$\mathcal{A}$ -矩阵及若当标准形的几何理论

首都师范大学石生明

2012.11.

§1 $\mathcal{A}$ -矩阵的概念与性质

任给数域 P 上 n 维空间 V 上线性变换 $\mathcal{A}$ ，已定义过 P 上 $\mathcal{A}$ 的多项式。即对任意 $f(\lambda) \in P[\lambda], f(\lambda) = a_k \lambda^k + a_{k-1} \lambda^{k-1} + \cdots + a_1 \lambda + a_0$ ，称

$$f(\mathcal{A}) = a_k \mathcal{A}^k + a_{k-1} \mathcal{A}^{k-1} + \cdots + a_1 \mathcal{A} + a_0 \mathcal{E}$$

为 $\mathcal{A}$ 的一个多项式，其中 $\mathcal{E}$ 为 V 上恒等变换。 $f(\mathcal{A})$ 仍是 V 上线性变换。引入

定义 1 对 P 上任意 λ -矩阵 $\mathbf{K}(\lambda) = (a_{ij}(\lambda))_{m \times l}$ ， $a_{ij}(\lambda)$ 是 $P[\lambda]$ 中多项式。可令 $\mathbf{K}(\mathcal{A}) = (a_{ij}(\mathcal{A}))$ ，称它为 P 上的一个 $\mathcal{A}$ -矩阵。

例如

$$\mathbf{K}(\lambda) = \begin{pmatrix} \lambda^2 + 1 & \lambda - 1 \\ 3 & \lambda^3 + \lambda \end{pmatrix}, \quad \mathbf{K}(\mathcal{A}) = \begin{pmatrix} \mathcal{A}^2 + \mathcal{E} & \mathcal{A} - \mathcal{E} \\ 3\mathcal{E} & \mathcal{A}^3 + \mathcal{A} \end{pmatrix}.$$

又如设 $\mathbf{F}(\lambda) = \lambda \mathbf{E} - \mathbf{B}$ ，称为 $\mathbf{B}$ 的特征矩阵，设 $\mathbf{B} = (b_{ij})_{m \times n}$ 则

$$\mathbf{F}(\lambda) = \lambda \mathbf{E} - \mathbf{B} = \begin{pmatrix} \lambda - b_{11} & -b_{12} & \cdots & -b_{1n} \\ -b_{21} & \lambda - b_{22} & \cdots & -b_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -b_{n1} & -b_{n2} & \cdots & \lambda - b_{nn} \end{pmatrix},$$

$$\mathbf{F}(\mathcal{A}) = \begin{pmatrix} \mathcal{A} - b_{11}\mathcal{E} & -b_{12}\mathcal{E} & \cdots & -b_{1n}\mathcal{E} \\ -b_{21}\mathcal{E} & \mathcal{A} - b_{22}\mathcal{E} & \cdots & -b_{2n}\mathcal{E} \\ \cdots & \cdots & \cdots & \cdots \\ -b_{n1}\mathcal{E} & -b_{n2}\mathcal{E} & \cdots & \mathcal{A} - b_{nn}\mathcal{E} \end{pmatrix}$$

由于矩阵运算及行列式定义中只用到元素的加法，乘法。而 $\mathcal{A}$ 的多项式有加法和乘法与 λ -矩阵一样， $\mathcal{A}$ -矩阵也有加法，乘法及“数量”乘法（用 $\mathcal{A}$ 的多项式作元素与 $\mathcal{A}$ -矩阵，甚至数字矩阵作“数量”乘法），也有与数字矩阵， λ -矩阵类似的运算性质。也能定义 $\mathcal{A}$ -矩阵的行列式，有着与数字行列式相同的性质。

用矩阵运算可以表达前面的

$$\mathbf{F}(\mathcal{A}) = \begin{pmatrix} \mathcal{A} - b_{11}\mathcal{E} & -b_{12}\mathcal{E} & \cdots & -b_{1n}\mathcal{E} \\ -b_{21}\mathcal{E} & \mathcal{A} - b_{22}\mathcal{E} & \cdots & -b_{2n}\mathcal{E} \\ \cdots & \cdots & \cdots & \cdots \\ -b_{n1}\mathcal{E} & -b_{n2}\mathcal{E} & \cdots & \mathcal{A} - b_{nn}\mathcal{E} \end{pmatrix} = \mathcal{A}\mathbf{E} - \mathcal{E}\mathbf{B}$$

我们知道对 $f(\lambda), g(\lambda) \in P[\lambda]$ ，令 $h(\lambda) = f(\lambda) + g(\lambda)$ ， $p(\lambda) = f(\lambda)g(\lambda)$ 就有 $h(\mathcal{A}) = f(\mathcal{A}) + g(\mathcal{A}), p(\mathcal{A}) = f(\mathcal{A})g(\mathcal{A})$

不仅如此，对 λ -矩阵 $\mathbf{K}(\lambda)_{m \times k}, \mathbf{L}(\lambda)_{m \times k}$ ，若 $\mathbf{H}(\lambda) = \mathbf{K}(\lambda) + \mathbf{L}(\lambda)$ ，则 $\mathbf{H}(\mathcal{A}) = \mathbf{K}(\mathcal{A})$

$+L(\mathcal{A})$. 又对 $K(\lambda)_{m \times s}, L(\lambda)_{s \times k}$, 若 $M(\lambda)_{m \times k} = K(\lambda)L(\lambda)$, 则 $M(\mathcal{A}) = K(\mathcal{A})L(\mathcal{A})$. (请读者验证).

对 V 上的一个线性变换 $\mathcal{A}$, 它在基 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 下有矩阵 A^T , 即

$$\mathcal{A}(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)A^T, \quad A = (a_{ij}), \quad (1)$$

或写成

$$\begin{aligned} \mathcal{A}\varepsilon_1 &= a_{11}\varepsilon_1 + a_{12}\varepsilon_2 + \dots + a_{1n}\varepsilon_n = a_{11}\varepsilon\varepsilon_1 + a_{12}\varepsilon\varepsilon_2 + \dots + a_{1n}\varepsilon\varepsilon_n \\ \mathcal{A}\varepsilon_2 &= a_{21}\varepsilon_1 + a_{22}\varepsilon_2 + \dots + a_{2n}\varepsilon_n = a_{21}\varepsilon\varepsilon_1 + a_{22}\varepsilon\varepsilon_2 + \dots + a_{2n}\varepsilon \\ &\dots\dots\dots \\ \mathcal{A}\varepsilon_n &= a_{n1}\varepsilon_1 + a_{n2}\varepsilon_2 + \dots + a_{nn}\varepsilon_n = a_{n1}\varepsilon\varepsilon_1 + a_{n2}\varepsilon\varepsilon_2 + \dots + a_{nn}\varepsilon \end{aligned}$$

这组式子可以用一个左形式表达式来表示

$$\mathcal{A} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = A \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} \quad (2)$$

或

$$\begin{pmatrix} \mathcal{A} - a_{11}\varepsilon & -a_{12}\varepsilon & \dots & -a_{1n}\varepsilon \\ -a_{21}\varepsilon & \mathcal{A} - a_{22}\varepsilon & \dots & -a_{2n}\varepsilon \\ & \dots & \dots & \\ -a_{n1}\varepsilon & -a_{n2}\varepsilon & \dots & \mathcal{A} - a_{nn}\varepsilon \end{pmatrix} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \mathbf{0}$$

这个式子可用下列 $\mathcal{A}$ -矩阵的形式表达式表示

$$(\mathcal{A}E - \varepsilon A) \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \mathbf{0} \quad (3)$$

(1) 与 (2) 相比, (2) 可以用 $\mathcal{A}$ -矩阵的形式表达式 (3) 写出, 而 (1) 则不能, 这是表达式 (2) 的好处.

