第二章 线性算子

§1 不同基底下线性映射的矩阵表示

§2 线性算子代数和矩阵相似

§3 单个算子生成的子环

§4 算子和矩阵的极小多项式

§5 $F[t]$ 中的最小公倍式(复习与加细)

§6 商算子

§7 不可分子空间

§8 特征向量和特征多项式

§9 对角化

§10 循环子空间

§11 广义特征子空间(根子空间)

§12 循环子空间分解

定理 12.9. 设 $\mathcal{A} \in \mathcal{L}(V)$. 则存在 $\mathcal{A}$-不可分子空间 $W_1, \ldots, W_k$ 使得

$$V = W_1 \oplus \cdots \oplus W_k$$

且 W_i 是 $\mathcal{A}$-循环的, $\mu_{\mathcal{A}_{W_i}}$ 是 $F[t]$ 中某个不可约多项式的幂次, $i = 1, 2, \ldots, k$.

证明. 结合引理 12.8 和第二章第三讲定理 7.2 即可. □

例 12.10 计算下列实矩阵对应的线性算子的不可分子空间分解:

$$A = \begin{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} & O \\ O & \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \end{pmatrix}, \quad B = \begin{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} & O \\ O & \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \end{pmatrix}, \quad C = \begin{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} & O \\ O & O \end{pmatrix}_{4\times 4}.$$

直接计算 $\mu_A = t^2(t-1)^2$. 由广义特征子空间分解(极小多项式版)可得

$$\mathbb{R}^4 = V(t) \oplus V(t-1)$$

且 $\mu_{A_{V(t)}} = t^2$ 且 $\mu_{A_{V(t-1)}} = (t-1)^2$.

因为 $\chi_A = t^2(t-1)^2$, 所以 $\chi_{A_{V(t)}} = t^2$ 且 $\chi_{A_{V(t-1)}} = (t-1)^2$ (由广义特征子空间分解(特征多项式版)). 于是, $\chi_{A_{V(t)}} = \mu_{A_{V(t)}}$ 且 $\chi_{A_{V(t-1)}} = \mu_{A_{V(t-1)}}$. 故 $V(t)$ 和 $V(t-1)$ 都是 $\mathcal{A}$ 不可分的.

直接计算 $\mu_B = t^2$. 由广义特征子空间分解(极小多项式版)可得

$$V = V(t)$$

且 V 不是 B-循环的. 设 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4$ 是 $\mathbb{R}^4$ 的一组基. 直接计算得

$$W_1 = F[A] \cdot \mathbf{e}_2 = \langle \mathbf{e}_1, \mathbf{e}_2 \rangle \quad 和 \quad W_2 = F[A] \cdot \mathbf{e}_4 = \langle \mathbf{e}_3, \mathbf{e}_4 \rangle.$$

我们有

$$\mathbb{R}^4 = W_1 \oplus W_2$$

且 $\mu_{\mathcal{B}_{W_1}} = \mu_{\mathcal{B}_{W_2}} = t^2$. 子空间 W_1 和 W_2 是 B-不可分的.

直接计算 $\mu_C = t^2$. 由广义特征子空间分解(极小多项式版)可得

$$V = V(t)$$

且 V 不是 C-循环的. 直接计算得 $W_1 = F[A] \cdot \mathbf{e}_2 = \langle \mathbf{e}_1, \mathbf{e}_2 \rangle$, $W_2 = F[A] \cdot \mathbf{e}_3 = \langle \mathbf{e}_3 \rangle$, $W_4 = F[A] \cdot \mathbf{e}_4 = \langle \mathbf{e}_4 \rangle$. 我们有

$$\mathbb{R}^4 = W_1 \oplus W_2 \oplus W_3$$

且 $\mu_{\mathcal{C}_{W_1}} = t^2$, $\mu_{\mathcal{C}_{W_2}} = \mu_{\mathcal{C}_{W_3}} = t$. 子空间 W_1, W_2 和 W_3 是 C-不可分的. $\square$

例 12.11. 设 $\mathcal{D}$ 是 $R[x]_n$ 上的导数算子. 则 $\mu_{\mathcal{D}} = t^n$ 且 $R[x]_n$ 是 $\mathcal{D}$-循环的(第二章第四讲例 *10.4*). 于是, $R[x]_n$ 是 $\mathcal{D}$-不可分的. $\square$

§13 复数域上的 Jordan 标准型 (存在性)

记号: 除非特别说明, 在本节中 V 是 $\mathbb{C}$ 上的 n 维线性空间.

