

## 第二章 $\lambda$ 矩阵与矩阵的 Jordan 标准型

### 2.1 $\lambda$ 矩阵及其 Smith 标准型

- $\mathbb{F}[\lambda]$  : 系数在  $\mathbb{F}$  中的  $\lambda$  的多项式的全体.
- $\mathbb{F}(\lambda)$  : 系数在  $\mathbb{F}$  中的  $\lambda$  的有理分式的全体. 显然,  $\mathbb{F}[\lambda] \subset \mathbb{F}(\lambda)$ .
- $\mathbb{F}[\lambda]$  是环, 称为多项式环.
- $\mathbb{F}(\lambda)$  是域, 称为有理分式域.

定义 (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  $\lambda$  矩阵. 记号

$\mathbb{F}^{m \times n}[\lambda]$  表示所有  $m$  行  $n$  列的  $\lambda$  矩阵的集合, 矩阵的元素是系数在  $\mathbb{F}$

中的  $\lambda$  的多项式. 也就是说,  $A(\lambda) \in \mathbb{F}^{m \times n}[\lambda]$  表示  $A(\lambda) = [a_{ij}(\lambda)]_{m \times n}$ ,

其中,  $a_{ij}(\lambda) \in \mathbb{F}[\lambda]$ .

注 (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

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

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。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式的值都是多项式。

定义（多项式矩阵的秩）

多项式矩阵的秩，也用  $\text{rank}$  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 (单位模阵)

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

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

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 (多项式矩阵的逆一般情况下会是有理分式矩阵)

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

$$(A(\lambda))^{-1} = \frac{\text{adj}(A(\lambda))}{\det(A(\lambda))}$$

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（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,  $\text{adj}(U(\lambda)) \in \mathbb{F}^{n \times n}[\lambda]$  为多项式矩阵 (不含分母为次数  $\geq 1$  的多项式的不可约分式为元素, 即  $\text{adj}(U(\lambda)) \notin \mathbb{F}^{n \times n}(\lambda) \setminus \mathbb{F}^{n \times n}[\lambda]$ ). 因此, 若  $\det(U(\lambda)) \in \mathbb{F}[\lambda]$  为非零常值多项式, 则  $(U(\lambda))^{-1} \in \mathbb{F}^{n \times n}[\lambda]$  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

$$\partial(|U(\lambda)|) + \partial(|V(\lambda)|) = \partial(1) = 0.$$

因此, 必有

$$\partial(|U(\lambda)|) = \partial(|V(\lambda)|) = 0.$$

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 (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  $(r_{i_1}) \leftrightarrow (r_{i_2})$ . 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  $c \cdot (r_i)$ .**
- (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  $(r_{i_2}) + f(\lambda) \cdot (r_{i_1})$ .**

**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，若是这样，则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（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（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  $\partial(b_{11}(\lambda)) < \partial(a_{11}(\lambda))$ .

## 定理 (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

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

其中  $d_i(\lambda), i = 1, 2, \dots, r$ , 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  $d_i(\lambda) \mid d_{i+1}(\lambda)$ ,  $i = 1, 2, \dots, r-1$ .

问题 (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

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

及

$$\left[ \begin{array}{ccc|c} e_1(\lambda) & & & 0 \\ & e_2(\lambda) & & \\ & & \ddots & \\ & & & e_t(\lambda) \\ \hline 0 & & & 0 \end{array} \right]_{m-t, n-t},$$

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

设  $f(\lambda), g(\lambda) \in \mathbb{F}[\lambda]$ . 回忆: 若  $f(\lambda)$  为非零多项式, 且  $f(\lambda) \mid g(\lambda)$  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子.

定义 (多项式矩阵的行列式因子)

设  $A(\lambda) \in \mathbb{F}^{m \times n}[\lambda]$ , 正整数  $k \leq \text{rank}(A(\lambda))$  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  $\text{rank}(A(\lambda)) < k \leq \min\{m, n\}$ ,  $A(\lambda)$  的  $k$  阶行列式因子规定为零多项式.

定理（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.

定理 (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

$$\begin{bmatrix} d_1(\lambda) & & & & 0 & & \\ & d_2(\lambda) & & & & & \\ & & \ddots & & & & \\ & & & d_r(\lambda) & & & \\ 0 & & & & & & \\ \hline & & 0 & & & & \\ & & & & & 0_{m-r, n-r} & \end{bmatrix}.$$

则

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

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

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

或

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

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  $d_i(\lambda), i = 1, 2, \dots, r$  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（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.

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

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

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

公因子为 1, 故  $D_1(\lambda) = 1$ .

## 2) 不为零的二阶子式有

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

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

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

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 (么模阵写为初等矩阵的乘积)

- (1) 么模阵的 Smith 标准型为单位矩阵.
- (2) 么模阵可写为有限个初等矩阵的乘积.

