第一章 λ 矩阵与矩阵的 Jordan 标准型

1.1 λ 矩阵及其 Smith 标准型

记号 $\mathbb{F}[\lambda]$ 表示系数在 $\mathbb{F}$ 中的 λ 的多项式的全体；而 $\mathbb{F}(\lambda)$ 表示系数在 $\mathbb{F}$ 中的 λ 的有理分式的全体. 显然，$\mathbb{F}[\lambda]\subset\mathbb{F}(\lambda)$. 我们知道，两个多项式相除，结果不一定是多项式(只在整除时才是). 因此，作为运算系统，加、减、乘三法在 $\mathbb{F}[\lambda]$ 中总能进行，而除法不是总能. 这样的运算系统称为环，而 $\mathbb{F}[\lambda]$ 称为多项式环. 另一方面，在 $\mathbb{F}(\lambda)$ 中，四则运算皆能进行，这样的运算系统称为域，$\mathbb{F}(\lambda)$ 称为有理分式域.

定义 1.1.1 (多项式矩阵)

以多项式为元素的矩阵称为多项式矩阵，简称为 λ 矩阵. 记号 $(\mathbb{F}[\lambda])^{m\times n}$ 表示所有 m 行 n 列的 λ 矩阵的集合，矩阵的元素是系数在 $\mathbb{F}$ 中的 λ 的多项式. 也就是说，$A(\lambda)\in(\mathbb{F}[\lambda])^{m\times n}$ 表示 $A(\lambda)=\lfloor a_{ij}(\lambda)\rfloor_{m\times n}$，其中，$a_{ij}(\lambda)\in\mathbb{F}[\lambda]$.

注 1.1.1 (以多项式为元素的矩阵和以矩阵为系数的多项式)

如

$$A(\lambda)=\begin{bmatrix}\lambda^2+\lambda+1 & \lambda-1\\ 2\lambda & 1\end{bmatrix}$$

可以写成

$$A(\lambda)=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\lambda^2+\begin{bmatrix}1 & 1\\ 2 & 0\end{bmatrix}\lambda+\begin{bmatrix}1 & -1\\ 0 & 1\end{bmatrix}.$$

记号 $\mathbb{F}^{m\times n}[\lambda]$ 表示以 $\mathbb{F}^{m\times n}$ 中元素为“系数”，λ 为未定元的多项式的集合. 记号 $(\mathbb{F}[\lambda])^{m\times n}$ 表示以元素在 $\mathbb{F}[\lambda]$ 中的 m 行 n 列的矩阵的集合. 如上所述，这两个集合是“同构”的.

多项式矩阵和通常矩阵的主要区别在于，其元素所在的运算系统—多项式环 $\mathbb{F}[\lambda]$—不是一个域，多项式矩阵是环上的矩阵. 所以通常矩阵的性质中，那些涉及到元素除法的，就可能不再成立了.

注意到矩阵的行列式的定义只涉及到加、减、乘三种运算，因此方形的多项式矩阵的行列式，以及不必方的多项式矩阵的各阶子行列的值都是多项式.

定义 1.1.2 (多项式矩阵的秩)

多项式矩阵的秩，也用 rank 表示，是指其值为非零多项式的子行列式的最大阶数. 换言之，多项式矩阵的秩为 r 是指：存在 r 阶子行列式，其值为非零多项式；且所有阶数 $\geq r+1$ 的子行列式的值均为零多项式.

定义 1.1.3 (单位模阵)

多项式矩阵 $U(\lambda)\in\mathbb{F}^{n\times n}[\lambda]$ 称为单位模阵(简称幺模阵)，若存在多项式矩阵 $V(\lambda)\in\mathbb{F}^{n\times n}[\lambda]$ 使得

$U(\lambda)V(\lambda)=V(\lambda)U(\lambda)=I_n$. (1.1.1)

换言之，多项式方阵为单位模阵，若有多项式矩阵是它的逆矩阵，或曰在多项式矩阵的范围 $U(\lambda)\in\mathbb{F}^{n\times n}[\lambda]$ 内可逆.

注 1.1.2 (多项式矩阵的逆一般情况下是有理分式矩阵)

只要 $A(\lambda)\in\mathbb{F}^{n\times n}[\lambda]$ 满足行列式 $\det(A(\lambda))=|A(\lambda)|\in\mathbb{F}[\lambda]$ 为非零多项式，就可以用公式

$$(A(\lambda))^{-1}=\frac{\operatorname{adj}(A(\lambda))}{\det(A(\lambda))}\quad(1.1.2)$$

计算逆矩阵. 然而这样求得的逆矩阵一般情况下会是一个有理分式矩阵.

定理 1.1.1 (单位模阵的行列式刻画)

多项式矩阵 $U(\lambda)\in\mathbb{F}^{n\times n}[\lambda]$ 是单位模阵，当且仅当行列式 $\det(A(\lambda))\in\mathbb{F}[\lambda]$ 为非零常值多项式.

证明: “当”部分. 易见, $\operatorname{adj}(U(\lambda))\in\mathbb{F}^{n\times n}[\lambda]$ 为多项式矩阵. 因此, 若 $\det(U(\lambda))\in\mathbb{F}[\lambda]$ 为非零常值多项式, 则由公式(1.1.2)知 $(U(\lambda))^{-1}\in\mathbb{F}^{n\times n}[\lambda]$ 为多项式矩阵, 即 $U(\lambda)$ 在 $\mathbb{F}^{n\times n}[\lambda]$ 中可逆.
“仅当”部分. 由定义, 有(1.1.1)成立. 由矩阵乘法的行列式等于行列式的乘法, 有
$|U(\lambda)|\cdot|V(\lambda)|=1$.
因此行列式 $|U(\lambda)|$ 及 $|V(\lambda)|$ 均为非零多项式, 且它们的次数满足
$$\deg(|U(\lambda)|)+\deg(|V(\lambda)|)=\deg(1)=0.$$
因此, 必有
$$\deg(|U(\lambda)|)=\deg(|V(\lambda)|)=0.$$
即 $|U(\lambda)|$ 和 $|V(\lambda)|$ 均为零次多项式, 也就是非零常值多项式. 证毕

定义 1.1.4 (多项式矩阵的三种初等行(列)变换)
(1) 互换矩阵的某两行, 例如 i_1 和 i_2 两行. 该变换记为 $(\mathrm{r}_{i_1})\leftrightarrow(\mathrm{r}_{i_2})$. 这里字母 r 表示行(row). 类似的, 用字母 c 表示列(column).
(2) 将某行, 例如 i 行, 乘以非零常数 $c\in\mathbb{F}, c\neq 0$. 记为 $c\cdot(\mathrm{r}_i)$.
(3) 将某行例如 i_1 行, 乘以一个多项式如 $f(\lambda)\in\mathbb{F}[\lambda]$, 加到另一不同行例如 $i_2\neq i_1$ 行上去. 记为 $(\mathrm{r}_{i_2})+f(\lambda)\cdot(\mathrm{r}_{i_1})$.

和通常的数域上的矩阵的初等变换完全类似, 初等行(列)变换可用左(相应地, 右)乘以相应的初等矩阵来实现. 易见, 三种初等矩阵都是单位模阵.

和通常的数域上的矩阵的初等变换相比, 只有类型(2)需要加以说明: 看上去合理的定义似乎应该是

(2') 将某行, 例如 i 行, 乘以非零多项式 $f(\lambda)\in\mathbb{F}[\lambda]$.

然而, 这样的变换在多项式的范围内就不可逆了, 相应于(2')的“初等矩阵”一般情况($f(\lambda)$ 不是非零常数)就不再是单位模阵了.

定义 1.1.5 (多项式矩阵的等价)
两个多项式矩阵 $A(\lambda),B(\lambda)\in\mathbb{F}^{m\times n}[\lambda]$ 称为等价, 如果 $A(\lambda)$ 可经有限个初等行、列变换化成 $B(\lambda)$.

下面研究一个给定的多项式矩阵, 可以化成什么样的“最简等价型”. 这就是所谓的多项式矩阵的 Smith 标准型.

