第三章　　λ矩阵与矩阵的Jordan标准形

3.1　一元多项式

- 定义3.1.1　设$\mathbb{P}$是数域，n是一非负整数，λ是一个文字，形式表达式

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

  其中$a_i\in\mathbb{P}\ (i=0,1,\cdots,n)$，称为数域$\mathbb{P}$上的一元多项式，数域$\mathbb{P}$上一元多项式的全体记为$\mathbb{P}[\lambda]$；$n$称为$f(\lambda)$的次数，记为$\partial(f(\lambda))=n$;若$a_n=1$，则称$f(\lambda)$ 为首一多项式。

- 一元多项式的相等、相加、相减、相乘、数乘及其运算规律

- 定理3.1.1　设$f(\lambda),g(\lambda)\in\mathbb{P}[\lambda]$且$g(\lambda)\neq 0$，则存在惟一的多项式$q(\lambda),r(\lambda)\in\mathbb{P}[\lambda]$使得

$$f(\lambda)=q(\lambda)g(\lambda)+r(\lambda)$$

  其中$r(\lambda)=0$或$r(\lambda)\neq 0$但$\partial(r(\lambda))<\partial(g(\lambda))$；$q(\lambda)$称为$g(\lambda)$除$f(\lambda)$的商，$r(\lambda)$称为$g(\lambda)$除$f(\lambda)$的余式。

- 定义3.1.2　设$f(\lambda), g(\lambda) \in \mathbb{P}[\lambda]$如果存在$h(\lambda) \in \mathbb{P}[\lambda]$使得

$$f(\lambda) = h(\lambda)g(\lambda)$$

则称多项式$g(\lambda)$整除$f(\lambda)$，记为$g(\lambda)|f(\lambda)$；$g(\lambda)$称为$f(\lambda)$的因式；$f(\lambda)$称为$g(\lambda)$的倍式。如果$g(\lambda)$不能整除$f(\lambda)$，记为$g(\lambda) \nmid f(\lambda)$

- 注意区别一元多项式的整除与整数的整除。如：在$\mathbb{P}[\lambda]$中，

$$\begin{array}{ll} 3|4, & 4 = \frac{4}{3} \times 3 \\ 3|5\lambda + 7, & 5\lambda + 7 = \left(\frac{5}{3}\lambda + \frac{7}{3}\right) \times 3 \\ 3|8\lambda^2 + \lambda + 2, & 8\lambda^2 + \lambda + 2 = \left(\frac{8}{3}\lambda^2 + \frac{1}{3}\lambda + \frac{2}{3}\right) \times 3 \end{array}$$

特别地，$g|f(\lambda)$，$\forall$非零$g \in \mathbb{P}, \forall f(\lambda) \in \mathbb{P}[\lambda]$.

- 性质　如果$g(\lambda)|f_i(\lambda)\ \ (i = 1, \cdots, m)$，则$g(\lambda)|(u_1(\lambda)f_1(\lambda) + \cdots + u_m(\lambda)f_m(\lambda))$，其中$u_i(\lambda) \in \mathbb{P}[\lambda]\ \ (i = 1, \cdots, m)$
- 公因式，最大公因式，互素，公倍式，最小公倍式

3.2 λ矩阵及其在相抵下的标准形

3.2.1 λ矩阵的基本概念

- 定义3.2.1 设$a_{ij}(\lambda)$ $(i=1,2,\cdots,m;j=1,2,\cdots,n)$是数域$\mathbb{P}$上的多项式，以$a_{ij}(\lambda)$为元素的$m\times n$矩阵

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

  称为多项式矩阵或λ矩阵，多项式$a_{ij}(\lambda)$ $(i=1,2,\cdots,m;j=1,2,\cdots,n)$中的最高次数称为$\mathbf{A}(\lambda)$的次数，数域$\mathbb{P}$上$m\times n$维$\lambda$矩阵的全体记为$\mathbb{P}[\lambda]^{m\times n}$

- 数字矩阵是λ矩阵的特例，$\mathbf{A}\in\mathbb{C}^{n\times n}$的特征矩阵$\lambda\mathbf{I}-\mathbf{A}$是1次$\lambda$矩阵

- λ矩阵的相等、加法、减法、乘法、数乘及其运算规则

- 如果$m\times n$维λ矩阵$\mathbf{A}(\lambda)$的次数为k，则$\mathbf{A}(\lambda)$可表示为

$$\mathbf{A}(\lambda)=\mathbf{A}_k\lambda^k+\mathbf{A}_{k-1}\lambda^{k-1}+\cdots+\mathbf{A}_1\lambda+\mathbf{A}_0$$

  其中$\mathbf{A}_i\ \ (i=0,1,\cdots,k)$是$m\times n$数字矩阵，并且$\mathbf{A}_k\neq 0$；例如

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

- λ矩阵的行列式、子式、代数余子式的运算和性质与数字矩阵类似。
  子式的定义：$\mathbf{A}(\lambda)\in\mathbb{P}[\lambda]^{m\times n}$，在$\mathbf{A}(\lambda)$中选取$r$行$r$列$(1\leq r\leq n)$，由这些行和列相交处的元素按原来相关位置构成的$r$阶行列式称为$\mathbf{A}(\lambda)$的$r$阶子式。

- 比如：

$$\begin{bmatrix} \lambda+1 & \lambda^2+1 & \lambda^2 \\ 3\lambda-1 & 3\lambda^2-1 & \lambda^2+2\lambda \end{bmatrix}$$

  有3个2阶子式

$$\begin{vmatrix} \lambda+1 & \lambda^2+1 \\ 3\lambda-1 & 3\lambda^2-1 \end{vmatrix} = 4\lambda^2-4\lambda$$
$$\begin{vmatrix} \lambda+1 & \lambda^2 \\ 3\lambda-1 & \lambda^2+2\lambda \end{vmatrix} = -2\lambda^3+4\lambda^2+2\lambda$$
$$\begin{vmatrix} \lambda^2+1 & \lambda^2 \\ 3\lambda^2-1 & \lambda^2+2\lambda \end{vmatrix} = -2\lambda^4+2\lambda^3+2\lambda^2+2\lambda$$

