n(1)$ 只有一个 Jordan 块, 则 $J = I_n + J_0$, 其中 $J_0 = J_n(0)$ 是特征值为 0 的 n 阶 Jordan 块. 注意到

$$J^k = (I_n + J_0)^k = I_n + C_k^1 J_0 + C_k^2 J_0^2 + \dots + J_0^k.$$

故 J^k 是一个上三角矩阵, 其主对角线上的元素全为 1, 上次对角线上的元素全为 k , 从而它的特征多项式为 $(\lambda - 1)^n$. 为了确定它的极小多项式, 我们可进行如下计算:

$$(J^k - I_n)^{n-1} = (C_k^1 J_0 + C_k^2 J_0^2 + \dots + J_0^k)^{n-1} = k^{n-1} J_0^{n-1} \neq O.$$

于是 J^k 的极小多项式为 $(\lambda - 1)^n$, 其不变因子组为 $1, \dots, 1, (\lambda - 1)^n$. 因此 J^k 与 J 有相同的不变因子, 从而 J^k 与 J 相似.

证法二: 显然 A^k 的特征值也全为 1. 注意到

$$(A^k - I_n)^l = (A - I_n)^l (A^{k-1} + A^{k-2} + \dots + I_n)^l, \quad l \geq 1.$$

由于 $A^{k-1} + A^{k-2} + \dots + I_n$ 的特征值全为 k , 故为可逆矩阵, 从而 $r((A^k - I_n)^l) = r((A - I_n)^l)$ 对任意的正整数 l 都成立. 由命题 1.38 可知, A^k 与 A 相似. \square

例题 1.12 设 n 阶矩阵 A 的特征值全为 1 或 -1, 求证: A^{-1} 与 A 相似.

注 证法一 是用“三段论法”和极小多项式来证明的 (当然用行列式因子和几何重数替代也可以); 后面利用命题 1.38 给出了第二种证法; 而命题 1.39 (当 $a = \pm 1$ 时) 给出了第三种证法.

证明 证法一: 设 P 为可逆矩阵, 使得 $P^{-1}AP = J = \text{diag}\{J_{r_1}(\lambda_1), \dots, J_{r_s}(\lambda_s)\}$ 为 Jordan 标准型, 其中 $\lambda_i = \pm 1$. 由于 $P^{-1}A^{-1}P = (P^{-1}AP)^{-1} = J^{-1}$, 故只要证明 J^{-1} 与 J 相似即可. 又因为 $J^{-1} = \text{diag}\{J_{r_1}(\lambda_1)^{-1}, \dots, J_{r_s}(\lambda_s)^{-1}\}$, 故问题可进一步归结到每个 Jordan 块, 即只要证明 $J_{r_i}(\lambda_i)^{-1}$ 与 $J_{r_i}(\lambda_i)$ 相似即可 (见 Jordan 块的性质 (4)). 因此不妨设 $J = J_n(\lambda_0)$ 只有一个 Jordan 块, 则 $J = \lambda_0 I_n + J_0$, 其中 $\lambda_0 = \pm 1, J_0 = J_n(0)$ 是特征值为 0 的 n 阶 Jordan 块. 注意到

$$\lambda_0^n I_n = (\lambda_0 I_n)^n - (-J_0)^n = (\lambda_0 I_n + J_0)(\lambda_0^{n-1} I_n - \lambda_0^{n-2} J_0 + \dots + (-1)^{n-1} J_0^{n-1}).$$

以及 $\lambda_0^{-1} = \lambda_0$, 故可得

$$J^{-1} = (\lambda_0 I_n + J_0)^{-1} = \lambda_0 I_n - \lambda_0^2 J_0 + \cdots + (-1)^{n-1} \lambda_0^n J_0^{n-1}.$$

因此 J^{-1} 是一个上三角矩阵, 其主对角线上的元素全为 λ_0 , 上次对角线上的元素全为 $-\lambda_0^2$, 从而它的特征多项式为 $(\lambda - \lambda_0)^n$. 为了确定它的极小多项式, 我们可进行如下计算:

$$(J^{-1} - \lambda_0 I)^{n-1} = (-\lambda_0^2 J_0 + \cdots + (-1)^{n-1} \lambda_0^n J_0^{n-1})^{n-1} = (-1)^{n-1} J_0^{n-1} \neq O$$

于是 J^{-1} 的极小多项式为 $(\lambda - \lambda_0)^n$, 其不变因子组为 $1, \cdots, 1, (\lambda - \lambda_0)^n$. 因此 J^{-1} 与 J 有相同的不变因子组, 从而 J^{-1} 与 J 相似.

证法二: 显然 A^{-1} 的特征值也全为 1 或 -1. 设 $\lambda_0 = \pm 1$, 则由 A 可逆以及 $(A^{-1} - \lambda_0 I_n)^l = (-\lambda_0)^l A^{-l} (A - \lambda_0 I_n)^l$ 可得 $r((A^{-1} - \lambda_0 I_n)^l) = r((A - \lambda_0 I_n)^l)$ 对任意的正整数 l 都成立. 由命题 1.38 可知, A^{-1} 与 A 相似. \square

1.9.4 矩阵相似的判定准则四: 有相同的初等因子组

定义 1.13 (准素因子)

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**. 所有 $f(\lambda)$ 的准素因子称为 $f(\lambda)$ 的**准素因子组**.

事实上, 若设 $f(\lambda)$ 在 \mathbb{K} 上的标准因式分解为

$$f(\lambda) = c P_1(\lambda)^{e_1} P_2(\lambda)^{e_2} \cdots P_t(\lambda)^{e_t}$$

其中 c 为非零常数, $P_i(\lambda)$ 为互异的首一不可约多项式, $e_i > 0 (1 \leq i \leq t)$, 则 $f(\lambda)$ 的所有准素因子为 $P_1(\lambda)^{e_1}, P_2(\lambda)^{e_2}, \cdots, P_t(\lambda)^{e_t}$.

定理 1.21 (λ -矩阵和初等因子的基本性质)

- (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:

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

- (2) 设 A 是数域 \mathbb{K} 上的 n 阶矩阵, 其特征矩阵 $\lambda I_n - A$ 经过初等变换可化为对角矩阵 $\text{diag}\{f_1(\lambda), f_2(\lambda), \cdots, f_n(\lambda)\}$, 其中 $f_i(\lambda)$ 是 \mathbb{K} 上的首一多项式. 求证: 矩阵 A 的初等因子组等于所有 $f_i(\lambda)$ 的准素因子组.



笔记 由 (2) 可知, 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. 实际上, (2) 就是引理 1.8 的一个推广.

证明

- (1) 由已知, 存在多项式 $u(\lambda), v(\lambda)$, 使得 $f(\lambda)u(\lambda) + g(\lambda)v(\lambda) = d(\lambda)$. 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.

- (2) 对任意的 $i < j$, 以下操作记为 $O(i, j)$: 设 $d(\lambda) = (f_i(\lambda), f_j(\lambda)), m(\lambda) = [f_i(\lambda), f_j(\lambda)]$ 分别是 $f_i(\lambda)$ 和 $f_j(\lambda)$ 的最大公因式和最小公倍式, 则用 $d(\lambda)$ 替代 $f_i(\lambda)$, 用 $m(\lambda)$ 替代 $f_j(\lambda)$. 我们先证明, 操作 $O(i, j)$ 可通过 λ -矩阵的初等变换来实现, 并且前后两个对角矩阵, 即 $\text{diag}\{f_1(\lambda), \cdots, f_i(\lambda), \cdots, f_j(\lambda), \cdots, f_n(\lambda)\}$ 与 $\text{diag}\{f_1(\lambda), \cdots, d(\lambda), \cdots, m(\lambda), \cdots, f_n(\lambda)\}$ 有相同的准素因子组.

由 (1) 即知 $O(i, j)$ 是 λ -矩阵的相抵变换. 设 $f_i(\lambda), f_j(\lambda)$ 的公共因式分解为

$$f_i(\lambda) = P_1(\lambda)^{e_{i1}} P_2(\lambda)^{e_{i2}} \cdots P_t(\lambda)^{e_{it}}, \quad f_j(\lambda) = P_1(\lambda)^{e_{j1}} P_2(\lambda)^{e_{j2}} \cdots P_t(\lambda)^{e_{jt}}$$

其中 $P_i(\lambda)$ 为互异的首一不可约多项式, $e_{ik} \geq 0, e_{jk} \geq 0 (1 \leq k \leq t)$, 令 $r_k = \min\{e_{ik}, e_{jk}\}, s_k = \max\{e_{ik}, e_{jk}\}$, 则有

$$d(\lambda) = P_1(\lambda)^{r_1} P_2(\lambda)^{r_2} \cdots P_t(\lambda)^{r_t}, \quad m(\lambda) = P_1(\lambda)^{s_1} P_2(\lambda)^{s_2} \cdots P_t(\lambda)^{s_t}$$

显然 $\{f_i(\lambda), f_j(\lambda)\}$ 和 $\{d(\lambda), m(\lambda)\}$ 有相同的准素因子组, 因此 $O(i, j)$ 操作前后的两个对角矩阵也有相同的准素因子组.

对对角矩阵 $\text{diag}\{f_1(\lambda), f_2(\lambda), \cdots, f_n(\lambda)\}$ 依次实施操作 $O(1, j) (2 \leq j \leq n)$, 则得到对角矩阵的第 $(1, 1)$ 元素的所有不可约因式的幂在主对角元素中都是最小的; 然后依次操作 $O(2, j) (3 \leq j \leq n); \cdots$; 最后操作 $O(n-1, n)$, 可得一个对角矩阵 $\Lambda = \text{diag}\{d_1(\lambda), d_2(\lambda), \cdots, d_n(\lambda)\}$. 由操作的性质可知, Λ 满足 $d_i(\lambda) \mid d_{i+1}(\lambda) (1 \leq i \leq n-1)$, 因此 Λ 就是矩阵 A 的法式. 又因为对角矩阵 $\text{diag}\{f_1(\lambda), f_2(\lambda), \cdots, f_n(\lambda)\}$ 与法式有相同的准素因子组, 故所有 $f_i(\lambda)$ 的准素因子组就是矩阵 A 的初等因子组.

□

命题 1.13

设 $A = \text{diag}\{A_1, A_2, \cdots, A_k\}$ 为分块对角矩阵, 求证: A 的初等因子组等于 $A_i (1 \leq i \leq k)$ 的初等因子组的无交并集. 又若交换各块的位置, 则所得的矩阵仍和 A 相似.

◆

证明

□

显然 $\lambda I - A$ 也是一个分块对角矩阵, 用 λ -矩阵的初等变换将每一块化为法式, 则由 **λ -矩阵和初等因子的基本性质 (2)** 可知, A 的初等因子组就是所有各块的初等因子组的无交并集. 又交换 A 的各块并不改变 A 的初等因子组, 因此所得之矩阵仍和 A 相似.

1.10 有理标准型的几何与应用

回顾有理标准型和循环子空间相关理论.

命题 1.14

设 n 阶矩阵 A 有 n 个不同的特征值, 求证: A 的特征多项式和极小多项式相等.

◆

证明 证法一: 设 A 的 n 个不同的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 则由推论?? 可知, 特征多项式 $f(\lambda)$ 和极小多项式 $m(\lambda)$ 有相同的根 (不计重数), 因此

$$f(\lambda) = m(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$$

证法二: 由于 A 有 n 个不同的特征值, 故 A 相似于对角矩阵. 又因为相似矩阵有相同的特征多项式和极小多项式, 所以只要对对角矩阵证明此结论即可. 设 $A = \text{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$, 则 $\lambda I_n - A = \text{diag}\{\lambda - \lambda_1, \lambda - \lambda_2, \cdots, \lambda - \lambda_n\}$, 这是一个主对角元素两两互素的对角矩阵, 由 **λ -矩阵和初等因子的基本性质 (1)** 以及数学归纳法可知其法式为 $\text{diag}\{1, \cdots, 1, (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)\}$. 因此, A 的特征多项式和极小多项式相等.

□

命题 1.15

设 n 阶矩阵 A 有 n 个不同的特征值, 且特征值 λ_i 对应的特征向量为 α_i , 由推论?? 可知 $\{\alpha_1, \cdots, \alpha_n\}$ 为 \mathbb{C}^n 的一组基. 则 $\alpha = \alpha_1 + \cdots + \alpha_n$ 是 A 的循环空间 \mathbb{C}^n 的循环向量, 即 $\mathbb{C}^n = L(\alpha, A\alpha, \cdots, A^{n-1}\alpha) = C(A, \alpha)$ 为循环空间, α 是循环向量.

◆

证明 事实上, 由 $A^k \alpha = \lambda_1^k \alpha_1 + \cdots + \lambda_n^k \alpha_n$, 利用 Vandermonde 行列式容易证明 $\{\alpha, A\alpha, \cdots, A^{n-1}\alpha\}$ 是 \mathbb{C}^n 的一组基, 从而 $\mathbb{C}^n = L(\alpha, A\alpha, \cdots, A^{n-1}\alpha) = C(A, \alpha)$ 为循环空间, α 是循环向量.

□

命题 1.16

设数域 \mathbb{K} 上的 n 阶矩阵 A 的特征多项式 $f(\lambda) = P_1(\lambda)P_2(\lambda)\cdots P_k(\lambda)$, 其中 $P_i(\lambda) (1 \leq i \leq k)$ 是 \mathbb{K} 上互异的首一不可约多项式. 求证: A 的有理标准型只有一个 Frobenius 块, 并且 A 在复数域上可对角化.

注 我们也可以利用定理 1.12 和初等因子证明这个命题. 若利用不变因子在基域扩张下的不变性, 则这个命题也可由命题 1.14 得到.

证明 设 A 的不变因子组为 $d_1(\lambda), d_2(\lambda), \dots, d_n(\lambda)$, 其中 $d_i(\lambda) \mid d_{i+1}(\lambda), i = 1, 2, \dots, n-1$, 则有

$$f(\lambda) = P_1(\lambda)P_2(\lambda)\cdots P_k(\lambda) = d_1(\lambda)d_2(\lambda)\cdots d_n(\lambda) \quad (1.29)$$

由于 $P_i(\lambda)$ 是不可约多项式, 故存在某个 j , 使得 $P_i(\lambda) \mid d_j(\lambda)$, 否则, 由不可约多项式的基本性质 (1) 可知 $(P_i(\lambda), d_j(\lambda)) = 1, j = 1, 2, \dots, n$. 再由互素多项式和最大公因式的基本性质 (5) 可知 $(P_i(\lambda), d_1(\lambda)d_2(\lambda)\cdots d_n(\lambda)) = 1$, 这与 (1.29) 矛盾! 从而 $P_i(\lambda) \mid d_n(\lambda) (1 \leq i \leq k)$. 由互素多项式和最大公因式的基本性质 (1) 可知, $P_1(\lambda)P_2(\lambda)\cdots P_k(\lambda) \mid d_n(\lambda)$, 因此只能是 $d_1(\lambda) = \cdots = d_{n-1}(\lambda) = 1, d_n(\lambda) = f(\lambda)$, 从而 A 的有理标准型只有一个 Frobenius 块. 由于特征多项式 $f(\lambda) = P_1(\lambda)P_2(\lambda)\cdots P_k(\lambda)$ 在 \mathbb{K} 上无重因式, 故 $(f(\lambda), f'(\lambda)) = 1$, 从而 $f(\lambda)$ 在复数域上无重根, 即 A 有 n 个不同的特征值, 于是 A 在复数域上可对角化. \square

推论 1.12

设数域 \mathbb{K} 上的 n 阶矩阵 A 的特征多项式 $f(\lambda) = P_1(\lambda)P_2(\lambda)\cdots P_k(\lambda)$, 其中 $P_i(\lambda) (1 \leq i \leq k)$ 是 \mathbb{K} 上互异的首一不可约多项式. 并且 α_i 为线性方程组 $P_i(A)x = 0$ 的非零解, 则 $\alpha = \alpha_1 + \cdots + \alpha_k$ 是 A 的循环空间 \mathbb{K}^n 的循环向量.

证明 由命题 1.16 及定理 1.10 可知 \mathbb{K}^n 就是一个循环空间. (未完成证明) \square

命题 1.17

设 φ 是数域 \mathbb{K} 上 n 维线性空间 V 上的线性变换, φ 的特征多项式为 $f(\lambda)$, 证明以下 3 个结论等价:

- (1) V 只有平凡的 φ -不变子空间;
- (2) V 中任一非零向量都是循环向量, 使 V 成为循环空间;
- (3) $f(\lambda)$ 是 \mathbb{K} 上的不可约多项式.

证明 (1) \Rightarrow (2): 任取 V 中非零向量 α , 则循环子空间 $C(\varphi, \alpha)$ 是非零 φ -不变子空间. 由于 V 只有平凡的 φ -不变子空间, 故 $C(\varphi, \alpha) = V$, 即 V 中任一非零向量都是循环向量, 使 V 成为循环空间.