引理 1.1.1 (用初等变换将左上角降次)
设多项式矩阵 $A(\lambda)$ 的左上角元素 $a_{11}(\lambda)\neq 0$ (不是零多项式, 下同), 并且 $A(\lambda)$ 中至少有一个元素不能被 $a_{11}(\lambda)$ 整除. 那么 $A(\lambda)$ 等价于一个多项式矩阵 $B(\lambda)$, 使得 $b_{11}(\lambda)\neq 0$, 且 $\deg(b_{11}(\lambda))<\deg(a_{11}(\lambda))$.
证明: 分三种情况
1) 在第一行中有一个元素, 记为 $a_{1j}(\lambda)$, $j\neq 1$, 不能被 $a_{11}(\lambda)$ 整除. 用 $a_{11}(\lambda)$ 除 $a_{1j}(\lambda)$, 有带余除法
$a_{1j}(\lambda)=a_{11}(\lambda)q(\lambda)+r(\lambda)$,
其中
$r(\lambda)\neq 0$ 且 $\deg(r(\lambda))<\deg(a_{11}(\lambda))$.
对 $A(\lambda)$ 实施列变换 $(\mathrm{c}_j)+(\mathrm{c}_1)\cdot(-q(\lambda))$ 后, 位于 $(1,j)$-位置的元素变为 $r(\lambda)$. 再作列变换 $(\mathrm{c}_j)\leftrightarrow(\mathrm{c}_1)$ 即得.
2) 在第一行中没有, 但在第一列中有一个元素不能被 $a_{11}(\lambda)$ 整除. 此情形完全类似 1).
3) 在第一行和第一列中都没有, 但其他位置有一个元素, 记为 $a_{ij}(\lambda)$, $i\neq 1, j\neq 1$, 不能被 $a_{11}(\lambda)$ 整除. 则由 $\frac{a_{1j}(\lambda)}{a_{11}}, \frac{a_{i1}(\lambda)}{a_{11}}$ 为多项式, 依次用如下初等变换作用在 $A(\lambda)$ 上:
$(\mathrm{c}_j)+(\mathrm{c}_1)\cdot(-\frac{a_{1j}(\lambda)}{a_{11}(\lambda)}))$, $(\mathrm{r}_i)+(-\frac{a_{i1}(\lambda)}{a_{11}(\lambda)}))\cdot(\mathrm{r}_1)$, $(\mathrm{r}_1)+1\cdot(\mathrm{r}_i)$.
则此时, 在 $(1,1)$-位置的元素依然是 $a_{11}(\lambda)$; 但 $(1,j)$-位置的元素变为 $a_{ij}(\lambda)-a_{i1}a_{ij}/a_{11}$. 由 a_{i1},a_{ij} 均被 a_{11} 整

除，而 $a_{ij}(\lambda)$ 不被 $a_{11}(\lambda)$ 整除，知 $a_{ij}(\lambda)-a_{i1}a_{ij}/a_{11}$ 不被 $a_{11}(\lambda)$ 整除．于是问题化为情形 1)而得证．证毕

定理 1.1.2 (多项式矩阵的 Smith 标准型)

任意多项式矩阵 $A(\lambda)\in\mathbb{F}^{m\times n}[\lambda]$ 等价于下面的 Smith 标准型

$$\left[\begin{array}{cccc|c} d_1(\lambda) & & & 0 & \\ & d_2(\lambda) & & & \\ & & \ddots & & 0 \\ 0 & & & d_r(\lambda) & \\ \hline & & 0 & & 0_{m-r,n-r} \end{array}\right], \quad (1.1.3)$$

其中 $d_i(\lambda), i=1,2,\ldots,r$，为首项系数为 1 的非零多项式，满足 $d_i(\lambda)$ 整除 $d_{i+1}(\lambda)$，记为 $d_i(\lambda)\,|\,d_{i+1}(\lambda)$，$i=1,2,\ldots,r-1$．

证明：若 $A(\lambda)$ 为零矩阵，则结论显然．下设 $A(\lambda)$ 为非零矩阵．其非零元素中次数最低者，不妨设为 $a_{11}(\lambda)$．记次数 $\partial(a_{11}(\lambda))=t_1$．为叙述方便，我们简称"$A(\lambda)$ 可经适当初等变换后化为 $A^1(\lambda)=\left[a_{ij}^1(\lambda)\right]_{m\times n}$，且 $a_{11}^1(\lambda)\neq 0$，$a_{11}^1(\lambda)$ 是 $A^1(\lambda)$ 中非零元素之次数最低者，$\deg(a_{11}^1(\lambda))<\deg(a_{11}(\lambda))$"为"可对 $A(\lambda)$ 进行一次规约"．由于非零多项式的次数总是 ≥ 0，而 $\partial(a_{11}(\lambda))=t_1$，因此不可能相继进行 t_1+1 次规约．也就是至多在 t_1+1 次以前，$A(\lambda)$ 经若干次，例如 k 次规约得到 $A^k(\lambda)$，而对 $A^k(\lambda)$ 则不再能进行一次规约了．由引理 1.1.1，这说明 $a_{11}^k(\lambda)\neq 0$，且 $a_{11}^k(\lambda)$ 整除 $A^k(\lambda)$ 所有其他元素．再利用第三类初等行，列变换，即可进一步化 $A^k(\lambda)$ 为

$$\left[\begin{array}{c|c} a_{11}^k(\lambda) & 0 \\ \hline 0 & \widetilde{A}(\lambda) \end{array}\right],$$

其中 $\widetilde{A}(\lambda)\in\mathbb{F}^{(m-1)\times(n-1)}[\lambda]$，且 $a_{11}^k(\lambda)$ 整除 $\widetilde{A}(\lambda)$ 的所有元素．易知，$a_{11}^k(\lambda)$ 也整除任一与 $\widetilde{A}(\lambda)$ 等价的多项式矩阵的所有元素．若 $\widetilde{A}(\lambda)$ 为零矩阵，则结束；若 $\widetilde{A}(\lambda)$ 不为零矩阵，则继续上述过程．这样一直进行下去，即得证定理．证毕

问题 1.1.1 (Smith 型的唯一性)

Smith 型是否唯一？确切的说，若同一多项式矩阵 $A(\lambda)\in\mathbb{F}^{m\times n}[\lambda]$ 分别被两个不同的初等变换序列化成 Smith 标准型

$$\left[\begin{array}{cccc|c} d_1(\lambda) & & & 0 & \\ & d_2(\lambda) & & & \\ & & \ddots & & 0 \\ 0 & & & d_r(\lambda) & \\ \hline & & 0 & & 0_{m-r,n-r} \end{array}\right] \text{ 及 } \left[\begin{array}{cccc|c} e_1(\lambda) & & & 0 & \\ & e_2(\lambda) & & & \\ & & \ddots & & 0 \\ 0 & & & e_t(\lambda) & \\ \hline & & 0 & & 0_{m-t,n-t} \end{array}\right],$$

那么，是否有 $r=t$？若 $r=t$，是否 $d_i(\lambda)=e_i(\lambda)$，$i=1,2,\ldots,r$？

设 $f(\lambda)$，$g(\lambda)\in\mathbb{F}[\lambda]$．回忆：若 $f(\lambda)$ 为非零多项式，且 $f(\lambda)\,|\,g(\lambda)$，则称 $f(\lambda)$ 为 $g(\lambda)$ 的一个因式，$g(\lambda)$ 称为 $f(\lambda)$ 的一个倍式．特别注意，任意非零多项式是零多项式的因式．

定义 1.1.6 (多项式矩阵的行列式因子)

设 $A(\lambda)\in\mathbb{F}^{m\times n}[\lambda]$，正整数 $k\leq \mathrm{rank}(A(\lambda))$ 的秩．$A(\lambda)$ 的 k 阶行列式因子是指 $A(\lambda)$ 的所有 k 阶子式—一共是 $\binom{m}{k}\binom{n}{k}$ 个多项式—的最高公因式(首项系数不妨规定成 1)．这里，符号

$$\binom{m}{k}=\frac{m!}{k!(m-k)!}.$$

另外，若 $\mathrm{rank}(A(\lambda))<k\leq\min\{m,n\}$，$A(\lambda)$ 的 k 阶行列式因子规定为零多项式．

定理 1.1.3 (初等行、列变换不改变多项式矩阵的行列式因子)

设两个多项式矩阵 $A(\lambda), B(\lambda) \in \mathbb{F}^{m\times n}[\lambda]$ 等价，则它们的各阶行列式因子分别相同.

证明：只证明 $A(\lambda)$ 经过一次第三类初等行变换，记为 $(\mathrm{r}_{i_2}) + f(\lambda)\cdot(\mathrm{r}_{i_1})$，变为 $B(\lambda)$ 的情形. 任给 $k \le \min\{m,n\}$，记

$\mathcal{A}_k = \{A(\lambda)\text{的}k\text{阶子式}\}$，$\mathcal{B}_k = \{B(\lambda)\text{的}k\text{阶子式}\}$.

注意在经变换 $(\mathrm{r}_{i_2}) + f(\lambda)\cdot(\mathrm{r}_{i_1})$ 后，只有第 i_2 行发生了变化. 任给 $b(\lambda) \in \mathcal{B}_k$，根据 k 阶行列子式 $b(\lambda)$ 在矩阵 $B(\lambda)$ 中的行是否含第 i_1，第 i_2 行分三种情况:

1) $b(\lambda)$ 在矩阵 $B(\lambda)$ 中的行不含第 i_2 行. 此时 $b(\lambda) \in \mathcal{A}_k$.
2) $b(\lambda)$ 在矩阵 $B(\lambda)$ 中的行含第 i_2 行，并同时含第 i_1 行. 此时由行列式的性质，亦有 $b(\lambda) \in \mathcal{A}_k$.
3) $b(\lambda)$ 在矩阵 $B(\lambda)$ 中的行含第 i_2 行，但不含有第 i_1 行. 此时由行列式的性质，有 $a_1(\lambda), a_2(\lambda) \in \mathcal{A}_k$ 使得 $b(\lambda) = a_1(\lambda) \pm f(\lambda)a_2(\lambda)$.