- 定义3.2.2　设$\mathbf{A}(\lambda) \in \mathbb{P}[\lambda]^{m\times n}$，如果$\mathbf{A}(\lambda)$中有一个$r(1 \le r \le \min\{m,n\})$阶子式不为零，而所有$r+1$阶子式（如果有的话）全为零，则称$\mathbf{A}(\lambda)$的秩为$r$，记为$\operatorname{rank}(\mathbf{A}(\lambda)) = r$
- 例　$\mathbf{A} \in \mathbb{C}^{n\times n}$，$|\lambda\mathbf{I}-\mathbf{A}|$是$\lambda$的$n$次首1多项式，因此$|\lambda\mathbf{I}-\mathbf{A}|$是$\lambda$的秩为$n$，总是满秩的。

- 定义3.2.3　设$\mathbf{A}(\lambda)\in\mathbb{P}[\lambda]^{n\times n}$，如果存在一个$n$阶矩阵$\mathbf{B}(\lambda)$使得

$$\mathbf{A}(\lambda)\mathbf{B}(\lambda)=\mathbf{B}(\lambda)\mathbf{A}(\lambda)=\mathbf{I}$$

则称λ矩阵$\mathbf{A}(\lambda)$是可逆的，并称$\mathbf{B}(\lambda)$为$\mathbf{A}(\lambda)$的逆矩阵，记作$\mathbf{A}(\lambda)^{-1}$

- 注意：上述定义仅指在$\mathbb{P}[\lambda]^{n\times n}$中的可逆。比如：

$$\begin{bmatrix}\lambda+1 & 1\\ 2 & 1\end{bmatrix}^{-1}=\begin{bmatrix}\frac{1}{\lambda-1} & \frac{-1}{\lambda-1}\\ \frac{-2}{\lambda-1} & \frac{\lambda+1}{\lambda-1}\end{bmatrix}$$

在λ有理分式矩阵意义下是可逆的，但在$\mathbb{P}[\lambda]^{2\times 2}$中不是可逆的。

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

在$\mathbb{P}[\lambda]^{2\times 2}$中是可逆的。

- 定理3.2.1　设$\mathbf{A}(\lambda) \in \mathbb{P}[\lambda]^{n\times n}$，则$\mathbf{A}(\lambda)$可逆的充分必要条件是$|\mathbf{A}(\lambda)|$为非零常数

证明　必要性。

$$\begin{gathered}\mathbf{A}(\lambda)\text{可逆}\\ \Downarrow\\ \mathbf{A}(\lambda)\mathbf{B}(\lambda)=\mathbf{B}(\lambda)\mathbf{A}(\lambda)=\mathbf{I}\\ \Downarrow\\ |\mathbf{A}(\lambda)\mathbf{B}(\lambda)|=|\mathbf{A}(\lambda)||\mathbf{B}(\lambda)|=1\\ \Downarrow\\ |\mathbf{A}(\lambda)|\text{是非零常数}\end{gathered}$$

充分性。

$$\begin{gathered}|\mathbf{A}(\lambda)|=d\text{是非零常数}\\ \Downarrow\\ \text{构造}\mathbf{A}(\lambda)\text{的伴随矩阵}\mathbf{A}(\lambda)^*\text{，则}\mathbf{A}(\lambda)\tfrac{1}{d}\mathbf{A}(\lambda)^*=\tfrac{1}{d}\mathbf{A}(\lambda)^*\mathbf{A}(\lambda)=\mathbf{I}\\ \Downarrow\\ \mathbf{A}(\lambda)^{-1}=\tfrac{1}{d}\mathbf{A}(\lambda)^*\end{gathered}$$

3.2.2 λ矩阵的初等变换与相抵

- 定义3.2.4 下列三种变换称为λ矩阵的初等变换

  （1）λ矩阵的两行（列）互换位置

  （2）λ矩阵的某一行（列）乘以非零常数k

  （3）λ矩阵的某一行（列）的$\varphi(\lambda)$倍加到另一行（列），其中$\varphi(\lambda)$是多项式

- 记号$[i,j]$表示第i,j行（列）互换位置

  记号$[i(k)]$表示用非零常数k乘第i行（列）

  记号$[i+j(\varphi)]$表示将第j行（列）的$\varphi(\lambda)$倍加到第i行（列）

- 三种λ矩阵的初等矩阵

$$\mathbf{P}(i,j)=\left[\begin{array}{ccccccccccc}
1 & & & & & & & & & & \\
 & \ddots & & & & & & & & & \\
 & & 1 & & & & & & & & \\
 & & & 0 & & \cdots & & 1 & & & \\
 & & & & 1 & & & & & & \\
 & & & \vdots & & \ddots & \vdots & & & & \\
 & & & & & & 1 & & & & \\
 & & & 1 & & \cdots & & 0 & & & \\
 & & & & & & & & 1 & & \\
 & & & & & & & & & \ddots & \\
 & & & & & & & & & & 1
\end{array}\right]\begin{array}{l}
 \\ \\ \\ \text{第}i\text{行} \\ \\ \\ \\ \text{第}j\text{行} \\ \\ \\ \\
\end{array}$$

$$\mathbf{P}(i(k))=\left[\begin{array}{ccccccc}
1 & & & & & & \\
 & \ddots & & & & & \\
 & & 1 & & & & \\
 & & & k & & & \\
 & & & & 1 & & \\
 & & & & & \ddots & \\
 & & & & & & 1
\end{array}\right]\begin{array}{l}
 \\ \\ \\ \text{第}i\text{行} \\ \\ \\ \\
\end{array},\mathbf{P}(i,j(\varphi))=\left[\begin{array}{ccccccc}
1 & & & & & & \\
 & \ddots & & & & & \\
 & & 1 & \cdots & \varphi(\lambda) & & \\
 & & & \ddots & \vdots & & \\
 & & & & 1 & & \\
 & & & & & \ddots & \\
 & & & & & & 1
\end{array}\right]\begin{array}{l}
 \\ \\ \text{第}i\text{行} \\ \\ \text{第}j\text{行} \\ \\ \\
\end{array}$$