(2) \Rightarrow (3): 用反证法, 假设 $f(\lambda) = g(\lambda)h(\lambda)$, 其中 $g(\lambda), h(\lambda)$ 是 \mathbb{K} 上次数小于 n 的首一多项式. 由 Cayley-Hamilton 定理可知 $0 = f(\varphi) = g(\varphi)h(\varphi)$, 故由命题 1.1(1) 的逆否命题可知 $g(\varphi), h(\varphi)$ 中至少有一个是奇异 (不可逆/非双射) 线性变换, 不妨设为 $g(\varphi)$, 由推论 1.12 可知 $\text{Ker} g(\varphi) \neq 0$. 任取 $\text{Ker} g(\varphi)$ 中的非零向量 α , 设 $\deg g(\lambda) = r$, 则不妨设

$$g(\varphi) = a_r \varphi^r + a_{r-1} \varphi^{r-1} + \cdots + a_1, \quad \text{其中 } a_r \neq 0.$$

由 $\alpha \in \text{Ker} g(\varphi)$ 可知

$$g(\varphi)(\alpha) = a_r \varphi^r(\alpha) + a_{r-1} \varphi^{r-1}(\alpha) + \cdots + a_1 \alpha = 0.$$

于是

$$\varphi^r(\alpha) = -\frac{a_{r-1}}{a_r} \varphi^{r-1}(\alpha) - \cdots - \frac{a_1}{a_r} \alpha. \quad (1.30)$$

假设对 $k \geq r$ 且 $k \in \mathbb{N}$, 成立 $\varphi^k(\alpha)$ 可由 $\{\alpha, \varphi(\alpha), \dots, \varphi^{r-1}(\alpha)\}$ 线性表示, 则对 (1.30) 式两边同时作用 φ^{k-r+1} 可得

$$\varphi^{k+1}(\alpha) = -\frac{a_{r-1}}{a_r} \varphi^k(\alpha) - \cdots - \frac{a_1}{a_r} \varphi^{k-r+1}(\alpha).$$

于是由归纳假设可知, $\varphi^{k+1}(\alpha)$ 可由 $\{\alpha, \varphi(\alpha), \dots, \varphi^{r-1}(\alpha)\}$ 线性表示. 故由数学归纳法可得, 对 $\forall k \geq r$ 且 $k \in \mathbb{N}$, 都有 $\varphi^k(\alpha)$ 可由 $\{\alpha, \varphi(\alpha), \dots, \varphi^{r-1}(\alpha)\}$ 线性表示. 因此 $C(\varphi, \alpha) = L(\alpha, \varphi(\alpha), \dots) = L(\alpha, \varphi(\alpha), \dots, \varphi^{r-1}(\alpha))$, 其维数

$\leq r < n$, 故 $C(\varphi, \alpha) \neq V$, 这与 V 中任一非零向量都是循环向量矛盾!

(3) \Rightarrow (1): 用反证法, 假设存在非平凡的 φ -不变子空间 $U, \dim U = r$, 则 φ 在一组基下的表示矩阵为分块上三角矩阵 $M = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$, 其中 A 是 $\varphi|_U$ 的表示矩阵. 于是特征多项式

$$f(\lambda) = |\lambda I_V - \varphi| = |\lambda I_n - M| = |\lambda I_r - A| \cdot |\lambda I_{n-r} - B|.$$

是两个低次多项式的乘积, 这与 $f(\lambda)$ 的不可约性矛盾! □

命题 1.18

设 φ 是数域 \mathbb{K} 上 n 维线性空间 V 上的线性变换, φ 的极小多项式为 $m(\lambda)$. 证明: $m(\lambda)$ 是 \mathbb{K} 上的不可约多项式的充要条件是 V 的任一非零 φ -不变子空间 U 必为如下形式:

$$U = C(\varphi, \alpha_1) \oplus C(\varphi, \alpha_2) \oplus \cdots \oplus C(\varphi, \alpha_k)$$

并且 $\varphi|_{C(\varphi, \alpha_i)}$ 的极小多项式都是 $m(\lambda)$. 此时, $\varphi|_U$ 的极小多项式也是 $m(\lambda)$. ♣

证明 必要性: 设 $\varphi|_U$ 的极小多项式为 $n(\lambda)$, 则 $m(\varphi|_U) = m(\varphi)|_U = 0$, 从而 $n(\lambda) \mid m(\lambda)$. 因为 $m(\lambda)$ 不可约, 所以 $n(\lambda) = m(\lambda)$. 又由于 $\varphi|_U$ 的所有不变因子都要整除 $m(\lambda)$ 且 $m(\lambda)$ 不可约, 故所有的非常数不变因子都等于 $m(\lambda)$. 最后, 由有理标准型的几何意义即得

$$U = C(\varphi, \alpha_1) \oplus C(\varphi, \alpha_2) \oplus \cdots \oplus C(\varphi, \alpha_k).$$

并且 $\varphi|_{C(\varphi, \alpha_i)}$ 在基 $\{\alpha_i, \varphi(\alpha_i), \dots, \varphi^{r_i-1}(\alpha_i)\}$ 下的表示矩阵就是友阵 $C(m(\lambda))$, 其中 $r_i = \dim C(\varphi, \alpha_i)$. 于是 $\varphi|_{C(\varphi, \alpha_i)}$ 的极小多项式就是其表示矩阵 $C(m(\lambda))$ 的极小多项式. 又由引理 1.5 可知, $C(m(\lambda))$ 的极小多项式就是 $m(\lambda)$, 并且 $n(\lambda) = m(\lambda)$, 故结论得证.

充分性: 用反证法, 设 $m(\lambda) = g(\lambda)h(\lambda)$, 其中 $g(\lambda), h(\lambda)$ 是 \mathbb{K} 上次数小于 $m(\lambda)$ 次数的首一多项式, 则 $0 = m(\varphi) = g(\varphi)h(\varphi)$, 故由命题 1.1(1) 的逆否命题可知 $g(\varphi), h(\varphi)$ 中至少有一个是奇异线性变换, 不妨设为 $g(\varphi)$, 于是由推论 1.1 可知 $\text{Ker} g(\varphi) \neq 0$. 任取 $\text{Ker} g(\varphi)$ 中的非零向量 α , 得到循环子空间 $U = C(\varphi, \alpha)$,

由 $g(\varphi)(\alpha) = 0$ 可知, 对 $\forall k \in \mathbb{N}$, 都有 $\varphi^k(g(\varphi)(\alpha)) = 0$. 从而对 $\forall \beta \in U = C(\varphi, \alpha)$, 存在不全为零的 a_i 使得

$$\begin{aligned} g(\varphi)(\beta) &= g(a_1\alpha + a_2\varphi(\alpha) + a_3\varphi^2(\alpha) + \cdots) \\ &= a_1g(\alpha) + a_2g(\varphi(\alpha)) + a_3g(\varphi^2(\alpha)) + \cdots \\ &= a_1g(\alpha) + a_2\varphi(g(\alpha)) + a_3\varphi^2(g(\alpha)) + \cdots \\ &= 0 + 0 + 0 + \cdots = 0. \end{aligned}$$

因此 $g(\varphi|_U) = g(\varphi)|_U = 0$, 于是 $\varphi|_U$ 的极小多项式 $m(\lambda)$ 整除 $g(\lambda)$, 从而其次数 $\leq \deg g(\lambda) < \deg m(\lambda)$, 这与条件矛盾! □

定理 1.22 (基于初等因子组的有理标准型)

设数域 \mathbb{K} 上的 n 阶矩阵 A 的初等因子组为 $P_1(\lambda)^{r_1}, P_2(\lambda)^{r_2}, \dots, P_k(\lambda)^{r_k}$, 证明: A 相似于分块对角矩阵

$$\begin{aligned} \tilde{F} &= \text{diag}\{F(P_1(\lambda)^{r_1}), F(P_2(\lambda)^{r_2}), \dots, F(P_k(\lambda)^{r_k})\} \\ \tilde{C} &= \text{diag}\{C(P_1(\lambda)^{r_1}), C(P_2(\lambda)^{r_2}), \dots, C(P_k(\lambda)^{r_k})\} \end{aligned}$$

称为 A 的基于初等因子组的有理标准型. ♥

证明 由 Frobenius 块和友阵的性质可知, $\lambda I_n - \tilde{F}$ 和 $\lambda I_n - \tilde{C}$ 都相抵于

$$\text{diag}\{1, \dots, 1, P_1(\lambda)^{r_1}; 1, \dots, 1, P_2(\lambda)^{r_2}; \dots; 1, \dots, 1, P_k(\lambda)^{r_k}\}$$

再由 λ -矩阵和初等因子的基本性质 (2) 可知, \tilde{F}, \tilde{C} 与 A 有相同的初等因子组, 从而它们相似. □

定理 1.23

设 φ 是数域 \mathbb{K} 上 n 维线性空间 V 上的线性变换, φ 的初等因子组为 $P_1(\lambda)^{r_1}, P_2(\lambda)^{r_2}, \dots, P_k(\lambda)^{r_k}$. 证明: 存在 $\alpha_1, \alpha_2, \dots, \alpha_k \in V$, 使得

$$V = C(\varphi, \alpha_1) \oplus C(\varphi, \alpha_2) \oplus \dots \oplus C(\varphi, \alpha_k)$$

♡

证明 由基于初等因子组的有理标准型和定理 1.10 即得. □

例题 1.13 求证: 存在 n 阶实方阵 A , 满足 $A^2 + 2A + 5I_n = O$ 的充要条件是 n 为偶数. 当 $n \geq 4$ 时, 验证满足上述条件的矩阵 A 有无限个不变子空间.

证明 **必要性:** 注意到 A 适合多项式 $g(\lambda) = \lambda^2 + 2\lambda + 5$, 故 A 的极小多项式 $m(\lambda) \mid g(\lambda)$, 又因为 $g(\lambda)$ 在实数域上不可约, 故只能是 $m(\lambda) = g(\lambda)$. 同理可证 A 所有的非常数不变因子都等于 $g(\lambda)$, 从而 A 的不变因子组为 $1, \dots, 1, g(\lambda), \dots, g(\lambda)$ (k 个 $g(\lambda)$). 因此 A 的特征多项式 $f(\lambda) = g(\lambda)^k$, 于是 $n = \deg f(\lambda) = 2k$ 为偶数.

充分性: 设 $n = 2k$ 为偶数, 则由必要性的证明可知, A 的不变因子组为 $1, \dots, 1, g(\lambda), \dots, g(\lambda)$ (k 个 $g(\lambda)$). 可用有理标准型构造满足条件的矩阵:

$$A = \text{diag} \left\{ \begin{pmatrix} 0 & -5 \\ 1 & -2 \end{pmatrix}, \dots, \begin{pmatrix} 0 & -5 \\ 1 & -2 \end{pmatrix} \right\} \quad (k \text{ 个二阶方阵}) \quad (1.31)$$

当 $n \geq 4$ 时, 设 $\{e_1, e_2, e_3, e_4\}$ 是前 4 个标准单位列向量, 则容易验证 (1.31) 式中的矩阵 A 满足 $Ae_1 = e_2, Ae_3 = e_4$, 于是构造循环子空间 $\{C_l := C(A, e_1 + le_3) = L(e_1 + le_3, e_2 + le_4), l \in \mathbb{R}\}$, 进一步容易验证循环子空间 $\{C_l := C(A, e_1 + le_3) = L(e_1 + le_3, e_2 + le_4), l \in \mathbb{R}\}$ 是两两互异的 A -不变子空间, 故 A 有无限个不变子空间. □

命题 1.19

设 A 是数域 \mathbb{K} 上的 n 阶方阵, 求证: A 的极小多项式的次数小于等于 $r(A) + 1$.

♣

证明 **证法一:** 设 A 的不变因子组为 $1, \dots, 1, d_1(\lambda), \dots, d_k(\lambda)$, 则极小多项式 $m(\lambda) = d_k(\lambda)$, 并且由定理 1.6 可知 A 相似于 $F = \text{diag}\{F(d_1(\lambda)), \dots, F(d_k(\lambda))\}$. 设 $\deg d_k(\lambda) = r$, 若 $d_k(0) \neq 0$, 则由 Frobenius 块的基本性质 (1) 可知 $F(d_k(\lambda))$ 非异; 若 $d_k(0) = 0$, 则由 Frobenius 块的基本性质 (1) 可知 $F(d_k(\lambda))$ 奇异且右上角的 $r-1$ 阶子式非零, 从而秩为 $r-1$. 因此, $r(A) = r(F) \geq r(F(d_k(\lambda))) \geq r-1 = \deg d_k(\lambda) - 1$.

证法二: 从 A 的极小多项式 $m(\lambda)$ 分离出来的初等因子中, 形如 λ^r 的初等因子至多只有 1 个, 对应于零特征值的 Jordan 块 $J_r(0)$, 其余的初等因子对应于非零特征值的 Jordan 块. 因此 $r(A)$ 大于等于这些 Jordan 块秩的和, 后者等于 $\deg m(\lambda) - 1$ 或 $\deg m(\lambda)$. □

命题 1.20

设数域 \mathbb{K} 上的 n 阶矩阵 A 的不变因子组是 $1, \dots, 1, d_1(\lambda), \dots, d_k(\lambda)$, 其中 $d_i(\lambda)$ 是非常数首一多项式, $d_i(\lambda) \mid d_{i+1}(\lambda)$ ($1 \leq i \leq k-1$). 求证: 对 A 的任一特征值 λ_0 ,

$$r(\lambda_0 I_n - A) = n - \sum_{i=1}^k \delta_{d_i(\lambda_0), 0}$$

其中记号 $\delta_{a,b}$ 表示: 若 $a = b$, 取值为 1; 若 $a \neq b$, 取值为 0.

♣

证明 **证法一:** 设 $\deg d_i(\lambda) = r_i$, 则由定理 1.6 可知 A 相似于 $F = \text{diag}\{F(d_1(\lambda)), \dots, F(d_k(\lambda))\}$, 而相似矩阵有相同的特征多项式, 故 $\lambda_0 I_n - A = \lambda_0 I_n - F$. 由 Frobenius 块的基本性质 (2) 可知 $|\lambda_0 I_{r_i} - F(d_i(\lambda))| = d_i(\lambda_0)$. 若 $d_i(\lambda_0) \neq 0$, 则 $\lambda_0 I_{r_i} - F(d_i(\lambda))$ 非异; 若 $d_i(\lambda_0) = 0$, 则 $\lambda_0 I_{r_i} - F(d_i(\lambda))$ 奇异且右上角的 r_i-1 阶子式非零, 从而秩为 r_i-1 . 因此,

$$\begin{aligned} r(\lambda_0 I_n - A) &= r(\lambda_0 I_n - F) = \sum_{i=1}^k r(\lambda_0 I_{r_i} - F(d_i(\lambda))) \\ &= \sum_{i=1}^k (r_i - \delta_{d_i(\lambda_0), 0}) = n - \sum_{i=1}^k \delta_{d_i(\lambda_0), 0} \end{aligned}$$

证法二: 由定理 1.6 可知存在可逆 λ -矩阵 $P(\lambda), Q(\lambda)$, 使得

$$P(\lambda)(\lambda I_n - A)Q(\lambda) = \text{diag}\{1, \dots, 1, d_1(\lambda), \dots, d_k(\lambda)\}$$

在上式中令 $\lambda = \lambda_0$, 注意到 $P(\lambda_0), Q(\lambda_0)$ 是 \mathbb{K} 上的可逆矩阵, 故 $\lambda_0 I_n - A$ 相抵于 $\text{diag}\{1, \dots, 1, d_1(\lambda_0), \dots, d_k(\lambda_0)\}$, 于是 $r(\lambda_0 I_n - A)$ 等于 n 减去等于零的 $d_i(\lambda_0)$ 的个数, 从而结论得证. \square

命题 1.21

设 A 是数域 \mathbb{K} 上的 n 阶矩阵, 求证: 若 $\text{tr}(A) = 0$, 则 A 相似于一个 \mathbb{K} 上主对角元全为零的矩阵.