引理 13.1 设 $\mathcal{A} \in \mathcal{L}(V)$. 则 V 是 $\mathcal{A}$-不可分的当且仅当存在 $\lambda \in \mathbb{C}$ 使得 $\mu_{\mathcal{A}} = (t-\lambda)^n$. 此时, $\mathcal{A}$ 在 $\mathcal{L}(V)$ 的某组基下的矩阵是

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}. \tag{1}$$

证明. 设 $\mu_{\mathcal{A}} = (t-\lambda)^n$. 则 V 是 $\mathcal{A}$-循环的(第二章第四讲定理 10.7). 由第二章第五讲引理 12.8 可知, V 是 $\mathcal{A}$-不可分的. 反之, 设 V 是 $\mathcal{A}$-不可分的. 同样的引理蕴含 $\mu_{\mathcal{A}}$ 是 $\mathbb{C}[t]$ 中某个首一的不可约多项式的幂次. 由代数学基本定理, $\mu_{\mathcal{A}} = (t-\lambda)^m$, 其中 $\lambda \in \mathbb{C}$ 且 $m \in \mathbb{Z}^+$. 该引理还蕴含 V 是 $\mathcal{A}$-循环的. 于是 $m = n$ (第二章第四讲定理 10.7).

设 V 是 $\mathcal{A}$-不可分的, $V = F[\mathcal{A}] \cdot \mathbf{v}$ 且 $\mu_{\mathcal{A}} = (t-\lambda)^n$.

断言. 令 $\epsilon_j = (\mathcal{A} - \lambda\mathcal{E})^{n-j}(\mathbf{v})$, $j = 1, 2, \ldots, n$, 是 V 的一组基.

断言的证明. 只要证明 $\epsilon_1, \ldots, \epsilon_n$ 线性无关即可. 设 $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ 使得

$$\alpha_1\epsilon_1 + \cdots + \alpha_n\epsilon_n = \mathbf{0}.$$

则

$$\alpha_1(\mathcal{A} - \lambda\mathcal{E})^{n-1}(\mathbf{v}) + \cdots + \alpha_{n-1}(\mathcal{A} - \lambda\mathcal{E})(\mathbf{v}) + \alpha_n\mathbf{v} = \mathbf{0}.$$

令 $f(t) = \alpha_1(t-\lambda)^{n-1} + \cdots + \alpha_{n-1}(t-\alpha) + \alpha_0$. 则 $f(\mathcal{A})(\mathbf{v}) = \mathbf{0}$. 设 $\mathbf{x} \in V$. 则存在 $g(t) \in \mathbb{C}[t]$ 使得 $\mathbf{x} = g(\mathcal{A})(\mathbf{v})$ (第二章第四讲命题 10.2 (i)). 于是,

$$f(\mathcal{A})(\mathbf{x}) = f(\mathcal{A})g(\mathcal{A})(\mathbf{v}) = g(\mathcal{A})f(\mathcal{A})(\mathbf{v}) = g(\mathcal{A})(\mathbf{0}) = \mathbf{0},$$

其中 $f(\mathcal{A})g(\mathcal{A}) = g(\mathcal{A})f(\mathcal{A})$ 可由 $F[\mathcal{A}]$ 是交换环得出(见第一学期第四章讲义二命题 3.4). 由 $\mathbf{x}$ 的任意性可知, $f(\mathcal{A}) = \mathcal{O}$. 因为 $\deg(f) < n$, 所以 $f(t) = 0$. 通过分析 f 的次数, 我们得到

$$\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0.$$

断言成立.