这里，加减号的选择取决于在行列子式 $b(\lambda)$ 中，将第 i_2 行替换为第 i_1 行后引起的逆序数的奇偶性. 上述三种情况都表明，$\mathcal{A}_k$ 的任一公因式都整除 $b(\lambda)$. 再由 $b(\lambda) \in \mathcal{B}_k$ 的任意性，即知 $\mathcal{A}_k$ 的任一公因式都是 $\mathcal{B}_k$ 的公因式. 由于 $B(\lambda)$ 也可经一次第三类初等行变换变为 $A(\lambda)$，同理也有 $\mathcal{B}_k$ 的任一公因式都是 $\mathcal{A}_k$ 的公因式. 于是，$\mathcal{A}_k$ 与 $\mathcal{B}_k$ 有完全相同的公因式集合. 证毕

定理 1.1.4 (多项式矩阵的 Smith 型，行列式因子，不变因子三者相互唯一决定)

设多项式矩阵 $A(\lambda) \in \mathbb{F}^{m\times n}[\lambda]$ 的 Smith 标准型为

$$\left[\begin{array}{cccc|c} d_1(\lambda) & & & 0 & \\ & d_2(\lambda) & & & 0 \\ & & \ddots & & \\ 0 & & & d_r(\lambda) & \\ \hline & & 0 & & 0_{m-r,n-r} \end{array}\right]. \quad (1.1.4)$$

则

(1) $r = \operatorname{rank}(A(\lambda))$.

(2) 记 $A(\lambda)$ 的 k 阶行列式因子为 $D_k(\lambda)$，则有

$D_k(\lambda) = d_1(\lambda)d_2(\lambda)\cdots d_k(\lambda)$，$k = 1,2,\ldots,r$，

或

$$d_1(\lambda) = D_1(\lambda),\ d_k(\lambda) = \frac{D_k(\lambda)}{D_{k-1}(\lambda)},\quad k = 2,\ldots,r.$$

由于行列式因子是由 $A(\lambda)$ 唯一决定的(与初等变换无关). 由此 $d_i(\lambda), i = 1,2,\ldots,r$ 也是由 $A(\lambda)$ 唯一决定的，与初等变换无关，称为 $A(\lambda)$ 的不变因子.

例 1.1.1 (利用行列式因子求多项式矩阵的不变因子和 Smith 标准型)

一般情况下，这种方法工作量巨大. 但有些特殊情况下，用行列式因子求不变因子和 Smith 标准型很方便.

$$\begin{bmatrix} \lambda(\lambda+1) & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & (\lambda+1)^2 \end{bmatrix}$$

方法一：用初等变换化 Smith 型.

$$\begin{bmatrix}\lambda(\lambda+1) & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & (\lambda+1)^2\end{bmatrix}\to\begin{bmatrix}\lambda(\lambda+1) & 0 & 0\\ 0 & \lambda & (\lambda+1)^2\\ 0 & 0 & (\lambda+1)^2\end{bmatrix}$$

$$\to\begin{bmatrix}\lambda(\lambda+1) & & \\ & \lambda & 1\\ & & (\lambda+1)^2\end{bmatrix}$$

$$\to\begin{bmatrix}\lambda(\lambda+1) & 0 & 0\\ 0 & 0 & 1\\ 0 & -\lambda(\lambda+1)^2 & (\lambda+1)^2\end{bmatrix}$$

$$\to\begin{bmatrix}\lambda(\lambda+1) & 0 & 0\\ 0 & 0 & 1\\ 0 & -\lambda(\lambda+1)^2 & 0\end{bmatrix}$$

$$\to\begin{bmatrix}1 & 0 & 0\\ 0 & \lambda(\lambda+1) & 0\\ 0 & 0 & \lambda(\lambda+1)^2\end{bmatrix}.$$

方法二: 先求出各阶行列式因子, 再求出不变因子和 Smith 型.

1) 不为零的一阶子式有

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

公因子为 1, 故 $D_1(\lambda)=1$.

2) 不为零的二阶子式有

$$\begin{vmatrix}\lambda(\lambda+1) & 0\\ 0 & \lambda\end{vmatrix}=\lambda^2(\lambda+1),\quad \begin{vmatrix}\lambda(\lambda+1) & 0\\ 0 & (\lambda+1)^2\end{vmatrix}=\lambda(\lambda+1)^3,\quad \begin{vmatrix}\lambda & 0\\ 0 & (\lambda+1)^2\end{vmatrix}=\lambda(\lambda+1)^2,$$

公因子为 $\lambda(\lambda+1)$, 故 $D_2(\lambda)=\lambda(\lambda+1)$.

3) 不为零的三阶子式有

$$\begin{vmatrix}\lambda(\lambda+1) & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & (\lambda+1)^2\end{vmatrix}=\lambda^2(\lambda+1)^3,$$

故 $D_3(\lambda)=\lambda^2(\lambda+1)^3$

于是,

$$d_1(\lambda)=D_1(\lambda)=1,$$
$$d_2(\lambda)=D_2(\lambda)/D_1(\lambda)=\lambda(\lambda+1),$$
$$d_3(\lambda)=D_3(\lambda)/D_2(\lambda)=\lambda(\lambda+1)^2.$$

下面研究单位模阵的 Smith 标准型.

定理 1.1.5 (单位模阵写为初等矩阵的乘积)

(1) 单位模阵的 Smith 标准型为单位矩阵.

(2) 单位模阵可写为有限个初等矩阵的乘积.

证明: 设 $U(\lambda)\in\mathbb{F}^{n\times n}$ 为么模阵.

(1) 由行列式 $|U(\lambda)|$ 为非零常数, 知 $U(\lambda)$ 的 n 阶行列式因子 $D_n(\lambda)=1$, 于是

$d_1(\lambda)d_2(\lambda)\cdots d_n(\lambda)=1$,

于是

$d_1(\lambda)=d_2(\lambda)=\cdots=d_n(\lambda)=1$,

于是 $U(\lambda)$ 的 Smith 标准型为单位矩阵 I_n.

(2) 设 $U(\lambda)$ 依次经过 s 次初等行变换, 相应的初等矩阵记为 $P_1(\lambda),\ldots,P_s(\lambda)$, t 次初等行变换, 相应的初等

矩阵记为$Q_1(\lambda),\ldots,Q_t(\lambda)$，化成了 Smith 标准型$I_n$，这相当于
$P_s(\lambda)\cdots P_1(\lambda)U(\lambda)Q_1(\lambda)\cdots Q_t(\lambda)=I_n$.
于是
$U(\lambda)=(P_1(\lambda))^{-1}\cdots(P_s(\lambda))^{-1}(Q_t(\lambda))^{-1}\cdots(Q_1(\lambda))^{-1}$.
由于初等矩阵本身也是幺模阵，且其逆矩阵也还是幺模阵，上述表面结论(2)成立．证毕

推论 1.1.1 (多项式矩阵等价的幺模阵表述)
(1) 两个多项式矩阵$A(\lambda),B(\lambda)\in\mathbb{F}^{m\times n}$相互等价，当且仅当存在两个幺模阵$U(\lambda)\in\mathbb{F}^{m\times m}$，$V(\lambda)\in\mathbb{F}^{n\times n}$使得$U(\lambda)A(\lambda)V(\lambda)=B(\lambda)$.
(2) 给定多项式矩阵$A(\lambda)\in\mathbb{F}^{m\times n}$，存在两个幺模阵$U(\lambda)\in\mathbb{F}^{m\times m}$，$V(\lambda)\in\mathbb{F}^{n\times n}$使得

$$U(\lambda)A(\lambda)V(\lambda)=\left[\begin{array}{cccc|c} d_1(\lambda) & & & 0 & \\ & d_2(\lambda) & & & \\ & & \ddots & & 0 \\ 0 & & & d_r(\lambda) & \\ \hline & & 0 & & 0_{m-r,n-r}\end{array}\right]$$

为$A(\lambda)$的 Smith 标准型.

1.2 数域上矩阵的特征矩阵

定义 1.2.1 (特征矩阵)
给定$A\in\mathbb{F}^{n\times n}$，多项式矩阵

$$\lambda I_n-A=\begin{bmatrix}\lambda-a_{11} & \cdots & -a_{1,n}\\ \vdots & \ddots & \vdots\\ -a_{n,1} & \cdots & \lambda-a_{n,n}\end{bmatrix}\in\mathbb{F}^{n\times n}[\lambda]\quad(1.2.1)$$

称为A的特征矩阵.

引入特征矩阵的好处在于，可将数域上矩阵的相似问题，转化为特征矩阵作为多项式矩阵的等价问题．因为有下面的定理:

定理 1.2.1 (两个矩阵相似当且仅当它们的特征矩阵等价)
两个矩阵$A,B\in\mathbb{F}^{n\times n}$相似，当且仅当两个特征矩阵$\lambda I-A,\lambda I-B\in\mathbb{F}^{n\times n}[\lambda]$作为多项式矩阵等价.

为证明此定理，我们需要下面的两个引理．为简化记号，我们有时不明显写出数域$\mathbb{F}$，矩阵集合$\mathbb{F}^{n\times n}$和多项式矩阵集合$\mathbb{F}^{n\times n}[\lambda]$．读者容易根据上下文断定．另外，根据注 1.1.1，以多项式为元素的矩阵(多项式矩阵)和以矩阵为系数的多项式(简称“矩阵多项式”)两种表达形式在本节中同时使用，读者应仔细体会．在式(1.2.1)的等号两边就用了两种不同的表达．可用明显的方式定义多项式矩阵$A(\lambda)$的次数$\deg(A(\lambda))$如下

定义 1.2.2 (多项式矩阵的次数)
(1) 零多项式矩阵的次数无意义，可规定为无穷.
(2) 设非零多项式矩阵$A(\lambda)\in\mathbb{F}^{m\times n}[\lambda]$的矩阵多项式表示如下
$A(\lambda)=A_0+A_1\lambda+\cdots+A_d\lambda^d$，
其中，矩阵$A_i\in\mathbb{F}^{m\times n}, i=0,1,\ldots,d$且$A_d\neq 0$．则称$A(\lambda)$的次数为$d$，并记为$\deg(A(\lambda))=d$．特别地，按定义，零次多项式矩阵就是普通的数域上的非零矩阵$A(\lambda)=A_0\neq 0$.