- 对$m\times n$的$\mathbf{A}(\lambda)$作一次初等行变换相当于在$\mathbf{A}(\lambda)$左边乘上相应的m阶初等矩阵；对$\mathbf{A}(\lambda)$作一次初等列变换相当于在$\mathbf{A}(\lambda)$右边乘上相应的n阶初等矩阵

- $\mathbf{P}(i,j)^{-1}=\mathbf{P}(i,j), \mathbf{P}(i(k))^{-1}=\mathbf{P}(i(k^{-1})), \mathbf{P}(i,j(\varphi))^{-1}=\mathbf{P}(i,j(-\varphi))$,
- 定义3.2.5　设$\mathbf{A}(\lambda),\mathbf{B}(\lambda)\in\mathbb{P}[\lambda]^{m\times n}$，如果$\mathbf{A}(\lambda)$经过有限次的初等变换化为$\mathbf{B}(\lambda)$，则称$\lambda$矩阵$\mathbf{A}(\lambda)$与$\mathbf{B}(\lambda)$相抵，记为$\mathbf{A}(\lambda)\cong\mathbf{B}(\lambda)$
- 定理3.2.2　设$\mathbf{A}(\lambda),\mathbf{B}(\lambda)\in\mathbb{P}[\lambda]^{m\times n}$，则$\mathbf{A}(\lambda)$与$\mathbf{B}(\lambda)$相抵的充分必要条件为存在$m$阶初等矩阵$\mathbf{P}_1(\lambda),\cdots,\mathbf{P}_l(\lambda)$与$n$阶初等矩阵$\mathbf{Q}_1(\lambda),\cdots,\mathbf{Q}_t(\lambda)$使得

$$\mathbf{A}(\lambda)=\mathbf{P}_l(\lambda)\cdots\mathbf{P}_1(\lambda)\mathbf{B}(\lambda)\mathbf{Q}_1(\lambda)\cdots\mathbf{Q}_t(\lambda)$$

- 定理3.3.5　设$\mathbf{A}(\lambda),\mathbf{B}(\lambda)\in\mathbb{P}[\lambda]^{m\times n}$，则$\mathbf{A}(\lambda)$与$\mathbf{B}(\lambda)$相抵的充分必要条件是存在两个可逆$\lambda$矩阵$\mathbf{P}(\lambda)\in\mathbb{P}[\lambda]^{m\times m}$与$\mathbf{Q}(\lambda)\in\mathbb{P}[\lambda]^{n\times n}$ 使得

$$\mathbf{A}(\lambda)=\mathbf{P}(\lambda)\mathbf{B}(\lambda)\mathbf{Q}(\lambda)$$

- 定理3.3.4　设$\mathbf{A}(\lambda)\in\mathbb{P}[\lambda]^{n\times n}$，则$\mathbf{A}(\lambda)$可逆的充分必要条件是$\mathbf{A}(\lambda)$可表示为一系列初等矩阵的乘积

3.2.3 λ矩阵在相抵下的标准形

- 引理3.2.1 设λ矩阵$\mathbf{A}(\lambda)=(a_{ij}(\lambda))$的左上角元素$a_{11}(\lambda)\neq 0$，并且$\mathbf{A}(\lambda)$中至少有一个元素不能被$a_{11}(\lambda)$整除，则存在一个与$\mathbf{A}(\lambda)$相抵的λ矩阵$\mathbf{B}(\lambda)=(b_{ij}(\lambda))$使得$b_{11}(\lambda)\neq 0$且$\partial(b_{11}(\lambda))<\partial(a_{11}(\lambda))$

证明：（1）$a_{11}(\lambda)\nmid a_{i1}(\lambda)$，即$a_{i1}(\lambda)=q(\lambda)a_{11}(\lambda)+r(\lambda), r(\lambda)\neq 0$;

$$\mathbf{A}(\lambda)=\begin{bmatrix} a_{11}(\lambda) & \cdots \\ \vdots & \vdots \\ q(\lambda)a_{11}(\lambda)+r(\lambda) & \cdots \\ \vdots & \vdots \end{bmatrix}\xRightarrow{[i+1(-q(\lambda))]}\begin{bmatrix} a_{11}(\lambda) & \cdots \\ \vdots & \vdots \\ r(\lambda) & \cdots \\ \vdots & \vdots \end{bmatrix}\xRightarrow{[1,i]}\begin{bmatrix} r(\lambda) & \cdots \\ \vdots & \vdots \\ a_{11}(\lambda) & \cdots \\ \vdots & \vdots \end{bmatrix}=\mathbf{B}(\lambda)$$

（2）$a_{11}(\lambda)\nmid a_{1j}(\lambda)$，类似（1）

（3）$a_{11}(\lambda)\nmid a_{ij}(\lambda)$ $(i>1,j>1)$，而$a_{1j}(\lambda)=\varphi(\lambda)a_{11}(\lambda)$，

$$\begin{bmatrix} a_{11}(\lambda) & \cdots & \varphi(\lambda)a_{11}(\lambda) & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1}(\lambda) & \cdots & a_{ij}(\lambda) & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}\underset{[j+1(-\varphi(\lambda))]}{\Longrightarrow}\begin{bmatrix} a_{11}(\lambda) & \cdots & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1}(\lambda) & \cdots & a_{ij}(\lambda)-\varphi(\lambda)a_{i1}(\lambda) & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

$$\underset{[1+j(1)]}{\Longrightarrow}\begin{bmatrix} a_{11}(\lambda) & \cdots & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ a_{ij}(\lambda)+[1-\varphi(\lambda)]a_{i1}(\lambda) & \cdots & a_{ij}(\lambda)-\varphi(\lambda)a_{i1}(\lambda) & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

