第十二周讲义从第四页开始

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

定义 9.1 设 $\mathcal{A} \in \mathcal{L}(V)$. 如果 $\mathcal{A}$ 在 V 的某组基下的矩阵是对角矩阵, 则称 $\mathcal{A}$ 是可对角化的. 如果 $A \in \mathrm{M}_n(\mathbb{R})$ 相似于一个对角矩阵, 则称 A 是可对角化的.

定理 9.2 *(可对角化判别法I)* 设 $n = \dim(V)$, $\mathcal{A} \in \mathcal{L}(V)$. 则 $\mathcal{A}$ 可对角化当且仅当 $\mathcal{A}$ 有 n 个线性无关得特征向量.

证明. 设 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 是 $\mathcal{A}$ 的 n 个线性无关的特征向量. 设 $\mathcal{A}(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$, $i = 1, 2, \ldots, n$. 注意到 $\lambda_1, \ldots, \lambda_n \in F$ 不一定两两不同. 此时, $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 是 V 的一组基, 且

$$(\mathcal{A}(\mathbf{v}_1), \mathcal{A}(\mathbf{v}_2), \ldots, \mathcal{A}(\mathbf{v}_n)) = (\lambda_1\mathbf{v}_1, \lambda_2\mathbf{v}_2, \ldots, \lambda_n\mathbf{v}_n) = (\mathbf{v}_1, \ldots, \mathbf{v}_n) \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

反之, 设 $\mathcal{A}$ 在基底 $\epsilon_1, \ldots, \epsilon_n$ 下的矩阵是 $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. 则 $\mathcal{A}(\epsilon_i) = \lambda_i \epsilon_i$. 于是, ϵ_i 是 $\mathcal{A}$ 的特征向量且 $\epsilon_1, \ldots, \mathbf{e}_n$ 线性无关. $\square$

推论 9.3 设 $A \in \mathrm{M}_n(F)$. 则 A 可对角化当且仅当 A 有 n 个线性无关的特征向量. 设 A 有 n 个线性无关的特征向量 $\mathbf{v}_1, \ldots, \mathbf{v}_n$. 令 $P = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$. 则 $P^{-1}AP$ 是对角矩阵.

证明. 把 A 看成 $\mathcal{L}(F^n)$ 上的线性算子满足 $\mathbf{x} \mapsto A\mathbf{x}$. 则由上述定理可知矩阵 A 相似于对角阵当且仅当 A 有 n 个线性无关的特征向量. 此时, P 是从标准基到基底 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 的转换矩阵. 于是, $P^{-1}AP$ 是对角阵. □

注解 9.4 线性算子 $\mathcal{A} \in \mathcal{L}(V)$ 可对角化当且仅当 $\mathcal{A}$ 在 V 任何一组基下的矩阵可对角化.

例 9.5 如上一讲例所示, $\mathbb{R}[x]_n$ 中关于导数算子 $\mathcal{D}$ 的特征向量是 $f \in \mathbb{R} \setminus \{0\}$. 不存在两个线性无关的特征向量. 于是, 当 $n > 1$ 是 $\mathcal{D}$ 不能对角化.

例 9.6 *(科斯特利金第一卷第 72 页)* 设

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \in \mathrm{M}_2(\mathbb{R}).$$

判断 A 是否能对角化. 如果可以, 求 $P \in \mathrm{M}_2(\mathbb{R})$ 使得 $P^{-1}\mathrm{M}_n P$ 是对角矩阵.

解. 计算

$$\chi_A(t) = \det \begin{pmatrix} t & -1 \\ -1 & t-1 \end{pmatrix} = t^2 - t - 1.$$

解方程得

$$\lambda_1 = \frac{1+\sqrt{5}}{2}, \quad \lambda_2 = \frac{1-\sqrt{5}}{2}.$$

下面计算特征向量. 特征值 λ_1 对应得特征向量是方程组

$$\begin{pmatrix} \lambda_1 & -1 \\ -1 & \lambda_1 - 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

的非零解. 因为 $\dim(V^{\lambda_1}) = 1$, 所以取 $(1, \lambda_1)^t$ 即可. 类似取 λ_2 对应的特征向量 $(1, \lambda_2)^2$. 因为 $\lambda_1 \neq \lambda_2$, 所以这两个特征向量线性无关. 则

$$P = \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix}.$$

于是

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

引理 9.7 设 $\mathcal{A} \in \mathcal{L}(V)$, $\lambda_1, \ldots, \lambda_k \in_F (\mathcal{A})$ 两两不同. 则

$$V^{\lambda_1} + \cdots + V^{\lambda_k}$$

是直和.