## 推论 (多项式矩阵等价的么模阵表述)

(1) 两个多项式矩阵  $A(\lambda), B(\lambda) \in \mathbb{F}^{m \times n}$  相互等价, 当且仅当存在两个么模阵  $U(\lambda) \in \mathbb{F}^{m \times m}$ ,  $V(\lambda) \in \mathbb{F}^{n \times n}$  使得

$$U(\lambda)A(\lambda)V(\lambda) = B(\lambda).$$

(2) 给定多项式矩阵  $A(\lambda) \in \mathbb{F}^{m \times n}$ , 存在两个么模阵  $U(\lambda) \in \mathbb{F}^{m \times m}$ ,  $V(\lambda) \in \mathbb{F}^{n \times n}$  使得

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

为  $A(\lambda)$  的 Smith 标准型.

## 2.2 特征矩阵

定义 (特征矩阵)

给定  $A \in \mathbb{F}^{n \times n}$ , 多项式矩阵

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

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, 因为有下列定理:

定理 (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.



## 定义 (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  $A_i \in \mathbb{F}^{m \times n}, i = 0, 1, \dots, d$  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 (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

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

## 引理 (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

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

且

$$R_1(\lambda) = 0 \text{ 或者 } \deg(R_1(\lambda)) < \deg(A(\lambda)).$$

这里, 下标字母“1”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.

## 定理 (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

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

则

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

推论 (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. 设  $d_i(\lambda)$ ,  $i = 1, 2, \dots, n$ , 中的非常数者(即次数  $\geq 1$  的)记为

$$h_1(\lambda), h_2(\lambda), \dots, h_p(\lambda),$$

次数分别为

$$n_1, n_2, \dots, n_p.$$

则  $d_i(\lambda)$ ,  $i = 1, 2, \dots, n$ , 中恰有  $n - p = (n_1 - 1) + (n_2 - 1) + \dots + (n_p - 1)$  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.

$$\left[ \begin{array}{c|c|c|c}
 \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}_{n_1 \times n_1} & & & \\
 \hline
 & \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}_{n_2 \times n_2} & & \\
 \hline
 & & \ddots & \\
 \hline
 & & & \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}_{n_p \times n_p}
 \end{array} \right]$$

下面将进一步通过不变因子的质因式分解, 将上述型式化成更特殊的结构. 为此先引入:

定义 (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.



记  $\lambda I - A$  的最后一个不变因子的质因式分解为

$$d_n(\lambda) = (e_1(\lambda))^{r_{n1}} \cdots (e_p(\lambda))^{r_{np}},$$

其中  $e_k(\lambda)$  为不可约多项式,  $r_{n_k} \geq 1$ ,  $k = 1, \dots, p$ . 则必有

$$d_i(\lambda) = (e_1(\lambda))^{r_{i1}} \cdots (e_p(\lambda))^{r_{ip}}, \quad i = 1, \dots, n-1.$$

且

$$0 \leq r_{i_k} \leq r_{i+1_k}, \quad i = 1, \dots, n-1, \quad k = 1, \dots, p.$$

则初等因子组为

$$\{(e_k(\lambda))^{r_{ik}} : r_{ik} \geq 1, i = 1, \dots, n-1; k = 1, \dots, p\},$$

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 (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) 设  $A \in \mathbb{C}^{4 \times 4}$ , 已知  $A$  的初等因子组为

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

求它的不变因子组 (Smith 型).

将所有初等因子按质因式分组; 每组至多 4 个, 将不够 4 个的用 1 补够 4 个, 并按升幂排列:

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

依次从每组中取元, 构成不变因子如下:

$$d_1(\lambda) = 1 \cdot 1 = 1$$

$$d_2(\lambda) = \lambda \cdot 1 = \lambda$$

$$d_3(\lambda) = \lambda \cdot (\lambda + 1)^2 = \lambda(\lambda + 1)^2$$

$$d_4(\lambda) = \lambda^2 \cdot (\lambda + 1)^2 = \lambda^2(\lambda + 1)^2.$$

下面基于初等因子将每一个子块进一步分解.

引理 (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

$$\begin{bmatrix} I_{n+1} & 0 \\ 0 & f_1(\lambda)f_2(\lambda) \end{bmatrix},$$
$$\left[ \begin{array}{cc|cc} I_{n_1} & 0 & & \\ 0 & f_1(\lambda) & & 0 \\ \hline & & I_{n_2} & 0 \\ 0 & & 0 & f_2(\lambda) \end{array} \right].$$