下面我们计算 $\mathcal{A}$ 在 $\epsilon_1, \ldots, \epsilon_n$ 下的矩阵. 设 $j = 2, \ldots, n$.

$$\mathcal{A}(\epsilon_j) = (\mathcal{A} - \lambda\mathcal{E} + \lambda\mathcal{E})(\epsilon_j) = (\mathcal{A} - \lambda\mathcal{E})((\mathcal{A} - \lambda\mathcal{E})^{n-j}(\mathbf{v})) + \lambda\mathcal{E}(\epsilon_j) = (\mathcal{A} - \lambda\mathcal{E})^{n-j+1}(\mathbf{v}) + \lambda\epsilon_j = \epsilon_{j-1} + \lambda\epsilon_j.$$

而

$$\mathcal{A}(\epsilon_1) = (\mathcal{A} - \lambda\mathcal{E} + \lambda\mathcal{E})(\epsilon_1) = (\mathcal{A} - \lambda\mathcal{E})((\mathcal{A} - \lambda\mathcal{E})^{n-1}(\mathbf{v})) + \lambda\mathcal{E}(\epsilon_1) = (\mathcal{A} - \lambda\mathcal{E})^{n}(\mathbf{v}) + \lambda\epsilon_1 = \lambda\epsilon_j.$$

于是

$$\mathcal{A}(\epsilon_1,\ldots,\epsilon_n)=(\epsilon_1,\ldots,\epsilon_n)\begin{pmatrix}\lambda & 1 & 0 & \cdots & 0 & 0\\ 0 & \lambda & 1 & \cdots & 0 & 0\\ 0 & 0 & \lambda & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & \lambda & 1\\ 0 & 0 & 0 & \cdots & 0 & \lambda\end{pmatrix}. \quad \square.$$

我们称 (1) 中的矩阵为 关于 λ 的 n 阶 *Jordan* 块. Jordan 块的若干基本性质如下.

注解 13.2 *(i)* 如果 $\lambda \neq 0$, 则 $\operatorname{rank}(J_n(\lambda))=n$. 而 $\operatorname{rank}(J_n(0))=n-1$;

(ii) $J_n(\lambda)=\lambda E_n+J_n(0)$;

(iii) $J_n(\lambda)$ 的极小多项式和特征多项式都等于 $(t-\lambda)^n$; 从而把 $J_n(\lambda)$ 看成 $\mathbb{C}^n$ 上的算子后, $\mathbb{C}^n$ 是 $J_n(\lambda)$-循环的;

(iv) $J_n(\lambda)$ 的唯一的特征值是 λ, 而对应的特征子空间的维数等于 1, 这是因为

$$J_n(\lambda)-\lambda E_n=J_n(0),$$

其秩等于 $n-1$.

(v) $J_n(\lambda)$ 可对角化当且仅当 $n=1$.

定理 13.3 设 $\mathcal{A}\in\mathcal{L}(V)$. 则 $\mathcal{A}$ 在 V 的某组基下的矩阵是

$$\begin{pmatrix}J_{d_1}(\lambda_1) & O & \cdots & O\\ O & J_{d_2}(\lambda_2) & \cdots & O\\ \vdots & \vdots & \ddots & \vdots\\ O & O & \cdots & J_{d_k}(\lambda_k)\end{pmatrix}, \tag{2}$$

其中 $d_1,\ldots,d_k\in\mathbb{Z}^+$, $\lambda_1,\ldots,\lambda_k\in\mathbb{C}$, 都不必两两不同.