证明. 对 k 归纳. 当 $k = 1$ 时结论显然成立. 设 $k > 1$ 且当 $k - 1$ 时结论成立.

设 $\mathbf{v}_i \in V^{\lambda_i}$, $i = 1, 2, \ldots, k$, 满足

$$\mathbf{0} = \mathbf{v}_1 + \cdots + \mathbf{v}_k.$$

则

$$\mathbf{0} = \mathcal{A}(\mathbf{v}_1 + \cdots + \mathbf{v}_{k-1} + \mathbf{v}_k) = \mathcal{A}(\mathbf{v}_1) + \cdots + \mathcal{A}(\mathbf{v}_{k-1}) + \mathcal{A}(\mathbf{v}_k) = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_{k-1}\mathbf{v}_{k-1} + \lambda_k \mathbf{v}_k.$$

第一式通乘 λ_k 与第二式相减得

$$\mathbf{0} = (\lambda_k - \lambda_1)\mathbf{v}_1 + \cdots + (\lambda_k - \lambda_{k-1})\mathbf{v}_{k-1}.$$

由归纳假设可知, $(\lambda_k - \lambda_1) = \cdots = (\lambda_k - \lambda_{k-1})\mathbf{v}_{k-1} = \mathbf{0}$. 因为 $(\lambda_k - \lambda_1), \ldots, (\lambda_k - \lambda_{k-1})$ 都非零, 所以 $\mathbf{v}_1 = \cdots = \mathbf{v}_{k-1} = \mathbf{0}$. 进而 $\mathbf{v}_k = \mathbf{0}$. 由第一章第一讲定理 1.12, $V^{\lambda_1} + \cdots + V^{\lambda_k}$ 是直和. □

第十二周讲义从下页开始

第十二周讲义

推论 9.8 设 $\mathcal{A}\in\mathcal{L}(V)$, $n=\dim(V)$. 如果 $\chi_{\mathcal{A}}$ 在 F 中有 n 个不同的根, 则 $\mathcal{A}$ 可对角化. 设 $A\in \mathrm{M}_n(F)$. 如果 χ_A 在 F 中有 n 个不同的根, 则 A 可对角化.

证明. 设 $\lambda_1,\ldots,\lambda_n$ 是 $\mathcal{A}$ 的互不相同的特征根. 任取 $\mathbf{v}_i\in V^{\lambda_i}$, $i=1,2,\ldots,n$. 因为 $V^{\lambda_1}+\cdots+V^{\lambda_n}$ 是直和(引理 9.7), 所以 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 线性无关(第一章第一讲定理 1.12 (ii)). 于是, 特征向量 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 是 V 的一组基. 由定理 9.2, $\mathcal{A}$ 可对角化.

对于矩阵情形, 把 A 看成 $\mathcal{L}(F^n)$ 上的线性算子满足 $\mathbf{x}\mapsto A\mathbf{x}$ 即可. □

例 9.9 设

$$A=\begin{pmatrix}1&0&1\\0&1&0\\-1&0&1\end{pmatrix}\in \mathrm{M}_3(\mathbb{C}).$$

判断 A 是否可以对角化.

解. 计算得 $\chi_A(t)=(t-1)(t^2-2t+2)$. 其导数是 $(t^2-2t+2)+2(t-1)^2$. 它们是互素的. 所以 $\chi_A(t)$ 在 $\mathbb{C}$ 中由三个互不相同的根*(代数学基本定理和第零章推论 5.2)*. 由上述推论, A 可以对角化. □

定理 9.10 *(可对角化判别法II)* 设 $\mathcal{A}\in\mathcal{L}(V)$ 且 $\mathrm{spec}_F(\mathcal{A})=\{\lambda_1,\cdots,\lambda_k\}$. 则 $\mathcal{A}$ 可对角化当且仅当 $V=V^{\lambda_1}+\cdots+V^{\lambda_k}$.

证明. 设 $\mathcal{A}$ 可对角化. 则存在特征向量 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 构成 V 的一组基(定理 9.2). 因为 $\mathbf{v}_1,\ldots,\mathbf{v}_n\in V^{\lambda_1}+\cdots+V^{\lambda_k}$, 所以 $V=\langle\mathbf{v}_1,\ldots,\mathbf{v}_n\rangle\subset V^{\lambda_1}+\cdots+V^{\lambda_k}$. 于是, $V=V^{\lambda_1}+\cdots+V^{\lambda_k}$.