通常多项式理论中，以多项式的次数概念为基础的比较系数法的推理方法占有核心地位．对矩阵多项式，情形是类似的.

引理 1.2.1 (矩阵多项式乘积的次数)
设$A(\lambda)\in\mathbb{F}^{m\times m}[\lambda]$，$B(\lambda),C(\lambda)\in\mathbb{F}^{m\times n}[\lambda]$，且$A(\lambda)B(\lambda)=C(\lambda)$．若

$A(\lambda)=A_0+A_1\lambda+\cdots+A_q\lambda^q$

满足 A_q 非奇异，$B(\lambda)\neq 0$，则有

$\deg(A(\lambda))+\deg(B(\lambda))=\deg(C(\lambda))$.

证明：设 $\deg(B(\lambda))=p$，即

$B(\lambda)=B_0+B_1\lambda+\cdots+B_p\lambda^p, B_p\neq 0$.

则由 A_q 非奇异知 $A_qB_p\neq 0$，从而

$C(\lambda)=A_0B_0+\cdots+A_qB_p\lambda^{p+q}, A_qB_p\neq 0$.

因此

$q+p=\deg(C(\lambda))$. 证毕

引理 1.2.2 (矩阵多项式的带余除法)

设

$A(\lambda)=A_0+A_1\lambda+\cdots+A_q\lambda^q\in\mathbb{F}^{m\times m}[\lambda]$,

$B(\lambda)\in\mathbb{F}^{m\times n}[\lambda]$. 若 A_q 非奇异，$q\geq 1$，则存在唯一的多项式矩阵 $Q_1(\lambda),R_1(\lambda)\in\mathbb{F}^{m\times n}[\lambda]$ 使得

$B(\lambda)=A(\lambda)Q_1(\lambda)+R_1(\lambda)$, (1.2.2)

且

$R_1(\lambda)=0$ 或者 $\deg(R_l(\lambda))<\deg(A(\lambda))$. (1.2.3)

这里，下标字母“l”表示“左边”，意思是“除数” $A(\lambda)$ 在“商” $Q_1(\lambda)$ 的左边，称为左除. 类似地，有关于右除的带余除法.

证明：存在性可由用类似多项式的长除法的算法得到. 或在 $\deg(B(\lambda))$ 有定义时，对 $\deg(B(\lambda))$ 用数学归纳法如下证明. 首先说明，若 $B(\lambda)$ 为零多项式矩阵(从而 $\deg(B(\lambda))$ 无定义)，或 $B(\lambda)$ 为非零多项式矩阵，但 $\deg(B(\lambda))<q$，则取

$Q_1(\lambda)=0, R_1(\lambda)=B(\lambda)$,

易知(1.1.2)和(1.1.3)均成立. 对任意 $p\geq 1$，现假设 $\deg(B(\lambda))<p$ 时，存在性结论已成立，下证 $\deg(B(\lambda))=p$ 的情形. 设

$B(\lambda)=B_0+B_1\lambda+\cdots+B_p\lambda^p\in\mathbb{F}^{m\times n}[\lambda]$.

记

$C(\lambda)=B(\lambda)-A(\lambda)(A_q^{-1}\lambda^{p-q})$, (1.2.4)

则易知 $C(\lambda)=0$ 或 $\deg(C(\lambda))\leq p-1$. 有归纳假设，存在多项式矩阵 $H(\lambda),L(\lambda)\in\mathbb{F}^{m\times n}[\lambda]$ 使得

$C(\lambda)=A(\lambda)H(\lambda)+L(\lambda)$, (1.2.5)

且

$\deg(L(\lambda))<\deg(A(\lambda))$. (1.2.6)

于是有(1.2.4)及(1.2.5)有

$B(\lambda)=A(\lambda)(A_q^{-1}\lambda^{p-q}+H(\lambda))+L(\lambda)$.

令

$Q_1(\lambda)=A_q^{-1}\lambda^{p-q}+H(\lambda), R_1(\lambda)=L(\lambda)$

即知结论成立.

唯一性. 设又有

$B(\lambda)=A(\lambda)\tilde{Q}_1(\lambda)+\tilde{R}_1(\lambda), \deg(\tilde{R}_1(\lambda))<\deg(A(\lambda))$.

则

$A(\lambda)(Q(\lambda)-\tilde{Q}_1(\lambda))=\tilde{R}_1(\lambda)-R(\lambda), \deg(\tilde{R}_1(\lambda)-R(\lambda))<\deg(A(\lambda))$.

于是由引理 1.2.1，$(Q(\lambda)-\tilde{Q}_1(\lambda))=0$，进而 $(\tilde{R}_1(\lambda)-R(\lambda))=0$.

定理 1.2.1 的证明：若 A,B 相似，即存在 n 阶可逆矩阵 $P\in\mathbb{F}^{n\times n}$ 使得

$P^{-1}AP=B$.

由此,

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

显然，P^{-1} 和 P 均可视为单位模阵，因此 $\lambda I-A,\lambda I-B\in\mathbb{F}^{n\times n}[\lambda]$ 作为多项式矩阵等价.

现假设 $\lambda I-A,\lambda I-B\in\mathbb{F}^{n\times n}[\lambda]$ 作为多项式矩阵等价，即存在单位模阵 $U(\lambda),V(\lambda)$ 使得

$U(\lambda)(\lambda I-A)V(\lambda)=\lambda I-B$. (1.2.7)

用 $\lambda I-B$ 左除 $U(\lambda)$，由于 $\deg(\lambda I-B)=1$，易知余式必为常数矩阵，即

$U(\lambda)=(\lambda I-B)Q(\lambda)+T$，$T\in\mathbb{F}^{n\times n}$. (1.2.8)

将此式代入(1.2.7)得

$T(\lambda I-A)=(\lambda I-B)[(V(\lambda))^{-1}-Q(\lambda)(\lambda I-A)]$. (1.2.9)

注意，由 $V(\lambda)$ 为单位模阵知 $(V(\lambda))^{-1}$ 也是多项式矩阵. 由(1.2.8)，$U(\lambda)$ 为单位模阵，及定理 1.1.1 知 T 非零. 于是由引理 1.2.1 有 $(V(\lambda))^{-1}-Q(\lambda)(\lambda I-A)$ 也非零，且

$\deg(T)+\deg(\lambda I-A)=\deg(\lambda I-B)+\deg((V(\lambda))^{-1}-Q(\lambda)(\lambda I-A))$.

于是 $\deg((V(\lambda))^{-1}-Q(\lambda)(\lambda I-A))=0$，即 $(V(\lambda))^{-1}-Q(\lambda)(\lambda I-A)$ 也为非零常数矩阵，记为

$(V(\lambda))^{-1}-Q(\lambda)(\lambda I-A)=S$.

于是, (1.2.9)式写为

$T(\lambda I-A)=(\lambda I-B)S$， (1.2.10)

此即(比较系数)

$T=S$

$TA=BS$.

为说明 A,B 相似，只需进一步证明 T 可逆. 由 $U(\lambda)$ 为单位模阵，$(U(\lambda))^{-1}$ 也为多项式矩阵. 用 $\lambda I-A$ 左除 $(U(\lambda))^{-1}$，由于 $\deg(\lambda I-A)=1$，易知余式必为常数矩阵，即

$(U(\lambda))^{-1}=(\lambda I-A)G(\lambda)+K$，$K\in\mathbb{F}^{n\times n}$.

$U(\lambda)=(\lambda I-B)Q(\lambda)+T$

于是由(1.2.8), (1.2.10)及上式，有

$$
\begin{aligned}
I&=U(\lambda)(U(\lambda))^{-1}\\
&=(\lambda I-B)Q(\lambda)(U(\lambda))^{-1}+T(\lambda I-A)G(\lambda)+TK\\
&=(\lambda I-B)Q(\lambda)(U(\lambda))^{-1}+(\lambda I-B)SG(\lambda)+TK\\
&=(\lambda I-B)[Q(\lambda)(U(\lambda))^{-1}+SG(\lambda)]+TK.
\end{aligned}
$$

于是

$(\lambda I-B)[Q(\lambda)(U(\lambda))^{-1}+SG(\lambda)]=I-TK$. (1.2.11)

由此及引理 1.2.1，必有 $Q(\lambda)(U(\lambda))^{-1}+SG(\lambda)=0$. 从而 $I-TK=0$，T 可逆. 证毕

定理 1.2.2 (特征矩阵的 Smith 标准型)

(1) $\lambda I-A$ 作为多项式矩阵，其 n 阶行列式因子，也就是特征多项式 $|\lambda I-A|$，为 n 次多项式. 因而 $\lambda I-A$ 作为多项式矩阵的秩为 n.