证明. 由定理 12.9 可知

$$V=W_1\oplus\cdots\oplus W_k,$$

其中 W_i 是 d_i-维 $\mathcal{A}$-不可分子空间, $i=1,2,\ldots,k$. 由第二章第五讲引理 12.8 和代数学基本定理, $\mathcal{A}_{W_i}$ 的极小多项式是 $(t-\lambda_i)^{d_i}$. 根据引理 13.1, $\mathcal{A}_{W_i}$ 在 W_i 的某组基下的矩阵

是 $J_{d_i}(\lambda_i)$. 再由第二章第二讲定理 6.9, $\mathcal{A}$ 在 V 的某组基下的矩阵是

$$\begin{pmatrix} J_{d_1}(\lambda_1) & O & \cdots & O \\ O & J_{d_2}(\lambda_2) & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & J_{d_k}(\lambda_k) \end{pmatrix}. \quad \square$$

例 13.4 设 $\mathcal{D}:\mathbb{C}[x]_n \longrightarrow \mathbb{C}[x]_n$ 是导数算子. 计算 $\mathcal{D}$ 的 *Jordan* 标准型.

解.由例 *12.11*, $\mathbb{C}[x]_n$ 是 $\mathcal{D}$-不可分的. 于是, $\mathcal{D}$ 的 *Jordan* 标准型是 $J_n(0)$. 注意到

$$\mathbb{C}[x]_n = \mathbb{C}[\mathcal{D}] \cdot x^{n-1}.$$

由引理 *13.1* 可知 $\mathcal{D}$ 在基底

$$\mathcal{D}^{n-j}(x^{n-1}) = \frac{(n-1)!}{j!} x^j, \quad j = 1, 2, \ldots, n$$

下的矩阵是 $J_n(0)$. $\square$

推论 13.5 设 $A \in \mathrm{M}_n(\mathbb{C})$. 则 A 相似于

$$J_A = \begin{pmatrix} J_{d_1}(\lambda_1) & O & \cdots & O \\ O & J_{d_2}(\lambda_2) & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & J_{d_k}(\lambda_k) \end{pmatrix},$$

其中 $d_1, \ldots, d_k \in \mathbb{Z}^+$, $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$, 都不必两两不同.

J_A 中的矩阵称为 $\mathcal{A}$ 的一个 *Jordan* 标准型. J_A 的基本性质如下:

注解 13.6 *(i)* $\mathrm{rank}(J_A) = \sum_{i=1}^{k} \mathrm{rank}(J_{d_i}(\lambda_i))$;

(ii) J_A 的(也是 A 的)极小多项式等于

$$\mathrm{lcm}((t-\lambda_1)^{d_1}, \ldots, (t-\lambda_k)^{d_k});$$

特征多项式等于

$$(t-\lambda_1)^{d_1} \cdots (t-\lambda_k)^{d_k};$$

(iii) 设 $\lambda \in \mathrm{spec}_{\mathbb{C}}(\mathcal{A})$. 则 J_A 中至少有有一个关于 λ 的 *Jordan* 块.

(iv) λ 的代数重数等于 λ 在 J_A 主对角线上出现的次数; λ 的几何重数等于关于 λ 的 *Jordan* 块在 J_A 中出现的次数;

(v) λ 在极小多项式中的重数等于 J_A 中关于 λ 的 *Jordan* 块在的最大阶数.

(vi) A 可对角化当且仅当 $d_1 = \cdots = d_k = 1$.

性质 (i) 来自 J_A 是分块对角矩阵.

性质 (ii) 成立是因为第二章第二讲定理 6.9 和第二章第三讲例 8.15.

性质 (iii) 成立是因为 $\lambda \in \{\lambda_1, \ldots, \lambda_k\}$.