反之, 我们有 $V=V^{\lambda_1}\oplus\cdots\oplus V^{\lambda_k}$ (引理 9.7). 设 $\mathbf{e}_{i,1},\ldots\mathbf{e}_{i,d_i}$ 是 V^{λ_i} 的一组基, $i=1,2,\ldots,k$. 基底中的元素都是特征向量. 由直和分解可知, $\mathbf{e}_{1,1},\ldots,\mathbf{e}_{1,d_1},\ldots,\mathbf{e}_{k,1},\ldots,\mathbf{e}_{k,d_k}$ 是 V 的一组基. 由定理 9.2, $\mathcal{A}$ 可对角化. □

例 9.11 设 $\mathcal{A}\in\mathcal{L}(V)$ 可对角化, V 的一组基底是 $\mathbf{e}_{1,1},\ldots,\mathbf{e}_{1,d_1},\ldots,\mathbf{e}_{k,1},\ldots,\mathbf{e}_{k,d_k}$ 如上述证明中给出. 则 $\mathcal{A}$ 在该基底下的矩阵是

$$\begin{pmatrix}\lambda_1E_{d_1}&O&\cdots&O\\O&\lambda_2E_{d_2}&\cdots&O\\\vdots&\vdots&\ddots&\vdots\\O&O&\cdots&\lambda_kE_{d_k}\end{pmatrix}.$$

推论 9.12 设 $A \in \mathrm{M}_n(F)$ 且 $\mathrm{spec}_F(A) = \{\lambda_1, \cdots, \lambda_k\}$. 则 A 可对角化当且仅当

$$F^n = V^{\lambda_1} + \cdots + V^{\lambda_k}.$$

证明. 把 A 看成 F^n 上在标准基下矩阵为 A 的线性算子即可. □

定理 9.13 *(可对角化判别法III)* 设 $\mathcal{A} \in \mathcal{L}(V)$ 且 $\mathrm{spec}_F(\mathcal{A}) = \{\lambda_1, \cdots, \lambda_k\}$. 则 $\mathcal{A}$ 可对角化当且仅当 $\dim(V^{\lambda_1}) + \cdots + \dim(V^{\lambda_k}) = \dim(V)$.

证明. 由引理 9.7 可知, $\dim(V_1^{\lambda} + \cdots + V_k^{\lambda}) = \dim(V^{\lambda_1}) + \cdots + \dim(V^{\lambda_k})$. 于是

$$\dim(V^{\lambda_1} + \cdots + V^{\lambda_k}) = \dim(V) \implies V_1 + \cdots + V_k = V.$$

由定理 9.10, $\mathcal{A}$ 可对角化当且仅当 $\dim(V^{\lambda_1}) + \cdots + \dim(V^{\lambda_k}) = \dim(V)$. □

推论 9.14 设 $A \in \mathrm{M}_n(F)$ 且 $\mathrm{spec}_F(A) = \{\lambda_1, \cdots, \lambda_k\}$. 则 A 可对角化当且仅当

$$\dim(V^{\lambda_1}) + \cdots + \dim(V^{\lambda_k}) = n.$$

证明. 把 A 看成 F^n 上在标准基下矩阵为 A 的线性算子即可. □

例 9.15 设 $A \in \mathrm{M}_n(F)$ 是非零的幂零矩阵. 证明 A 不可对角化.
证明. 设 $k \in \mathbb{Z}^+$ 使得 $A^k = O$. 假设 $\lambda \in (F \setminus \{0\})$ 是 A 的特征根且其对应的特征向量是 $\mathbf{y}$. 则

$$A^k(\mathbf{y}) = A^{k-1}A\mathbf{y} = A^{k-1}(\lambda\mathbf{y}) = \lambda A^{k-1}\mathbf{y} = \lambda^2 A^{k-2}\mathbf{y} = \cdots = \lambda^k \mathbf{y}.$$

因为 $A^k = O$, 所以 $\lambda^k \mathbf{y} = \mathbf{0}$. 矛盾. 由此得出, $\mathrm{spec}_F(A) = \{0\}$. 假设 A 可对角化. 则 $\dim(V^0) = n$ *(定理 9.13.* 于是, $\mathrm{rank}(A) = 0$, 即 $A = O$. 矛盾. □

定义 9.16 设 $\mathcal{A} \in \mathcal{L}(V)$, $\lambda \in \mathrm{spec}_F(\mathcal{A})$. 特征根 λ 在 $\chi_{\mathcal{A}}(t)$ 中的重数称为 λ 的代数重数. 特征子空间 V^{λ} 的维数称为 λ 的几何重数. 类似地, 我们可以定义矩阵特征根的代数和几何重数.