(2) 设 $\lambda I-A$ 的 Smith 标准型为

$$
\begin{bmatrix} d_1(\lambda) & & \\ & \ddots & \\ & & d_n(\lambda) \end{bmatrix}. \quad (1.2.12)
$$

则

$d_1(\lambda)\cdots d_n(\lambda)=|\lambda I-A|$，$\partial(d_1(\lambda))+\cdots+\partial(d_n(\lambda))=n$.

证明：由

$|\lambda I-A|=\lambda^n+(-\operatorname{trace}(A))\lambda^{n-1}+\cdots+(-1)^n\det(A)$

对任意 A 均为 n 次多项式可得结论. 证毕

推论 1.2.1 (特征矩阵的不变因子的次数规律)
考虑特征矩阵 $\lambda I - A$ 的 Smith 型(1.2.12). 设 $d_i(\lambda)$, $i=1,2,\ldots,n$, 中的非常数者(即次数 ≥ 1 的)记为
$h_1(\lambda), h_2(\lambda), \ldots, h_p(\lambda)$,
次数分别为
$n_1, n_2, \ldots, n_p$.
则 $d_i(\lambda)$, $i=1,2,\ldots,n$, 中恰有 $n-p=(n_1-1)+(n_2-1)+\cdots+(n_p-1)$ 个为 1.

由此推论, 通过一系列的行交换及列交换, 可将特征矩阵 $\lambda I - A$ 的 Smith 标准型, 化为一种特殊的块对角形式. 每个子块相应于一个非常数不变因子, 并配以若干个常数不变因子, 使子块的阶数恰为该不变因子的次数. 即 Smith 型(1.2.12)等价于

$$\begin{bmatrix} \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & h_1(\lambda) \end{bmatrix}_{n_1\times n_1} & & & \\ & \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & h_2(\lambda) \end{bmatrix}_{n_2\times n_2} & & \\ & & \ddots & \\ & & & \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & h_p(\lambda) \end{bmatrix}_{n_p\times n_p} \end{bmatrix}. \quad (1.2.13)$$

为说话方便, 我们称(1.2.13)为特征矩阵 $\lambda I - A$ 的第二等价规范型.

下面进一步通过不变因子的质因式分解, 将第二规范型化成更特殊的结构.

定义 1.2.3 (特征矩阵的初等因子组)
设 $A \in \mathbb{F}^{n\times n}$. 特征矩阵 $\lambda I - A$ 的(亦有时简称为 A 的)初等因子是指将 $\lambda I - A$ 的所有不变因子在 $\mathbb{F}[\lambda]$ 中作质因式分解时出现的质因式的方幂. 同一质因式的方幂若出现多次, 则算作是多个初等因子. 所有初等因子的全体称为初等因子组.

设 $\lambda I - A$ 的 Smith 标准型如(1.2.1). 记最后一个不变因子的质因式分解为
$d_n(\lambda) = (e_1(\lambda))^{r_{n_1}} \cdots (e_p(\lambda))^{r_{n_p}}$,
其中 $e_k(\lambda)$ 为不可约多项式, $r_{n_k} \geq 1$, $k=1,\ldots,p$.
则必有
$d_i(\lambda) = (e_1(\lambda))^{r_{i_1}} \cdots (e_p(\lambda))^{r_{i_p}}$, $i=1,\ldots,n-1$.
且
$0 \leq r_{i_k} \leq r_{i+1_k}$, $i=1,\ldots,n-1$, $k=1,\ldots,p$.
则初等因子组为
$\{(e_k(\lambda))^{r_{i_k}} : r_{i_k} \geq 1, i=1,\ldots,n-1; k=1,\ldots,p\}$,
其中, 多次出现的重复计入.

根据上述规则, 初等因子组也可完全决定不变因子组和 Smith 型. 下面通过例子来说明如何决定.

例 1.2.1 (特征矩阵的不变因子组和初等因子组相互决定)
(1) 设 $A \in \mathbb{C}^{5\times 5}$, $\lambda I - A$ 的 Smith 型为

$$\begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & (\lambda+1)^2 & \\ & & & & \lambda(\lambda+1)^2 \end{bmatrix}.$$

则初等因子组为

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

(2) 已知复矩阵 A 的初等因子组为

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

求它的不变不变因子组(Smith 型).

首先可以知道，A 必是 8 阶的.

将所有初等因子按质因式分组；每组至多 8 个，将不够 8 个的用 1 补够，并按升幂排列:

$\{1,1,1,1,1,\lambda,\lambda,\lambda^2\}$；$\{1,1,1,1,1,1,(\lambda+1)^2,(\lambda+1)^2\}$；

依次从每组中取元，构成不变因子如下:

$d_5(\lambda)=\cdots=d_1(\lambda)=1$

$d_6(\lambda)=\lambda\cdot 1=\lambda$

$d_7(\lambda)=\lambda\cdot(\lambda+1)^2=\lambda(\lambda+1)^2$

$d_8(\lambda)=\lambda^2\cdot(\lambda+1)^2=\lambda^2(\lambda+1)^2.$

下面基于初等因子将第二规范型(1.2.13)中每一个子块进一步分解.

引理 1.2.3 (不变因子的互质分解引起的分解)

设 $f_1(\lambda), f_2(\lambda)\in\mathbb{F}[\lambda]$ 互质，整数 $n_1+n_2=n$. 则作为 $\mathbb{F}$ 上的多项式矩阵，下列两矩阵等价

$$\begin{bmatrix} I_{n+1} & 0 \\ 0 & f_1(\lambda)f_2(\lambda) \end{bmatrix}, \quad (1.2.14)$$

$$\left[\begin{array}{cc:cc} I_{n_1} & 0 & \multicolumn{2}{c}{\multirow{2}{*}{0}} \\ 0 & f_1(\lambda) & & \\ \hdashline \multicolumn{2}{c:}{\multirow{2}{*}{0}} & I_{n_2} & 0 \\ & & 0 & f_2(\lambda) \end{array}\right]. \quad (1.2.15)$$

证明: 只需证明它们的各阶行列式因子相同. 首先，它们的 $n+2$ 阶行列式因子，也就是它们的行列式，均为 $f_1(\lambda)f_2(\lambda)$. 易知，多项式矩阵(1.2.14)的 $n+1$ 阶行列式因子为 1. 下面看(1.2.15)的 n 阶行列式因子. 其所有 $n+1$ 阶非零子行列式中，有一个为 $f_1(\lambda)$，一个为 $f_2(\lambda)$，n 个为 $f_1(\lambda)f_2(\lambda)$，由 $f_1(\lambda)$ 与 $f_2(\lambda)$ 互质，易见这些 n 阶非零子行列式的最高公因式为 1. 因此, (1.2.14)和(1.2.15)的 $n+1$ 阶行列式因子相同，都为 1. 显然，其他各阶的行列式因子也都相同，都为 1. 证毕

反复运用此引理，可将特征矩阵的第二规范型(1.2.13)中每一个子块进一步分解为若干个子子块，每一个子子块相应于一个初等因子.

定理 1.2.3 (基于初等因子组的规范型)

设特征矩阵 $\lambda I-A$ 的初等因子组为

$(e_k(\lambda))^{r_k}: k=1,\ldots,q$,

次数分别为

$m_1=r_1\cdot\deg(e_1(\lambda)), m_2=r_2\cdot\deg(e_2(\lambda)),\ldots, m_q=r_q\cdot\deg(e_q(\lambda))$.

则 $\lambda I-A$ 等价于

$$
\left[\begin{array}{c:c:c:c}
\begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & (e_1(\lambda))^{r_1} \end{bmatrix}_{m_1\times m_1} & & & \\ \hdashline
 & \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & (e_2(\lambda))^{r_2} \end{bmatrix}_{m_2\times m_2} & & \\ \hdashline
 & & \ddots & \\ \hdashline
 & & & \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & (e_q(\lambda))^{r_q} \end{bmatrix}_{m_q\times m_q}
\end{array}\right]. \quad (1.2.16)
$$

称(1.2.16)为特征矩阵 $\lambda I - A$ 的第三等价规范型.

下面我们总结矩阵相似的条件

定理 1.2.4 (矩阵相似的各种刻画)

给定两个矩阵 $A, B \in \mathbb{F}^{n\times n}$. 下列条件等价:

(1) A 与 B 相似.

(2) $\lambda I - A$ 与 $\lambda I - B$ 作为多项式矩阵等价.

(3) $\lambda I - A$ 与 $\lambda I - B$ 作为多项式矩阵有相同的 Smith 标准型.

(4) $\lambda I - A$ 与 $\lambda I - B$ 作为多项式矩阵有相同的第二规范型.

(5) $\lambda I - A$ 与 $\lambda I - B$ 作为多项式矩阵有相同的第三规范型.

(6) $\lambda I - A$ 与 $\lambda I - B$ 作为多项式矩阵有相同的各阶行列式因子.

(7) $\lambda I - A$ 与 $\lambda I - B$ 作为多项式矩阵有相同的各阶不变因子.

(8) $\lambda I - A$ 与 $\lambda I - B$ 作为多项式矩阵有相同的初等因子组.

例 1.2.2 (任一矩阵与其转置相似)

设矩阵 $A \in \mathbb{F}^{n\times n}$. 则 A 的转置 A^{T} 与 A 相似.