性质 (iv) 中的第一部分可由 (i) 中特征多项式的形式直接得出. 下面来验证性质 (iv) 的第二部分. 注意到

$$J_A - \lambda E_n = \begin{pmatrix} J_{d_1}(\lambda_1) - \lambda E_{d_1} & O & \cdots & O \\ O & J_{d_2}(\lambda_2) - \lambda E_{d_2} & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & J_{d_k}(\lambda_k) - \lambda E_{d_k} \end{pmatrix}.$$

因为

$$\operatorname{rank}(J_{d_i}(\lambda_i) - \lambda_{\lambda_i} E_{d_i}) = \begin{cases} d_i, & \lambda \neq \lambda_i, \\ d_i - 1, & \lambda \neq \lambda_i, \end{cases}$$

$i = 1, 2, \ldots, k$. 于是,

$$\operatorname{rank}(J_A - \lambda E_n) = n - (J_A \text{ 中关于 } \lambda \text{ 的 Jordan 块出现的次数}).$$

由此和矩阵的秩和解空间维数的关系得出 $\dim(V^\lambda)$ 等于 J_A 中关于 λ 的 Jordan 块出现的次数. 于是, (iv) 成立.

性质 (v) 来自于 (i) 中极小多项式的形式.

最后, 我们来验证性质 (vi). 如果 $d_1 = \cdots = d_k = 1$, 则 J_A 可对角化. 于是, A 可对角化. 反之, A 可对角化蕴含 μ_A 中每个因子的重数都等于 1 (对角化判别法 V). 由性质 (v), J_A 是对角阵.

例 13.7 设 $\alpha \in \mathbb{C}$,

$$A = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ \alpha & 0 & \alpha \end{pmatrix}.$$

计算 J_A.

解. 直接计算得 $\chi_A = (t - \alpha)^3$. 于是, α 是 A 唯一的特征根, 其代数重数等于 *3*.

$$\operatorname{rank}(A - \alpha E) = \operatorname{rank} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \alpha & 0 & 0 \end{pmatrix}.$$

由此看出, 当 $\alpha \neq 0$ 时, $\operatorname{rank}(A-\alpha E)=1$. 故 α 的几何重数等于 *2*. 由基本性质 *(iv)*,

$$J_A=\begin{pmatrix} J_2(\alpha) & O \\ O & J_1(\alpha) \end{pmatrix}=\begin{pmatrix} \alpha & 1 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix}.$$

当 $\alpha=0$ 时, $\operatorname{rank}(A-\alpha E)=0$. 故 α 的几何重数等于 *3*. 由基本性质 *(iv)*,

$$J_A=\begin{pmatrix} J_1(0) & 0 & 0 \\ 0 & J_1(0) & 0 \\ 0 & 0 & J_1(0) \end{pmatrix}=O_{3\times 3}. \quad \square$$

例 13.8 设

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

计算 J_A.

解. 直接计算得 $\chi_A=(t-2)(t-1)^2$. 于是, 2 的代数和几何重数都等于 *1*. 故 $J_1(2)$ 在 J_A 中出现一次. 注意到 1 的几何重数等于

$$3-\operatorname{rank}(A-E)=3-\operatorname{rank}\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix}=1.$$

故 $J_2(1)$ 在 J_A 中出现一次. 由此得出

$$J_A=\begin{pmatrix} J_1(2) & O \\ O & J_2(1) \end{pmatrix}=\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}. \quad \square$$

例 13.9 设

$$A=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & -5 & 2 \\ 0 & 5 & -7 & 3 \\ 0 & 6 & -9 & 4 \end{pmatrix}.$$

计算 J_A.

解. 直接计算得 $\chi_A=(t-1)^2t^2$. 注意到 1 的几何重数等于

$$4-\operatorname{rank}(A-E)=4-\operatorname{rank}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 3 & -5 & 2 \\ 0 & 5 & -8 & 3 \\ 0 & 6 & -9 & 3 \end{pmatrix}=2.$$

故 $J_1(1)$ 在 J_A 中出现两次(*1* 的代数重数是 *2*). 0 的几何重数等于

$$4 - \mathrm{rank}(A) = 1.$$

故 $J_2(0)$ 在 J_A 中出现一次.

$$J_A = \begin{pmatrix} J_1(1) & & \\ & J_1(1) & \\ & & J_2(0) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad \square$$