引理 9.17 设 $\mathcal{A} \in \mathcal{L}(V)$, $\lambda \in \mathrm{spec}_F(\mathcal{A})$. 则 λ 的代数重数不低于它的几何重数. 对矩阵也有类似的结论.

证明. 设 d 是 λ 的几何重数. 则 V^{λ} 是 $\mathcal{A}$ 的 d-维不变子空间(第二章第三讲第四页第一段验证了特征子空间是 $\mathcal{A}$ 不变的). 于是, 在 V 的某组基下 $\mathcal{A}$ 的矩阵是

$$\begin{pmatrix} B & * \\ O & C \end{pmatrix},$$

其中 B 是 $\mathcal{A}_{V^\lambda}$ 在 V^λ 某组基下的矩阵(第二章第二讲命题 6.3). 因为 $\mathcal{A}_{V^\lambda}=\lambda\mathcal{E}_d$, 所以 $B=\lambda E_d$. 于是,

$$\chi_{\mathcal{A}}(t)=\chi_B(t)\chi_C(t)=(t-\lambda)^d\chi_C(t)$$

(见第二章第三讲例 8.15). 我们有 $(t-\lambda)^d|\chi_{\mathcal{A}}(t)$. 而 λ 得代数重数是最大的整数 m 使得 $(t-\lambda)^m|\chi_{\mathcal{A}}(t)$. 故 $d\le m$. □

定理 9.18 *(可对角化判别法IV) 设 $\mathcal{A}\in\mathcal{L}(V)$. 则$\mathcal{A}$可对角化当且仅当以下两个条件成立*

(i) $\chi_{\mathcal{A}}(t)$ 在 $F[t]$ 中可以分解为一次因子之积, 即 $\chi_{\mathcal{A}}(t)$ 的所有根都在 F 中;

(ii) 对任意 $\lambda\in\mathrm{spec}_F(\mathcal{A})$, λ 得几何重数等于它的代数重数.

证明. 设 $\mathrm{spec}_F(A)=\{\lambda_1,\cdots,\lambda_k\}$.

设 $\mathcal{A}$ 可对角化. 由例 9.11, $\mathcal{A}$ 在某组基下得矩阵是

$$\begin{pmatrix}\lambda_1E_{d_1} & O & \cdots & O\\ O & \lambda_2E_{d_2} & \cdots & O\\ \vdots & \vdots & \ddots & \vdots\\ O & O & \cdots & \lambda_kE_{d_k}\end{pmatrix},$$

其中 d_i 是 λ_i 的几何重数, $i=1,2,\ldots,k$. 于是, $\chi_{\mathcal{A}}(t)=(t-\lambda_1)^{d_1}(t-\lambda_2)^{d_2}\cdots(t-\lambda_k)^{d_k}$. 条件 (i) 和 (ii) 都成立.

反之, 设条件 (i) 和 (ii) 成立. 则 $\chi_{\mathcal{A}}(t)=(t-\lambda_1)^{d_1}(t-\lambda_2)^{d_2}\cdots(t-\lambda_k)^{d_k}$, 其中 d_i 是 λ_i 的几何重数, $i=1,2,\ldots,k$. 于是, $d_1+\cdots+d_k=\deg(\chi_{\mathcal{A}})=\dim(V)$. 根据定理 9.13, $\mathcal{A}$ 可对角化. □

推论 9.19 *设 $A\in\mathrm{M}_n(F)$. 则 A 可对角化当且仅当以下两个条件成立*

(i) $\chi_A(t)$ 在 $F[t]$ 中可以分解为一次因子之积, 即 $\chi_A(t)$ 得所有根都在 F 中;

(ii) 对任意 $\lambda\in\mathrm{spec}_F(A)$, λ 得几何重数等于它的代数重数.

证明. 把 A 看成 F^n 上在标准基下矩阵为 A 的线性算子即可. □

例 9.20 设

$$A=\begin{pmatrix}a & * & \cdots & *\\ 0 & a & \cdots & *\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & a\end{pmatrix}\in\mathrm{M}_n(F),$$

其中 $*$ 部分不全为零. 证明 A 不可对角化.