证明 对阶数进行归纳. 当 $n = 1$ 时, $A = O$, 结论显然成立. 设阶数小于 n 时结论成立, 现证 n 阶的情形. 由于题目的条件和结论在相似关系下不改变, 故不妨从一开始就假设 A 是有理标准型

$$F = \text{diag}\{F(d_1(\lambda)), \dots, F(d_k(\lambda))\}$$

其中 $d_i(\lambda)$ 是 A 的非常数不变因子, $d_i(\lambda) \mid d_{i+1}(\lambda) (1 \leq i \leq k-1)$, $\deg d_i(\lambda) = r_i$. 若 r_i 都为 1, 则 $d_1(\lambda) = \dots = d_n(\lambda) = \lambda - c$, 从而 $A = cI_n$. 又 $\text{tr}(A) = 0$, 故 $c = 0$, 从而 $A = O$, 结论成立. 以下假设存在某个 $r_i > 1$, 将第 $(1, 1)$ 分块与第 (i, i) 分块对换, 这是一个相似变换, 此时矩阵的第 $(1, 1)$ 元为零, 故不妨设 A 的第 $(1, 1)$ 元为零. 注意到矩阵 $A = \begin{pmatrix} 0 & \alpha' \\ \beta & B \end{pmatrix}$, 其中 $\alpha, \beta \in \mathbb{K}^{n-1}, B \in M_{n-1}(\mathbb{K}), \text{tr}(B) = 0$. 由归纳假设, 存在 \mathbb{K} 上的 $n-1$ 阶非异阵 Q , 使得 $Q^{-1}BQ$

的主对角元全为零, 令 $P = \begin{pmatrix} 1 & O \\ O & Q \end{pmatrix}$ 为 \mathbb{K} 上的 n 阶非异阵, 则 $P^{-1}AP = \begin{pmatrix} 0 & \alpha'Q \\ Q^{-1}\beta & Q^{-1}BQ \end{pmatrix}$ 的主对角元全为零, 结论得证. \square

命题 1.22

设 C 是数域 \mathbb{K} 上的 n 阶矩阵, 求证: 存在 \mathbb{K} 上的 n 阶矩阵 A, B , 使得 $AB - BA = C$ 的充要条件是 $\text{tr}(C) = 0$.

证明 必要性由矩阵迹的线性和交换性即得, 下证充分性. 由于题目的条件和结论在同时相似变换 $A \mapsto P^{-1}AP, B \mapsto P^{-1}BP, C \mapsto P^{-1}CP$ 下不改变, 故由命题 1.21 不妨从一开始就假设 $C = (c_{ij})$ 的主对角元 $c_{ii} = 0 (1 \leq i \leq n)$. 取定 $A = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ 为 \mathbb{K} 上的主对角元互异的对角矩阵. 设 $B = (x_{ij})$, 则 $AB - BA = C$ 等价于方程 $\lambda_i x_{ij} - \lambda_j x_{ij} = c_{ij}$. 当 $i = j$ 时, 上式恒成立, 故 x_{ii} 可任取. 当 $i \neq j$ 时, $x_{ij} = \frac{c_{ij}}{\lambda_i - \lambda_j}$ 被唯一确定. 因此, 一定存在 \mathbb{K} 上的矩阵 A, B , 使得 $AB - BA = C$ 成立. \square

1.11 乘法交换性诱导的多项式表示

定理 1.24

设 φ 是数域 \mathbb{K} 上 n 维线性空间 V 上的线性变换, 则对 V 上任一与 φ 乘法可交换的线性变换 ψ , 都存在不超过 $n-1$ 次的多项式 $g(x) \in \mathbb{K}[x]$, 使得 $\psi = g(\varphi)$ 成立的充要条件是 φ 的极小多项式等于其特征多项式.

注 本题充分性证明的关键点是: $V = C(\varphi, e_1)$ 是一个循环空间, 循环向量 e_1 经过 φ 的 $n-1$ 次作用, 生成了 V 的一组基 $\{e_1, e_2, \dots, e_n\}$. 因此, 只要验证了 ψ 和 $g(\varphi)$ 在循环向量 e_1 上的取值相同, 那么由 φ, ψ 的乘法交换性可知 ψ 和 $g(\varphi)$ 在上述基上的取值也相同, 从而它们必相等.

证明 充分性: 设 φ 的极小多项式等于其特征多项式 $f(\lambda) = \lambda^n + a_1\lambda^{n-1} + \dots + a_{n-1}\lambda + a_n$, 则 φ 只有一个非常数不

变因子. 由有理标准型理论, 存在 V 的一组基 $\{e_1, e_2, \dots, e_n\}$, 使得 φ 在这组基下的表示矩阵为友阵

$$C(f(\lambda)) = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_n \\ 1 & 0 & \cdots & 0 & -a_{n-1} \\ 0 & 1 & \cdots & 0 & -a_{n-2} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_1 \end{pmatrix}$$

即有

$$\varphi(e_1) = e_2, \varphi(e_2) = e_3, \dots, \varphi(e_{n-1}) = e_n, \varphi(e_n) = -a_n e_1 - a_{n-1} e_2 - \cdots - a_1 e_n$$

任取 V 上满足 $\varphi\psi = \psi\varphi$ 的线性变换 ψ , 设

$$\psi(e_1) = b_n e_1 + b_{n-1} e_2 + \cdots + b_1 e_n \quad (1.32)$$

令 $g(x) = b_1 x^{n-1} + \cdots + b_{n-1} x + b_n$, 我们来证明: $\psi = g(\varphi)$. 首先由 $e_k = \varphi^{k-1}(e_1) (k \geq 2)$ 以及 (1.32) 式可知 $\psi(e_1) = g(\varphi)(e_1)$ 成立. 其次由 φ, ψ 乘法可交换, 故对任意的 $e_k (k \geq 2)$ 有

$$\begin{aligned} \psi(e_k) &= \psi(\varphi^{k-1}(e_1)) = \varphi^{k-1}(\psi(e_1)) = \varphi^{k-1}(g(\varphi)(e_1)) \\ &= g(\varphi)(\varphi^{k-1}(e_1)) = g(\varphi)(e_k) \end{aligned}$$

最后, 注意到 ψ 与 $g(\varphi)$ 在基向量 $\{e_1, e_2, \dots, e_n\}$ 上的取值都相等, 故由线性扩张定理可知 $\psi = g(\varphi)$ 成立.

必要性: 设 φ 的不变因子组为 $1, \dots, 1, d_1(\lambda), \dots, d_k(\lambda)$, 其中 $d_i(\lambda)$ 为非常数首一多项式, $d_i(\lambda) \mid d_{i+1}(\lambda) (1 \leq i \leq k-1)$, 则 φ 的有理标准型 $F = \text{diag}\{F_1, F_2, \dots, F_k\}$, 其中 $F_i = F(d_i(\lambda))$ 为 n_i 阶矩阵. 若 φ 的极小多项式不等于其特征多项式, 则 $k \geq 2$. 构造分块对角矩阵

$$B = \text{diag}\{I_{n_1}, O_{n_2}, \dots, O_{n_k}\}$$

显然 $BF = FB$. 用反证法, 若存在多项式 $g(x)$, 使得 $B = g(F)$, 即

$$B = \text{diag}\{g(F_1), g(F_2), \dots, g(F_k)\}$$

则 $g(F_1) = I_{n_1}, g(F_i) = O (i \geq 2)$. 由于 $d_k(\lambda)$ 是 F_k 的极小多项式 (也是特征多项式), 故 $d_k(\lambda) \mid g(\lambda)$, 从而 $d_1(\lambda) \mid g(\lambda)$, 于是 $g(F_1) = O$, 矛盾! 因此 B 不能表示为 F 的多项式, 从而由 B 定义的线性变换 ψ 符合题目要求. \square

推论 1.13

设 φ 是数域 \mathbb{K} 上 n 维线性空间 V 上的线性变换, $\mathbb{K}[\varphi] = \{f(\varphi) \mid f(x) \in \mathbb{K}[x]\}$, $C(\varphi) = \{\psi \in \mathcal{L}(V) \mid \varphi\psi = \psi\varphi\}$, 则 V 是关于 φ 的循环空间的充要条件是 $C(\varphi) = \mathbb{K}[\varphi]$. 此时, $C(\varphi)$ 的一组基为 $\{I_V, \varphi, \dots, \varphi^{n-1}\}$. ♥



笔记 定理 1.12 证明了: 线性变换 φ 的极小多项式等于其特征多项式当且仅当 V 是关于 φ 的循环空间. 因此, 作为定理 1.24 的推论, 我们给出了循环空间的另一刻画.

证明

\square

例题 1.14 设数域 \mathbb{K} 上的 n 阶矩阵 A 的特征多项式 $f(\lambda) = P_1(\lambda)P_2(\lambda)\cdots P_k(\lambda)$, 其中 $P_i(\lambda) (1 \leq i \leq k)$ 是 \mathbb{K} 上互异的首一不可约多项式. 设 \mathbb{K} 上的 n 阶矩阵 B 满足 $AB = BA$, 求证: 存在 \mathbb{K} 上次数不超过 $n-1$ 的多项式 $f(x)$, 使得 $B = f(A)$.

证明 由定理 1.12 及命题 1.14 可知, \mathbb{K}^n 是关于 A 的循环空间, 再由定理 1.24 即得结论. \square

例题 1.15 设 A 是数域 \mathbb{K} 上的 2 阶矩阵, 试求 $C(A) = \{X \in M_2(\mathbb{K}) \mid AX = XA\}$.

证明 若 A 的极小多项式等于特征多项式, 则由定理 1.24 可知 $C(A) = \mathbb{K}[A]$. 若极小多项式不等于特征多项式, 则

极小多项式必为一次多项式 $x - c$, 从而 $A = cI_2$, 于是 $C(A) = M_2(\mathbb{K})$. □

例题 1.16 设数域 \mathbb{K} 上的 n 阶矩阵

$$A = \begin{pmatrix} a_1 & b_1 & 0 & \cdots & 0 \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ & * & & \ddots & b_{n-1} \\ & & & & a_n \end{pmatrix}$$

其中 b_1, \dots, b_{n-1} 均不为零. 记 $C(A) = \{X \in M_n(\mathbb{K}) \mid AX = XA\}$, 证明: 线性空间 $C(A)$ 的一组基为 $\{I_n, A, \dots, A^{n-1}\}$.


证明 题目中的 A 是类下三角矩阵, 上次对角元全部非零, 比如 Frobenius 块、Jordan 块和三对角矩阵都满足这样的特点. 考虑特征矩阵 $\lambda I_n - A$ 的前 $n-1$ 行、后 $n-1$ 列构成的下三角行列式, 其值为 $(-1)^{n-1}b_1 \cdots b_{n-1} \neq 0$, 故 A 的行列式因子组为 $1, \dots, 1, f(\lambda)$, 从而 \mathbb{K}^n 是关于 A 的循环空间, 再由定理 1.24 即得结论. □

命题 1.23

设有 n 阶分块对角矩阵

$$A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{pmatrix}, \quad B = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_k \end{pmatrix}$$

其中 A_i 和 B_i 是同阶方阵. 设 A_i 适合非零多项式 $g_i(x)$, 且 $g_i(x) (1 \leq i \leq k)$ 两两互素. 求证: 若对每个 i , 存在多项式 $f_i(x)$, 使得 $B_i = f_i(A_i)$, 则必存在次数不超过 $n-1$ 的多项式 $f(x)$, 使得 $B = f(A)$. ◆

 **笔记** 这个命题告诉我们, 在什么条件下可以将分块对角矩阵的多项式表示问题归结为对每一分块的讨论.

证明 因为 $g_i(x)$ 两两互素, 故由 theorem: 中国剩余定理可知, 存在多项式 $h(x)$ 满足 $h(x) = g_i(x)q_i(x) + f_i(x)$. 将 $x = A_i$ 代入上式, 可得 $h(A_i) = f_i(A_i) = B_i$, 从而

$$h(A) = \text{diag}\{h(A_1), \dots, h(A_k)\} = \text{diag}\{B_1, \dots, B_k\} = B$$

设 A 的特征多项式为 $g(x)$, 作带余除法 $h(x) = g(x)q(x) + f(x)$, 其中 $\deg f(x) < n$. 将 $x = A$ 代入上式, 则由 Cayley-Hamilton 定理可得 $B = h(A) = f(A)$. □

命题 1.24

设 n 阶矩阵 A 的秩等于 $n-1$, B 是同阶非零矩阵且 $AB = BA = O$, 求证: 存在次数不超过 $n-1$ 的多项式 $f(x)$, 使得 $B = f(A)$. ◆

注 对适合 $AB = BA$ 的矩阵, 由于 $AB = BA$ 当且仅当 $(P^{-1}AP)(P^{-1}BP) = (P^{-1}BP)(P^{-1}AP)$, 因此我们可以通过同时相似变换, 把问题归结为其中一个矩阵是相似标准型 (或分块对角型矩阵) 的情形来证明.

证明 证法一: 由于题目的条件和结论在同时相似变换 $A \mapsto P^{-1}AP, B \mapsto P^{-1}BP$ 下保持不变, 故不妨从一开始就假设 A 为 Jordan 标准型. 因为 $r(A) = n-1$, 故 A 关于特征值 0 的几何重数为 1, 从而属于特征值 0 的 Jordan 块只有一个, 记为 J_0 ; 将属于其他非零特征值的 Jordan 块合在一起, 记为 J_1 , 于是 $A = \text{diag}\{J_0, J_1\}$. 设 $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ 为相应的分块, 则由 $AB = BA = O$ 可得 B_{12}, B_{21}, B_{22} 都是零矩阵, 于是 $B = \text{diag}\{B_{11}, O\}$ 且 $J_0 B_{11} = B_{11} J_0$. 由于 J_0 是幂零矩阵而 J_1 是可逆矩阵, 故 J_0 的特征多项式 $g_0(x)$ 和 J_1 的特征多项式 $g_1(x)$ 互素; 又由 Jordan 块的性质 (5) 可知, 存在多项式 $f_0(x)$, 使得 $B_{11} = f_0(J_0)$; 再取 $f_1(x) = 0$, 则 $O = f_1(J_1)$; 最后由命题 1.23 即得结论.

证法二: 也可以用线性方程组的求解理论和极小多项式来做. 由于 $r(A) = n-1$, 故线性方程组 $Ax = 0$ 解空间的维数为 1, 再由 $AB = O$ 可知, B 的列向量都是解空间的向量, 从而它们成比例, 于是 $r(B) = 1$. 设 $B = \alpha\beta'$, 其中 α, β 为 n 维非零列向量, 由 $AB = O$ 可推出 $A\alpha = 0$, 即 α 是线性方程组 $Ax = 0$ 的基础解系. 同理, 由 $BA = O$

可推出 β 是 $A'x = 0$ 的基础解系. 设 $m(x)$ 是 A 的极小多项式, 由于 A 不是可逆矩阵, 故 $m(x)$ 的常数项等于零, 即 $m(x) = xg(x)$, 于是 $Ag(A) = O$ 但 $g(A) \neq O$. 由类似于矩阵 B 的讨论可得, $g(A)$ 也可以写为 $g(A) = \eta\xi'$, 其中 η 是 $Ax = 0$ 的解, 故 $\eta = k\alpha$. 同理可得 $\xi = t\beta$, 于是 $g(A) = ktB$, 从而 $B = \frac{1}{kt}g(A)$ 可表示为 A 的次数不超过 $n-1$ 的多项式. \square

1.12 可对角化的判断 (二)

1.12.1 极小多项式无重根

例题 1.17 求适合下列条件的 n 阶矩阵 A 的 Jordan 标准型:

$$(1) A^2 = A; \quad (2) A^k = I_n.$$

解 (1) 矩阵 A 适合 $g(x) = x^2 - x$ 且 $g(x)$ 无重根, 故由命题??可知 A 可对角化, 并且由命题??可知 A 的特征值也适合 $g(x)$, 故只能是 0, 1. 因此, A 的 Jordan 标准型为 $\text{diag}\{1, \dots, 1, 0, \dots, 0\}$, 其中有 $r(A)$ 个 1.

(2) 矩阵 A 适合 $g(x) = x^k - 1$ 且 $g(x)$ 无重根, 故由命题??可知 A 可对角化, 并且由命题??可知 A 的特征值也适合 $g(x)$, 故只能是 1 的 k 次方根. 因此, A 的 Jordan 标准型为 $\text{diag}\{\omega_1, \omega_2, \dots, \omega_n\}$, 其中 $\omega_i^k = 1 (1 \leq i \leq n)$. \square

例题 1.18 设 A 是有理数域上的 n 阶矩阵, 其特征多项式的所有不可约因式为 $\lambda^2 + \lambda + 1, \lambda^2 - 2$. 又 A 的极小多项式是四次多项式, 求证: A 在复数域上可对角化.

证明 因为 A 的极小多项式 $m(\lambda)$ 和特征多项式 $f(\lambda)$ 有相同的根 (不计重数), 且 $\deg m(\lambda) = 4$, 所以 $m(\lambda) = (\lambda^2 + \lambda + 1)(\lambda^2 - 2)$. 注意到 $m(\lambda)$ 在复数域内无重根, 故 A 在复数域上可对角化. \square

命题 1.25

设 φ 是复线性空间 V 上的线性变换, V_0 是 φ 的不变子空间. 求证: 若 φ 可对角化, 则 φ 在 V_0 上的限制变换和 φ 在 V/V_0 上的诱导变换都可对角化.

证明 证法一: 由命题??的几何版本可知, 限制变换 $\varphi|_{V_0}$ 和诱导变换 $\bar{\varphi}$ 都有完全的特征向量系, 从而可对角化.