这就化为（1）

- 定理3.2.3　设$\mathbf{A}(\lambda)=(a_{ij}(\lambda))\in\mathbb{P}[\lambda]^{m\times n}$, $\mathrm{rank}(\mathbf{A}(\lambda))=r$，则$\mathbf{A}(\lambda)$相抵于Smith标准形

$$\begin{bmatrix} d_1(\lambda) & & & & & & \\ & d_2(\lambda) & & & & & \\ & & \ddots & & & & \\ & & & d_r(\lambda) & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix}\in\mathbb{P}[\lambda]^{m\times n}$$

其中$d_i(\lambda)\ \ (i=1,\cdots,r)$是首1多项式，且$d_i(\lambda)|d_{i+1}(\lambda)\ \ (i=1,\cdots,r-1)$

证明：设$r>0$且$a_{11}(\lambda)\neq 0$，由引理3.2.1知，$\mathbf{A}(\lambda)\cong\mathbf{B}(\lambda)=(b_{ij}(\lambda))$，$b_{11}(\lambda)$是首1多项式，$b_{11}(\lambda)|b_{ij}(\lambda)\ \ (i=1,\cdots,m;j=1,\cdots,n)$

$$\therefore\quad \mathbf{B}(\lambda)\cong\begin{bmatrix} d_1(\lambda) & \\ & \mathbf{A}_1(\lambda)\end{bmatrix}$$

$d_1(\lambda)=b_{11}(\lambda)$，$d_1(\lambda)$整除$\mathbf{A}_1(\lambda)$的所有元素。若$\mathbf{A}_1(\lambda)\neq\mathbf{0}$，对$\mathbf{A}_1(\lambda)$重复上述过程，有

$$\mathbf{B}(\lambda)\cong\begin{bmatrix} d_1(\lambda) & & \\ & d_2(\lambda) & \\ & & \mathbf{A}_2(\lambda)\end{bmatrix}$$

$d_1(\lambda),d_2(\lambda)$都是首1多项式，且$d_1(\lambda)|d_2(\lambda)$，$d_2(\lambda)$整除$\mathbf{A}_2(\lambda)$的所有元素。继续上述过程，最后得Smith标准形

- 定理3.3.2　λ矩阵$\mathbf{A}(\lambda)$的Smith标准形是惟一的.(用λ矩阵的行列式因子来证明)
- 定义3.2.6　λ矩阵$\mathbf{A}(\lambda) \in \mathbb{P}[\lambda]^{m\times n}$的Smith标准形"主对角线"上非零元$d_1(\lambda)$，$d_2(\lambda)$，$\cdots$，$d_r(\lambda)$ 称为$\mathbf{A}(\lambda)$的不变因子
- 定理3.3.3　设$\mathbf{A}(\lambda), \mathbf{B}(\lambda) \in \mathbb{P}[\lambda]^{m\times n}$，则$\mathbf{A}(\lambda)$与$\mathbf{B}(\lambda)$相抵的充分必要条件是它们有相同的不变因子
- 例3.2.2

$$
\begin{bmatrix} -\lambda+1 & \lambda^2 & \lambda \\ \lambda & \lambda & -\lambda \\ \lambda^2+1 & \lambda^2 & -\lambda^2 \end{bmatrix}
\underset{[1+3(1)]}{\Longrightarrow}
\begin{bmatrix} 1 & \lambda^2 & \lambda \\ 0 & \lambda & -\lambda \\ 1 & \lambda^2 & -\lambda^2 \end{bmatrix}
\overset{[3+1(-1)]}{\Longrightarrow}
\begin{bmatrix} 1 & \lambda^2 & \lambda \\ 0 & \lambda & -\lambda \\ 0 & 0 & -\lambda^2-\lambda \end{bmatrix}
$$

$$
\underset{\substack{[2+1(-\lambda^2)] \\ [3+1(-\lambda)]}}{\Longrightarrow}
\begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & -\lambda^2-\lambda \end{bmatrix}
\overset{[3(-1)]}{\underset{[3+2(1)]}{\Longrightarrow}}
\begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda^2+\lambda \end{bmatrix}
$$

3.3 λ矩阵的初等因子

- 设λ矩阵$\mathbf{A}(\lambda)$的不变因子为$d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda)$，在复数域内将它们分解成一次因式的幂的乘积

$$\begin{cases} d_1(\lambda) = (\lambda-\lambda_1)^{e_{11}}(\lambda-\lambda_2)^{e_{12}}\cdots(\lambda-\lambda_s)^{e_{1s}} \\ d_2(\lambda) = (\lambda-\lambda_1)^{e_{21}}(\lambda-\lambda_2)^{e_{22}}\cdots(\lambda-\lambda_s)^{e_{2s}} \\ \cdots\cdots \\ d_r(\lambda) = (\lambda-\lambda_1)^{e_{r1}}(\lambda-\lambda_2)^{e_{r2}}\cdots(\lambda-\lambda_s)^{e_{rs}} \end{cases}$$

因为$d_i(\lambda)|d_{i+1}(\lambda)\ \ (i=1,\cdots,r-1)$，所以

$$\begin{cases} 0 \le e_{11} \le e_{21} \le \cdots \le e_{r1} \\ 0 \le e_{12} \le e_{22} \le \cdots \le e_{r2} \\ \cdots\cdots \\ 0 \le e_{1s} \le e_{2s} \le \cdots \le e_{rs} \end{cases}$$