例 13.10 设 $S \in \mathrm{M}_n(\mathbb{C})$ 满足 $S^2 - nS = O$. 求 J_S.
解. 设 $f(t) = t(t-n)$. 则 $f(A) = O$.
情形 1**:** $\mu_S = t$**.** $S = O = J_S$.
情形 2**:** $\mu_S = t - n$**.** $S = nE_n = J_S$.
情形 3**:** $\mu_S = t(t-n)$**.** 因为 $n \neq 0$, 所以 S 可对角化(判别法 *V*). 于是

$$J_S = \begin{pmatrix} nE_r & O \\ O & O \end{pmatrix},$$

其中 $r = \mathrm{rank}(S)$. 当

$$S = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}$$

时, 当

$$J_S = \begin{pmatrix} n & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}. \quad \square$$

关于 $\mathbb{C}$ 上 Jordan 标准型尚不知道的秘密:

(i) 唯一性是否成立?

(ii) 关于特征值 λ 的 Jordan 块由多少个, 每个的阶是多少?

§14 初等因子组

记号: 除非特别说明, 在本节中 V 是 F 上的 n 维线性空间, 其中 F 是任意域.

重集 是指中的元素可以重复出现的集合.

例 14.1 设 $S=\{a,a,b\}$ 和 $T=\{a,b\}$. 它们作为重集不相等. 元素 a 在 S 中的重数等于 2, 在 T 中等于 1.

例 14.2 我们由素分解 $24=2^3 3$. 利用重集表示 *24* 的素因子为 $\{2,2,2,3\}$.

定义 14.3 设 $\mathcal{A}\in\mathcal{L}(V)$,

$$V=U_1\oplus\cdots\oplus U_\ell \tag{3}$$

是 V 的 $\mathcal{A}$-不可分子空间直和分解. 令 $\mathcal{A}_i=\mathcal{A}_{U_i}$ 和 $\mu_i=\mu_{\mathcal{A}_i}$. 重集 $\{\mu_1,\ldots,\mu_\ell\}$ 称为 $\mathcal{A}$ 关于 (3) 的初等因子组.

由上一讲引理 12.8 可知, 初等因子组中的元素都是 $F[t]$ 中(首一)不可约多项式的幂次. 类似地, 我们可以定义矩阵的初等因子组.

例 14.4 设 $\mathcal{E}$ 是 V 上的恒同算子, $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的一组基. 则

$$V=\langle\mathbf{e}_1\rangle\oplus\cdots\oplus\langle\mathbf{e}_n\rangle,$$

是 V 的一个 $\mathcal{E}$-不可分子空间的直和分解. 算子 $\mathcal{E}$ 关于上述直和分解的初等因子组是

$$\{\underbrace{t-1,\ldots,t-1}_{n}\}.$$

例 14.5 设 $\mathcal{D}$ 是 $\mathbb{R}[x]_n$ 上的导数算子. 则 $\mathbb{R}[x]_n$ 是 $\mathcal{D}$-不可分的. 于是, $\mathcal{D}$ 的初等因子组是 $\{t^n\}$.

在本节中, 我们将说明以下结论:

(i) 对于任何 V 的 $\mathcal{A}$-不可分子空间直和分解, 初等因子组相同;

(ii) 当 $F=\mathbb{C}$ 时, 初等因子组唯一确定 Jordan 标准型;

(iii) 初等因子组可以通过计算若干矩阵的秩得到.

引理 14.6 设 $\mathcal{A}\in\mathcal{L}(V)$, V 是 $\mathcal{A}$-循环的. 设 $\mu_{\mathcal{A}}=pq$, 其中 $p,q\in F[t]$ 首一. 再设 $\mathbf{v}$ 是 V 中的 $\mathcal{A}$-循环向量. 令 $\mathbf{w}=q(\mathcal{A})(\mathbf{v})$. 则 $\mu_{\mathcal{A},\mathbf{w}}=p$.