证法二: 设线性变换 φ 、限制变换 $\varphi|_{V_0}$ 和诱导变换 $\bar{\varphi}$ 的极小多项式分别为 $m(\lambda)$, $g(\lambda)$ 和 $h(\lambda)$, 则容易验证 $\varphi|_{V_0}$ 和 $\bar{\varphi}$ 都适合多项式 $m(\lambda)$, 从而 $g(\lambda) | m(\lambda)$ 且 $h(\lambda) | m(\lambda)$. 由于 φ 可对角化, 故 $m(\lambda)$ 无重根, 从而 $g(\lambda), h(\lambda)$ 也无重根, 于是 $\varphi|_{V_0}$ 和 $\bar{\varphi}$ 都可对角化. \square

命题 1.26

设 φ 是 n 维复线性空间 V 上的线性变换, 求证: φ 可对角化的充要条件是对任一 φ -不变子空间 U , 均存在 φ -不变子空间 W , 使得 $V = U \oplus W$. 这样的 W 称为 U 的 φ -不变补空间.

证明 先证充分性: 假设 φ 不能对角化, 则 φ 只有 m 个线性无关的特征向量, 其中 $1 \leq m < n$. 设由这些特征向量张成的子空间为 U , 由条件可知, U 存在非零的 φ -不变补空间 W . 考虑限制变换 $\varphi|_W$, 它在 W 上必存在特征值和特征向量, 这些也是 φ 的特征值和特征向量, 于是 φ 有多于 m 个线性无关的特征向量, 矛盾!

再证必要性: 设 φ 可对角化, U 是 φ -不变子空间, 则由命题 1.25 可知, $\varphi|_U$ 仍可对角化, 故存在 U 的一组基 $\alpha_1, \dots, \alpha_r$, 它们是 $\varphi|_U$ 也是 φ 的线性无关的特征向量. 又因为 φ 可对角化, 故存在 n 个线性无关的特征向量 $\{e_1, e_2, \dots, e_n\}$, 再基底扩张定理可知, 可从这组基中取出 $n-r$ 个向量和 $\alpha_1, \dots, \alpha_r$ 一起组成 V 的一组基. 设这 $n-r$ 个向量张成的子空间为 W , 则 W 是 U 的 φ -不变补空间. \square

命题 1.27

设 n 阶矩阵 A 的极小多项式 $m(\lambda)$ 的次数为 s , $B = (b_{ij})$ 为 s 阶矩阵, 其中 $b_{ij} = \text{tr}(A^{i+j-2})$ (约定 $b_{11} = n$), 求证: A 可对角化的充要条件是 B 为可逆矩阵.

注 本题主要利用的方法是: 设矩阵 A 的全体不同特征值为 $\lambda_1, \lambda_2, \dots, \lambda_k$, 定义

$$g(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_k)$$

若 A 可对角化, 则 A 的极小多项式就是 $g(\lambda)$ (参考常见矩阵的极小多项式 (2)). 反之, 若 A 适合多项式 $g(\lambda)$, 则由极小多项式的性质可知, $g(\lambda)$ 就是 A 的极小多项式. 特别地, 由于 $g(\lambda)$ 无重根, 故 A 可对角化.

证明 设 A 的全体不同特征值为 $\lambda_1, \lambda_2, \dots, \lambda_k$, 其代数重数分别为 m_1, m_2, \dots, m_k , 则 $\text{tr}(A^i) = m_1\lambda_1^i + m_2\lambda_2^i + \dots + m_k\lambda_k^i$. 定义 $g(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_k)$, 则 $g(\lambda) \mid m(\lambda)$, 从而 $s \geq k$. 若 A 可对角化, 则 $m(\lambda) = g(\lambda)$, 从而 $s = k$. 若 A 不可对角化, 则 $m(\lambda)$ 有重根, 从而 $s > k$. 考虑矩阵 B 的如下分解:

$$B = \begin{pmatrix} m_1 & m_2 & \cdots & m_k \\ m_1\lambda_1 & m_2\lambda_2 & \cdots & m_k\lambda_k \\ \vdots & \vdots & & \vdots \\ m_1\lambda_1^{s-1} & m_2\lambda_2^{s-1} & \cdots & m_k\lambda_k^{s-1} \end{pmatrix} \begin{pmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{s-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{s-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_k & \cdots & \lambda_k^{s-1} \end{pmatrix}$$

其中上式右边第一个矩阵是 $s \times k$ 矩阵, 第二个矩阵是 $k \times s$ 矩阵. 若 $s = k$, 则由 Vandermonde 行列式可知

$$|B| = m_1 m_2 \cdots m_k \prod_{1 \leq i < j \leq k} (\lambda_i - \lambda_j)^2 \neq 0$$

即 B 是可逆矩阵. 若 $s > k$, 则由 Cauchy - Binet 公式可得 $|B| = 0$, 即 B 不可逆. □

1.12.2 初等因子都是一次多项式, 或 Jordan 块都是一阶矩阵

命题 1.28

设 n 阶复方阵 A 的特征多项式为 $f(\lambda)$, 复系数多项式 $g(\lambda)$ 满足 $(f(\lambda), g'(\lambda)) = 1$. 证明: A 可对角化的充要条件是 $g(A)$ 可对角化.

证明 必要性显然成立, 下证充分性. 用反证法, 设 A 不可对角化, 则存在可逆矩阵 P , 使得 $P^{-1}AP = J = \text{diag}\{J_{r_1}(\lambda_1), \dots, J_{r_k}(\lambda_k)\}$ 为 Jordan 标准型, 其中 $r_1 > 1$. 注意到

$$P^{-1}g(A)P = g(P^{-1}AP) = g(J) = \text{diag}\{g(J_{r_1}(\lambda_1)), \dots, g(J_{r_k}(\lambda_k))\}$$

其中

$$g(J_{r_1}(\lambda_1)) = \begin{pmatrix} g(\lambda_1) & g'(\lambda_1) & \cdots & * \\ & g(\lambda_1) & \ddots & \vdots \\ & & \ddots & g'(\lambda_1) \\ & & & g(\lambda_1) \end{pmatrix}$$

由 $(f(\lambda), g'(\lambda)) = 1$ 可知 $g'(\lambda_1) \neq 0$, 于是 $g(J_{r_1}(\lambda_1))$ 的特征值全为 $g(\lambda_1)$, 其几何重数为 $r_1 - \text{r}(g(J_{r_1}(\lambda_1)) - g(\lambda_1)I_{r_1}) = 1$, 因此 $g(J_{r_1}(\lambda_1))$ 的 Jordan 标准型为 $J_{r_1}(g(\lambda_1))$, 其阶数 $r_1 > 1$. 由于 $J_{r_1}(g(\lambda_1))$ 也是 $g(A)$ 的一个 Jordan 块, 故 $g(A)$ 不可对角化, 矛盾! □

命题 1.29

设 φ 是 n 维复线性空间 V 上的线性变换, 求证: φ 可对角化的充要条件是对 φ 的任一特征值 λ_0 , 总有 $\text{Ker}(\varphi - \lambda_0 I_V) \cap \text{Im}(\varphi - \lambda_0 I_V) = 0$.

注 这个命题 1.29 是这个命题 1.31 的特例.

证明 先证必要性: 若 φ 可对角化, 则存在一组基 $\{e_1, e_2, \dots, e_n\}$, 使得 φ 在这组基下的表示矩阵为 $\text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$. 适当调整基向量的顺序, 不妨设 $\lambda_0 = \lambda_1 = \dots = \lambda_r, \lambda_0 \neq \lambda_j (j > r)$, 则容易验证 $\text{Ker}(\varphi - \lambda_0 I_V) = L(e_1, \dots, e_r)$, $\text{Im}(\varphi - \lambda_0 I_V) = L(e_{r+1}, \dots, e_n)$, 从而 $\text{Ker}(\varphi - \lambda_0 I_V) \cap \text{Im}(\varphi - \lambda_0 I_V) = 0$.

再证充分性: 用反证法, 设 φ 不可对角化, 则存在 V 的一组基 $\{e_1, e_2, \dots, e_n\}$, 使得 φ 在这组基下的表示矩阵为 Jordan 标准型 $J = \text{diag}\{J_{r_1}(\lambda_1), \dots, J_{r_k}(\lambda_k)\}$, 其中 $r_1 > 1$. 由表示矩阵的定义可得 $\varphi(e_1) = \lambda_1 e_1, \varphi(e_2) = e_1 + \lambda_1 e_2$, 于是 $(\varphi - \lambda_1 I_V)(e_1) = 0, (\varphi - \lambda_1 I_V)(e_2) = e_1$, 从而 $0 \neq e_1 \in \text{Ker}(\varphi - \lambda_1 I_V) \cap \text{Im}(\varphi - \lambda_1 I_V)$, 这与假设矛盾. □

命题 1.30

求证: n 阶复矩阵 A 可对角化的充要条件是对 A 的任一特征值 $\lambda_0, (\lambda_0 I_n - A)^2$ 和 $\lambda_0 I_n - A$ 的秩相同.

注 这个命题 1.30 是这个命题 1.31 的特例.

证明 先证必要性: 若 A 可对角化, 则存在可逆矩阵 P , 使得 $P^{-1}AP = \Lambda = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$. 适当调整 P 的列向量的顺序, 不妨设 $\lambda_0 = \lambda_1 = \dots = \lambda_r, \lambda_0 \neq \lambda_j (j > r)$, 则 $r(\lambda_0 I_n - A) = r(\lambda_0 I_n - \Lambda) = n - r, r((\lambda_0 I_n - A)^2) = r((\lambda_0 I_n - \Lambda)^2) = n - r$, 于是结论成立.

再证充分性: 用反证法, 若 A 不可对角化, 则存在可逆矩阵 P , 使得 $P^{-1}AP = J = \text{diag}\{J_{r_1}(\lambda_1), \dots, J_{r_k}(\lambda_k)\}$ 为 Jordan 标准型, 其中 $r_1 > 1$. 注意到

$$r((\lambda_1 I_n - A)^j) = r((\lambda_1 I_n - J)^j) = \sum_{i=1}^k r((\lambda_1 I_{r_i} - J_{r_i}(\lambda_i))^j), j \geq 1$$

又 $r(\lambda_1 I_{r_1} - J_{r_1}(\lambda_1)) = r_1 - 1, r((\lambda_1 I_{r_1} - J_{r_1}(\lambda_1))^2) = r_1 - 2$, 因此 $r((\lambda_1 I_n - A)^2) < r(\lambda_1 I_n - A)$, 这与假设矛盾. \square

命题 1.31

设 φ 是 n 维复线性空间 V 上的线性变换, 求证: φ 可对角化的充要条件是对 φ 的任一特征值 λ_0 , 下列条件之一成立:

- (1) $V = \text{Ker}(\varphi - \lambda_0 I_V) + \text{Im}(\varphi - \lambda_0 I_V)$;
- (2) $V = \text{Ker}(\varphi - \lambda_0 I_V) \oplus \text{Im}(\varphi - \lambda_0 I_V)$;
- (3) $\text{Ker}(\varphi - \lambda_0 I_V) \cap \text{Im}(\varphi - \lambda_0 I_V) = 0$;
- (4) $\dim \text{Ker}(\varphi - \lambda_0 I_V) = \dim \text{Ker}(\varphi - \lambda_0 I_V)^2$;
- (5) $\text{Ker}(\varphi - \lambda_0 I_V) = \text{Ker}(\varphi - \lambda_0 I_V)^2 = \text{Ker}(\varphi - \lambda_0 I_V)^3 = \dots$;
- (6) $r(\varphi - \lambda_0 I_V) = r((\varphi - \lambda_0 I_V)^2)$;
- (7) $\text{Im}(\varphi - \lambda_0 I_V) = \text{Im}(\varphi - \lambda_0 I_V)^2 = \text{Im}(\varphi - \lambda_0 I_V)^3 = \dots$;
- (8) $\text{Ker}(\varphi - \lambda_0 I_V)$ 存在 φ -不变补空间, 即存在 φ -不变子空间 U , 使得 $V = \text{Ker}(\varphi - \lambda_0 I_V) \oplus U$;
- (9) $\text{Im}(\varphi - \lambda_0 I_V)$ 存在 φ -不变补空间, 即存在 φ -不变子空间 W , 使得 $V = \text{Im}(\varphi - \lambda_0 I_V) \oplus W$.

注 命题 1.29 与命题 1.30 都是这个命题 1.31 的特例.



笔记 由命题??可知条件 (1)~(9) 是相互等价的, 因此本题的结论由命题 1.29 (与条件 (3) 对应) 或命题 1.30 (与条件 (6) 对应) 即得. 事实上, 对充分性而言, 我们还可以从其他条件出发来证明 φ 可对角化, 下面是 3 种证法.

证明 证法一: 对任一特征值 λ_0 , 由 $\text{Ker}(\varphi - \lambda_0 I_V) = \text{Ker}(\varphi - \lambda_0 I_V)^2 = \dots = \text{Ker}(\varphi - \lambda_0 I_V)^n$, 取维数之后可得特征值 λ_0 的几何重数等于代数重数, 从而 φ 有完全的特征向量系, 于是 φ 可对角化.

证法二: 对任一特征值 λ_0 , 由 $\text{Ker}(\varphi - \lambda_0 I_V) = \text{Ker}(\varphi - \lambda_0 I_V)^2 = \dots = \text{Ker}(\varphi - \lambda_0 I_V)^n$ 可知, 特征子空间等于根子空间, 再由根子空间的直和分解可知, 全空间等于特征子空间的直和, 从而 φ 可对角化.

证法三: 设 φ 的全体不同特征值为 $\lambda_1, \lambda_2, \dots, \lambda_k$, 特征多项式 $f(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \dots (\lambda - \lambda_k)^{m_k}$, 则对任意的 $\alpha \in V$, 由 Cayley - Hamilton 定理可得

$$(\varphi - \lambda_1 I_V)^{m_1}(\varphi - \lambda_2 I_V)^{m_2} \dots (\varphi - \lambda_k I_V)^{m_k}(\alpha) = 0,$$

即有 $(\varphi - \lambda_2 I_V)^{m_2} \dots (\varphi - \lambda_k I_V)^{m_k}(\alpha) \in \text{Ker}(\varphi - \lambda_1 I_V)^{m_1} = \text{Ker}(\varphi - \lambda_1 I_V)$, 从而

$$(\varphi - \lambda_1 I_V)(\varphi - \lambda_2 I_V)^{m_2} \dots (\varphi - \lambda_k I_V)^{m_k}(\alpha) = 0.$$

不断这样做下去, 最终可得对任意的 $\alpha \in V$, 总有

$$(\varphi - \lambda_1 I_V)(\varphi - \lambda_2 I_V) \dots (\varphi - \lambda_k I_V)(\alpha) = 0,$$

即 φ 适合多项式 $g(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_k)$, 从而 φ 可对角化. \square

例题 1.19 若 $n(n \geq 2)$ 阶矩阵 B 相似于 $R = \text{diag}\left\{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, I_{n-2}\right\}$, 则称 B 为反射矩阵. 证明: 任一对合矩阵 A (即 $A^2 = I_n$) 均可分解为至多 n 个两两乘法可交换的反射矩阵的乘积.

证明 由例 1.17 可知, 对合矩阵 A 可对角化, 即存在可逆矩阵 P , 使得 $P^{-1}AP = \text{diag}\{-I_r, I_{n-r}\}$, 其中 $0 \leq r \leq n$. 当 $r = 0$ 时, $A = I_n = R^2$, 结论成立. 当 $r \geq 1$ 时, 设 $B_i = P \text{diag}\{1, \dots, 1, -1, 1, \dots, 1\} P^{-1}$, 其中 -1 在主对角线上的第 i 个位置, 则 $B_i (1 \leq i \leq r)$ 两两乘法可交换, 并且 $A = B_1 B_2 \cdots B_r$. 由于 $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ 的特征值是 $-1, 1$, 故其相似于 $\text{diag}\{-1, 1\}$, 因此矩阵 B 是反射矩阵当且仅当 B 相似于 $\text{diag}\{-1, 1, \dots, 1\}$. 因为对角矩阵的两个主对角元素交换是一个相似变换, 所以上述 B_i 都是反射矩阵, 于是 A 可以分解为 r 个两两乘法可交换的反射矩阵的乘积. \square

1.13 Jordan 标准型的求法

分析矩阵结构的方法:

- (1) 计算行列式因子对于某些具有简单结构的矩阵 (如上 (下) 三角矩阵、类上 (下) 三角矩阵), 可以通过选取适当的子式, 计算出行列式因子, 再得到不变因子和初等因子. 比如, Frobenius 块和 Jordan 块就是利用这种方法的典型例子.
- (2) 计算极小多项式因为矩阵的极小多项式是整除关系下最大的不变因子, 所以极小多项式确定了最大 Jordan 块的阶数.
- (3) 计算特征值的几何重数因为特征值的几何重数等于其 Jordan 块的个数, 所以计算几何重数有助于 Jordan 标准型的确定.

命题 1.32

设 n 阶矩阵 A 的不变因子组为 $d_1(\lambda), d_2(\lambda), \dots, d_n(\lambda)$, 其中 $d_i(\lambda) \mid d_{i+1}(\lambda) (1 \leq i \leq n-1)$, 又 λ_0 是 A 的特征值. 求证: $r(\lambda_0 I_n - A) = r$ 的充要条件是 $(\lambda - \lambda_0) \nmid d_r(\lambda)$ 但 $(\lambda - \lambda_0) \mid d_{r+1}(\lambda)$.