- 定义3.3.2　$(\lambda-\lambda_j)^{e_{ij}}\ \ (i=1,\cdots,r; j=1,\cdots,s)$称为$\lambda$矩阵$\mathbf{A}(\lambda)$的初等因子
- 例如 若λ矩阵$\mathbf{A}(\lambda)$的不变因子为

$$\begin{cases} d_1(\lambda) = 1 \\ d_2(\lambda) = \lambda(\lambda-1) \\ d_3(\lambda) = \lambda(\lambda-1)^2(\lambda+1)^2 \\ d_4(\lambda) = \lambda^2(\lambda-1)^3(\lambda+1)^3(\lambda-2) \end{cases}$$

则$\mathbf{A}(\lambda)$的初等因子为$\lambda, \lambda, \lambda^2, \lambda-1, (\lambda-1)^2, (\lambda-1)^3, (\lambda+1)^2, (\lambda+1)^3, \lambda-2$

- 反过来，如果知道一个λ矩阵的秩r和初等因子，也可惟一确定其不变因子：r确定了不变因子的个数，同一个一次因式的方幂作成的初等因子中，方次最高的必在$d_r(\lambda)$的分解中，方次次高的必在$d_{r-1}(\lambda)$的分解中，如此顺推。
- 例如5×6维λ矩阵$\mathbf{A}(\lambda)$的秩为4，其初等因子

$$\lambda, \lambda, \lambda^2, \lambda-1, (\lambda-1)^2, (\lambda-1)^3, (\lambda+\mathrm{i})^3, (\lambda-\mathrm{i})^3$$

则不变因子

$$\begin{aligned}
&d_4(\lambda)=\lambda^2(\lambda-1)^3(\lambda+\mathrm{i})^3(\lambda-\mathrm{i})^3=\lambda^2(\lambda-1)^3(\lambda^2+1)^3\\
&d_3(\lambda)=\lambda(\lambda-1)^2\\
&d_2(\lambda)=\lambda(\lambda-1)\\
&d_1(\lambda)=1
\end{aligned}$$

Smith标准形

$$\mathbf{A}(\lambda)\cong\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0\\
0 & \lambda(\lambda-1) & 0 & 0 & 0 & 0\\
0 & 0 & \lambda(\lambda-1)^2 & 0 & 0 & 0\\
0 & 0 & 0 & \lambda^2(\lambda-1)^3(\lambda^2+1)^3 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

- 定理3.3.6　设$\mathbf{A}(\lambda), \mathbf{B}(\lambda) \in \mathbb{P}[\lambda]^{m\times n}$，则$\mathbf{A}(\lambda)$与$\mathbf{B}(\lambda)$相抵的充分必要条件是它们有相同的秩和相同的初等因子
- 定理3.3.7　设λ矩阵
$$\mathbf{A}(\lambda) = \begin{bmatrix} \mathbf{B}(\lambda) & \\ & \mathbf{C}(\lambda) \end{bmatrix}$$
为块对角矩阵，则$\mathbf{B}(\lambda)$与$\mathbf{C}(\lambda)$的初等因子的全体是$\mathbf{A}(\lambda)$的全部初等因子
- 定理3.3.8　若λ矩阵$\mathbf{A}(\lambda)$相抵于块对角矩阵
$$\mathbf{A}(\lambda) \cong \begin{bmatrix} \mathbf{A}_1(\lambda) & & & \\ & \mathbf{A}_2(\lambda) & & \\ & & \ddots & \\ & & & \mathbf{A}_t(\lambda) \end{bmatrix}$$
则$\mathbf{A}_1(\lambda), \mathbf{A}_2(\lambda), \cdots, \mathbf{A}_t(\lambda)$各个初等因子的全体就是$\mathbf{A}(\lambda)$的全部初等因子

3.4 矩阵相似的条件

- 引理3.4.1 $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}$，如果存在$\mathbf{P}, \mathbf{Q} \in \mathbb{C}^{n\times n}$使得$\lambda\mathbf{I} - \mathbf{A} = \mathbf{P}(\lambda\mathbf{I} - \mathbf{B})\mathbf{Q}$，则$\mathbf{A}$与$\mathbf{B}$相似。

  证明：比较$\lambda\mathbf{I} - \mathbf{A} = \mathbf{PQ}\lambda - \mathbf{PBQ}$两边的系数矩阵，得$\mathbf{PQ} = \mathbf{I}, \mathbf{A} = \mathbf{PBQ}$，由此

$$\mathbf{Q} = \mathbf{P}^{-1}, \mathbf{A} = \mathbf{PBP}^{-1}$$

- 定理3.4.1 n阶矩阵$\mathbf{A}$与$\mathbf{B}$相似的充分必要条件是$\lambda\mathbf{I} - \mathbf{A}$和$\lambda\mathbf{I} - \mathbf{B}$相抵。

- 定理 3.4.2 n 阶矩阵 $\mathbf{A}$ 与 $\mathbf{B}$ 相似的充分必要条件是$\lambda\mathbf{I} - \mathbf{A}$和$\lambda\mathbf{I} - \mathbf{B}$ 有相同的不变因子。

- 定理 3.4.3 n 阶矩阵 $\mathbf{A}$ 与 $\mathbf{B}$ 相似的充分必要条件是$\lambda\mathbf{I} - \mathbf{A}$和$\lambda\mathbf{I} - \mathbf{B}$有相同的秩和相同的初等因子。

3.5 矩阵的Jordan标准形

- 定义3.5.1 形状为

$$\mathbf{J}_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_i & 1 \\ & & & \lambda_i \end{bmatrix} \in \mathbb{C}^{n_i \times n_i}$$

的矩阵称为Jordan块，其中λ_i为复数。由若干个Jordan块为对角块组成的块对角矩阵称为Jordan形矩阵，如：

$$\begin{bmatrix} 1 & 1 & & & & & & \\ & 1 & & & & & & \\ & & 4 & & & & & \\ & & & 4 & 1 & & & \\ & & & & 4 & & & \\ & & & & & -\mathrm{i} & 1 & \\ & & & & & & -\mathrm{i} & 1 \\ & & & & & & & -\mathrm{i} \end{bmatrix}$$