证明. 因为 $\chi_A(t)=(t-a)^n$, 所以 $\chi_A(t)$ 在 $F[t]$ 中可以分解为一次因子之积且 a 是 A 的唯一特征根. 假设 A 可对角化. 则 a 的几何重数等于 n, 即 $\dim(V^a)=n$(定理 9.18). 于是,

$$aE-A=\begin{pmatrix}0 & -* & \cdots & -* \\ 0 & 0 & \cdots & -* \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0\end{pmatrix}$$

对应的齐次线性方程组的解空间维数等于 n. 我们有 $\operatorname{rank}(aE-A)=0$, 即 $*$ 部分全为零. 矛盾. □

定理 9.21 *(可对角化判别法V)* 设 $\mathcal{A}\in\mathcal{L}(V)$. 则 $\mathcal{A}$ 可对角化当且仅当 $\mu_{\mathcal{A}}(t)$ 在 $F[t]$ 中可以分解为两两互素一次因子之积.

证明. 设 $\mathcal{A}$ 在 V 的某组基下的矩阵是 $A=\operatorname{diag}(\lambda_1,\ldots,\lambda_n)$. 则

$$\mu_{\mathcal{A}}(t)=\mu_A(t)=\operatorname{lcm}(t-\lambda_1,\ldots,t-\lambda_n)=\prod_{\alpha\in\{\lambda_1,\ldots,\lambda_n\}}(t-\alpha).$$

(见第二章第二讲定理 6.9). 于是, $\mu_{\mathcal{A}}(t)$ 在 $F[t]$ 中可以分解为两两互素一次因子之积.

反之, 设 $\mu_{\mathcal{A}}(t)=(t-\beta_1)\cdots(t-\beta_k)$, 其中 $\beta_1,\ldots,\beta_k\in F$ 两两不同. 则 $t-\beta_i,t-\beta_j$ 互素, $i\neq j$. 由扩展的核核分解定理(第二章第二讲习题课讲义定理 4.3)

$$V=U_1\oplus\cdots\oplus U_k,$$

其中 $U_i=\ker(\mathcal{A}-\beta_i\mathcal{E})$, $i=1,\ldots,k$. 根据第二章第二讲命题 6.5, U_i 是 $\mathcal{A}$- 不变的. 由 U_i 的定义可知，对任意 $\mathbf{x}\in U_i$, $\mathcal{A}(\mathbf{x})=\beta_i\mathbf{x}$, 即限制算子 $\mathcal{A}_{U_i}$ 是的数乘算子 $\beta_i\mathcal{E}_{d_i}$, 其中 $d_i=\dim(U_i)$. 由第二章第二讲定理 6.9, $\mathcal{A}$ 在 V 的某组基下的矩阵是

$$\begin{pmatrix}\beta_1E_{d_1} & O & \cdots & O \\ O & \beta_2E_{d_2} & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & \beta_kE_{d_k}\end{pmatrix}. \quad \square$$

推论 9.22 *(可对角化判别法V)* 设 $A\in\mathrm{M}_n(F)$. 则 A 可对角化当且仅当 $\mu_A(t)$ 在 $F[t]$ 中可以分解为两两互素一次因子之积.

用上述定理推论考虑例 9.15. 因为 A 是非零的幂零矩阵, 所以 $\mu_A=t^k$ 且 $k>1$. 于是, A 不能对角化.

例 9.23 设 F 的特征不等于 *2*. 证明: V 上的对合算子 $\mathcal{A}$, 即满足 $\mathcal{A}^2=\mathcal{E}$, 是可对角化.

证明. 设 $f(t)=t^2-1$. 则 $f(\mathcal{A})=\mathcal{O}$. 由第二章第二讲引理 4.2, $\mu_{\mathcal{A}}(t)|f(t)$. 于是, $\mu_{\mathcal{A}}(t)=t-1$ 或 $\mu_{\mathcal{A}}(t)=t+1$ 或 $\mu_{\mathcal{A}}(t)=(t-1)(t+1)$. 它的不可约因子都是一次的且两两互素. 于是, A 可对角化. □

注解 9.24 设 F 的特征等于 2. 则对合算子可对角化当且仅当 $\mathcal{A}=\mathcal{E}$. 这是因为 $\mathcal{A}\neq\mathcal{E}$ 当且仅当 $\mu_{\mathcal{A}}=(t-1)^2$.

注解 9.25 当算子或矩阵是通过多项式关系给出时, 第五个判别法比较容易应用.