证明 证法一: $r(\lambda_0 I_n - A) = r$ 当且仅当特征值 λ_0 的几何重数为 $n - r$; 这当且仅当特征值 λ_0 的 Jordan 块有 $n - r$ 个; 由不变因子之间的整除关系可知, 这当且仅当后 $n - r$ 个不变因子能被 $\lambda - \lambda_0$ 整除, 而前 r 个不变因子不能被 $\lambda - \lambda_0$ 整除.

证法二: 由命题 1.20 可知, $r(\lambda_0 I_n - A) = r$ 当且仅当 $\sum_{i=1}^n \delta_{d_i(\lambda_0), 0} = n - r$; 由不变因子之间的整除关系可知, 这当且仅当 $d_i(\lambda_0) \neq 0 (1 \leq i \leq r)$ 且 $d_i(\lambda_0) = 0 (r + 1 \leq i \leq n)$; 最后由余数定理即得结论. \square

命题 1.33

设 φ 是 n 维线性空间 V 上的线性变换, U 是 V 的非零 φ -不变子空间. 设 λ_0 是限制变换 $\varphi|_U$ 的特征值, 证明: $\varphi|_U$ 的属于特征值 λ_0 的 Jordan 块的个数不超过 φ 的属于特征值 λ_0 的 Jordan 块的个数.

证明 Jordan 块的个数等于特征值的几何重数, 即线性无关的特征向量的个数. 设 $\varphi|_U$ 的属于特征值 λ_0 的 Jordan 块的个数为 r , 则 $\varphi|_U$ 关于特征值 λ_0 有 r 个线性无关的特征向量, 它们也都是 φ 关于特征值 λ_0 的线性无关的特征向量, 从而 φ 的属于特征值 λ_0 的 Jordan 块至少有 r 个. 也可用纯代数的方法 (矩阵的秩) 进行证明, 请读者自行思考完成. \square

例题 1.20 求下列 n 阶矩阵的 Jordan 标准型, 其中 $a \neq 0$:

$$A = \begin{pmatrix} a & a & a & \cdots & a \\ & a & a & \cdots & a \\ & & a & \cdots & a \\ & & & \ddots & \vdots \\ & & & & a \end{pmatrix}$$

注 本题同时利用分析矩阵结构的三种方法计算其 Jordan 标准型.

解 解法一: 由命题 1.10 可知, A 的行列式因子组为 $1, \dots, 1, (\lambda - a)^n$, 这也是 A 的不变因子组, 从而 A 的 Jordan 标准型为 $J_n(a)$.

解法二:显然 A 的特征多项式为 $(\lambda - a)^n$, 故 A 的极小多项式是 $\lambda - a$ 的某个幂. 设 $N = J_n(0)$, 即特征值为 0 的 n 阶 Jordan 块, 它满足 $N^{n-1} \neq O$ 但 $N^n = O$, 则 $A = a(I_n + N + N^2 + \cdots + N^{n-1})$. 注意到

$$(A - aI_n)^{n-1} = a^{n-1}(N + N^2 + \cdots + N^{n-1})^{n-1} = a^{n-1}N^{n-1} \neq O$$

故 A 不适合多项式 $(\lambda - a)^{n-1}$, 于是 A 的极小多项式只能是 $(\lambda - a)^n$. 因此 A 的不变因子组是 $1, \cdots, 1, (\lambda - a)^n$, 从而 A 的 Jordan 标准型为 $J_n(a)$.

解法三:显然 A 的特征值全为 a , 我们来计算它的几何重数. 注意到 $r(aI_n - A) = n - 1$, 故特征值 a 的几何重数为 $n - r(aI_n - A) = 1$, 于是 A 的 Jordan 标准型中关于特征值 a 的 Jordan 块只有一个, 因此 A 的 Jordan 标准型为 $J_n(a)$. \square

命题 1.34 (秩一阵的 Jordan 标准型)

设 $n(n > 1)$ 阶矩阵 A 的秩为 1, 试求 A 的 Jordan 标准型.

解 解法一:由 $r(A) = 1$ 可知, 存在非零列向量 α, β , 使得 $A = \alpha\beta'$. 由特征值的降价公式可得 $|\lambda I_n - A| = \lambda^{n-1}(\lambda - \beta'\alpha)$, 再由所有特征值之和等于矩阵的迹可得 $\text{tr}(A) = \beta'\alpha$. 若 $\text{tr}(A) \neq 0$, 则特征值 $\text{tr}(A)$ 的几何重数等于 1, 特征值 0 的几何重数等于 $n - r(A) = n - 1$, 因此 A 的 Jordan 标准型为 $\text{diag}\{0, \cdots, 0, \text{tr}(A)\}$. 若 $\text{tr}(A) = 0$, 则特征值 0 的代数重数是 n , 几何重数是 $n - 1$, 因此 A 的 Jordan 标准型为 $\text{diag}\{0, \cdots, 0, J_2(0)\}$.

解法二:特征多项式的计算同解法 1, 又由常见矩阵的极小多项式 (4) 可知, A 的极小多项式 $m(\lambda) = \lambda(\lambda - \text{tr}(A))$, 于是 A 的不变因子组为 $1, \lambda, \cdots, \lambda, m(\lambda)$. 若 $\text{tr}(A) \neq 0$, 则 A 的 Jordan 标准型为 $\text{diag}\{0, \cdots, 0, \text{tr}(A)\}$. 若 $\text{tr}(A) = 0$, 则 A 的 Jordan 标准型为 $\text{diag}\{0, \cdots, 0, J_2(0)\}$.

解法三:直接利用 Jordan 标准型来解最为简单. 设 A 的 Jordan 标准型 $J = \text{diag}\{J_{r_1}(0), \cdots, J_{r_k}(0), J_{s_1}(\lambda_1), \cdots, J_{s_l}(\lambda_l)\}$, 其中 $\lambda_j \neq 0 (1 \leq j \leq l)$. 由于相似关系不改变矩阵的秩, 故 J 的秩也为 1, 即有 $(r_1 - 1) + \cdots + (r_k - 1) + s_1 + \cdots + s_l = 1$. 于是只有以下两种情况成立: 第一种情况是 $l = 1, s_1 = 1, \lambda_1 = \text{tr}(A) \neq 0$, 且所有的 $r_i = 1$, 此时 A 的 Jordan 标准型为 $\text{diag}\{0, \cdots, 0, \text{tr}(A)\}$. 第二种情况是某个 $r_i = 2$, 其余的 $r_i = 1$ 且 $l = 0$, 此时 A 的 Jordan 标准型为 $\text{diag}\{0, \cdots, 0, J_2(0)\}$. \square

命题 1.35

设 $n(n > 1)$ 阶矩阵 A 的秩为 1, 求证:

- (1) A 是幂等矩阵的充要条件是 $\text{tr}(A) = 1$,
- (2) A 是幂零矩阵的充要条件是 $\text{tr}(A) = 0$.

证明 由命题 1.34 的证明过程即得结论. \square

例题 1.21 设 $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ a+2 & 1 & 0 & 0 \\ 5 & 3 & 1 & 0 \\ 7 & 6 & b+4 & 1 \end{pmatrix}$, 求 A 的 Jordan 标准型.

解 显然 A 的特征值全为 1, 首先我们来计算特征值 1 的几何重数. 考虑矩阵

$$A - I_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ a+2 & 0 & 0 & 0 \\ 5 & 3 & 0 & 0 \\ 7 & 6 & b+4 & 0 \end{pmatrix}.$$

(1) 当 $a+2 \neq 0$ 且 $b+4 \neq 0$ 时, $r(A - I_4) = 3$, 于是特征值 1 的几何重数等于 1, 从而只有一个 Jordan 块, 因此 A 的 Jordan 标准型是 $J_4(1)$.

(2) 当 $a+2 = 0$ 或 $b+4 = 0$ 时, $r(A - I_4) = 2$, 于是特征值 1 的几何重数等于 2, 从而有两个 Jordan 块. 进一步我们来计算 A 的极小多项式.

(2.1) 若 $a+2 = 0$ 和 $b+4 = 0$ 中只有一个成立, 容易验证 $(A - I_4)^2 \neq O$, 但 $(A - I_4)^3 = O$, 于是 A 的极小多项式是 $(\lambda - 1)^3$, 从而不变因子组为 $1, 1, \lambda - 1, (\lambda - 1)^3$, 因此 A 的 Jordan 标准型为 $\text{diag}\{1, J_3(1)\}$.

(2.2) 若 $a+2=0$ 和 $b+4=0$ 都成立, 容易验证 $(A-I_4)^2=O$, 于是 A 的极小多项式是 $(\lambda-1)^2$, 从而不变因子组为 $1, 1, (\lambda-1)^2, (\lambda-1)^2$, 因此 A 的 Jordan 标准型为 $\text{diag}\{J_2(1), J_2(1)\}$. \square

命题 1.36

设 $J = J_n(0)$ 是特征值为零的 $n(n \geq 2)$ 阶 Jordan 块, 求 J^2 的 Jordan 标准型.

解 显然 J^2 的特征值全为 0 且 $r(J^2) = n-2$, 于是特征值 0 的几何重数等于 2, 从而有两个 Jordan 块. 接下去计算 J^2 的极小多项式, 注意到 $J^n = O, J^{n-1} \neq O$.

(1) 当 $n = 2m$ 时, λ^m 是 J^2 的极小多项式, 于是 J^2 的不变因子组为 $1, \dots, 1, \lambda^m, \lambda^m$, 因此 J^2 的 Jordan 标准型为 $\text{diag}\{J_m(0), J_m(0)\}$.

(2) 当 $n = 2m+1$ 时, λ^{m+1} 是 J^2 的极小多项式, 于是 J^2 的不变因子组为 $1, \dots, 1, \lambda^m, \lambda^{m+1}$, 因此 J^2 的 Jordan 标准型为 $\text{diag}\{J_m(0), J_{m+1}(0)\}$.

另外, 也可以用行列式因子的讨论来替代几何重数的讨论. 注意到 $\lambda I_n - J^2$ 的右上角有一个 $n-2$ 阶子式等于 $(-1)^{n-2}$, 故 J^2 的 $n-2$ 阶行列式因子为 1, 从而前 $n-2$ 个不变因子都是 1, 后面再用极小多项式的讨论即可得到结论. \square

例题 1.22 求下列 $n(n \geq 2)$ 阶矩阵的 Jordan 标准型:

$$A = \begin{pmatrix} c & 0 & 1 & 0 & \cdots & 0 \\ c & 0 & 1 & \cdots & 0 & \\ & \ddots & \ddots & \ddots & \vdots & \\ & & \ddots & \ddots & 1 & \\ & & & \ddots & 0 & \\ & & & & & c \end{pmatrix}.$$

解 利用例题 1.36 的记号和结论, 显然 $A = cI_n + J^2$. 设 P 是可逆矩阵, 使得 $P^{-1}J^2P$ 是 J^2 的 Jordan 标准型, 则 $P^{-1}AP = cI_n + P^{-1}J^2P$ 就是 A 的 Jordan 标准型. 具体地, 当 $n = 2m$ 时, A 的 Jordan 标准型是 $\text{diag}\{J_m(c), J_m(c)\}$; 当 $n = 2m+1$ 时, A 的 Jordan 标准型是 $\text{diag}\{J_m(c), J_{m+1}(c)\}$. \square

我们可以自然地考虑如下问题: 如果已知 n 阶矩阵 A 的 Jordan 标准型, 那么对任意的正整数 m , A^m 的 Jordan 标准型应该有怎样的形状呢?(后续的命题 1.39 和命题 1.40 完美解决了这个问题) 首先, 我们可以把这个问题化约到 Jordan 块的情形. 设 A 的 Jordan 标准型为 $J = \text{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \dots, J_{r_s}(\lambda_s)\}$, 则 A^m 相似于 $J^m = \text{diag}\{J_{r_1}(\lambda_1)^m, J_{r_2}(\lambda_2)^m, \dots, J_{r_s}(\lambda_s)^m\}$, 因此要求 A^m 的 Jordan 标准型, 只要求每一个 $J_{r_i}(\lambda_i)^m$ 的 Jordan 标准型即可. 若 $\lambda_i \neq 0$, 则由例题 1.20 类似的讨论可知, $J_{r_i}(\lambda_i)^m$ 的 Jordan 标准型为 $J_{r_i}(\lambda_i^m)$. 若 $\lambda_i = 0$, 则例题 1.36 处理了 $m=2$ 的情形, 不过类似的讨论很难推广到 $m \geq 3$ 的情形, 换言之, 只依靠几何重数和极小多项式还不能完全确定 $J_{r_i}(0)^m$ 的 Jordan 标准型. 解决这个问题可以有代数和几何两种方法, 几何方法 (利用 Jordan 标准型的几何意义), 而代数方法 (利用矩阵的秩) 则需要下面的命题 1.37.

命题 1.37

设 λ_0 是 n 阶矩阵 A 的特征值, 证明: 对任意的正整数 k , 特征值为 λ_0 的 k 阶 Jordan 块 $J_k(\lambda_0)$ 在 A 的 Jordan 标准型 J 中出现的个数为

$$r((A - \lambda_0 I_n)^{k-1}) + r((A - \lambda_0 I_n)^{k+1}) - 2r((A - \lambda_0 I_n)^k),$$

其中约定 $r((A - \lambda_0 I_n)^0) = n$.

注 这个命题 1.37 告诉我们, n 阶矩阵 A 的 Jordan 标准型被若干个非负整数, 即 $\{r((A - \lambda_i I_n)^j) \mid \lambda_i \text{ 为 } A \text{ 的特征值}, 1 \leq j \leq n\}$ 完全决定. 因此从理论上说, 我们可以不计算矩阵 A 的不变因子或初等因子, 改为计算上述若干个矩阵的秩, 也可以求出 A 的 Jordan 标准型. 进一步, 我们还可以得到如下矩阵相似的判定准则.

证明 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. 注意到

$$(A - \lambda_0 I_n)^k = P \text{diag}\{J_{r_1}(\lambda_1 - \lambda_0)^k, J_{r_2}(\lambda_2 - \lambda_0)^k, \dots, J_{r_s}(\lambda_s - \lambda_0)^k\} P^{-1},$$

故 $r((A - \lambda_0 I_n)^k) = \sum_{i=1}^s r(J_{r_i}(\lambda_i - \lambda_0)^k)$. 当 $\lambda_i \neq \lambda_0$ 时, $r(J_{r_i}(\lambda_i - \lambda_0)^k) = r_i$. 当 $\lambda_i = \lambda_0$ 时, 若 $r_i < k$, 则 $r(J_{r_i}(\lambda_i - \lambda_0)^k) = 0$; 若 $r_i \geq k$, 则 $r(J_{r_i}(\lambda_i - \lambda_0)^k) = r_i - k$. 因此 $r((A - \lambda_0 I_n)^{k-1}) - r((A - \lambda_0 I_n)^k)$ 等于特征值为 λ_0 且阶数大于等于 k 的 Jordan 块的个数. 同理, $r((A - \lambda_0 I_n)^k) - r((A - \lambda_0 I_n)^{k+1})$ 等于特征值为 λ_0 且阶数大于等于 $k+1$ 的 Jordan 块的个数, 从而特征值为 λ_0 的 k 阶 Jordan 块 $J_k(\lambda_0)$ 在 A 的 Jordan 标准型 J 中出现的个数为

$$\begin{aligned} & (r((A - \lambda_0 I_n)^{k-1}) - r((A - \lambda_0 I_n)^k)) - (r((A - \lambda_0 I_n)^k) - r((A - \lambda_0 I_n)^{k+1})) \\ &= r((A - \lambda_0 I_n)^{k-1}) + r((A - \lambda_0 I_n)^{k+1}) - 2r((A - \lambda_0 I_n)^k). \end{aligned}$$

□

命题 1.38

设 A, B 为 n 阶矩阵, 证明: 它们相似的充要条件是对 A 或 B 的任一特征值 λ_0 以及任意的 $1 \leq k \leq n$, 有 $r((A - \lambda_0 I_n)^k) = r((B - \lambda_0 I_n)^k)$.

▲

证明 必要性显然, 现证充分性. 由已知条件及命题??可知,

$$r((A - \lambda_0 I_n)^{n+1}) = r((A - \lambda_0 I_n)^n) = r((B - \lambda_0 I_n)^n) = r((B - \lambda_0 I_n)^{n+1}).$$

因此由命题 1.37 可知, 特征值为 λ_0 的 k 阶 Jordan 块 $J_k(\lambda_0)$ 在 A, B 的 Jordan 标准型中出现的个数相同, 从而 A, B 有相同的 Jordan 标准型, 于是它们相似. □

命题 1.39

设 $J = J_n(a)$ 是特征值为 $a \neq 0$ 的 n 阶 Jordan 块, 求 J^m 的 Jordan 标准型, 其中 m 为非零整数.