证明. 首先, $p(\mathcal{A})(\mathbf{w}) = p(\mathcal{A})q(\mathcal{A})(\mathbf{v}) = \mu_{\mathcal{A}}(\mathbf{v}) = \mathbf{0}$. 于是, $\mu_{\mathcal{A},\mathbf{w}}(t)|p(t)$. 反之, 因为 $\mu_{\mathcal{A},\mathbf{w}}(\mathbf{w}) = \mathbf{0}$, 所以 $(\mu_{\mathcal{A},\mathbf{w}}q)(\mathcal{A})(\mathbf{v}) = \mathbf{0}$. 对任意 $\mathbf{x} \in V$, 存在 $f \in F[t]$ 使得 $\mathbf{x} = f(\mathcal{A})(\mathbf{v})$. 于是, $(\mu_{\mathcal{A},\mathbf{w}}q)(\mathcal{A})(\mathbf{x}) = f(\mathcal{A})(\mu_{\mathcal{A},\mathbf{w}}q)(\mathcal{A})(\mathbf{v}) = \mathbf{0}$. 于是, $(\mu_{\mathcal{A},\mathbf{w}}q)(\mathcal{A}) = \mathcal{O}$. 由此得出,

$$\mu_{\mathcal{A}}|(\mu_{\mathcal{A},\mathbf{w}}q) \implies (pq)|(\mu_{\mathcal{A},\mathbf{w}}q) \implies p|\mu_{\mathcal{A},\mathbf{w}}.$$

故 $p = \mu_{\mathcal{A},\mathbf{w}}$. □

引理 14.7 设 $\mathcal{A} \in \mathcal{L}(V)$, V 是 $\mathcal{A}$-循环的. 设 $\mu_{\mathcal{A}} = p^m$, 其中 $p \in F[t] \setminus F$. 则

$$\operatorname{rank}(p(\mathcal{A})^k) = \begin{cases} (m-k)\deg(p), & 0 \le k \le m \\ \\ 0, & k > m. \end{cases}$$

证明. 设 $\mathbf{v} \in V$ 是 $\mathcal{A}$-循环向量. 令 $\mathbf{w}_k = p(\mathcal{A})^k(\mathbf{v})$.

断言. 对任意 $k \in \mathbb{Z}^+$, $\operatorname{im}(p^k(\mathcal{A})) = F[\mathcal{A}] \cdot \mathbf{w}_k$.

断言的证明. 设 $\mathbf{x} \in V$. 则存在 $f \in F[t]$ 使得 $\mathbf{x} = f(\mathcal{A})(\mathbf{v})$ (第二章第四讲命题 10.2 (i)). 于是,

$$p^k(\mathcal{A})(\mathbf{x}) = p^k(\mathcal{A})f(\mathcal{A})(\mathbf{v}) = f(\mathcal{A})p(\mathcal{A})^k(\mathbf{v}) = f(\mathcal{A})(\mathbf{w}_k) \in F[\mathcal{A}] \cdot \mathbf{w}_k.$$

由此得出, $\operatorname{im}(p^k(\mathcal{A})) \subset F[\mathcal{A}] \cdot \mathbf{w}_k$. 反之, 由 $\mathbf{w}_k$ 的定义可知, $\mathbf{w}_k \in \operatorname{im}(p^k(\mathcal{A}))$. 因为 $\operatorname{im}(p^k(\mathcal{A}))$ 是 $\mathcal{A}$-不变的(第二章第二讲命题 6.5), 所以 $F[\mathcal{A}] \cdot \mathbf{w}_k \subset \operatorname{im}(p^k(\mathcal{A}))$. 断言成立.