- $\mathbf{J}_i$ 具有一个 n_i 重特征值 λ_i，对应于特征值 λ_i 仅有一个线性无关的特征向量。
- $\mathbf{J}_i$ 的乘幂有明显的表示式

$$\mathbf{J}_i^p = \begin{bmatrix} f_p(\lambda_i) & f_p^{'}(\lambda_i) & \frac{1}{2!}f_p^{''}(\lambda_i) & \cdots & \cdots & \frac{1}{(n_i-1)!}f_p^{n_i-1}(\lambda_i) \\ & & & \ddots & & \vdots \\ & f_p(\lambda_i) & f_p^{'}(\lambda_i) & & \ddots & \vdots \\ & & \ddots & \ddots & & \frac{1}{2!}f_p^{''}(\lambda_i) \\ & & & & \ddots & \\ & 0 & & & \ddots & f_p^{'}(\lambda_i) \\ & & & & & f_p(\lambda_i) \end{bmatrix}, p = 1, 2, \cdots$$

其中 $f_p(\lambda) = \lambda^p$

-

$$\lambda\mathbf{I} - \mathbf{J}_i \cong \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & (\lambda-\lambda_i)^{n_i} \end{bmatrix} \in \mathbb{P}[\lambda]^{n_i \times n_i}$$

所以$\mathbf{J}_i$ 的不变因子为$d_1(\lambda) = \cdots = d_{n_i-1}(\lambda) = 1, d_{n_i}(\lambda) = (\lambda-\lambda_i)^{n_i}$，从而$\mathbf{J}_i$的初等因子为$(\lambda-\lambda_i)^{n_i}$

- 设

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_1 & & & \\ & \mathbf{J}_2 & & \\ & & \ddots & \\ & & & \mathbf{J}_s \end{bmatrix}$$

是Jordan形矩阵，其中$\mathbf{J}_i \in \mathbb{C}^{n_i \times n_i}$是Jordan块。$\mathbf{J}$的特征矩阵为

$$\lambda\mathbf{I} - \mathbf{J} = \begin{bmatrix} \lambda\mathbf{I} - \mathbf{J}_1 & & & \\ & \lambda\mathbf{I} - \mathbf{J}_2 & & \\ & & \ddots & \\ & & & \lambda\mathbf{I} - \mathbf{J}_s \end{bmatrix}$$

由定理3.3.8知，Jordan形矩阵$\mathbf{J}$的初等因子为

$$(\lambda - \lambda_1)^{n_1}, (\lambda - \lambda_2)^{n_2}, \cdots, (\lambda - \lambda_s)^{n_s}$$

- 定理3.5.1 设$\mathbf{A}\in\mathbb{C}^{n\times n}$,则**A**与一个Jordan形矩阵相似，并且Jordan形矩阵除去其中Jordan块的排列次序外是被矩阵**A**唯一决定的。

证明：设 **A** 的初等因子为

$$(\lambda-\lambda_1)^{n_1},(\lambda-\lambda_2)^{n_2},\cdots,(\lambda-\lambda_s)^{n_s}, \tag{1}$$

其中 $\lambda_1,\cdots,\lambda_s$ 可能有相同的， $n_1,\cdots,n_s$ 也可能有相同的。每个初等因子 $(\lambda-\lambda_i)^{n_i}$ 对应于一个 Jordan 块

$$\mathbf{J}_i=\begin{bmatrix}\lambda_i & 1 & & & 0\\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1\\ 0 & & & & \lambda_i\end{bmatrix}\in\mathbb{C}^{n_i\times n_i},i=1,\cdots,s$$

这些 Jordan 块构成一个 Jordan 形矩阵

$$\mathbf{J}=\mathrm{diag}(\mathbf{J}_1,\mathbf{J}_2,\cdots,\mathbf{J}_s) \tag{2}$$

其初等因子也是 (1)。因为 **J** 与 **A** 有相同的初等因子， **J** 与 **A** 相似。Jordan 形矩阵 (2) 称为矩阵**A**的Jordan标准形。

若有另一个 Jordan 形矩阵 $\mathbf{J}'$ 与 **A** 相似，则 $\mathbf{J}'$ 与 **A** 有相同的初等因子。因此， $\mathbf{J}'$ 与 **J** 除去其中 Jordan 块排列的次序外是相同的，这就证明了唯一性。

- 定理3.5.3　设$\mathbf{A} \in \mathbb{C}^{n\times n}$, 则$\mathbf{A}$与一个对角矩阵相似的充分必要条件是$\mathbf{A}$的初等因子都是一次的。
- 例3.5.1　求矩阵

$$\mathbf{A} = \begin{bmatrix} -1 & -2 & 6 \\ -1 & 0 & 3 \\ -1 & -1 & 4 \end{bmatrix}$$

的Jordan标准形。

解：

$$\lambda\mathbf{I} - \mathbf{A} = \begin{bmatrix} \lambda+1 & 2 & -6 \\ 1 & \lambda & -3 \\ 1 & 1 & \lambda-4 \end{bmatrix} \cong \begin{bmatrix} 1 & & \\ & \lambda-1 & \\ & & (\lambda-1)^2 \end{bmatrix}$$

初等因子$\lambda-1, (\lambda-1)^2$，Jordan标准形

$$\mathbf{J} = \begin{bmatrix} 1 & & \\ & 1 & 1 \\ & & 1 \end{bmatrix}$$

- 例3.5.2　对例3.5.1的$\mathbf{A}$，求3阶可逆阵$\mathbf{P}$使得$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{J}$