表达式 (1) 中的矩阵 A^T 是 $\mathcal{A}$ 在基 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 下的矩阵, 而表达式 (2) 中的矩阵 A 称为 $\mathcal{A}$ 在基 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 下的左矩阵(出现在基 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 的左边). $\mathcal{A}$ 在同一基下的矩阵 A^T 和左矩阵 A 互为转置. 它们有相同的特征多项式. 由于 $\mathcal{A}$ 在不同基下矩阵是相似的, 则它们的转置也是相似的. 即 $\mathcal{A}$ 在不同基下的左矩阵也相似. 故 $\mathcal{A}$ 在任何基下的矩阵及左矩阵的特征多项式都相等, 都是 $\mathcal{A}$ 的特征多项式.

(3) 的左端的写法是一种形式写法. 这是 $\mathcal{A}$ -矩阵与向量元素的列矩阵的“乘法”. 这不是 $\mathcal{A}$ -矩阵的乘法. 数字矩阵与向量元素矩阵也有过类似的乘法, 当时称为形式写法. 现在类似地引入 $\mathcal{A}$ -矩阵的形式写法. 对任意 $\mathcal{A}$ -矩阵 $K(\mathcal{A}) = (k_{ij}(\mathcal{A}))_{m \times n}$ 将下列等式

$$\mathbf{K}(\mathcal{A}) \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} k_{11}(\mathcal{A})\xi_1 + \cdots + k_{1n}(\mathcal{A})\xi_n \\ k_{21}(\mathcal{A})\xi_1 + \cdots + k_{2n}(\mathcal{A})\xi_n \\ \dots \dots \\ k_{m1}(\mathcal{A})\xi_1 + \cdots + k_{mn}(\mathcal{A})\xi_n \end{pmatrix}. \quad (4)$$

也称为 $\mathcal{A}$ -矩阵的形式写法. 与数字矩阵的形式写法一样有下列性质:

$$(1) (\mathbf{K}(\mathcal{A})_{m \times s} \mathbf{L}_{s \times k}(\mathcal{A})) \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_k \end{pmatrix} = \mathbf{K}(\mathcal{A}) \left[\mathbf{L}(\mathcal{A}) \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_k \end{pmatrix} \right]$$

$$(2) [\mathbf{K}(\mathcal{A})_{m \times k} + \mathbf{L}_{m \times k}(\mathcal{A})] \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_k \end{pmatrix} = \mathbf{K}(\mathcal{A})_{m \times k} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_k \end{pmatrix} + \mathbf{L}_{m \times k}(\mathcal{A}) \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_k \end{pmatrix}$$

$$(3) f(\mathcal{A}) \text{ 是 } \mathcal{A} \text{ - 多项式, 则 } (f(\mathcal{A})\mathbf{K}(\mathcal{A})) \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_k \end{pmatrix} = \mathbf{K}(\mathcal{A}) \left[f(\mathcal{A}) \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_k \end{pmatrix} \right]$$

但是 $\mathcal{A}$ -矩阵的形式表达式与数字矩阵形式表达式也有不同性质. 例如数字可逆阵

$\mathbf{T}_{n \times n}$, 对基 $\varepsilon_1, \dots, \varepsilon_n$ 作

$$\mathbf{T}_{n \times n} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix}.$$

则 $\eta_1, \eta_2, \dots, \eta_n$ 仍为一组基. 但对于可逆的 $\mathcal{A}$ -矩阵 $\mathbf{K}(\mathcal{A})_{n \times n}$ (即存在 $\mathcal{A}$ -矩阵 $\mathbf{L}(\mathcal{A})_{n \times n}$ 使

$\mathbf{K}(\mathcal{A})\mathbf{L}(\mathcal{A}) = \mathbf{L}(\mathcal{A})\mathbf{K}(\mathcal{A}) = \mathbf{E}$) 对基 $\varepsilon_1 \cdots \varepsilon_n$ 作形式写法

$$\mathbf{K}(\mathcal{A}) \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix}.$$

则 $\eta_1, \eta_2, \dots, \eta_n$ 不一定是基, 甚至有些 η_i 可以为零.

例. 设

$$\mathcal{A} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix} = \begin{pmatrix} 4 & 6 & -15 \\ 1 & 3 & -5 \\ 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}, \quad \mathbf{Q}(\lambda) = \begin{pmatrix} 1 & -\lambda + 3 & -5 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

$|\mathbf{Q}(\lambda)|=1$. 故是可逆 λ - 阵. 设 $\mathbf{Q}(\lambda)\mathbf{L}(\lambda)=\mathbf{L}(\lambda)\mathbf{Q}(\lambda)=\mathbf{E}$ 则 $\mathbf{Q}(\mathcal{A})\mathbf{L}(\mathcal{A})=\mathbf{L}(\mathcal{A})\mathbf{Q}(\mathcal{A})=\mathbf{E}$. 即 $\mathbf{Q}(\mathcal{A})$ 是可逆 $\mathcal{A}$ - 矩阵. 但

$$\begin{aligned} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix} &= \mathbf{Q}(\mathcal{A}) \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix} = \begin{pmatrix} \varepsilon_1 + (-\mathcal{A} + 3\varepsilon)\varepsilon_2 - 5\varepsilon_3 \\ \varepsilon_2 - \varepsilon_3 \\ \varepsilon_3 \end{pmatrix} \\ &= \begin{pmatrix} \varepsilon_1 + (-\varepsilon_1 + 3\varepsilon_2 - 5\varepsilon_3) + 3\varepsilon_2 - 5\varepsilon_3 \\ \varepsilon_2 - \varepsilon_3 \\ \varepsilon_3 \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \varepsilon_2 - \varepsilon_3 \\ \varepsilon_3 \end{pmatrix} \end{aligned}$$

得 $\eta_1 = \mathbf{0}$.

§2 若当标准形的几何理论

先对若当形作一些几何刻画.

引理 1 对复数域上 n 维空间 V 上线性变换 $\mathcal{A}$, 下列是等价的:

(1) $\mathcal{A}$ 在某基 $\varepsilon_0, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_{n-1}$ 下的矩阵是若当块

$$\mathbf{J}(\lambda_0, n) = \begin{pmatrix} \lambda_0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & \lambda_0 & \cdots & 0 & 0 & 0 \\ & \cdots & & \cdots & & \\ 0 & 0 & \cdots & 1 & \lambda_0 & 0 \\ 0 & 0 & \cdots & 0 & 1 & \lambda_0 \end{pmatrix}_{n \times n},$$

(2) $\varepsilon_0, (\mathcal{A} - \lambda_0\varepsilon)\varepsilon_0, \dots, (\mathcal{A} - \lambda_0\varepsilon)^{n-1}\varepsilon_0$ 是基, 且 $(\lambda - \lambda_0)^n$ 是 ε_0 对于 $\mathcal{A}$ 的最小零化多项式 (最低次的使 $f(\mathcal{A})\varepsilon_0 = \mathbf{0}$ 的首 1 多项式)

(3) $V = P[\mathcal{A}]\varepsilon_0 = \{f(\mathcal{A})\varepsilon_0 \mid f(\lambda) \in P[\lambda]\}$, 且 $(\lambda - \lambda_0)^n$ 是 ε_0 对于 $\mathcal{A}$ 的最小零化多项式.