根据第二章第一讲推论 1.14,

$$\operatorname{rank}(p^k(\mathcal{A})) = \dim(\operatorname{im}(p^k(\mathcal{A}))).$$

由上述断言,

$$\operatorname{rank}(p^k(\mathcal{A})) = \dim(F[\mathcal{A}] \cdot \mathbf{w}_k).$$

再由第二章第四讲命题 10.2 (iii),

$$\operatorname{rank}(p^k(\mathcal{A})) = \deg(\mu_{\mathcal{A},\mathbf{w}_k}).$$

由引理 14.6, 当 $0 \le k \le m$ 时, $\mu_{\mathcal{A},\mathbf{w}_k} = p^{m-k}$. 于是,

$$\operatorname{rank}(p^k(\mathcal{A})) = (m-k)\deg(p).$$

当 $k > m$ 时, $\mathbf{w}_k = \mathbf{0}$, 从而 $\mu_{\mathcal{A},\mathbf{w}_k} = 1$. 于是,

$$\operatorname{rank}(p^k(\mathcal{A})) = 0. \quad \square$$

引理 14.8 设 $\mathcal{A} \in \mathcal{L}(V)$, $f \in F[t]$. 设

$$V = U_1 \oplus \cdots \oplus U_\ell, \tag{4}$$

其中 $U_1, \ldots, U_\ell$ 是 $\mathcal{A}$-不变的子空间. 则

$$f(\mathcal{A})(V) = f(\mathcal{A})(U_1) \oplus \cdots \oplus f(\mathcal{A})(U_\ell).$$

证明. 设 $W = f(\mathcal{A})(U_1) + \cdots + f(\mathcal{A})(U_\ell)$. 因为 $U_i \subset V$, 所以 $f(\mathcal{A})(U_i) \subset f(\mathcal{A})(V)$. 于是, $W \subset f(\mathcal{A})(V)$. 反之, 设 $\mathbf{x} \in f(\mathcal{A})(V)$. 则存在 $\mathbf{v} \in V$ 使得 $\mathbf{x} = f(\mathcal{A})(\mathbf{v})$. 由 (4) 可知

$$\mathbf{v} = \mathbf{v}_1 + \cdots + \mathbf{v}_\ell,$$

其中 $\mathbf{v}_i \in U_i, i = 1, 2, \ldots, \ell$. 由此得出

$$\mathbf{x} = f(\mathcal{A})(\mathbf{v}) = f(\mathcal{A})(\mathbf{v}_1) + \cdots + f(\mathcal{A})(\mathbf{v}_\ell).$$

因为 U_i 是 $\mathcal{A}$-不变的, 所以 $f(\mathcal{A})(\mathbf{v}_i) \in U_i$. 故 $\mathbf{x} \in W$. 我们得到 $f(\mathcal{A})(V) = W$.

下面验证: $f(\mathcal{A})(U_1) + \cdots + f(\mathcal{A})(U_\ell)$ 是直和. 设

$$\mathbf{0} = \mathbf{w}_1 + \cdots + \mathbf{w}_\ell,$$

其中 $\mathbf{w}_i \in f(\mathcal{A})(U_i)$. 则 $\mathbf{w}_i \in U_i$. 由 (4) 可知, $\mathbf{w}_i = \mathbf{0}$, $i = 1, 2, \ldots, \ell$ (第一章第二讲命题 4.16). 进而, 同样的命题蕴含 $f(\mathcal{A})(U_1) + \cdots + f(\mathcal{A})(U_\ell)$ 是直和. □

定理 14.9 设 $\mathcal{A} \in \mathcal{L}(V)$, $\mu_{\mathcal{A}} = p^m$, 其中 $p \in F[t] \setminus F$ 不可约. 对任意 $\ell \in \mathbb{Z}^+$, 设 n_ℓ 是 p^ℓ 在 V 的某个 $\mathcal{A}$-不可分子空间分解的初等因子组中 p^ℓ 的重数. 令 $d = \deg(p)$ 和 $r_i = \operatorname{rank}(p^i(\mathcal{A})), i \in \mathbb{N}$. 则

$$n_\ell = \frac{1}{d}(r_{\ell-1} + r_{\ell+1} - 2r_\ell). \tag{5}$$