▲

解 先处理 $m \geq 1$ 的情形, 采用几何重数的方法来做, 行列式因子和极小多项式的方法也可以做, 请读者自行补充完成. 显然 J^m 的所有特征值都为 a^m . 作分解 $J = aI_n + N$, 其中 $N = J_n(0)$, 则有

$$J^m = (aI_n + N)^m = a^m I_n + C_m^1 a^{m-1} N + \dots + N^m,$$

于是 $r(J^m - a^m I_n) = r(C_m^1 a^{m-1} N + \dots + N^m) = n - 1$, 从而特征值 a^m 的几何重数等于 1, 因此 J^m 的 Jordan 标准型中只有一个 Jordan 块, 即 J^m 的 Jordan 标准型为 $J_n(a^m)$.

再处理 $m = -1$ 的情形. 显然 J^{-1} 的所有特征值都为 a^{-1} . 注意到

$$J^{-1} = (aI_n + N)^{-1} = a^{-1} I_n - a^{-2} N + \dots + (-1)^{n-1} a^{-n} N^{n-1},$$

故 $r(J^{-1} - a^{-1} I_n) = r(-a^{-2} N + \dots + (-1)^{n-1} a^{-n} N^{n-1}) = n - 1$, 从而特征值 a^{-1} 的几何重数等于 1, 因此 J^{-1} 的 Jordan 标准型中只有一个 Jordan 块, 即 J^{-1} 的 Jordan 标准型为 $J_n(a^{-1})$.

最后处理 $m \leq -1$ 的情形. 注意到 $J^m = (J^{-1})^{-m}$, 故由前面两个结论即得 J^m 的 Jordan 标准型为 $J_n((a^{-1})^{-m}) = J_n(a^m)$. □

命题 1.40

设 $J = J_n(0)$ 是特征值为零的 n 阶 Jordan 块, 求 $J^m (m \geq 1)$ 的 Jordan 标准型.

▲

注 这个命题是命题 1.36 的推广.

解 若 $m \geq n$, 则 $J^m = O$, 这就是它的 Jordan 标准型. 下设 $m < n$, 并作带余除法: $n = mq + r$, 其中 $0 \leq r < m$. 我们先来计算 J^m 的幂的秩, 再利用命题 1.37 来计算 Jordan 块的个数. 注意到

$$r((J^m)^k) = n - mk, \quad 0 \leq k \leq q; \quad r((J^m)^k) = 0, \quad k \geq q + 1.$$

(1) 当 $1 \leq k < q$ 时, $J_k(0)$ 的个数为 $r((J^m)^{k-1}) + r((J^m)^{k+1}) - 2r((J^m)^k) = (n - m(k-1)) + (n - m(k+1)) - 2(n - mk) = 0$;

(2) $J_q(0)$ 的个数为 $r((J^m)^{q-1}) + r((J^m)^{q+1}) - 2r((J^m)^q) = (n - m(q-1)) + 0 - 2(n - mq) = m - r$;

(3) $J_{q+1}(0)$ 的个数为 $r((J^m)^q) + r((J^m)^{q+2}) - 2r((J^m)^{q+1}) = (n - mq) + 0 - 0 = r$;

(4) 当 $k > q + 1$ 时, $J_k(0)$ 的个数为 0.

因此 J^m 的 Jordan 标准型为 $\text{diag}\{J_q(0), \dots, J_q(0), J_{q+1}(0), \dots, J_{q+1}(0)\}$, 其中有 $m - r$ 个 $J_q(0)$, r 个 $J_{q+1}(0)$. \square

命题 1.41

设 m 阶矩阵 A 与 n 阶矩阵 B 没有公共的特征值, 且 A, B 的 Jordan 标准型分别为 J_1, J_2 , 又 C 为 $m \times n$ 矩阵, 求证: $M = \begin{pmatrix} A & C \\ O & B \end{pmatrix}$ 的 Jordan 标准型为 $\text{diag}\{J_1, J_2\}$.

注 这个命题可用来化简矩阵, 消去其非主对角块, 使其剩下低阶的主对角块.

证明 证法一: 设 $P_1(\lambda), P_2(\lambda), Q_1(\lambda), Q_2(\lambda)$ 是可逆 λ -矩阵, 使得

$$P_1(\lambda)(\lambda I_m - A)Q_1(\lambda) = \Lambda_1 = \text{diag}\{f_1(\lambda), f_2(\lambda), \dots, f_m(\lambda)\},$$

$$P_2(\lambda)(\lambda I_n - B)Q_2(\lambda) = \Lambda_2 = \text{diag}\{g_1(\lambda), g_2(\lambda), \dots, g_n(\lambda)\}$$

分别是 A, B 的法式. 考虑如下 λ -矩阵的初等变换:

$$\begin{pmatrix} P_1 & O \\ O & P_2 \end{pmatrix} \begin{pmatrix} \lambda I_m - A & -C \\ O & \lambda I_n - B \end{pmatrix} \begin{pmatrix} Q_1 & O \\ O & Q_2 \end{pmatrix} = \begin{pmatrix} \Lambda_1 & D \\ O & \Lambda_2 \end{pmatrix},$$

其中 $D = -P_1 C Q_2 = (d_{ij}(\lambda))$ 是 $m \times n$ λ -矩阵. 由于 A, B 没有公共的特征值, 故对任意的 $1 \leq i \leq m, 1 \leq j \leq n$, $(f_i(\lambda), g_j(\lambda)) = 1$, 从而存在 $u_{ij}(\lambda), v_{ij}(\lambda)$, 使得 $f_i(\lambda)u_{ij}(\lambda) + g_j(\lambda)v_{ij}(\lambda) = 1$. 将 λ -矩阵 $\begin{pmatrix} \Lambda_1 & D \\ O & \Lambda_2 \end{pmatrix}$ 的第 i 列乘以 $-u_{ij}(\lambda)d_{ij}(\lambda)$ 加到第 $m + j$ 列上, 再将第 $m + j$ 行乘以 $-v_{ij}(\lambda)d_{ij}(\lambda)$ 加到第 i 行上, 则可以消去 D 的第 (i, j) 元素, 因此 M 的特征矩阵相抵于对角矩阵 $\text{diag}\{\Lambda_1, \Lambda_2\}$. 再由 λ -矩阵和初等因子的基本性质 (2) 可知, M 的初等因子组是 $f_1(\lambda), \dots, f_m(\lambda), g_1(\lambda), \dots, g_n(\lambda)$ 的准素因子组, 而 $f_1(\lambda), \dots, f_m(\lambda)$ 的准素因子组是 A 的初等因子组, $g_1(\lambda), \dots, g_n(\lambda)$ 的准素因子组是 B 的初等因子组, 因此 M 的初等因子组是 A, B 的初等因子组的无交并集, 于是 M 的 Jordan 标准型为 $\text{diag}\{J_1, J_2\}$.

证法二: 由命题 ?? 可知, 矩阵方程 $AX - XB = C$ 存在唯一解 $X = X_0$. 考虑如下相似变换:

$$\begin{pmatrix} I_m & X_0 \\ O & I_n \end{pmatrix} \begin{pmatrix} A & C \\ O & B \end{pmatrix} \begin{pmatrix} I_m & -X_0 \\ O & I_n \end{pmatrix} = \begin{pmatrix} A & -AX_0 + X_0B + C \\ O & B \end{pmatrix} = \begin{pmatrix} A & O \\ O & B \end{pmatrix},$$

因此 M 的 Jordan 标准型为 $\text{diag}\{J_1, J_2\}$. \square

例题 1.23 设 $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ b & a+1 & 0 & 0 \\ 3 & b & 2 & 0 \\ 5 & 4 & a & 2 \end{pmatrix}$, 求 A 的 Jordan 标准型.

解 显然, A 的特征值为 $1, a+1, 2, 2$. 对 A 进行分块 $A = \begin{pmatrix} A_{11} & O \\ A_{21} & A_{22} \end{pmatrix}$, 其中所有的分块都是二阶方阵. 下面按 $a+1$ 是否等于 $1, 2$ 进行分类讨论.

(1) 若 $a \neq 0$ 及 $a \neq 1$, 则可有两种方法来处理. **命题 1.41 证法一**(几何重数): 经计算可知特征值 2 的几何重数等于 1 , 因此 A 的 Jordan 标准型为 $\text{diag}\{1, a+1, J_2(2)\}$. **命题 1.41 证法二:** 显然 A_{11} 可对角化, A_{22} 不可对角化, 且 A_{11}, A_{22} 无公共特征值, 故可消去 A_{21} , 因此 A 的 Jordan 标准型为 $\text{diag}\{1, a+1, J_2(2)\}$.

(2) 若 $a = 0$ 及 $b \neq 0$, 则利用 **命题 1.41 证法二** 可得, A 的 Jordan 标准型为 $\text{diag}\{J_2(1), 2, 2\}$.

(3) 若 $a = 0$ 及 $b = 0$, 则利用 **命题 1.41 证法二** 可得, A 的 Jordan 标准型为 $\text{diag}\{1, 1, 2, 2\}$.

(4) 若 $a = 1$ 及 $b \neq 0$, 则利用 **命题 1.41 证法一**(几何重数) 可得, A 的 Jordan 标准型为 $\text{diag}\{1, J_3(2)\}$.

(5) 若 $a = 1$ 及 $b = 0$, 则利用 **命题 1.41 证法一**(几何重数) 可得, A 的 Jordan 标准型为 $\text{diag}\{1, 2, J_2(2)\}$. \square

1.14 过渡矩阵的求法

1.14.1 计算特征矩阵之间的相抵变换

命题 1.42

设 A 是 n 阶数字矩阵, $P(\lambda)$ 及 $Q(\lambda)$ 是同阶可逆 λ -矩阵, 且

$$Q(\lambda)(\lambda I_n - A)P(\lambda) = \lambda I_n - J,$$

其中 J 是 A 的 Jordan 标准型. 又

$$P(\lambda) = T(\lambda)(\lambda I_n - J) + P,$$

其中 P 是数字矩阵, 求证: $P^{-1}AP = J$.

注 由??可知, 两个数字矩阵相似当且仅当它们的特征矩阵作为 λ -矩阵相抵.

证明 由已知可得 $(\lambda I_n - A)P(\lambda) = Q(\lambda)^{-1}(\lambda I_n - J)$. 代入 $P(\lambda)$, 可得

$$(\lambda I_n - A)(T(\lambda)(\lambda I_n - J) + P) = Q(\lambda)^{-1}(\lambda I_n - J).$$

整理可得

$$(\lambda I_n - A)P = (Q(\lambda)^{-1} - (\lambda I_n - A)T(\lambda))(\lambda I_n - J).$$

比较 λ 的次数可知, $Q(\lambda)^{-1} - (\lambda I_n - A)T(\lambda)$ 必须是数字矩阵, 记之为 R , 于是

$$(\lambda I_n - A)P = R(\lambda I_n - J).$$

去括号再次比较次数可得 $P = R, AP = RJ$. 若可证明 P 是可逆矩阵, 即有 $P^{-1}AP = J$. 由 $Q(\lambda)^{-1} - (\lambda I_n - A)T(\lambda) = R$ 可得

$$I_n = Q(\lambda)(\lambda I_n - A)T(\lambda) + Q(\lambda)R.$$

注意到 $Q(\lambda)(\lambda I_n - A) = (\lambda I_n - J)P(\lambda)^{-1}$, 故

$$I_n = (\lambda I_n - J)P(\lambda)^{-1}T(\lambda) + Q(\lambda)R.$$

设 $Q(\lambda) = (\lambda I_n - J)M(\lambda) + N$, 其中 N 是数字矩阵, 于是

$$I_n = (\lambda I_n - J)(P(\lambda)^{-1}T(\lambda) + M(\lambda)R) + NR.$$

比较次数可得 $NR = I_n$, 即 R 可逆, 也即 P 可逆. □

1.14.2 计算特征向量和广义特征向量

例题 1.24 设复四维空间上的线性变换 φ 在基 $\{e_1, e_2, e_3, e_4\}$ 下的表示矩阵为

$$A = \begin{pmatrix} 4 & -1 & 1 & -7 \\ 9 & -2 & -7 & -1 \\ 0 & 0 & 5 & -8 \\ 0 & 0 & 2 & -3 \end{pmatrix}$$

求一组新基, 使 φ 在这组新基下的表示矩阵是 A 的 Jordan 标准型, 并求过渡矩阵.

注 在例题 1.24 中, 任取 $(A - I)x = 0$ 的两个线性无关的解作为特征向量 α_1, α_3 , 都可以解出对应的广义特征向量 α_2, α_4 , 即线性方程组 $(A - I)x = \alpha_1$ 和 $(A - I)x = \alpha_3$ 的可解性不依赖于 α_1, α_3 的选取 (请读者自行思考其中的原因), 但这并非普遍的情形. 一般来说, 我们总可以取到 $(A - \lambda_0 I)x = 0$ 的一个非零解 α_1 (即特征值 λ_0 的特征向量), 但若 α_1 选取不当, 线性方程组 $(A - \lambda_0 I)x = \alpha_1$ 有可能是无解的 (即求不出对应的广义特征向量). 因此在选取特征向量时, 需要我们仔细观察或设立参数, 这样才能保证最终得到正确的结果. 让我们来看下面两个例题中的具体分析.

证明 证法一: 通过计算可知 $\lambda I_4 - A$ 的法式为 $\text{diag}\{1, 1, (\lambda - 1)^2, (\lambda - 1)^2\}$, 故 A 的初等因子组为 $(\lambda - 1)^2, (\lambda - 1)^2$, 从而 A 的 Jordan 标准型为

$$J = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

设过渡矩阵为 $P = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$, 则 $P^{-1}AP = J$, 即

$$AP = (A\alpha_1, A\alpha_2, A\alpha_3, A\alpha_4) = PJ = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)J$$

从而得到线性方程组:

$$(A - I)\alpha_1 = 0, (A - I)\alpha_2 = \alpha_1, (A - I)\alpha_3 = 0, (A - I)\alpha_4 = \alpha_3$$

求解 $(A - I)x = 0$ 得到两个线性无关的解, 将它们分别作为 α_1 和 α_3 :

$$\alpha_1 = (1, 3, 0, 0)', \alpha_3 = (5, 0, 6, 3)'$$

再求解方程组 $(A - I)x = \alpha_1, (A - I)x = \alpha_3$, 得到

$$\alpha_2 = (\frac{1}{3}, 0, 0, 0)', \alpha_4 = (\frac{7}{6}, 0, \frac{3}{2}, 0)'$$

因此过渡矩阵

$$P = \begin{pmatrix} 1 & \frac{1}{3} & 5 & \frac{7}{6} \\ 3 & 0 & 0 & 0 \\ 0 & 0 & 6 & \frac{3}{2} \\ 0 & 0 & 3 & 0 \end{pmatrix}$$

新基为 $(f_1, f_2, f_3, f_4) = (e_1, e_2, e_3, e_4)P$, 即

$$f_1 = e_1 + 3e_2, f_2 = \frac{1}{3}e_1, f_3 = 5e_1 + 6e_3 + 3e_4, f_4 = \frac{7}{6}e_1 + \frac{3}{2}e_3$$

证法二: A 的初等因子组的计算同解法 1, 可得 A 的 Jordan 标准型 $J = \text{diag}\{J_2(1), J_2(1)\}$. 注意到 $(A - I_4)^2 = O$ 且 $r(A - I_4) = 2$, 故可取 $A - I_4$ 的第 1 列和第 3 列作为其列向量的极大无关组. 因此 $e_1 = (1, 0, 0, 0)', e_3 = (0, 0, 1, 0)'$ 为广义特征向量, 使得 $(A - I_4)e_1 = (3, 9, 0, 0)', (A - I_4)e_3 = (1, -7, 4, 2)'$ 为线性无关的特征向量, 则过渡矩阵 $P = ((A - I_4)e_1, e_1, (A - I_4)e_3, e_3)$ 满足 $P^{-1}AP = J$. \square

例题 1.25 设

$$A = \begin{pmatrix} 2 & 6 & -15 \\ 1 & 1 & -5 \\ 1 & 2 & -6 \end{pmatrix}$$

求非异阵 P , 使 $P^{-1}AP$ 为 Jordan 标准型.