证 (1) $\Rightarrow$ (2): 由假设 $\varepsilon_0 = \varepsilon_0$, $\varepsilon_1 = (\mathcal{A} - \lambda_0\varepsilon)\varepsilon_0$, $\varepsilon_2 = (\mathcal{A} - \lambda_0\varepsilon)\varepsilon_1 = (\mathcal{A} - \lambda_0\varepsilon)^2\varepsilon_0 \cdots$,

$$\varepsilon_{n-1} = (\mathcal{A} - \lambda_0\varepsilon)^{n-1}\varepsilon_0. \quad (\mathcal{A} - \lambda_0\varepsilon)^n\varepsilon_0 = (\mathcal{A} - \lambda_0\varepsilon)\varepsilon_{n-1} = \mathbf{0}$$

$$\text{设任何} \quad l_0 + l_1\lambda + \cdots + l_{n-1}\lambda^{n-1} = k_0 + k_1(\lambda - \lambda_0) + \cdots + k_{n-1}(\lambda - \lambda_0)^{n-1} \quad (5)$$

$$\text{若} \quad (l_0\varepsilon + l_1\mathcal{A} + \cdots + l_{n-1}\mathcal{A}^{n-1})\varepsilon_0 = k_0\varepsilon_0 + k_1(\mathcal{A} - \lambda_0\varepsilon)\varepsilon_0 + \cdots + k_{n-1}(\mathcal{A} - \lambda_0\varepsilon)^{n-1}\varepsilon_0 = \mathbf{0}$$

$$\because \varepsilon_0, \dots, (\mathcal{A} - \lambda_0\varepsilon)^{n-1}\varepsilon_0 \text{ 是基} \Rightarrow k_{n-1} = \cdots = k_0 = 0$$

$$\because k_{n-1} = l_{n-1} \therefore l_{n-1} = 0 \quad \text{则} \quad l_0 + l_1\lambda + \cdots + l_{n-2}\lambda^{n-2} = k_0 + k_1(\lambda - \lambda_0) + \cdots + k_{n-2}(\lambda - \lambda_0)^{n-2} = 0$$

$\therefore l_{n-2} = k_{n-2} = 0$, 继续下去可有 $l_0 = l_1 = \cdots = l_{n-1} = 0$.

故 $(\lambda - \lambda_0)^n$ 是使 $f(\mathcal{A})\varepsilon_0 = \mathbf{0}$ 的最低次多项式, 即为 ε_0 对于 $\mathcal{A}$ 的最小多项式.

(2) $\Rightarrow$ (3). 任意 $v \in V$,

$$v = l_0\varepsilon_0 + l_1(\mathcal{A} - \lambda_0\varepsilon)\varepsilon_0 + \cdots + l_{n-1}(\mathcal{A} - \lambda_0\varepsilon)^{n-1}\varepsilon_0 = (l_0\varepsilon + \cdots + l_{n-1}(\mathcal{A} - \lambda_0\varepsilon)^{n-1})\varepsilon_0 \in P[\mathcal{A}]\varepsilon_0.$$

即 $V \subset P[\mathcal{A}]\varepsilon_0$. 反包含是显然的, $\therefore V = P[\mathcal{A}]\varepsilon_0$.

(3) $\Rightarrow$ (1). $(\lambda - \lambda_0)^n$ 是 ε_0 对于 $\mathcal{A}$ 的最小零化多项式.

任意 $v \in P[\mathcal{A}]\varepsilon_0$, 有 $V = f(\mathcal{A})\varepsilon_0$. 作带余除法,

$$f(\lambda) = g(\lambda)(\lambda - \lambda_0)^n + \gamma(\lambda), \quad \gamma(\lambda) = k_0 + k_1(\lambda - \lambda_0) + \cdots + k_{n-1}(\lambda - \lambda_0)^{n-1}$$

$v = f(\mathcal{A})\varepsilon_0 = \gamma(\mathcal{A})\varepsilon_0 \therefore v$ 是 $\varepsilon_0, (\mathcal{A} - \lambda_0\varepsilon)\varepsilon_0, \cdots, (\mathcal{A} - \lambda_0\varepsilon)^{n-1}\varepsilon_0$ 的线性组合. 又 V 是 n 维的,

它们必是基. 又由 $(\mathcal{A} - \lambda_0\varepsilon)[(\mathcal{A} - \lambda_0\varepsilon)^{n-1}\varepsilon_0] = \mathbf{0}$. 故 $\mathcal{A}$ 在上述基下矩阵为 $J(\lambda_0, n)$

定理 1 $V, \mathcal{A}$ 如上. 则 $\mathcal{A}$ 在某基下矩阵为若当形

$$\mathbf{A} = \begin{pmatrix} \mathbf{J}(\lambda_1, k_1) & & & \\ & \mathbf{J}(\lambda_2, k_2) & & \\ & & \ddots & \\ & & & \mathbf{J}(\lambda_s, k_s) \end{pmatrix} \quad (6)$$

的充要条件是有 $\eta_1, \cdots, \eta_s$ 使得

$$V = P[\mathcal{A}]\eta_1 \oplus \cdots \oplus P[\mathcal{A}]\eta_s, \quad (7)$$

见每个 η_i 对 $\mathcal{A}$ 的最小零化多项式是 $(\lambda - \lambda_i)^{k_i}$.

证 $\mathbf{A}$ 有形式 (6) $\Leftrightarrow V = W_1 \oplus \cdots \oplus W_s$, 是不变子空间的直和, 且每 $\mathcal{A}|_{W_i}$ 在 W_i 上有

基使它的矩阵是 $\mathbf{J}(\lambda_i, k_i) \Leftrightarrow V = W_1 \oplus \cdots \oplus W_s, W_i = P[\mathcal{A}]\eta_i, \eta_i$ 对 $\mathcal{A}$ 的最小零化多项式是

$(\lambda - \lambda_i)^{k_i}$.

由这个定理, 要证明若当形存在, 只要证明存在定理中的 $\mathcal{A}$ -循环子空间的直和分解.

定义 2 设 P 上线性空间 V 的线性变换 $\mathcal{A}$, W 是 V 的一个不变子空间, 若有 $\eta \in W$ 使 $W = P[\mathcal{A}]\eta$ 就称 W 为 V 的一个 $\mathcal{A}$ 循环子空间.

引理 2 $V, \mathcal{A}$ 如上, $W = P[\mathcal{A}]\eta, \eta$ 的最小零化多项式为 $p(\lambda)$. 则 $\dim W = \partial(p(\lambda))$

证 设 $\partial(p(\lambda)) = k$. 任何 $w \in W, w = f(\mathcal{A})\eta$. 作带余除法,

$$f(\lambda) = q(\lambda)p(\lambda) + \gamma(\lambda), \quad \gamma(\lambda) = l_0 + l_1\lambda + \cdots + l_{k-1}\lambda^{k-1}.$$

则

$$w = f(\mathcal{A})\eta = q(\mathcal{A})p(\mathcal{A})\eta + \gamma(\mathcal{A})\eta = \gamma(\mathcal{A})\eta = l_0\eta + l_1\mathcal{A}\eta + \cdots + l_{k-1}\mathcal{A}^{k-1}\eta.$$

$\therefore \eta$ 是 $\eta, \mathcal{A}\eta, \cdots, \mathcal{A}^{k-1}\eta$ 的组合.