要直接利用矩阵相似定义来证明此结论, 并不容易. 下面利用定理 1.2.4 来证明. 设 $\lambda I - A$ 的 Smith 标准型为

$$
U(\lambda)(\lambda I - A)V(\lambda) = \begin{bmatrix} d_1(\lambda) & & & \\ & d_2(\lambda) & & \\ & & \ddots & \\ & & & d_n(\lambda) \end{bmatrix},
$$

这里 $U(\lambda), V(\lambda)$ 为单位模阵. 于是

$$
(V(\lambda))^{\mathrm{T}}(\lambda I - A^{\mathrm{T}})(U(\lambda))^{\mathrm{T}} = \begin{bmatrix} d_1(\lambda) & & & \\ & d_2(\lambda) & & \\ & & \ddots & \\ & & & d_n(\lambda) \end{bmatrix}.
$$

由于单位模阵的转置仍为单位模阵, 因此上两式说明 $\lambda I - A$ 与 $\lambda I - A^{\mathrm{T}}$ 有相同的 Smith 标准型.

1.3 有理标准型和 Jordan 标准型

问题 1.3.1 (构造其特征矩阵具有给定的单个非常数不变因子或给定的单个初等因子的简单矩阵)

(1) 设 $d(\lambda) \in \mathbb{F}[\lambda]$ 为首项系数为 1 的多项式, $\deg(d(\lambda)) = n > 0$. 求"尽可能简单的"矩阵 $D \in \mathbb{F}^{n\times n}$, 使得

$\lambda I-D$ 只有一个非常数不变因子 $d(\lambda)$，即 $\lambda I-D$ 等价于

$$\begin{bmatrix}1&&&\\&\ddots&&\\&&1&\\&&&d(\lambda)\end{bmatrix}_{n\times n}$$

(2) 设 $e(\lambda)\in\mathbb{F}[\lambda]$ 为首项系数为 1 的不可约多项式，r 为正整数，$r\cdot\deg(e(\lambda))=n>0$. 求“尽可能简单的”矩阵 $J\in\mathbb{F}^{n\times n}$，使得 $\lambda I-J$ 只有一个初等因子 $(e(\lambda))^r$ (同时也是唯一的非常数不变因子)，即 $\lambda I-J$ 等价于

$$\begin{bmatrix}1&&&\\&\ddots&&\\&&1&\\&&&(e(\lambda))^r\end{bmatrix}_{n\times n}$$

定理 1.3.1 (其特征矩阵以给定多项式为单个非常数不变因子的矩阵)

给定 $\mathbb{F}$ 上的多项式

$$d(\lambda)=a_0+a_1\lambda+\cdots+a_{n-1}\lambda^{n-1}+\lambda^n.$$

由它决定的 $\mathbb{F}$ 上的 n 阶矩阵

$$D_{d(\lambda)}=\left[\begin{array}{c|ccc}0&1&&\\\vdots&&\ddots&\\0&&&1\\\hline -a_0&-a_1&\cdots&-a_{n-1}\end{array}\right]$$

称为 $d(\lambda)$ 的友矩阵，它满足：$\lambda I-D_{d(\lambda)}$ 等价于(Smith 标准型是)

$$\begin{bmatrix}1&&&\\&\ddots&&\\&&1&\\&&&d(\lambda)\end{bmatrix}. \quad (1.3.1)$$

证明：计算 $\lambda I-D_{d(\lambda)}$ 的行列式因子. 由

$$\lambda I-D_{d(\lambda)}=\left[\begin{array}{c|cccc}\lambda&-1&&&\\0&\lambda&-1&&\\\vdots&&\ddots&\ddots&\\0&&&\lambda&-1\\\hline a_0&a_1&&a_{n-2}&\lambda+a_{n-1}\end{array}\right]$$

知它有一个 $n-1$ 阶子行列式为 $(-1)^{n-1}$，故 $n-1$ 阶行列式因子为 1，从而所有低于 $n-1$ 阶的行列式因子也为 1. 利用行列式性质容易求出 $|\lambda I-D_{d(\lambda)}|=d(\lambda)$. 因此 $\lambda I-D_{d(\lambda)}$ 的 n 阶的行列式因子为 $d(\lambda)$，从而 $\lambda I-D_{d(\lambda)}$ 与(1.3.1)的各阶行列式因子都相同. 证毕

注 1.3.1 (首一多项式的友矩阵的两种常见形式)

$d(\lambda)=a_0+a_1\lambda+\cdots+a_{n-1}\lambda^{n-1}+\lambda^n$ 的友矩阵也常写成如下形式

$$\left[\begin{array}{ccc|c}0&\cdots&0&-a_0\\\hline 1&&&-a_1\\&\ddots&&\vdots\\&&1&-a_{n-1}\end{array}\right],$$

即上述 $D_{d(\lambda)}$ 的转置 $(D_{d(\lambda)})^{\mathrm{T}}$ (列体制下应该用此作为友矩阵定义). 类似定理 1.3.1 的证明，或由例 1.2.2，易知 $(D_{d(\lambda)})^{\mathrm{T}}$ 的特征矩阵也以 $d(\lambda)$ 为单个非常数不变因子. 由例 1.2.2 知 $D_{d(\lambda)}$ 与 $(D_{d(\lambda)})^{\mathrm{T}}$ 相似，下面我们具体写出从 $(D_{d(\lambda)})^{\mathrm{T}}$ 到 $D_{d(\lambda)}$ 的相似变换所用的可逆矩阵.

可以验证: 令

$$P=\begin{bmatrix}1 & & & & a_1\\ & \ddots & & & \vdots\\ & & 1 & & a_{n-2}\\ & & & 1 & a_{n-1}\\ & & & & 1\end{bmatrix}\left[\begin{array}{cccc|c}1 & & & a_2 & \\ & \ddots & & \vdots & \\ & & 1 & a_{n-1} & \\ & & & 1 & \\ \hline & & & & 1\end{array}\right]\cdots\left[\begin{array}{cc|ccc}1 & a_{n-1} & & & \\ & 1 & & & \\ \hline & & 1 & & \\ & & & \ddots & \\ & & & & 1\end{array}\right]=\begin{bmatrix}1 & a_{n-1} & a_{n-2} & \cdots & a_1\\ & 1 & a_{n-1} & \ddots & \vdots\\ & & 1 & \ddots & a_{n-2}\\ & & & \ddots & a_{n-1}\\ & & & & 1\end{bmatrix},$$

则

$$P^{-1}(D_{d(\lambda)})^{\mathrm{T}}P=\left[\begin{array}{ccc|c}-a_{n-1} & \cdots & -a_1 & -a_0\\ \hline 1 & & & 0\\ & \ddots & & \vdots\\ & & 1 & 0\end{array}\right]$$

再令

$$Q=\begin{cases}(\mathrm{c}_1\leftrightarrow \mathrm{c}_{2k-0})(\mathrm{c}_2\leftrightarrow \mathrm{c}_{2k-1})\cdots(\mathrm{c}_k\leftrightarrow \mathrm{c}_{2k-(k-1)}), & n=2k\text{为偶数;}\\ (\mathrm{c}_1\leftrightarrow \mathrm{c}_{(2k+1)-0})(\mathrm{c}_2\leftrightarrow \mathrm{c}_{(2k+1)-1})\cdots(\mathrm{c}_k\leftrightarrow \mathrm{c}_{(2k+1)-(k-1)}), & n=2k+1\text{为奇数.}\end{cases}$$

这里, 符号 $(\mathrm{c}_i\leftrightarrow \mathrm{c}_j)$ 表示第 i 列与第 j 列互换的初等变换所对应的初等矩阵. (变换 $Q^{-1}BQ$ 的效果相当于对矩阵 B 行和列关于中心线各作一次镜像变换, 此变换亦可给出下面要讲的 Jordan 块和其转置之间的相似变换. 为看出此点, 画两次镜像反射的效果图.) 于是

$$(PQ)^{-1}(D_{d(\lambda)})^{\mathrm{T}}PQ=D_{d(\lambda)}.$$

方形矩阵可按明显的方式规定其代入多项式的含义. 给定 $\mathbb{F}$ 上的 n 阶矩阵 A, 及 $\mathbb{F}$ 上的多项式

$$f(\lambda)=a_0+a_1\lambda+\cdots+a_{t-1}\lambda^{t-1}+a_t\lambda^t\in\mathbb{F}[\lambda],$$

规定

$$f(A)=a_0\cdot I+a_1\cdot A+\cdots+a_{t-1}\cdot A^{t-1}+a_t\cdot A^t\in\mathbb{F}^{n\times n}.$$

定义 1.3.1 (矩阵的最低多项式)

给定矩阵 $A\in\mathbb{F}^{n\times n}$, 及首一多项式 $h(\lambda)=a_0+a_1\lambda+\cdots+a_{t-1}\lambda^{t-1}+\lambda^t\in\mathbb{F}[\lambda]$. 称 $h(\lambda)$ 是 A 的最低多项式, 如果它是使得矩阵 A 代入后为零的次数最低的多项式, 即

(1) $h(A)=0$,

(2) 若有非零多项式 $g(\lambda)\in\mathbb{F}[\lambda]$ 满足 $g(A)=0$, 则必有 $\deg(g(\lambda))\geq\deg(h(\lambda))$.

定理 1.3.2 (首一多项式是它的友矩阵的最低多项式)

给定 $\mathbb{F}$ 上的多项式 $d(\lambda)=a_0+a_1\lambda+\cdots+a_{n-1}\lambda^{n-1}+\lambda^n$ 及其友矩阵 $D_{d(\lambda)}$. 则 $d(\lambda)$ 是 $D_{d(\lambda)}$ 的最低多项式.