证明 证法一: 通过计算可知 $\lambda I_3 - A$ 的法式为 $\text{diag}\{1, \lambda + 1, (\lambda + 1)^2\}$, 故 A 的初等因子组为 $\lambda + 1, (\lambda + 1)^2$, 从而 A 的 Jordan 标准型为

$$J = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$$

设非异阵 $P = (\alpha_1, \alpha_2, \alpha_3)$, 使 $P^{-1}AP = J$, 则 $AP = (A\alpha_1, A\alpha_2, A\alpha_3) = PJ = (\alpha_1, \alpha_2, \alpha_3)J$, 从而得到线性方程组:

$$(A + I_3)\alpha_1 = 0, (A + I_3)\alpha_2 = 0, (A + I_3)\alpha_3 = \alpha_2$$

求解 $(A + I_3)x = 0$ 得到两个线性无关的解 $\beta_1 = (-2, 1, 0)'$ 和 $\beta_2 = (5, 0, 1)'$. 注意到 $(A + I_3)x = \beta_i (i = 1, 2)$ 都是无解的, 故不能将 β_1 或 β_2 直接作为 α_2 来求广义特征向量 α_3 . 一般地, 可设 $\alpha_2 = k_1\beta_1 + k_2\beta_2 = (-2k_1 + 5k_2, k_1, k_2)'$, 代入 $(A + I_3)x = \alpha_2$ 中, 利用 $r(A + I_3, \alpha_2) = r(A + I_3)$ 可得 $k_1 = k_2$. 因此, 可取 $\alpha_1 = \beta_1 = (-2, 1, 0)', \alpha_2 = \beta_1 + \beta_2 = (3, 1, 1)'$,

此时可解出 $\alpha_3 = (1, 0, 0)'$, 于是

$$P = \begin{pmatrix} -2 & 3 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

证法二: A 的初等因子组的计算同解法 1, 可得 A 的 Jordan 标准型 $J = \text{diag}\{-1, J_2(-1)\}$. 注意到 $(A + I_3)^2 = O$ 且 $r(A + I_3) = 1$, 故可取 $A + I_3$ 的第 1 列作为其列向量的极大无关组. 因此 $e_1 = (1, 0, 0)'$ 为循环向量 (即广义特征向量), 使得 $e_1, (A + I_3)e_1 = (3, 1, 1)'$ 构成了 $J_2(-1)$ 的循环轨道. 再取线性无关的特征向量 $\xi_1 = (-2, 1, 0)'$, 则过渡矩阵 $P = (\xi_1, (A + I_3)e_1, e_1)$ 满足 $P^{-1}AP = J$. \square

例题 1.26 设

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

求非异阵 P , 使 $P^{-1}AP$ 为 Jordan 标准型.

证明 证法一: 通过计算可知 $\lambda I_4 - A$ 的法式为 $\text{diag}\{1, 1, \lambda - 1, (\lambda - 1)^3\}$, 故 A 的初等因子组为 $\lambda - 1, (\lambda - 1)^3$, 从而 A 的 Jordan 标准型为 $J = \text{diag}\{1, J_3(1)\}$. 设非异阵 $P = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$, 使 $P^{-1}AP = J$, 则 $AP = (A\alpha_1, A\alpha_2, A\alpha_3, A\alpha_4) = PJ = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)J$, 从而得到线性方程组:

$$(A - I_4)\alpha_1 = 0, (A - I_4)\alpha_2 = 0, (A - I_4)\alpha_3 = \alpha_2, (A - I_4)\alpha_4 = \alpha_3.$$

求解 $(A - I_4)x = 0$ 得到两个线性无关的解 $\beta_1 = (-1, 0, 1, 0)'$ 和 $\beta_2 = (0, 1, 0, 1)'$. 设 $\alpha_2 = k_1\beta_1 + k_2\beta_2$, 代入 $(A - I_4)x = \alpha_2$ 中, 利用 $r(A - I_4, \alpha_2) = r(A - I_4)$ 可得 $k_1 = 0$. 于是可取 $\alpha_2 = k_2\beta_2$, 解出 $\alpha_3 = k_2e_1 + k_3\beta_1 + k_4\beta_2$, 其中 $e_1 = (1, 0, 0, 0)'$. 再代入 $(A - I_4)x = \alpha_3$ 中, 利用 $r(A - I_4, \alpha_3) = r(A - I_4)$ 可得 $k_2 = 2k_3$. 于是可取 $k_2 = 2, k_3 = 1, k_4 = 0$, 最终得到特征向量 $\alpha_1 = \beta_1 = (-1, 0, 1, 0)'$, $\alpha_2 = 2\beta_2 = (0, 2, 0, 2)'$, 1 级广义特征向量 $\alpha_3 = 2e_1 + \beta_1 = (1, 0, 1, 0)'$, 2 级广义特征向量 $\alpha_4 = (0, 0, 0, 1)'$, 从而

$$P = \begin{pmatrix} -1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{pmatrix}$$

证法二: A 的初等因子组的计算同解法 1, 可得 A 的 Jordan 标准型 $J = \text{diag}\{1, J_3(1)\}$. 注意到 $(A - I_4)^3 = O$ 且 $r((A - I_4)^2) = 1$, 故可取 $(A - I_4)^2$ 的第 4 列作为其列向量的极大无关组. 因此 $e_4 = (0, 0, 0, 1)'$ 为循环向量 (即 2 级广义特征向量), 使得 $e_4, (A - I_4)e_4 = (1, 0, 1, 0)', (A - I_4)^2e_4 = (0, 2, 0, 2)'$ 构成了 $J_3(1)$ 的循环轨道. 再取线性无关的特征向量 $\xi_1 = (-1, 0, 1, 0)'$, 则过渡矩阵 $P = (\xi_1, (A - I_4)^2e_4, (A - I_4)e_4, e_4)$ 满足 $P^{-1}AP = J$. \square

1.14.3 计算循环子空间的循环向量

根据 § 7.10 中所述 Jordan 标准型的几何意义, 全空间可分解为不同特征值的根子空间的直和, 每个根子空间可分解为若干个循环子空间的直和, 每个循环子空间对应于一条循环轨道, 这条轨道由循环向量 (即最高级的广义特征向量) 生成. 下面以幂零根子空间为例, 说明如何确定所有的循环向量, 从而确定所有的基向量 (等价于求过渡矩阵 P).

例题 1.27 设 9 阶幂零矩阵 A 的 Jordan 标准型 $J = \text{diag}\{0, J_2(0), J_3(0), J_3(0)\}$, 求非异阵 P , 使 $P^{-1}AP = J$.

注 这个例题采用的方法可以推广到一般的情形, 其原理是: 设 n 阶幂零矩阵 A 的极小多项式为 λ^k , 则依次选取第 i 级广义特征向量 ξ_i ($i = k - 1, \dots, 0$), 使得所有的 $A^i\xi_i$ ($i = k - 1, \dots, 0$) 在 $\text{Ker } A$ 中线性无关即可.

解 由已知条件 $A^3 = O, r(A^2) = 2$ 且 $r(A) = 5$, 可设 $A^2x = 0$ 的基础解系为 $\{\eta_i, 1 \leq i \leq 7\}$. 由于 A^2 的列秩为 2, 故不妨设 A^2 的第 1 列和第 2 列是 A^2 列向量的极大无关组, 即 A^2e_1, A^2e_2 线性无关, 其中 e_1, e_2 是 9 维标准单位列向量的前两个. 考虑限制映射 $A|_{\text{Ker } A^2} : \text{Ker } A^2 \rightarrow \text{Ker } A$, 容易验证 $\text{Ker}(A|_{\text{Ker } A^2}) = \text{Ker } A, \text{Im}(A|_{\text{Ker } A^2}) = \text{Ker } A \cap \text{Im } A$. 由

$\dim \operatorname{Ker} A^2 = 7, \dim \operatorname{Ker} A = 4$ 可知 $\dim(\operatorname{Ker} A \cap \operatorname{Im} A) = 3$, 且 $\operatorname{Ker} A \cap \operatorname{Im} A = L(A\eta_i, 1 \leq i \leq 7)$. 注意到 $A^2 e_1, A^2 e_2$ 是 $\operatorname{Ker} A \cap \operatorname{Im} A$ 中两个线性无关的向量, 故可从其生成元中取出一个向量, 不妨设为 $A\eta_1$, 使得 $A^2 e_1, A^2 e_2, A\eta_1$ 线性无关. 再次注意到 $\dim \operatorname{Ker} A = 4$, 且 $A^2 e_1, A^2 e_2, A\eta_1$ 是 $\operatorname{Ker} A$ 中 3 个线性无关的向量, 故可从其一组基 (即 $Ax = 0$ 的基础解系) 中取出一个向量 ξ_1 , 使得 $A^2 e_1, A^2 e_2, A\eta_1, \xi_1$ 线性无关.

下面证明: $\{e_1, Ae_1, A^2 e_1, e_2, Ae_2, A^2 e_2, \eta_1, A\eta_1, \xi_1\}$ 构成 \mathbb{C}^9 的一组基. 只要证明它们线性无关即可. 设 $c_1, \dots, c_9 \in \mathbb{C}$, 使得

$$c_1 e_1 + c_2 Ae_1 + c_3 A^2 e_1 + c_4 e_2 + c_5 Ae_2 + c_6 A^2 e_2 + c_7 \eta_1 + c_8 A\eta_1 + c_9 \xi_1 = 0 \quad (1.33)$$

将(1.33)式作用 A^2 可得

$$c_1 A^2 e_1 + c_4 A^2 e_2 = 0$$

由 $A^2 e_1, A^2 e_2$ 线性无关可知 $c_1 = c_4 = 0$. 将(1.33)式作用 A 可得

$$c_2 A^2 e_1 + c_5 A^2 e_2 + c_7 A\eta_1 = 0$$

由 $A^2 e_1, A^2 e_2, A\eta_1$ 线性无关可知 $c_2 = c_5 = c_7 = 0$. (1.33)式最后变成

$$c_3 A^2 e_1 + c_6 A^2 e_2 + c_8 A\eta_1 + c_9 \xi_1 = 0$$

由 $A^2 e_1, A^2 e_2, A\eta_1, \xi_1$ 线性无关可知 $c_3 = c_6 = c_8 = c_9 = 0$. 有了上面这组基, 我们可以把 4 个循环子空间的循环轨道全部确定如下:

轨道1	轨道2	轨道3	轨道4	
		e_1	e_2	2级广义特征向量
		\downarrow	\downarrow	
	η_1	Ae_1	Ae_2	1级广义特征向量
	\downarrow	\downarrow	\downarrow	
ξ_1	$A\eta_1$	$A^2 e_1$	$A^2 e_2$	0级广义特征向量
\downarrow	\downarrow	\downarrow	\downarrow	
0	0	0	0	

最后, 令 $P = (\xi_1, A\eta_1, \eta_1, A^2 e_1, Ae_1, e_1, A^2 e_2, Ae_2, e_2)$ 即为所求. \square

例题 1.28 设

$$A = \begin{pmatrix} 3 & -4 & 0 & 2 \\ 4 & -5 & -2 & 4 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 2 & -1 \end{pmatrix}$$

求非异阵 P , 使 $P^{-1}AP$ 为 Jordan 标准型.

注 下面的例题也与过渡矩阵有关, 它告诉我们: 满足基础矩阵乘法性质的矩阵类与基础矩阵类之间存在着一个相似变换. 利用这一结论可以证明: n 阶矩阵环 $M_n(\mathbb{K})$ 的任一自同构都是内自同构.

解 经计算可知 A 的初等因子组为 $(\lambda + 1)^2, (\lambda - 1)^2$, 于是 A 的 Jordan 标准型为 $J = \operatorname{diag}\{J_2(-1), J_2(1)\}$. 由命题??可知, $\mathbb{C}^4 = \operatorname{Ker}(A + I_4)^2 \oplus \operatorname{Ker}(A - I_4)^2$, 且 $\operatorname{Ker}(A + I_4)^2 = \operatorname{Im}(A - I_4)^2, \operatorname{Ker}(A - I_4)^2 = \operatorname{Im}(A + I_4)^2$. 经计算可取 $(A - I_4)^2$ 的第二列 $\alpha = (A - I_4)^2 e_2 = (16, 20, 0, 0)'$ 作为根子空间 $\operatorname{Ker}(A + I_4)^2$ 中的循环向量 (即广义特征向量), 于是 $\alpha, (A + I_4)\alpha = (-16, -16, 0, 0)'$ 构成根子空间 $\operatorname{Ker}(A + I_4)^2$ 中的循环轨道. 经计算可取 $(A + I_4)^2$ 的第三列 $\beta = (A + I_4)^2 e_3 = (12, 8, 12, 8)'$ 作为根子空间 $\operatorname{Ker}(A - I_4)^2$ 中的循环向量 (即广义特征向量), 于是 $\beta, (A - I_4)\beta = (8, 8, 8, 8)'$ 构成根子空间 $\operatorname{Ker}(A - I_4)^2$ 中的循环轨道. 因此, 过渡矩阵 $P = ((A + I_4)\alpha, \alpha, (A - I_4)\beta, \beta)$ 满足 $P^{-1}AP = J$. \square

定理 1.25

设有 n^2 个 n 阶非零矩阵 A_{ij} ($1 \leq i, j \leq n$), 适合

$$A_{ij}A_{jk} = A_{ik}, \quad A_{ij}A_{lk} = O \quad (j \neq l).$$

求证: 存在可逆矩阵 P , 使得对任意的 $i, j, P^{-1}A_{ij}P = E_{ij}$, 其中 E_{ij} 是基础矩阵.

♥

证明 因为 $A_{11} \neq O$, 故存在 α , 使得 $A_{11}\alpha \neq 0$. 令 $\alpha_1 = A_{11}\alpha$, 由 $A_{11}A_{11} = A_{11}$ 可得 $A_{11}\alpha_1 = \alpha_1$. 再令 $\alpha_i = A_{i1}\alpha_1$, 由 $A_{i1}A_{11} = A_{11}$ 可知 $\alpha_i \neq 0$. 我们得到了 n 个非零向量 $\alpha_1, \alpha_2, \dots, \alpha_n$, 由已知条件容易验证这 n 个向量适合下列性质:

$$A_{ij}\alpha_j = \alpha_i, \quad A_{ij}\alpha_k = 0 \quad (j \neq k)$$

由此不难证明这 n 个向量线性无关. 令 $P = (\alpha_1, \alpha_2, \dots, \alpha_n)$, 则 P 是可逆矩阵, 且

$$A_{ij}P = (A_{ij}\alpha_1, A_{ij}\alpha_2, \dots, A_{ij}\alpha_n) = (0, \dots, 0, \alpha_i, 0, \dots, 0).$$

其中上式中的 α_i 在第 j 列. 另一方面, 有

$$PE_{ij} = (\alpha_1, \alpha_2, \dots, \alpha_n)E_{ij} = (0, \dots, 0, \alpha_i, 0, \dots, 0).$$

因此, 对任意的 $i, j, A_{ij}P = PE_{ij}$, 即 $P^{-1}A_{ij}P = E_{ij}$. □

1.15 Jordan 标准型的应用

1.15.1 利用 Jordan 标准型研究矩阵的性质

命题 1.43

设 A 是 n 阶复矩阵, 求证: A 相似于分块对角矩阵 $\text{diag}\{B, C\}$, 其中 B 是幂零矩阵, C 是可逆矩阵.

▲

注 这个命题告诉我们: 在相似的意义下, 对复方阵的研究可归结为对幂零矩阵和可逆矩阵这两类特殊矩阵的研究, 它们的刻画分别是: 特征值全为零以及特征值全不为零. 这也是前面很多例题都处理这两类矩阵的深层次原因.

证明 我们发现 A 的初等因子分离开了零特征值和非零特征值, 从而 A 的 Jordan 标准型满足题目要求. 此时, 可将零特征值的 Jordan 块 $J_r(0)$ (幂零矩阵) 放入 B 中, 将非零特征值的 Jordan 块 $J_r(\lambda_0)$ (可逆矩阵) 放入 C 中, 即得结论. □

定理 1.26

设 λ_0 是 n 阶矩阵 A 的特征值, 其代数重数为 m . 设属于特征值 λ_0 的最大 Jordan 块的阶数为 k , 求证:

$$r(A - \lambda_0 I_n) > \dots > r((A - \lambda_0 I_n)^k) = r((A - \lambda_0 I_n)^{k+1}) = \dots = n - m$$

♥

证明 设 P 为可逆矩阵, 使 $P^{-1}AP = J = \text{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \dots, J_{r_s}(\lambda_s)\}$ 为 Jordan 标准型, 则对任意的正整数 j ,

$$r((A - \lambda_0 I_n)^j) = r(P^{-1}(A - \lambda_0 I_n)^j P) = r((J - \lambda_0 I_n)^j) = \sum_{i=1}^s r(J_{r_i}(\lambda_i - \lambda_0)^j)$$

若 $\lambda_i \neq \lambda_0$, 则 $r(J_{r_i}(\lambda_i - \lambda_0)^j) = r_i$. 若 $\lambda_i = \lambda_0$, 则当 $1 \leq j \leq r_i$ 时, $r(J_{r_i}(0)^j) = r_i - j$; 当 $j \geq r_i$ 时, $r(J_{r_i}(0)^j) = 0$. 注意到 A 至少有一个 Jordan 块 $J_k(\lambda_0)$, 并且属于特征值 λ_0 的所有 Jordan 块阶数之和等于 m , 故当 $1 \leq j \leq k$ 时, $r((A - \lambda_0 I_n)^j)$ 严格递减; 当 $j \geq k$ 时, $r((A - \lambda_0 I_n)^j) = n - m$. □

命题 1.44

设 λ_0 是 n 阶矩阵 A 的特征值, 其代数重数为 m . 设属于特征值 λ_0 的最大 Jordan 块的阶数为 k , 求证:

$$r(A - \lambda_0 I_n) > \dots > r((A - \lambda_0 I_n)^k) = r((A - \lambda_0 I_n)^{k+1}) = \dots = n - m.$$

▲

证明 设 P 为可逆矩阵, 使 $P^{-1}AP = J = \text{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \dots, J_{r_s}(\lambda_s)\}$ 为 Jordan 标准型, 则对任意的正整数 j ,

$$r((A - \lambda_0 I_n)^j) = r(P^{-1}(A - \lambda_0 I_n)^j P) = r((J - \lambda_0 I_n)^j) = \sum_{i=1}^s r(J_{r_i}(\lambda_i - \lambda_0)^j)$$

若 $\lambda_i \neq \lambda_0$, 则 $r(J_{r_i}(\lambda_i - \lambda_0)^j) = r_i$. 若 $\lambda_i = \lambda_0$, 则当 $1 \leq j \leq r_i$ 时, $r(J_{r_i}(0)^j) = r_i - j$; 当 $j \geq r_i$ 时, $r(J_{r_i}(0)^j) = 0$. 注意到 A 至少有一个 Jordan 块 $J_k(\lambda_0)$, 并且属于特征值 λ_0 的所有 Jordan 块阶数之和等于 m , 故当 $1 \leq j \leq k$ 时, $r((A - \lambda_0 I_n)^j)$ 严格递减; 当 $j \geq k$ 时, $r((A - \lambda_0 I_n)^j) = n - m$. \square

命题 1.45

设 λ_0 是 n 阶复矩阵 A 的特征值, 并且属于 λ_0 的初等因子都是次数大于等于 2 的多项式. 求证: 特征值 λ_0 的任一特征向量 α 均可表示为 $A - \lambda_0 I_n$ 的列向量的线性组合.