由于 η 的最小零化多项式为 $p(\lambda)$, 它是 k 次的. 没有小于等于 $k-1$ 次多项式 $\gamma(\lambda)$, 使 $\gamma(\mathcal{A})\eta = \mathbf{0}$. 这等价于 $\eta, \mathcal{A}\eta, \cdots, \mathcal{A}^{k-1}\eta$ 是无关的, 即为 W 的基. 故 W 为 $k = \partial(p(\lambda))$ 维的.

下面先证明 V 是 $\mathcal{A}$ -循环子空间的直和的一个条件. 然后再进一步证明 V 为定理 1 中要求 $\mathcal{A}$ -循环子空间的直和.

引理 3 $V, \mathcal{A}$ 如上. 若有 $\eta_1, \cdots, \eta_s \in V$ 使得

$$(1) \quad V = P[\mathcal{A}]\eta_1 + \cdots + P[\mathcal{A}]\eta_s$$

$$(2) \quad \text{设每个 } \eta_i \text{ 的最小零化多项式为 } p_i(\lambda), \text{ 且 } \sum_{i=1}^s \partial(p_i(\lambda)) = \text{维 } V \text{ 则 (1) 的和是直和.}$$

证 维 $P[\mathcal{A}]\eta_i = \partial(p_i(\lambda))$ (引理 2). 于是

$$\text{维 } V = \sum_{i=1}^s \partial(p_i(\lambda)) = \sum_{i=1}^s \text{维 } (P[\mathcal{A}]\eta_i)$$

由第六章定理*11 知 (1) 的和是直和.

定理 2 $V, \mathcal{A}$ 如上, 则 V 是 $\mathcal{A}$ -循环子空间的直和

证 取 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 为基, 设

$$\mathcal{A} \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \mathbf{A} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} \quad \text{或} \quad (\mathcal{A}\mathbf{E} - \varepsilon\mathbf{A}) \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \mathbf{0} \quad (8)$$

对 $\lambda\mathbf{E} - \mathbf{A}$, 有 $\mathbf{P}(\lambda), \mathbf{Q}(\lambda)$ 可逆使

$$\mathbf{P}(\lambda)(\lambda\mathbf{E} - \mathbf{A})\mathbf{Q}(\lambda) = \begin{pmatrix} h_1(\lambda) & & & \\ & h_2(\lambda) & & \\ & & \ddots & \\ & & & h_n(\lambda) \end{pmatrix} \quad (9)$$

由 (8)

* 见北大编高等代数

$$\begin{aligned} \mathbf{O} &= P(\mathcal{A})(\mathcal{A}E - \varepsilon A) \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = [P(\mathcal{A})(\mathcal{A}E - \varepsilon A)\mathbf{Q}(\mathcal{A})]\mathbf{Q}^{-1}(\mathcal{A}) \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} \\ &= \begin{pmatrix} h_1(\mathcal{A}) & & & \\ & h_2(\mathcal{A}) & & \\ & & \ddots & \\ & & & h_n(\mathcal{A}) \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix}, \quad \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix} = \mathbf{Q}^{-1}(\mathcal{A}) \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}, \end{aligned} \quad (10)$$

由 (10)

$$h_i(\mathcal{A})\eta_i = \mathbf{0}.$$

$h_i(\lambda)$ 是 η_i 的零化多项式. 设 $p_i(\lambda)$ 是最小零化多项式则易知 $p_i(\lambda) \mid h_i(\lambda), \partial(p_i(\lambda)) \leq \partial(h_i(\lambda))$.

再比较 (9) 两端的行列式, 还有 $\sum_{i=1}^n \partial(h_i(\lambda)) = n = \text{维 } V$

又

$$\begin{aligned} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} &= \mathbf{Q}(\mathcal{A}) \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n q_{1j}\eta_j \\ \sum_{j=2}^n q_{2j}\eta_j \\ \vdots \\ \sum_{j=1}^n q_{nj}\eta_j \end{pmatrix} \\ \therefore \varepsilon_i &= \sum_{j=1}^n q_{ij}(\mathcal{A})\eta_j \in P[\mathcal{A}]\eta_1 + \cdots + P[\mathcal{A}]\eta_n. \end{aligned}$$

于是任 $v \in V$, v 是 $\varepsilon_1, \dots, \varepsilon_n$ 的线性组合, 故 $v \in P[\mathcal{A}]\eta_1 + \cdots + P[\mathcal{A}]\eta_n$ 由此易知 $V =$

$$P[\mathcal{A}]\eta_1 + \cdots + P[\mathcal{A}]\eta_n.$$

即得

$$\text{维 } V \leq \sum_{i=1}^n \text{维 } P[\mathcal{A}]\eta_i = \sum_{i=1}^n \partial(p_i(\lambda)) \leq \sum_{i=1}^n \partial(h_i(\lambda)) = n = \text{维 } V$$

于是所有 “ $\leq$ ” 成为 “ $=$ ”.

由此

$$\text{维 } V = \sum_{i=1}^n \text{维 } P[\mathcal{A}]\eta_i,$$

得直和 $V = P[\mathcal{A}]\eta_1 \oplus \cdots \oplus P[\mathcal{A}]\eta_n$.

又 $\sum_{i=1}^n \partial(p_i(\lambda)) = \sum_{i=1}^n \partial(h_i(\lambda))$ 及 $\partial(h_i(\lambda)) \geq \partial(p_i(\lambda)), i=1, 2, \dots, n$ 得 $\partial p_i(\lambda) = \partial(h_i(\lambda))$. 又因为

$p_i(\lambda) \mid h_i(\lambda)$, 加上皆为首 1, 故 $h_i(\lambda) = p_i(\lambda)$ 是 η_i 对 $\mathcal{A}$ 的最小零化多项式.

最后若某 $h_i(\lambda) = 1$, 则 $\eta_i = 1 \cdot \eta_i = \mathbf{0}$. 删去这些 η_i 将剩下的 η_i , 重新编号为 $\eta_1, \eta_2, \dots, \eta_s$. 则

$$V = P[\mathcal{A}] \eta_1 \oplus \cdots \oplus P[\mathcal{A}] \eta_s$$

且各 η_i 对 $\mathcal{A}$ 的最小多项式次数 ≥ 1 .

定理 3 复数域上, $V, \mathcal{A}$ 如上, 则有 $\xi_1, \dots, \xi_s$ 使

$$V = P[\mathcal{A}] \xi_1 \oplus \cdots \oplus P[\mathcal{A}] \xi_s$$

且 ξ_i 对 $\mathcal{A}$ 的最小零化多项式为 $(\lambda - \lambda_i)^{k_i}, k_i \geq 1$.

证 由定理 2 的结果. 只要对每个 $P[\mathcal{A}] \eta_i$ 继续分解直至达到本定理的要求. 我们先证明下述引理.

引理 4 P 上线性空间 V , $\mathcal{A}$ 是 V 上线性变换. W 是 $\mathcal{A}$ -循环子空间, $W = P[\mathcal{A}] \eta$. η 的最小零化多项式为

$$d(\lambda) = p_1^{l_1}(\lambda) p_2^{l_2}(\lambda) \cdots p_s^{l_s}(\lambda).$$

$p_1(\lambda), p_2(\lambda), \dots, p_s(\lambda)$ 是两两互素, 首 1 不可约多项式. 则有 $\xi_1, \xi_2, \dots, \xi_s$ 使

$$W = P[\mathcal{A}] \xi_1 \oplus \cdots \oplus P[\mathcal{A}] \xi_s,$$

且 ξ_i 的对于 $\mathcal{A}$ 的最小零化多项式为 $p_i^{l_i}(\lambda)$.