定理（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

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

次数分别为

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

则  $\lambda I - A$  等价于

$$\left[ \begin{array}{c|c|c|c}
 \begin{array}{c} \left[ \begin{array}{ccc} 1 & & \\ & \ddots & \\ & & 1 \end{array} \right]_{m_1 \times m_1} \\ (e_1(\lambda))^{r_1} \end{array} & & & \\
 \hline
 & \begin{array}{c} \left[ \begin{array}{ccc} 1 & & \\ & \ddots & \\ & & 1 \end{array} \right]_{m_2 \times m_2} \\ (e_2(\lambda))^{r_2} \end{array} & & \\
 \hline
 & & \ddots & \\
 \hline
 & & & \begin{array}{c} \left[ \begin{array}{ccc} 1 & & \\ & \ddots & \\ & & 1 \end{array} \right]_{m_q \times m_q} \\ (e_q(\lambda))^{r_q} \end{array}
 \end{array} \right]$$

## 定理 (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各阶行列式因子.
- (5)  $\lambda I - A$  与  $\lambda I - B$  作为多项式矩阵有相同的各阶不变因子.
- (6)  $\lambda I - A$  与  $\lambda I - B$  作为多项式矩阵有相同的初等因子组.

例 (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. 下面利用上述定理来证明. 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

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

由于单位模阵的转置仍为单位模阵, 因此.....



## 2.3 复数域上 Jordan 标准型

问题（构造其特征矩阵具有给定的单个初等因子的简单矩阵）

给定复数  $\mu$ ，求“尽可能简单的”复矩阵  $J \in \mathbb{C}^{n \times n}$ ，使得  $\lambda I - J$  只有一个初等因子  $(\lambda - \mu)^n$ ，即  $\lambda I - J$  等价于

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

特别地, 若不可约因式为 1 次多项式

定理 (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  $\mu$  的  $n$  阶 Jordan 块. 则  $\lambda I - J_n(\mu)$  等价于 (Smith 标准型)

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

定理 (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  $(\lambda - \lambda_1)^{m_1}, (\lambda - \lambda_2)^{m_2}, \dots, (\lambda - \lambda_q)^{m_q}$ . 则  $A$  相似于下面的 Jordan 标准型

$$J = \text{diag}\{J_{m_1}(\lambda_1), J_{m_2}(\lambda_2), \dots, J_{m_q}(\lambda_q)\} = \begin{bmatrix} J_{m_1}(\lambda_1) & & & \\ & J_{m_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{m_q}(\lambda_q) \end{bmatrix}$$

设

$$T^{-1}AT = J .$$

下面研究相似矩阵  $T$  的几何意义.

已知：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  $J$ .

进而，我们研究与 Jordan 标准型  $J$  相应的基矩阵  $T$  的特殊结构.

考虑矩阵  $T$  与 Jordan 标准型相应的列分块

$$T = \left[ \underbrace{t_1 \cdots t_{m_1}}_{T_1} \mid \underbrace{t_{m_1+1} \cdots t_{m_1+m_2}}_{T_2} \mid \cdots \mid \underbrace{t_{m_1+\cdots+m_{q-1}+1} \cdots t_n}_{T_q} \right].$$

则

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

由于  $J$  为块对角矩阵, 因而每个子空间  $\text{im } T_j$  都是  $A$  的不变子空间, 且

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

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

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

注意到

$$J_{m_j}(\lambda_j) = \begin{bmatrix} \lambda_j & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_j \end{bmatrix}_{m_j \times m_j}$$
$$= \lambda_j I_{m_j} + \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix},$$

于是

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



即

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

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  $t_i$ ,  $1 \leq i \leq m_1$ , 满足

$$\begin{aligned}(A - \lambda_1 I)^i t_i &= 0, \\(A - \lambda_1 I)^{i-1} t_i &= t_1 \neq 0.\end{aligned}$$

定义 (广义特征向量)

我们称满足条件

$$\begin{aligned}(A - \mu I)^k p &= 0, \\ (A - \mu I)^{k-1} p &\neq 0\end{aligned}$$

的向量  $p$  是矩阵  $A$  的相应于特征值  $\mu$  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  $\lambda_1$  的指标为  $i$  的广义特征向量,  
 $1 \leq i \leq m_1$ , 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

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

是矩阵  $A$  的相应于特征值  $\lambda_1$  的一个长度为  $m_1$  的广义特征向量链.

一般地, 有定义

定义 (广义特征向量链)  
向量组

$$p_1, p_2, \dots, p_m$$

称为矩阵  $A$  的相应于特征值  $\mu$  的一个长度为  $m$  的广义特征向量链,  
如果

$$\begin{aligned}(A - \mu I)p_i &= p_{i-1}, i = 2, \dots, m, \\ (A - \mu I)p_1 &= 0, p_1 \neq 0.\end{aligned}$$

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.

定义 (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  $\mu$  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  $\mu$  作为多项式  $|\lambda I - A|$  的根的重数.

(2)  $A$  的特征值  $\mu$  的几何重数

=  $\mu$  对应的的线性无关的特 征向量的个数

= 方程  $(\mu I - A)x = 0$  解空间的维数

=  $\dim(\ker(\mu I - A))$

=  $n - \text{rank}(\mu I - A)$

=  $A$  的 Jordan 标准型中对角线上为  $\mu$  的 Jordan 块的个数 .