证明: (1) $d(D_{d(\lambda)})=0$. 将单位矩阵分成行向量组

$$\boldsymbol{I}=\begin{bmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\\ \vdots\\ \boldsymbol{\alpha}_n\end{bmatrix},$$

可以验证 $D_{d(\lambda)}$ 的如下性质:

$$\begin{aligned}\alpha_1D_{d(\lambda)}&=\alpha_2\\ \alpha_2D_{d(\lambda)}&=\alpha_3\\ &\vdots\\ \alpha_{n-1}D_{d(\lambda)}&=\alpha_n\\ \alpha_nD_{d(\lambda)}&=-a_0\alpha_1-a_1\alpha_2-\cdots-a_{n-2}\alpha_{n-1}-a_{n-1}\alpha_n.\end{aligned}$$

于是

$$\alpha_1(D_{d(\lambda)})^{n-1}D_{d(\lambda)} = -a_0\alpha_1 - a_1\alpha_1 D_{d(\lambda)} - \cdots - a_{n-2}\alpha_1(D_{d(\lambda)})^{n-2} - a_{n-1}\alpha_1(D_{d(\lambda)})^{n-1},$$

即

$$\begin{aligned} 0 &= \alpha_1\{a_0 \cdot I + a_1 \cdot D_{d(\lambda)} + \cdots + a_{n-2}\cdot(D_{d(\lambda)})^{n-2} + a_{n-1}\cdot(D_{d(\lambda)})^{n-1} + (D_{d(\lambda)})^n\} \\ &= \alpha_1 d(D_{d(\lambda)}). \end{aligned}$$

进而

$$\begin{aligned} &\alpha_2 d(D_{d(\lambda)}) = \alpha_1 D_{d(\lambda)} d(D_{d(\lambda)}) \\ &= \{\alpha_1 d(D_{d(\lambda)})\}D_{d(\lambda)} \\ &= 0D_{d(\lambda)} \\ &= 0, \end{aligned}$$

类似地,

$$\alpha_3 d(D_{d(\lambda)}) = \cdots = \alpha_n d(D_{d(\lambda)}) = 0.$$

从而,

$$d(D_{d(\lambda)}) = Id(D_{d(\lambda)}) = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} d(D_{d(\lambda)}) = \begin{bmatrix} \alpha_1 d(D_{d(\lambda)}) \\ \alpha_2 d(D_{d(\lambda)}) \\ \vdots \\ \alpha_n d(D_{d(\lambda)}) \end{bmatrix} = 0.$$

(2) 若有非零多项式 $g(\lambda) \in \mathbb{F}[\lambda]$ 满足 $\deg(g(\lambda)) = m < \deg(d(\lambda)) = n$，则必有 $g(D_{d(\lambda)}) \neq 0$. 令

$$g(\lambda) = b_0 + b_1\lambda + \cdots + b_m\lambda^m, \quad b_m \neq 0.$$

则

$$\begin{aligned} \alpha_1 g(D_{d(\lambda)}) &= b_0\alpha_1 + b_1\alpha_1 D_{d(\lambda)} + \cdots + b_m\alpha_1(D_{d(\lambda)})^m \\ &= b_0\alpha_1 + b_1\alpha_2 + \cdots + b_m\alpha_{m+1}. \end{aligned}$$

由向量组 $\alpha_1, \alpha_2, \ldots, \alpha_{m+1}$ 线性无关($m+1 \le n$)及 $b_m \neq 0$ 知

$$\alpha_1 g(D_{d(\lambda)}) = b_0\alpha_1 + b_1\alpha_2 + \cdots + b_m\alpha_{m+1} \neq 0.$$

所以有 $g(D_{d(\lambda)}) \neq 0$. 证毕

定理 1.3.3 (任意域上矩阵的有理标准型)

任给域 $\mathbb{F}$，及 $\mathbb{F}$ 上的矩阵 $A \in \mathbb{F}^{n\times n}$. 设 $\lambda I - A \in \mathbb{F}^{n\times n}[\lambda]$ 的非常数不变因子为 $h_1(\lambda), h_2(\lambda), \ldots, h_p(\lambda)$. 则 A 相似于下面的“有理标准型”

$$D = \operatorname{diag}\{D_{h_1(\lambda)}, D_{h_2(\lambda)}, \ldots, D_{h_p(\lambda)}\} = \left[\begin{array}{c:c:c:c} D_{h_1(\lambda)} & & & \\ \hdashline & D_{h_2(\lambda)} & & \\ \hdashline & & \ddots & \\ \hdashline & & & D_{h_p(\lambda)} \end{array}\right]. \quad (1.3.2)$$

证明: 由定理 1.3.1 易知，$\lambda I - D$ 与 $\lambda I - A$ 有相同的第二规范型. 再由定理 1.2.4 (4)即得结论. 证毕

注 1.3.2 (何以谓之“有理”?)

“有理”的含义是指: 矩阵 D 的所有元素，都可由 A 的元素经加减乘除四则运算得到. 为看出此点，只需说明 $\lambda I - A$ 的每个不变因子作为多项式，其系数可由 A 的元素经加减乘除四则运算得到. $\lambda I - A$ 的每个行列子式作为多项式，其系数显然可由 A 的元素经加减乘三种运算得到; 而根据辗转相除法，多项式们的公因式的系数可由多项式们的系数经四则运算得到. 因而，$\lambda I - A$ 的每个行列式因子作为多项式，其系数终究可由 A 的元素经加减乘除四则运算得到. 另一方面，商多项式的系数总是可由被除式多项式的系数和除式多项式的系数经四则运算得到. 因此，不变因子的系数总可由 A 的元素经加减乘除四则运算得到.

定理 1.3.4 (任意域上矩阵的特征矩阵的最后一个不变因子是该矩阵的最低多项式)

在定理 1.3.3 中，设 $h_p(\lambda)$ 为最后一个不变因子. 则 $h_p(\lambda)$ 是 A 的最小多项式.

证明: 易验证，相似的两个矩阵有相同的最小多项式. 因此，只需证明 $h_p(\lambda)$ 是 D 的最小多项式.

(1) $h_p(D)=0$. 首先

$h_p(D)=\operatorname{diag}\{h_p(D_{h_1(\lambda)}),h_p(D_{h_2(\lambda)}),\ldots,h_p(D_{h_p(\lambda)})\}$.

又由 $h_i(\lambda)\mid h_p(\lambda)$, 有多项式 $\tilde{h}_i(\lambda)$ 使得 $\tilde{h}_i(\lambda)h_i(\lambda)=h_p(\lambda)$, $i=1,\ldots,p-1$. 再由定理 1.3.2,

$h_p(D_{h_i(\lambda)})=\tilde{h}_i(D_{h_i(\lambda)})h_i(D_{h_i(\lambda)})=\tilde{h}_i(D_{h_i(\lambda)})0=0$.

(2) 若有非零多项式 $g(\lambda)$ 满足 $\deg(g(\lambda))<\deg(h_p(\lambda))$, 则 $g(D)\neq 0$. 首先由定理 1.3.2, $g(D_{h_p(\lambda)})\neq 0$. 于是 $g(D)=\operatorname{diag}\{g(D_{h_1(\lambda)}),g(D_{h_2(\lambda)}),\ldots,g(D_{h_p(\lambda)})\}\neq 0$. 证毕

注 1.3.3 (化有理标准型所用的相似变换的变换矩阵的求法)
化化有理标准型所用的相似变换的变换矩阵的存在性已由定理 1.3.2 保证. 变换矩阵的求法原则上可由定理 1.2.1 的证明过程给出如下(记号见定理 1.3.2)
(1) 求出单位模阵 $U_1(\lambda),V_1(\lambda)$ 使得

$$U_1(\lambda)(\lambda I-A)V_1(\lambda)=\begin{bmatrix} I_{n-p} & & & \\ & h_1(\lambda) & & \\ & & \ddots & \\ & & & h_p(\lambda)\end{bmatrix}.$$

(2) 求出单位模阵 $U_2(\lambda),V_2(\lambda)$ 使得

$$U_2(\lambda)(\lambda I-D)V_2(\lambda)=\begin{bmatrix} I_{n-p} & & & \\ & h_1(\lambda) & & \\ & & \ddots & \\ & & & h_p(\lambda)\end{bmatrix}.$$

(3) 由

$((U_2(\lambda))^{-1}U_1(\lambda))(\lambda I-A)(V_1(\lambda)(V_2(\lambda))^{-1})=(\lambda I-D)$,

计算带余除法

$V_1(\lambda)(V_2(\lambda))^{-1}=Q(\lambda)(\lambda I-D)+P$,

则有 P 可逆, 且

$P^{-1}AP=D$.