证 令 $m_i(\lambda) = p_1^{l_1}(\lambda) \cdots p_{i-1}^{l_{i-1}}(\lambda) p_{i+1}^{l_{i+1}}(\lambda) \cdots p_s^{l_s}(\lambda), 1 \leq i \leq s$. 作 $\xi_i = m_i(\mathcal{A}) \eta, 1 \leq i \leq s$ 则 ξ_i 的最小零化多项式是 $p_i^{l_i}(\lambda)$.

又 $m_1(\lambda), m_2(\lambda), \dots, m_s(\lambda)$ 是互素, 有 $u_1(\lambda), u_2(\lambda), \dots, u_s(\lambda)$ 使

$$u_1(\lambda) m_1(\lambda) + u_2(\lambda) m_2(\lambda) + \cdots + u_s(\lambda) m_s(\lambda) = 1$$

就有

$$\begin{aligned} \eta &= u_1(\mathcal{A}) m_1(\mathcal{A}) \eta + \cdots + u_s(\mathcal{A}) m_s(\mathcal{A}) \eta \\ &= u_1(\mathcal{A}) \xi_1 + u_2(\mathcal{A}) \xi_2 + \cdots + u_s(\mathcal{A}) \xi_s \in P[\mathcal{A}] \xi_1 + \cdots + P[\mathcal{A}] \xi_s \end{aligned}$$

推出

$$W = P[\mathcal{A}] \eta = P[\mathcal{A}] \xi_1 + P[\mathcal{A}] \xi_2 + \cdots + P[\mathcal{A}] \xi_s.$$

又

$$\sum_{i=1}^s (P[\mathcal{A}] \xi_i) \text{ 维} = \sum_{i=1}^s \partial(p_i^{l_i}(\lambda)) = \partial(p_1^{l_1}(\lambda) \cdots p_s^{l_s}(\lambda)) = \partial(d(\lambda)) = \text{维 } W \quad (\text{引理 2}),$$

再由引理 3, 得

$$W = P[\mathcal{A}] \xi_1 \oplus \cdots \oplus P[\mathcal{A}] \xi_s.$$

再回到定理 3 的证明. 对每个 $P[\mathcal{A}] \eta_i$ 应用引理 4 由于是复数域, 不可约多项式是一次

的. 引理 4 中各 $p_i^l(\lambda)$ 必形为 $(\lambda - \lambda_i)^l$. 故每个 $P[\mathcal{A}] \eta_i$ 可继续分解成定理 3 要求的直和. 于是 V 也能分解成定理 3 要求的直和.

定理 4 V 是复数域上 n 维线性空间, $\mathcal{A}$ 是 V 上线性变换. 则有 V 的基使 $\mathcal{A}$ 在该基下矩阵为若当形. 且除去各若当块排列顺序外是唯一决定的.

证 存在性是引理 1 与定理 3 的结果, 下证唯一性. 首先可计算任一若当块 $\mathbf{J}(\lambda_0, k)$ 的不变因子和初等因子. 对于

$$\mathbf{J}(\lambda_0, k) = \begin{pmatrix} \lambda_0 & & & & & \\ 1 & \lambda_0 & & & & \\ & 1 & \lambda_0 & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \lambda_0 & \\ & & & & 1 & \lambda_0 \end{pmatrix}_{k \times k}.$$

$\lambda \mathbf{E} - \mathbf{J}(\lambda_0, k)$ 的 k 级行列式因子为 $(\lambda - \lambda_0)^k$, 有 $k-1$ 级子式为 $(-1)^k$ 故它的不变因子为 $1, 1, \dots, 1, (\lambda - \lambda_0)^k$, 唯一初等因子为 $(\lambda - \lambda_0)^k$. 即若当块被它的初等因子唯一决定. 它设 $\mathcal{A}$ 的若当形为

$$\mathbf{A} = \begin{pmatrix} \mathbf{J}(\lambda_1, k_1) & & & \\ & \mathbf{J}(\lambda_2, k_2) & & \\ & & \ddots & \\ & & & \mathbf{J}(\lambda_s, k_s) \end{pmatrix}$$

$$\lambda \mathbf{E} - \mathbf{A} = \begin{pmatrix} \lambda \mathbf{E}_1 - \mathbf{J}(\lambda_1, k_1) & & & \\ & \lambda \mathbf{E}_2 - \mathbf{J}(\lambda_2, k_2) & & \\ & & \ddots & \\ & & & \lambda \mathbf{E}_s - \mathbf{J}(\lambda_s, k_s) \end{pmatrix}$$

它必等价于

$$\begin{pmatrix} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & (\lambda - \lambda_1)^{k_1} & & & & & & & \\ & & & & 1 & & & & & & \\ & & & & & \ddots & & & & & \\ & & & & & & 1 & & & & \\ & & & & & & & (\lambda - \lambda_2)^{k_2} & & & \\ & & & & & & & & \ddots & & \\ & & & & & & & & & 1 & \\ & & & & & & & & & & \ddots \\ & & & & & & & & & & & 1 \\ & & & & & & & & & & & & (\lambda - \lambda_s)^{k_s} \end{pmatrix}$$

由第六章定理 11, 后者的初等因子集是 $\{(\lambda - \lambda_i)^{k_i}, 1 \leq i \leq s\}$ 它也是 $\mathbf{A}$ 的, 因而是 $\mathcal{A}$ 的初等因

子集. 这样 $\mathcal{A}$ 的若当形中若当块的集合由 $\mathcal{A}$ 的初等因子集决定的, 即 $\mathcal{A}$ 的若当形中除去若当块的顺序外是由 $\mathcal{A}$ 唯一决定的.

推论 复矩阵 $\mathbf{A}$ 一定相似于一个若当形矩阵, 且其中各若当块除去顺序外由 $\mathbf{A}$ 唯一决定.

定理 5 复矩阵 $\mathbf{A}_{n \times n}$ 与 $\mathbf{B}_{n \times n}$ 相似的充要条件为 $\mathbf{A}$ 与 $\mathbf{B}$ 有相同的初等因子.

证 充分性设 $\mathbf{A}$ 相似于若当形

$$\mathbf{J}_A = \begin{pmatrix} \mathbf{J}(\lambda_1, k_1) & & & \\ & \mathbf{J}(\lambda_2, k_2) & & \\ & & \ddots & \\ & & & \mathbf{J}(\lambda_s, k_s) \end{pmatrix}$$

由于 $\mathbf{A}$ 与 $\mathbf{B}$ 有相同初等因子组, 故与 $\mathbf{B}$ 相似的若当形

$$\mathbf{J}_B = \begin{pmatrix} \mathbf{J}(\lambda_{i_1}, k_{i_1}) & & & \\ & \mathbf{J}(\lambda_{i_2}, k_{i_2}) & & \\ & & \ddots & \\ & & & \mathbf{J}(\lambda_{i_s}, k_{i_s}) \end{pmatrix}$$

其中 $i_1, i_2, \dots, i_s$ 是 $1, 2, \dots, s$ 的一个排列. 任取 n 维复空间 V 及它的基 $\varepsilon_{11}, \varepsilon_{12}, \dots, \varepsilon_{1k_1}, \varepsilon_{21}, \varepsilon_{22}, \dots, \varepsilon_{2k_2}, \dots, \varepsilon_{s1}, \varepsilon_{s2}, \dots, \varepsilon_{sk_s}$. 作线性变换 $\mathcal{A}$, 使它在这组基下的矩阵为 $\mathbf{J}_A$. 易见它在新基 $\varepsilon_{i_1 1}, \varepsilon_{i_1 2}, \dots, \varepsilon_{i_1 k_{i_1}}, \varepsilon_{i_2 1}, \varepsilon_{i_2 2}, \dots, \varepsilon_{i_2 k_{i_2}}, \dots, \varepsilon_{i_s 1}, \varepsilon_{i_s 2}, \dots, \varepsilon_{i_s k_{i_s}}$ 下的矩阵就是 $\mathbf{J}_B$. 因此 $\mathbf{J}_A$ 与 $\mathbf{J}_B$ 相似, 于是 $\mathbf{A}$ 与 $\mathbf{B}$ 相似.