解：记$\mathbf{P}=[p_1,p_2,p_3]$,则$[\mathbf{A}p_1,\mathbf{A}p_2,\mathbf{A}p_3]=[p_1,p_2,p_3]\begin{bmatrix}1&0&0\\0&1&1\\0&0&1\end{bmatrix}$，比较两边得

$$\mathbf{A}p_1=p_1,\mathbf{A}p_2=p_2,\mathbf{A}p_3=p_2+p_3$$

p_1,p_2是特征值1的两个线性无关特征向量。解$(\mathbf{I}-\mathbf{A})x=0$，得线性无关特征向量

$$\xi=\begin{bmatrix}-1&1&0\end{bmatrix}^{\mathrm{T}},\eta=\begin{bmatrix}3&0&1\end{bmatrix}^{\mathrm{T}}$$

可以取$p_1=\xi$，但若取$p_2=\eta$，下列求p_3的方程组无解

$$(\mathbf{I}-\mathbf{A})p_3=\left(\mathbf{I}-\begin{bmatrix}-1&-2&6\\-1&0&3\\-1&-1&4\end{bmatrix}\right)p_3=\begin{bmatrix}2&2&-6\\1&1&-3\\1&1&-3\end{bmatrix}p_3=\begin{bmatrix}-3\\0\\-1\end{bmatrix}=-p_2$$

p_2的选取应保证$(\mathbf{I}-\mathbf{A})p_2=0$，$p_2$与$p_1$线性无关且$(\mathbf{I}-\mathbf{A})p_3=-p_2$有解，因此

$$p_2=k_1\xi+k_2\eta=\begin{bmatrix}-k_1+3k_2&k_1&k_2\end{bmatrix}^{\mathrm{T}},\ \ k_2\neq 0$$

且方程组

$$\begin{bmatrix}2&2&-6\\1&1&-3\\1&1&-3\end{bmatrix}p_3=\begin{bmatrix}k_1-3k_2\\-k_1\\-k_2\end{bmatrix}$$

有解。容易看出，当$k_1=k_2$时上述方程组有解。取$k_1=k_2=1$，可得

$$p_2=\begin{bmatrix}2&1&1\end{bmatrix}^{\mathrm{T}},p_3=\begin{bmatrix}2&0&1\end{bmatrix}^{\mathrm{T}},\mathbf{P}=\begin{bmatrix}-1&2&2\\1&1&0\\0&1&1\end{bmatrix}$$

- $\mathbf{A} \in \mathbb{C}^{n\times n}$，则存在$n$阶可逆矩阵$\mathbf{P}$使得

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{J} = \begin{bmatrix} \mathbf{J}_1 & & & \\ & \mathbf{J}_2 & & \\ & & \ddots & \\ & & & \mathbf{J}_s \end{bmatrix}$$

其中$\mathbf{J}_i \in \mathbb{C}^{n_i\times n_i}$为Jordan块。记

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_1 & \mathbf{P}_2 & \cdots & P_s \end{bmatrix}$$

其中$\mathbf{P}_i \in \mathbb{C}^{n\times n_i}$，于是

$$\begin{bmatrix} \mathbf{A}\mathbf{P}_1 & \mathbf{A}\mathbf{P}_2 & \cdots & \mathbf{A}\mathbf{P}_s \end{bmatrix} = \begin{bmatrix} \mathbf{P}_1\mathbf{J}_1 & \mathbf{P}_2\mathbf{J}_2 & \cdots & \mathbf{P}_s\mathbf{J}_s \end{bmatrix}$$

比较上式两边得

$$\mathbf{A}\mathbf{P}_i = \mathbf{P}_i\mathbf{J}_i, i \in \{1, 2, \cdots, s\}$$

记$\mathbf{P}_i = \begin{bmatrix} p_1^{(i)}, p_2^{(i)}, \cdots, p_{n_i}^{(i)} \end{bmatrix}$，则上式可写为$\forall i = 1, 2, \cdots, s$

$$\begin{cases} \mathbf{A}p_1^{(i)} = \lambda_i p_1^{(i)} \\ \mathbf{A}p_2^{(i)} = \lambda_i p_2^{(i)} + p_1^{(i)} \\ \cdots\cdots\cdots\cdots \\ \mathbf{A}p_{n_i}^{(i)} = \lambda_i p_{n_i}^{(i)} + p_{n_i-1}^{(i)} \end{cases}$$

$p_1^{(i)}$是矩阵$\mathbf{A}$对应于特征值λ_i的特征向量，且由$p_1^{(i)}$可依次求得$p_2^{(i)}, \cdots, p_{n_i}^{(i)}$.特征向量$p_1^{(i)}$的选取应保证$p_2^{(i)}$可以求出，类似地$p_2^{(i)}$ 的选取（因为$p_2^{(i)}$的选取一般不惟一，只要适当选取一个即可）也应保证$p_3^{(i)}$可以求出，依次类推，并且使$p_1^{(i)}, p_2^{(i)}, \cdots, p_{n_i}^{(i)}$线性无关。

3.6 Caylay-Hamilton定理与最小多项式

- 定理3.6.1(Caylay-Hamilton定理) $\mathbf{A} \in \mathbb{C}^{n\times n}$,

$$f(\lambda) = \det(\lambda\mathbf{I} - \mathbf{A}) = \lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \cdots + a_{n-1}\lambda + a_n$$

是$\mathbf{A}$的特征多项式，则

$$f(\mathbf{A}) = \mathbf{A}^n + a_1\mathbf{A}^{n-1} + a_2\mathbf{A}^{n-2} + \cdots + a_{n-1}\mathbf{A} + a_n\mathbf{I} = 0$$