定理 1.3.5 (以一次多项式的方幂为单个初等因子的矩阵—Jordan 块)
给定 $\mu\in\mathbb{F}$. 矩阵

$$J_n(\mu)=\begin{bmatrix} \mu & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \mu\end{bmatrix}_{n\times n}$$

称为特征值为 μ 的 n 阶 Jordan 块. 这里, 未标出的位置上的元素均为 0. 它满足 $\lambda I-J_n(\mu)$ 等价于(Smith 标准型是)

$$\begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & (\lambda-\mu)^n\end{bmatrix}. \quad (1.3.3)$$

证明: 计算 $\lambda I-J_n(\mu)$ 的行列式因子. 由

$$\lambda I-J_n(\mu)=\begin{bmatrix}\lambda-\mu & -1 & & \\ & \ddots & \ddots & \\ & & \ddots & -1 \\ & & & \lambda-\mu\end{bmatrix}$$

知它有一个$n-1$阶子行列式为$(-1)^{n-1}$，故$n-1$阶行列式因子为 1，从而所有低于$n-1$阶的行列式因子也为 1. 显然，$|\lambda I-J_n(\mu)|=(\lambda-\mu)^n$. 因此$\lambda I-J_n(\mu)$的$n$阶的行列式因子为$(\lambda-\mu)^n$，从而$\lambda I-J_n(\mu)$与(1.3.3)的各阶行列式因子都相同. 证毕

下面研究复数域上的矩阵的相似标准型. 复数域的特点是复系数多项式总有复数根，换言之，复数域上的不可约多项式都是一次的. 特别地，由定理 1.2.2 及定义 1.2.3，对复数域上的矩阵A，其特征矩阵$\lambda I-A$的初等因子组作为多项式的全部根恰为矩阵A的全部特征值. 换言之，在下面定理 1.3.4，有

$(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_q)^{m_q}=|\lambda I-A|$, (1.3.4)

进而，

$m_1+m_2+\cdots+m_q=n$. (1.3.5)

定理 1.3.6 (复数域上矩阵的 Jordan 标准型)

任给复数域$\mathbb{C}$上的矩阵$A\in\mathbb{C}^{n\times n}$. 设$\lambda I-A\in\mathbb{C}^{n\times n}[\lambda]$的初等因子组为$(\lambda-\lambda_1)^{m_1},(\lambda-\lambda_2)^{m_2},\ldots,(\lambda-\lambda_q)^{m_q}$. 则$A$相似于下面的 Jordan 标准型

$$J=\operatorname{diag}\{J_{m_1}(\lambda_1),J_{m_2}(\lambda_2),\ldots,J_{m_q}(\lambda_q)\}=\left[\begin{array}{c|c|c|c} J_{m_1}(\lambda_1) & & & \\ \hline & J_{m_2}(\lambda_2) & & \\ \hline & & \ddots & \\ \hline & & & J_{m_q}(\lambda_q)\end{array}\right]. \quad (1.3.6)$$

证明：由定理 1.3.3 易知，$\lambda I-J$ 与 $\lambda I-A$有相同的第三规范型. 再由定理 1.2.4 (5)即得结论. 证毕

1.4 复数域上矩阵的特征结构

本节的数域都是复数域$\mathbb{C}$. 由 1.3 节定理 1.3.4，我们知道，对任意n阶矩阵A，存在可逆矩阵T，使得A相似于其 Jordan 标准型J，即

$T^{-1}AT=J$. (1.4.1)

我们要研究可逆矩阵T 的几何意义. 由第一??章第??节，我们知道，将A视为$\mathbb{C}^n$上的线性变换:

$A: x\mapsto Ax$,

则当入口基和出口基都取为可逆矩阵T 的列向量组时，线性变换A在该入口基和出口基下的表示即为J. 进一步，我们研究与 Jordan 标准型J 相应的基矩阵T 的特殊结构.

设J 如(1.3.4)所示. 考虑矩阵T 相应的列分块

$$T=\left[\begin{array}{c|c|c|c}\underbrace{t_1\ \cdots\ t_{m_1}}_{T_1} & \underbrace{t_{m_1+1}\ \cdots\ t_{m_1+m_2}}_{T_2} & \cdots & \underbrace{t_{m_1+\cdots+m_{q-1}+1}\ \cdots\ t_n}_{T_q}\end{array}\right].$$

则(1.4.1)成为

$$A\begin{bmatrix}T_1 & T_2 & \cdots & T_q\end{bmatrix}=\begin{bmatrix}T_1 & T_2 & \cdots & T_q\end{bmatrix}\begin{bmatrix}J_{m_1}(\lambda_1) & & & \\ & J_{m_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{m_q}(\lambda_q)\end{bmatrix}.$$

由于J 为块对角矩阵，由第一??章的定理??，每个子空间$\operatorname{im}T_j$都是A的不变子空间，且

$\operatorname{im}T_1\oplus\operatorname{im}T_2\oplus\cdots\oplus\operatorname{im}T_q=\mathbb{C}^n$.

只需研究一个不变子空间，例如$\operatorname{im}T_j$，上的基矩阵T_j的情况. 观察

$$AT_j = \begin{bmatrix} T_1 & \cdots & T_{j-1} & T_j & T_{j+1} & \cdots & T_q \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ J_{m_j}(\lambda_j) \\ 0 \\ \vdots \\ 0 \end{bmatrix} = T_j J_{m_j}(\lambda_j). \quad (1.4.2)$$

注意到

$$J_{m_j}(\lambda_j) = \begin{bmatrix} \lambda_j & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_j \end{bmatrix}_{m_j \times m_j}$$

$$= \lambda_j I_{m_j} + \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix},$$

(1.4.2)可写为

$$AT_j = T_j J_{m_j}(\lambda_j)$$

$$= T_j \left(\lambda_j I_{m_j} + \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix} \right)$$

$$= (\lambda_j I_n) T_j + T_j \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix},$$

即

$$(A - \lambda_j I) T_j = T_j \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}.$$

再观察 T_j 的列向量组. 为简化符号, 不妨取 $j = 1$.

$$(A - \lambda_1 I) \begin{bmatrix} t_1 & t_2 & \cdots & t_{m_1 - 1} & t_{m_1} \end{bmatrix} = \begin{bmatrix} t_1 & t_2 & \cdots & t_{m_1 - 1} & t_{m_1} \end{bmatrix} \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}.$$

按列分块, 得

$$\begin{aligned} (A - \lambda_1 I) t_1 &= 0 \\ (A - \lambda_1 I) t_2 &= t_1 \\ &\vdots \\ (A - \lambda_1 I) t_{m_1 - 1} &= t_{m_1 - 2} \\ (A - \lambda_1 I) t_{m_1} &= t_{m_1 - 1}. \end{aligned}$$

因此，每个 t_i，$1 \le i \le m_1$，满足

$$(A-\lambda_1 I)^i t_i = 0,$$

$$(A-\lambda_1 I)^{i-1} t_i = t_1 \ne 0.$$

定义 1.4.1 (广义特征向量)

我们称满足条件

$$(A-\mu I)^k p = 0,$$

$$(A-\mu I)^{k-1} p \ne 0$$

的向量 p 是矩阵 A 的相应于特征值 μ 的指标为 k 的广义特征向量. 指标为 1 的广义特征向量(按规定，$(A-\mu I)^0 = I$)就是通常的特征向量.

于是，t_i 恰为矩阵 A 的相应于特征值 λ_1 的指标为 i 的广义特征向量，$1 \le i \le m_1$，而这些广义特征向量之间的关系如下图

$$0 \xleftarrow{(A-\lambda_1 I)} t_1 \xleftarrow{(A-\lambda_1 I)} t_2 \xleftarrow{(A-\lambda_1 I)} \cdots\cdots\cdots \xleftarrow{(A-\lambda_1 I)} t_{m_1-1} \xleftarrow{(A-\lambda_1 I)} t_{m_1}.$$

据此关系图，我们形象地称向量组

$$t_1, t_2, \ldots, t_{m_1}$$

是矩阵 A 的相应于特征值 λ_1 的一个长度为 m_1 的广义特征向量链. 一般地，有定义

定义 1.4.2 (广义特征向量链)

向量组

$$p_1, p_2, \ldots, p_m$$

称为矩阵 A 的相应于特征值 μ 的一个长度为 m 的广义特征向量链，如果

$$(A-\mu I) p_i = p_{i-1}, i = 2, \ldots, m,$$

$$(A-\mu I) p_1 = 0, \ p_1 \ne 0.$$

因此，相应于每一个 Jordan 块的不变子空间的基向量组，是一个广义特征向量链.

(利用 Zariski 拓扑说明找广义特征向量链的难度)

下面引入几个常用的术语.

定义 1.4.3 (特征值的代数重数与几何重数)

设 $\mu \in \mathbb{C}$ 是 n 阶复矩阵 A 的一个特征值，即 $|\mu I - A| = 0$. 则

(1) A 的特征值 μ 的代数重数是指因式 $\lambda - \mu$ 在特征多项式 $|\lambda I - A|$ 的质因式分解中出现的次数，也就是 μ 作为多项式 $|\lambda I - A|$ 的根的重数.

(2) A 的特征值 μ 的几何重数

$$\begin{aligned}
&= \text{以}(\lambda-\mu)\text{为质因式的}\lambda I - A\text{的初等因子的个数} \\
&= A\text{的 Jordan 标准型中对角线上为}\mu\text{的 Jordan 块的个数} \\
&= \mu\text{对应的以线性无关的特征向量的个数} \\
&= \text{齐次线性方程组}(\mu I - A)x = 0\text{解空间的维数} \\
&= \dim(\ker(\mu I - A)) \\
&= n - \operatorname{rank}(\mu I - A).
\end{aligned}$$

上面一连串的等于关系请读者证明.