必要性 $\mathbf{A}, \mathbf{B}$ 相似, 有 $\mathbf{T}$ 使 $\mathbf{T}^{-1}\mathbf{A}\mathbf{T} = \mathbf{B}$, 即有 $\mathbf{T}^{-1}(\lambda\mathbf{E} - \mathbf{A})\mathbf{T} = \lambda\mathbf{E} - \mathbf{B}$. 于是 $\lambda\mathbf{E} - \mathbf{A}$ 与 $\lambda\mathbf{E} - \mathbf{B}$ 等价. 故 $\mathbf{A}$ 与 $\mathbf{B}$ 有相同的初等因子.

推论 对 n 级复矩阵 $\mathbf{A}$ 与 $\mathbf{B}$, 下列为等价: (1) $\mathbf{A}$ 与 $\mathbf{B}$ 相似, (2) $\lambda\mathbf{E} - \mathbf{A}$ 与 $\lambda\mathbf{E} - \mathbf{B}$ 等价, (3) $\mathbf{A}, \mathbf{B}$ 有相同不变因子, (4) $\mathbf{A}, \mathbf{B}$ 有相同的初等因子.

前面关于若当形存在性的证明是构造性的. 它提供了计算若当形的具体步骤. 说明如下.

问题: 给定复矩阵 $\mathbf{B}_{n \times n}$, 找 $\mathbf{B}$ 的相似变换使它变成若当形.

步骤 1 取任意 n 维复线性空间 V 及它的任一组基 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$. 作线性变换 $\mathcal{A}$ 使它在该基下矩阵为 $\mathbf{B}$, 即 $\mathcal{A}(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)\mathbf{B}$. 令 $\mathbf{A} = \mathbf{B}^T$ 就有

$$\mathcal{A} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \mathbf{A} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}, \quad \text{或} \quad (\mathcal{A}\mathbf{E} - \mathbf{A}) \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \mathbf{0} \quad (11)$$

步骤 2 按照定理 2 的证明找 $\mathbf{P}(\lambda), \mathbf{Q}(\lambda)$ 可逆使

计算举例. 对矩阵

$$\mathbf{B} = \begin{pmatrix} 0 & -1 & 2 \\ 3 & 8 & -14 \\ 3 & 6 & -10 \end{pmatrix}$$

求 $\mathbf{T}$, 使 $\mathbf{T}^{-1}\mathbf{B}\mathbf{T}$ 为若当形

解 (I) 化为线性变换的问题

取三维复线性空间 V 及它的一组基 $\varepsilon_1, \varepsilon_2, \varepsilon_3$. 作 $\mathcal{A}$ 使

$$\mathcal{A}(\varepsilon_1 \ \varepsilon_2 \ \varepsilon_3) = (\varepsilon_1, \varepsilon_2, \varepsilon_3).$$

令 $\mathbf{A} = \mathbf{B}^T$, 可写成

$$\mathcal{A} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix} = \mathbf{A} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix} = \begin{pmatrix} 0 & 3 & 3 \\ -1 & 8 & 6 \\ 2 & -14 & -10 \end{pmatrix} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix} \quad (12)$$

II, 化 $\lambda\mathbf{E} - \mathbf{A}$ 为对角形. 可作多次初等变换来实现. 为得到最后的 $\mathbf{Q}(\lambda)$, 我们记录下每次所作的列变换.

$$\begin{aligned} \lambda\mathbf{E} - \mathbf{A} &= \begin{pmatrix} \lambda & -3 & -3 \\ 1 & \lambda-8 & -6 \\ -2 & 14 & \lambda+10 \end{pmatrix} \xrightarrow{\text{行变换}} \begin{pmatrix} 1 & \lambda-8 & -6 \\ \lambda & -3 & -3 \\ -2 & 14 & \lambda+10 \end{pmatrix} \xrightarrow{\text{行变换}} \begin{pmatrix} 1 & \lambda-8 & -6 \\ 0 & -3-\lambda(\lambda-8) & -3+6\lambda \\ 0 & 14+2(\lambda-8) & \lambda+10-12 \end{pmatrix} \\ &= \begin{pmatrix} 1 & \lambda-8 & -6 \\ 0 & -\lambda^2+8\lambda-3 & 6\lambda-3 \\ 0 & 2\lambda-2 & \lambda-2 \end{pmatrix} \xrightarrow{\text{右乘}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\lambda^2+8\lambda-3 & 6\lambda-3 \\ 0 & 2\lambda-2 & \lambda-2 \end{pmatrix} \xrightarrow{\text{右乘}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\lambda^2+8\lambda-3 & 6\lambda-3 \\ 0 & 2\lambda-2 & \lambda-2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \\ &\begin{pmatrix} 1 & 0 & 0 \\ 0 & -\lambda(\lambda-2) & 6\lambda-3 \\ 0 & \lambda & \lambda-2 \end{pmatrix} \xrightarrow{\text{行变换}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & \lambda-2 \\ 0 & 0 & \lambda^2+2\lambda+1 \end{pmatrix} \xrightarrow{\text{行变换}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2}\lambda & \frac{1}{2}\lambda-1 \\ 0 & 0 & (\lambda+1)^2 \end{pmatrix} \\ &\xrightarrow{\text{右乘}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{1}{2}\lambda-1 \\ 0 & -(\lambda+1)^2 & (\lambda+1)^2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -(\frac{1}{2}\lambda-1) \end{pmatrix} \xrightarrow{\text{右乘}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -(\lambda+1)^2 & \frac{1}{2}(\lambda+1)^2\lambda \end{pmatrix} \\ &\xrightarrow{\text{行变换}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda(\lambda+1)^2 \end{pmatrix} \end{aligned}$$

可得

$$\begin{aligned}
\mathbf{Q}(\lambda) &= \begin{pmatrix} 1 & -(\lambda-8) & 6 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \end{pmatrix}^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{1}{2}(\lambda-1) \\ 0 & 0 & 1 \end{pmatrix} \\
\mathbf{Q}^{-1}(\lambda) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{1}{2}\lambda-1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}^2 \begin{pmatrix} 1 & \lambda-8 & -6 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & \lambda-8 & -6 \\ 0 & \lambda-1 & \frac{1}{2}\lambda-1 \\ 0 & 2 & 1 \end{pmatrix} \\
\begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix} &= \mathbf{Q}^{-1}(\mathcal{A}) \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}
\end{aligned}$$

由于 η_1, η_2 的最小零化多项式为 1, 故 $\eta_1 = \eta_2 = \mathbf{0}$. 而 $\eta_3 = 2\varepsilon_2 + \varepsilon_3$. 于是 $V = P[\mathcal{A}] \eta_3$.