证明：

$$(\lambda\mathbf{I} - \mathbf{A})^* = \mathbf{C}_1\lambda^{n-1} + \mathbf{C}_2\lambda^{n-2} + \cdots + \mathbf{C}_{n-1}\lambda + \mathbf{C}_n$$

$$(\lambda\mathbf{I} - \mathbf{A})(\lambda\mathbf{I} - \mathbf{A})^* = f(\lambda)\mathbf{I}$$

$$\Downarrow$$

$$(\lambda\mathbf{I} - \mathbf{A})(\mathbf{C}_1\lambda^{n-1} + \mathbf{C}_2\lambda^{n-2} + \cdots + \mathbf{C}_{n-1}\lambda + \mathbf{C}_n) = \mathbf{I}\lambda^n + a_1\mathbf{I}\lambda^{n-1} + \cdots + a_{n-1}\mathbf{I}\lambda + a_n\mathbf{I}$$

$$\Downarrow$$

$$\mathbf{C}_1 = \mathbf{I}, \mathbf{C}_2 - \mathbf{A}\mathbf{C}_1 = a_1\mathbf{I}, \cdots, \mathbf{C}_n - \mathbf{A}\mathbf{C}_{n-1} = a_{n-1}\mathbf{I}, -\mathbf{A}\mathbf{C}_n = a_n\mathbf{I}$$

$$\Downarrow$$

$$\begin{aligned} 0 &= \mathbf{A}^n\mathbf{C}_1 + \mathbf{A}^{n-1}(\mathbf{C}_2 - \mathbf{A}\mathbf{C}_1) + \cdots + \mathbf{A}(\mathbf{C}_n - \mathbf{A}\mathbf{C}_{n-1}) - \mathbf{A}\mathbf{C}_n \\ &= \mathbf{A}^n + a_1\mathbf{A}^{n-1} + \cdots + a_{n-1}\mathbf{A} + a_n\mathbf{I} = f(\mathbf{A}) \end{aligned}$$

- 定义3.6.1　设$\mathbf{A} \in \mathbb{C}^{n\times n}$，如果存在多项式$\varphi(\lambda)$使得$\varphi(\mathbf{A}) = 0$,则称$\varphi(\lambda)$为$\mathbf{A}$的化零多项式
- 定义3.6.2　$\mathbf{A} \in \mathbb{C}^{n\times n}$的所有化零多项式中，次数最低的首1多项式称为$\mathbf{A}$的最小多项式
- 定理3.6.3　相似的矩阵具有相同的最小多项式

  证明：设矩阵 $\mathbf{A}$ 与 $\mathbf{B}$ 相似，则存在非奇异矩阵 $\mathbf{P}$ 使得

$$\begin{aligned} &\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \\ &\mathbf{B}^k = \mathbf{P}^{-1}\mathbf{A}^k\mathbf{P}, \quad k = 1, 2, \cdots \end{aligned}$$

对于任意多项式

$$g(\lambda) = g_0\lambda^m + g_1\lambda^{m-1} + \cdots + g_{m-1}\lambda + g_m$$

恒有

$$\begin{aligned} g(\mathbf{B}) &= g_0\mathbf{B}^m + g_1\mathbf{B}^{m-1} + \cdots + g_{m-1}\mathbf{B} + g_m\mathbf{I} \\ &= g_0\mathbf{P}^{-1}\mathbf{A}^m\mathbf{P} + g_1\mathbf{P}^{-1}\mathbf{A}^{m-1}\mathbf{P} + \cdots + g_{m-1}\mathbf{P}^{-1}\mathbf{A}\mathbf{P} + g_m\mathbf{P}^{-1}\mathbf{P} \\ &= \mathbf{P}^{-1}g(\mathbf{A})\mathbf{P} \end{aligned}$$

可见， $\mathbf{A}$ 与 $\mathbf{B}$ 有相同的化零多项式，从而它们具有相同的最小多项式.

- 定理3.6.4　块对角矩阵$\mathbf{A}=\mathrm{diag}(\mathbf{A}_1,\cdots,\mathbf{A}_s)$的最小多项式等于其诸对角块的最小多项式的最小公倍式

证明：设$\mathbf{A}_i$的最小多项式为$m_i(\lambda)$ $(i=1,\cdots,s)$，由于对任意多项式$\varphi(\lambda)$

$$\varphi(\mathbf{A})=\varphi(\begin{bmatrix}\mathbf{A}_1 & & \\ & \ddots & \\ & & \mathbf{A}_s\end{bmatrix})=\begin{bmatrix}\varphi(\mathbf{A}_1) & & \\ & \ddots & \\ & & \varphi(\mathbf{A}_s)\end{bmatrix}$$

如果$\varphi(\lambda)$为$\mathbf{A}$的化零多项式，则$\varphi(\lambda)$必为$\mathbf{A}_i$ $(i=1,\cdots,s)$的化零多项式，从而$m_i(\lambda)|\varphi(\lambda)$ $(i=1,\cdots,s)$，因此$\varphi(\lambda)$为$m_1(\lambda),\cdots,m_s(\lambda)$的公倍式。反过来，如果$\varphi(\lambda)$为$m_1(\lambda),\cdots,m_s(\lambda)$的任一公倍式，则

$$\varphi(\mathbf{A}_i)=0\ \ (i=1,\cdots,s)$$

从而$\varphi(\mathbf{A})=0$，因此，$\mathbf{A}$的最小多项式为$m_1(\lambda),\cdots,m_s(\lambda)$ 的公倍式中次数最低者，即它们的最小公倍式。

- 定理3.6.5　设 $\mathbf{A} \in \mathbb{C}^{n\times n}$, 则 $\mathbf{A}$ 的最小多项式为 $\mathbf{A}$ 的第 n 个不变因子 $d_n(\lambda)$.

- 例:$\mathbf{A} = \begin{bmatrix} 3 & 1 & -1 \\ 0 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix}$

  $\lambda\mathbf{I} - \mathbf{A} = \begin{bmatrix} \lambda-3 & -1 & 1 \\ 0 & \lambda-2 & 0 \\ -1 & -1 & \lambda-1 \end{bmatrix} \cong \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda-2 & 0 \\ 0 & 0 & (\lambda-2)^2 \end{bmatrix}$

  其最小多项式为第3个不变因子$d_3(\lambda) = (\lambda-2)^2$

- 定理3.6.6　n阶矩阵 $\mathbf{A}$ 相似于对角矩阵的充分必要条件是 $\mathbf{A}$ 的最小多项式 $m(\lambda)$ 没有重零点

Thank you!