注 若特征值 λ_0 有一个初等因子为一次多项式, 则必存在特征向量 α , 它不能表示为 $A - \lambda_0 I_n$ 的列向量的线性组合. 证明的细节留给读者完成. 一个极端的例子就是 $A = I_n$, 其特征值 1 的初等因子都是一次的, 并且任一特征向量都不是 $A - I_n = O$ 的列向量的线性组合. 命题 1.45 与命题 1.29 (可对角化的判定) 有着密切的联系, 请读者思考两者之间的关系.

证明 由 Jordan 标准型理论可知, 属于特征值 λ_0 的每个 Jordan 块的特征向量均只有一个, 并且特征值 λ_0 的任一特征向量都是这些特征向量的线性组合, 因此我们只要证明 A 的 Jordan 标准型只含一个 Jordan 块 $J_n(\lambda_0)$ 的情形即可. 设 $P = (\alpha_1, \alpha_2, \dots, \alpha_n)$ 为非异阵, 使得 $P^{-1}AP = J_n(\lambda_0)$, 即 $AP = PJ_n(\lambda_0)$, 利用分块矩阵的乘法可得

$$\begin{aligned} A\alpha_1 &= \lambda_0 \alpha_1, \\ A\alpha_2 &= \alpha_1 + \lambda_0 \alpha_2, \\ &\vdots \\ A\alpha_n &= \alpha_{n-1} + \lambda_0 \alpha_n \end{aligned}$$

注意到 $n \geq 2$, 故有 $\alpha_1 = (A - \lambda_0 I_n)\alpha_2$, 从而特征向量 α_1 可表示为 $A - \lambda_0 I_n$ 的列向量的线性组合. \square

1.15.2 运用 Jordan 标准型进行相似问题的化简

命题 1.46

设 A, B 为 n 阶矩阵, 满足 $AB = BA = O, r(A) = r(A^2)$, 求证: $r(A + B) = r(A) + r(B)$.

证明 注意到问题的条件和结论在同时相似变换: $A \mapsto P^{-1}AP, B \mapsto P^{-1}BP$ 下不改变, 故不妨从一开始就假设 A 为 Jordan 标准型. 设 $A = \text{diag}\{A_0, A_1\}$, 其中 A_0 由零特征值的 Jordan 块构成, A_1 由非零特征值的 Jordan 块构成. 由 $r(A) = r(A^2)$ 可知, 零特征值的 Jordan 块都是一阶的, 即 $A_0 = O$. 将 B 进行对应的分块 $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$, 则由 $AB = BA = O$ 以及 A_1 非异可知, B_{12}, B_{21} 和 B_{22} 都是零矩阵. 于是

$$r(A + B) = r \begin{pmatrix} B_{11} & O \\ O & A_1 \end{pmatrix} = r(B_{11}) + r(A_1) = r(B) + r(A).$$

\square

命题 1.47

设 A, B 分别是 m, n 阶矩阵, 求证: 矩阵方程 $AX = XB$ 只有零解的充要条件是 A, B 无公共的特征值.

证明 先做两步化简. 注意到问题的条件和结论在矩阵变换: $B \mapsto P^{-1}BP, X \mapsto XP$ 下不改变, 故不妨从一开始就假设 B 为 Jordan 标准型. 设 $X = (\alpha_1, \alpha_2, \dots, \alpha_n)$ 为列分块, 则有

$$AX = (A\alpha_1, A\alpha_2, \dots, A\alpha_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)B = XB.$$

若 B 有 k 个 Jordan 块, 则方程组(??)可分解为 k 个独立方程组. 注意到:

- (i) 方程组(??)只有零解当且仅当这 k 个独立方程组都只有零解;
 (ii) 方程组(??)有非零解当且仅当这 k 个独立方程组中至少有一个有非零解.

因此,不妨进一步假设 $B = J_n(\lambda_0)$ 为 Jordan 块. 此时, 方程组(??)等价于下列方程组:

$$A\alpha_1 = \lambda_0\alpha_1, A\alpha_2 = \alpha_1 + \lambda_0\alpha_2, \dots, A\alpha_n = \alpha_{n-1} + \lambda_0\alpha_n.$$

充分性: 假设 A, B 没有公共的特征值, 则 λ_0 不是 A 的特征值, 从而由 $A\alpha_1 = \lambda_0\alpha_1$ 只能得到 $\alpha_1 = \mathbf{0}$. 代入第二个方程可得 $A\alpha_2 = \lambda_0\alpha_2$, 相同的理由可推出 $\alpha_2 = \mathbf{0}$. 不断这样做下去, 最后可得 $\alpha_i = \mathbf{0} (1 \leq i \leq n)$, 即 $X = O$, 从而矩阵方程 $AX = XB$ 只有零解.

必要性: 假设 A 和 $B = J_n(\lambda_0)$ 有公共的特征值 λ_0 , 在上述方程组中令 $\alpha_1 = \dots = \alpha_{n-1} = \mathbf{0}$. 因为 λ_0 也是 A 的特征值, 所以 $A\alpha_n = \lambda_0\alpha_n$ 有非零解 $\alpha_n = \alpha$, 于是 $X_0 = (\mathbf{0}, \dots, \mathbf{0}, \alpha)$ 是上述方程组的非零解, 从而矩阵方程 $AX = XB$ 有非零解. \square

命题 1.48

设 A, B 分别是 m, n 阶矩阵, C 是 $m \times n$ 矩阵, 求证: 矩阵方程 $AX - XB = C$ 存在唯一解的充要条件是 A, B 无公共的特征值.

证明 先做两步化简. 注意到问题的条件和结论在矩阵变换 $B \mapsto P^{-1}BP, C \mapsto CP, X \mapsto XP$ 下不改变, 故不妨从一开始就假设 B 为 Jordan 标准型. 设 $X = (\alpha_1, \alpha_2, \dots, \alpha_n), C = (\beta_1, \beta_2, \dots, \beta_n)$ 为列分块, 则 $AX - XB = C$ 即为:

$$(A\alpha_1, A\alpha_2, \dots, A\alpha_n) - (\alpha_1, \alpha_2, \dots, \alpha_n)B = (\beta_1, \beta_2, \dots, \beta_n).$$

若 B 有 k 个 Jordan 块, 则方程组(??)可分解为 k 个独立方程组. 注意到:

- (i) 方程组(??)无解当且仅当这 k 个独立方程组中至少有一个无解;
 (ii) 方程组(??)有唯一解当且仅当这 k 个独立方程组都只有唯一解;
 (iii) 方程组(??)有无穷个解当且仅当这 k 个独立方程组都有解, 且至少有一个有无穷个解.

进一步, 若假设 $B = J_n(\lambda_0)$ 为 Jordan 块, 则方程组(??)等价于下列方程组:

$$(A - \lambda_0 I_n)\alpha_1 = \beta_1, (A - \lambda_0 I_n)\alpha_2 = \alpha_1 + \beta_2, \dots, (A - \lambda_0 I_n)\alpha_n = \alpha_{n-1} + \beta_n.$$

充分性: 假设 A, B 没有公共的特征值, 则 λ_0 不是 A 的特征值, 从而 $A - \lambda_0 I_n$ 是可逆矩阵. 从第一个方程可解得 $\alpha_1 = (A - \lambda_0 I_n)^{-1}\beta_1$, 代入第二个方程可解得 $\alpha_2 = (A - \lambda_0 I_n)^{-1}(\alpha_1 + \beta_2)$, \dots , 代入最后一个方程可解得 $\alpha_n = (A - \lambda_0 I_n)^{-1}(\alpha_{n-1} + \beta_n)$, 从而上述方程组有唯一解, 因此矩阵方程 $AX - XB = C$ 也有唯一解.

必要性: 假设 A, B 有公共的特征值 λ_0 , 若这 k 个独立方程组中有一个无解, 则矩阵方程 $AX - XB = C$ 无解, 从而结论成立. 若这 k 个独立方程组都有解, 则不妨设 $B = J_n(\lambda_0)$ 为 Jordan 块. 由于 λ_0 是 A 的特征值, 故 $(A - \lambda_0 I_n)x = \mathbf{0}$ 有无穷个解. 注意到, 若 $(\alpha_1, \alpha_2, \dots, \alpha_n)$ 是上述方程组的一个解, 则对 $(A - \lambda_0 I_n)x = \mathbf{0}$ 的任一解 $\alpha_0, (\alpha_1, \alpha_2, \dots, \alpha_n + \alpha_0)$ 也是上述方程组的解, 因此矩阵方程 $AX - XB = C$ 有无穷个解. \square

命题 1.49

设 A, B 分别是 m, n 阶矩阵, $M_{m \times n}(\mathbb{C})$ 上的线性变换 φ 定义为 $\varphi(X) = AX - XB$, 则下列 3 个结论等价:

- (1) φ 是单映射;
 (2) φ 是自同构;
 (3) 对某个给定的 $m \times n$ 矩阵 C , 存在唯一的 X_0 , 使得 $\varphi(X_0) = C$.

证明 事实上, (1) \Rightarrow (2) 以及 (2) \Rightarrow (3) 显然都成立. 用反证法来证明 (3) \Rightarrow (1): 若 $\text{Ker} \varphi \neq \mathbf{0}$, 则 $\text{Ker} \varphi$ 中任一非零元 X_1 都满足 $\varphi(X_0 + X_1) = C$, 这与唯一性矛盾.

因此, 命题 1.47 和命题 1.48 都等价于命题 ??, 并且命题 1.47 和命题 1.48 都给出了它们的 Jordan 标准型证法. \square

1.15.3 应用 Jordan 标准型的三段论法

如果矩阵问题的条件和结论在相似关系下不改变,则可以先证明结论对 Jordan 块成立,再证明对 Jordan 标准型成立,最后证明对一般的矩阵也成立,这就是所谓的“三段论法”.事实上,我们已经利用三段论法证明过例题 1.11 和例题 1.12,下面再来看一些典型的例题.

首先,我们来看计算矩阵乘幂的问题.设 A 为 n 阶矩阵, P 为 n 阶可逆矩阵,使得

$$P^{-1}AP = J = \text{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \dots, J_{r_k}(\lambda_k)\}$$

为 Jordan 标准型.注意任一 Jordan 块 $J_{r_i}(\lambda_i)$ 都有分解 $J_{r_i}(\lambda_i) = \lambda_i I_{r_i} + N$, 其中 $N = J_{r_i}(0)$ 是特征值为零的 r_i 阶 Jordan 块,故对任意的正整数 m ,

$$J_{r_i}(\lambda_i)^m = (\lambda_i I_{r_i} + N)^m = \lambda_i^m I_{r_i} + C_m^1 \lambda_i^{m-1} N + \dots + C_m^{m-1} \lambda_i N^{m-1} + N^m.$$

于是 $J^m = \text{diag}\{J_{r_1}(\lambda_1)^m, J_{r_2}(\lambda_2)^m, \dots, J_{r_k}(\lambda_k)^m\}$, 从而 $A^m = (PJP^{-1})^m = PJ^mP^{-1}$ 便可计算出来了.

例题 1.29 设 $A = \begin{pmatrix} 2 & 6 & -15 \\ 1 & 1 & -5 \\ 1 & 2 & -6 \end{pmatrix}$, 求 $A^m (m \geq 1)$.

注 本题是例题 1.25 的延拓.

解 由例题 1.25, 我们已经计算出过渡矩阵 P , 使得 $P^{-1}AP = J = \text{diag}\{-1, J_2(-1)\}$, 于是可进一步计算出

$$\begin{aligned} A^m &= PJ^mP^{-1} = \begin{pmatrix} -2 & 3 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} (-1)^m & 0 & 0 \\ 0 & (-1)^m & (-1)^{m-1}m \\ 0 & 0 & (-1)^m \end{pmatrix} \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 1 \\ 1 & 2 & -5 \end{pmatrix} \\ &= (-1)^{m-1} \begin{pmatrix} 3m-1 & 6m & -15m \\ m & 2m-1 & -5m \\ m & 2m & -5m-1 \end{pmatrix}. \end{aligned}$$

□

例题 1.30 求矩阵 B , 使得 $A = B^2$, 其中 $A = \begin{pmatrix} 3 & 1 \\ -1 & 5 \end{pmatrix}$.

解 利用过渡矩阵的求法的方法, 可求出过渡矩阵 $P = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, 使得 $P^{-1}AP = J = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}$ 为 Jordan 标准型. 用

待定元素法不难求得 $C = \pm \begin{pmatrix} 2 & \frac{1}{4} \\ 0 & 2 \end{pmatrix}$, 使得 $C^2 = J$.

注意到 $(PCP^{-1})^2 = PC^2P^{-1} = PJ^mP^{-1} = A$, 故可取 $B = PCP^{-1}$. 经计算可得

$$B = \pm \begin{pmatrix} \frac{7}{4} & \frac{1}{4} \\ -\frac{1}{4} & \frac{5}{4} \end{pmatrix}.$$

□

定义 1.14 (非异阵的平方根)

设 A 为 n 阶非异复矩阵, 称 B 为 A 的 k 次方根, 若 B 满足 $A = B^k$, 其中 $k \in \mathbb{N}_1$.

♣

命题 1.50 (非异阵存在任意次的方根)

设 A 为 n 阶非异复矩阵, 证明: 对任一正整数 m , 存在 n 阶复矩阵 B , 使得 $A = B^m$.

♠

注 命题 1.50 的结论对奇异矩阵一般并不成立. 例如, 设 $A = J_n(0)^{m-1}$, 其中 $n = mq - r$, $m \geq 2$ 且 $0 \leq r < m$, 则不存在 B , 使得 $A = B^m$. 我们用反证法来证明这个结论. 若存在满足条件的 B , 则 B 的特征值全为零, 从而 B 也是幂零矩阵, 即有 $B^n = O$. 于是 $O = B^{n+r} = (B^m)^q = A^q = J_n(0)^{(m-1)q} \neq O$, 这就导出了矛盾.

证明 设 P 为非异阵, 使得 $P^{-1}AP = J = \text{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \dots, J_{r_k}(\lambda_k)\}$ 为 A 的 Jordan 标准型. 由于 A 非异, 故 A 的所有特征值都非零. 对 A 的任一 Jordan 块 $J_{r_i}(\lambda_i)$, 取定 λ_i 的某个 m 次方根 μ_i , 即 $\mu_i^m = \lambda_i$, 则由例 7.54 可知, $J_{r_i}(\mu_i)^m$ 相似于 $J_{r_i}(\lambda_i)$, 即存在非异阵 Q_i , 使得 $J_{r_i}(\lambda_i) = Q_i^{-1}J_{r_i}(\mu_i)^m Q_i = (Q_i^{-1}J_{r_i}(\mu_i)Q_i)^m$, 即结论对 Jordan 块成立. 令

$$C = \text{diag}\{Q_1^{-1}J_{r_1}(\mu_1)Q_1, Q_2^{-1}J_{r_2}(\mu_2)Q_2, \dots, Q_k^{-1}J_{r_k}(\mu_k)Q_k\},$$

则 $J = C^m$, 即结论对 Jordan 标准型也成立. 最后,

$$A = PJP^{-1} = PC^mP^{-1} = (PCP^{-1})^m.$$

令 $B = PCP^{-1}$, 则有 $A = B^m$, 即结论对一般的矩阵也成立. □