III. 继续分解 $P[\mathcal{A}] \eta_3$.

η_3 的最小零化多项式 $\lambda(\lambda+1)^2$. $P[\mathcal{A}] \eta_3 = P[\mathcal{A}] \eta'_1 \oplus P[\mathcal{A}] \eta'_2$ 其中 $\eta'_1 = (\mathcal{A} + \varepsilon)^2 \eta_3$, 它的最小零化多项式为 λ . 计算

$$\eta'_1 = (\mathcal{A} + \varepsilon)^2 (2\varepsilon_2 + \varepsilon_3) = 2(\mathcal{A} + \varepsilon)^2 \varepsilon_2 + (\mathcal{A} + \varepsilon)^2 \varepsilon_3.$$

$$(\mathcal{A} + \varepsilon)^2 \Rightarrow \begin{pmatrix} 1 & 3 & 3 \\ -1 & 9 & 6 \\ 2 & -14 & -9 \end{pmatrix}^2 = \begin{pmatrix} * & * & * \\ 2 & -6 & -3 \\ -2 & 6 & 3 \end{pmatrix}$$

$$\begin{aligned}
\eta'_1 &= 2(\mathcal{A} + \varepsilon)^2 \varepsilon_2 + (\mathcal{A} + \varepsilon)^2 \varepsilon_3 = 2(2\varepsilon_1 - 6\varepsilon_2 - 3\varepsilon_3) + (-2\varepsilon_1 + 6\varepsilon_2 + 3\varepsilon_3) \\
&= 2\varepsilon_1 - 6\varepsilon_2 - 3\varepsilon_3
\end{aligned}$$

$P[\mathcal{A}] \eta'_1$ 是一维的. 基就是 η'_1 , 且 $\mathcal{A} \eta'_1 = 0$. $\mathcal{A} |_{P[\mathcal{A}] \eta'_1}$ 在基 η'_1 下的矩阵是一阶矩阵 (0).

η'_2 的最小零化多项式为 $(\lambda+1)^2$, $\eta'_2 = \mathcal{A} \eta_3 = \mathcal{A}(2\varepsilon_2 + \varepsilon_3) = 2\mathcal{A} \varepsilon_2 + \mathcal{A} \varepsilon_3 \stackrel{\text{由(12)}}{=} 2(-\varepsilon_1 + 8\varepsilon_2 + 6\varepsilon_3) + 2\varepsilon_1 - 14\varepsilon_2 - 10\varepsilon_3 = 2\varepsilon_2 + 2\varepsilon_3$.

$P[\mathcal{A}] \eta'_2$ 是二维的, η'_2 的最小零化多项式为 $(\lambda+1)^2$, 则它的基是 η'_2 , $\eta'_3 = (\mathcal{A} + \varepsilon) \eta'_2 = 2(\mathcal{A} \varepsilon_2 + \mathcal{A} \varepsilon_3 + \varepsilon_2 + \varepsilon_3) \stackrel{\text{由(12)}}{=} 2\varepsilon_1 - 10\varepsilon_2 - 6\varepsilon_3 = 2(\varepsilon_2 + \varepsilon_3)$ 又 $(\mathcal{A} + \varepsilon)^2 \eta'_2 = (\mathcal{A} + \varepsilon)^2 \eta'_3 = 0$. 于是 $\mathcal{A} |_{P[\mathcal{A}] \eta'_2}$ 在基 η'_2, η'_3 下矩阵为 $\mathbf{J}(-1, 2)$.

最后 $\mathcal{A}$ 在基 $\eta'_1, \eta'_2, (\mathcal{A} + \varepsilon)\eta'_2 = \eta'_3$ 下矩阵为

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

基变换

$$(\eta'_1, \eta'_2, \eta'_3) = (\varepsilon_1 \quad \varepsilon_2 \quad \varepsilon_3) \mathbf{T} = (\varepsilon_1 \quad \varepsilon_2 \quad \varepsilon_3) \begin{pmatrix} 2 & 0 & 2 \\ -6 & 2 & -10 \\ -3 & 2 & -6 \end{pmatrix},$$

相似变换为

$$\mathbf{J} = \mathbf{T}^{-1} \mathbf{B} \mathbf{T},$$

即

$$\begin{pmatrix} 2 & 0 & 2 \\ -6 & 2 & -10 \\ -3 & 2 & -6 \end{pmatrix}^{-1} \begin{pmatrix} 0 & -1 & 2 \\ 3 & 8 & -14 \\ 3 & 6 & -10 \end{pmatrix} \begin{pmatrix} 2 & 0 & 2 \\ -6 & 2 & -10 \\ -3 & 2 & -6 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

§3 $\mathcal{A}$ -矩阵的两个应用

这一节利用 $\mathcal{A}$ -矩阵工具证明两个已知结果.

定理 6 (哈密尔顿-凯莱定理) 设数域 P 上 n 维线性空间 V 上线性变换 $\mathcal{A}$ 的特征多项式为 $f(\lambda)$, 则 $f(\mathcal{A}) = \mathbf{0}$.

证任给 V 的一组基 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$. 设 $\mathcal{A}$ 在 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 下的左矩阵为 $\mathbf{A}$, 即

$$(\mathcal{A}\mathbf{E} - \varepsilon\mathbf{A}) \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \mathbf{0}$$

则 $|\lambda\mathbf{E} - \mathbf{A}|$ 也是 $\mathcal{A}$ 的特征多项式. 即 $|\lambda\mathbf{E} - \mathbf{A}| = f(\lambda)$. 设 $\lambda\mathbf{E} - \mathbf{A}$ 的伴随矩阵为 $\mathbf{B}(\lambda)$, 它的元素是 $\lambda\mathbf{E} - \mathbf{A}$ 的 $n-1$ 级子式差一个符号, 仍是 λ 的多项式, 即 $\mathbf{B}(\lambda)$ 是 λ -矩阵. 由行列式理论

$$\mathbf{B}(\lambda)(\lambda\mathbf{E} - \mathbf{A}) = |\lambda\mathbf{E} - \mathbf{A}| \mathbf{E} = f(\lambda)\mathbf{E}$$

将 $\mathcal{A}$ 代替上式中的 λ , 两端都作用到 $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^T$ 上,

$$\text{左端} = \mathbf{B}(\mathcal{A})(\mathcal{A}E - \mathcal{A}) \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = 0$$

则

$$\text{右端} = f(\mathcal{A}) \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \begin{pmatrix} f(\mathcal{A})\varepsilon_1 \\ f(\mathcal{A})\varepsilon_2 \\ \vdots \\ f(\mathcal{A})\varepsilon_n \end{pmatrix} = 0.$$

即 $f(\mathcal{A})\varepsilon_i = \mathbf{0}, i=1,2,\dots,n$. 于是 $f(\mathcal{A})$ 是零变换,

$$f(\mathcal{A}) = 0.$$

引理 5 设 $\mathbf{A}_{n \times n}$ 是数字阵, $\mathbf{R}(\lambda)$ 是 n 级 λ -阵, 则存在 n 级 λ -阵 $\mathbf{M}_l(\lambda)$, $\mathbf{M}_r(\lambda)$ 及数字阵 $\mathbf{R}_l$ 及 $\mathbf{R}_r$ 使

$$\left. \begin{aligned} \mathbf{R}(\lambda) &= \mathbf{M}_l(\lambda)(\lambda E - \mathbf{A}) + \mathbf{R}_l \\ \mathbf{R}(\lambda) &= (\lambda E - \mathbf{A})\mathbf{M}_r(\lambda) + \mathbf{R}_r \end{aligned} \right\}$$

且这样的 $\mathbf{M}_l(\lambda)$, $\mathbf{M}_r(\lambda)$, $\mathbf{R}_l$ 及 $\mathbf{R}_r$ 由 $\mathbf{R}(\lambda)$ 及 $\mathbf{A}$ 唯一决定.

证 只证明 $\mathbf{M}_l(\lambda)$ 及 $\mathbf{R}_l$ 的存在及唯一性. 对 $\mathbf{M}_r(\lambda)$ 及 $\mathbf{R}_r$ 的存在及唯一性, 其证明是类似的. 将 $\mathbf{R}(\lambda)$ 写成

$$\mathbf{R}(\lambda) = \lambda^m \mathbf{R}_m + \lambda^{m-1} \mathbf{R}_{m-1} + \dots + \lambda \mathbf{R}_1 + \mathbf{R}_0,$$

$\mathbf{R}_m \neq \mathbf{O}$, 所有 $\mathbf{R}_m, \mathbf{R}_{m-1}, \dots, \mathbf{R}_1, \mathbf{R}_0$ 皆为 $n \times n$ 数字阵. 又写待定的 $\mathbf{M}_l(\lambda) = \lambda^{m-1} \mathbf{M}_{m-1} + \lambda^{m-2} \mathbf{M}_{m-2} + \dots + \lambda \mathbf{M}_1 + \mathbf{M}_0$, $\mathbf{M}_{m-1}, \dots, \mathbf{M}_0$ 皆为待定的数字阵. 则 $\mathbf{R}(\lambda) = \mathbf{M}_l(\lambda)(\lambda E - \mathbf{A}) + \mathbf{M}_l$ 的充分必要条件是

$$\begin{array}{ll} \mathbf{R}_m = \mathbf{M}_{m-1} & \text{或者} \quad \mathbf{M}_{m-1} = \mathbf{R}_m \\ \mathbf{R}_{m-1} = \mathbf{M}_{m-2} - \mathbf{M}_{m-1}\mathbf{A} & \mathbf{M}_{m-2} = \mathbf{R}_{m-1} + \mathbf{M}_{m-1}\mathbf{A} \\ \dots & \dots \\ \mathbf{R}_2 = \mathbf{M}_1 - \mathbf{M}_2\mathbf{A} & \mathbf{M}_1 = \mathbf{R}_2 + \mathbf{M}_2\mathbf{A} \\ \mathbf{R}_1 = \mathbf{M}_0 - \mathbf{M}_1\mathbf{A} & \mathbf{M}_0 = \mathbf{R}_1 + \mathbf{M}_1\mathbf{A} \\ \mathbf{R}_0 = \mathbf{M}_l - \mathbf{M}_0\mathbf{A} & \mathbf{R}_l = \mathbf{R}_0 + \mathbf{M}_0\mathbf{A} \end{array}$$

由右边这组式子可逐次求出 $\mathbf{M}_{m-1}, \mathbf{M}_{m-2}, \dots, \mathbf{M}_0, \mathbf{R}_l$ 且是唯一决定的.

定理 7 $\mathbf{A}_{n \times n}$ 与 $\mathbf{B}_{n \times n}$ 相似的充分必要条件是 $\lambda E - \mathbf{A}$ 与 $\lambda E - \mathbf{B}$ 等价.

证 必要性 设 $\mathbf{B} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$ ，则 $\lambda\mathbf{E} - \mathbf{B} = \mathbf{T}^{-1}(\lambda\mathbf{E} - \mathbf{A})\mathbf{T}$ ，故等价。

充分性 设 $\mathbf{P}(\lambda)$ 与 $\mathbf{Q}(\lambda)$ 可逆使 $\mathbf{P}(\lambda)(\lambda\mathbf{E} - \mathbf{A})\mathbf{Q}(\lambda) = \lambda\mathbf{E} - \mathbf{B}$ 。取 P 上 n 维空间 V ，基 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 。作线性变换 $\mathcal{A}$ 使得

$$\mathcal{A} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \mathbf{A} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} \quad \text{或} \quad ({}_{\mathcal{A}}\mathbf{E} - \varepsilon\mathbf{A}) \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \mathbf{O}. \quad (13)$$

由假设有

$$\mathbf{P}({}_{\mathcal{A}}\mathbf{E} - \varepsilon\mathbf{A}) = ({}_{\mathcal{A}}\mathbf{E} - \varepsilon\mathbf{B})\mathbf{Q}^{-1}({}_{\mathcal{A}}). \quad (14)$$

两端都作用到 $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^\top$ 上，由 (13) 左端为零，于是右端为零，即得

$$({}_{\mathcal{A}}\mathbf{E} - \varepsilon\mathbf{B})\mathbf{Q}^{-1}({}_{\mathcal{A}}) \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \mathbf{O}. \quad (15)$$

令

$$\begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix} = \mathbf{Q}^{-1}({}_{\mathcal{A}}) \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix},$$

$$\text{则 } ({}_{\mathcal{A}}\mathbf{E} - \varepsilon\mathbf{B}) \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix} = \mathbf{O}.$$

由引理 5 有 λ -阵 $\mathbf{R}(\lambda)$ 及数字阵 V_0 使

$$\mathbf{Q}^{-1}(\lambda) = \mathbf{R}(\lambda)(\lambda\mathbf{E} - \mathbf{A}) + V_0 \quad (16)$$

代入 (14) 式后，用 $\mathcal{A}$ 代替 λ ，则

$$\mathbf{P}({}_{\mathcal{A}}\mathbf{E} - \varepsilon\mathbf{A}) = ({}_{\mathcal{A}}\mathbf{E} - \varepsilon\mathbf{B})\mathbf{R}({}_{\mathcal{A}})({}_{\mathcal{A}}\mathbf{E} - \varepsilon\mathbf{A}) + ({}_{\mathcal{A}}\mathbf{E} - \varepsilon\mathbf{B})V_0$$

将两端都作用于 $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^\top$ ，用 (13) 得

$$({}_{\mathcal{A}}\mathbf{E} - \varepsilon\mathbf{B})V_0 \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \mathbf{O}. \quad (17)$$

若 V_0 是可逆的数字阵，则

$$\begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix} = V_0 \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} \quad (17')$$

是一组基且

$$({}^{\mathcal{A}}E - \varepsilon B) \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix} = \mathbf{O} \quad (18)$$

这说明 η 是 $\mathcal{A}$ 在基 $\eta_1, \eta_2, \dots, \eta_n$ 下的左矩阵. A 与 B 是 $\mathcal{A}$ 在不同基下的左矩阵, 故 A 与 B 相似.

下面来证 V_0 是可逆数字阵.

用 $Q(\lambda)$ 乘 (16) 的两端得

$$E = Q(\lambda)R(\lambda)(\lambda E - A) + Q(\lambda)V_0.$$

用 $\mathcal{A}$ 代替 λ 后, 两端同作用于 $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^T$, 再用上 (13), 则有

$$E \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = Q({}^{\mathcal{A}})V_0 \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = Q({}^{\mathcal{A}}) \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix} \quad (19)$$

用引理 5, 有 λ -阵 $N(\lambda)$ 及数字阵 U_0 使

$$Q(\lambda) = N(\lambda)(\lambda E - B) + U_0$$

将 $\mathcal{A}$ 替代上式中的 λ , 再将 (19) 式中 $Q({}^{\mathcal{A}})$ 中用替代 $\mathcal{A}$ 后的上式右端代入, 则 (19) 式成为

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = N({}^{\mathcal{A}})({}^{\mathcal{A}}E - \varepsilon B) \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix} + U_0 \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix}.$$

用 (17') 及 (18) 则上式就是

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = U_0 V_0 \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 是一组基, $U_0 V_0$ 又是数字阵, 必有 $U_0 V_0 = E$. 因此 V_0 可逆, 定理得证.