§10 循环子空间

定义 10.1 设 $\mathcal{A}\in\mathcal{L}(V)$ 和 $\mathbf{v}\in V$. 由

$$\mathbf{v},\mathcal{A}(\mathbf{v}),\mathcal{A}^2(\mathbf{v}),\ldots$$

生成的子空间称为由 $\mathcal{A}$ 和 $\mathbf{v}$ 生成的循环子空间. 记为 $F[\mathcal{A}]\cdot\mathbf{v}$

命题 10.2 设 $\mathcal{A}\in\mathcal{L}(V)$ 和 $\mathbf{v}\in V$.

(i) $F[\mathcal{A}]\cdot\mathbf{v}=\{p(\mathcal{A})(\mathbf{v})|p(t)\in F[t]\}$;

(ii) $F[\mathcal{A}]\cdot\mathbf{v}$ 是 $\mathcal{A}$-不变的;

(iii) 设 $d=\deg(\mu_{\mathcal{A},\mathbf{v}})$. 则 $\mathbf{v},\mathcal{A}(\mathbf{v}),\ldots,\mathcal{A}^{d-1}(\mathbf{v})$ 是 $F[\mathcal{A}]\cdot\mathbf{v}$ 的基. 特别地, $d=\dim(F[\mathcal{A}]\cdot\mathbf{v})$.

证明. (i) 设 $p=p_kt^k+p_{k-1}t^{k-1}+\cdots+p_0$, 其中 $p_k,p_{k-1},\ldots,p_0\in F$. 则

$$p(\mathcal{A})(\mathbf{v})=(p_k\mathcal{A}^k+p_{k-1}\mathcal{A}^{k-1}+\cdots+p_0\mathcal{E})(\mathbf{v})=p_k\mathcal{A}^k(\mathbf{v})+p_{k-1}\mathcal{A}^{k-1}(\mathbf{v})+\cdots+p_0\mathbf{v}\in F[\mathcal{A}]\cdot\mathbf{v}.$$

反之, 设 $\mathbf{w}\in F[\mathcal{A}]\cdot\mathbf{v}$. 则存在 $\ell\in\mathbb{N}$, $\alpha_0,\alpha_1,\ldots,\alpha_\ell\in F$ 使得 $\mathbf{w}=\sum_{i=0}^{\ell}\alpha_i\mathcal{A}^i(\mathbf{v})$. 令 $q(t)=\sum_{i=0}^{\ell}\alpha_it^i$. 则 $\mathbf{w}=q(\mathcal{A})(\mathbf{v})\in\{p(\mathcal{A})(\mathbf{v})|p(t)\in F[t]\}$.

(ii) 设 $\mathbf{w}\in F[\mathcal{A}]\cdot\mathbf{v}$. 根据 (i), 存在 $p\in F[t]$ 使得 $\mathbf{w}=p(\mathcal{A})(\mathbf{v})$. 于是,

$$\mathcal{A}(\mathbf{w})=\mathcal{A}p(\mathcal{A})(\mathbf{v}).$$

令 $q=tp$. 则 $\mathcal{A}(\mathbf{w})=q(\mathcal{A})(\mathbf{v})\in F[\mathcal{A}]\cdot\mathbf{v}$.

(iii) 设 $\mathbf{w}\in F[\mathcal{A}]\cdot\mathbf{v}$. 根据 (i), 存在 $p\in F[t]$ 使得 $\mathbf{w}=p(\mathcal{A})(\mathbf{v})$. 由多项式带余除法, 存在 $q,r\in F[t]$ 满足 $\deg(r)<d$ 和

$$p(t)=q(t)\mu_{\mathcal{A},\mathbf{v}}(t)+r(t)\implies p(\mathcal{A})=q(\mathcal{A})\mu_{\mathcal{A},\mathbf{v}}(\mathcal{A})+r(\mathcal{A}).$$

作用到 $\mathbf{v}$ 上得

$$\mathbf{w} = p(\mathcal{A})(\mathbf{v}) = q(\mathcal{A})\mu_{\mathcal{A},\mathbf{v}}(\mathcal{A})(\mathbf{v}) + r(\mathcal{A})(\mathbf{v}) = r(\mathcal{A})(\mathbf{v}).$$

因为 $\deg(r) < d$, 所以 $\mathbf{w}$ 是 $\mathbf{v}, \mathcal{A}(\mathbf{v}), \ldots, \mathcal{A}^{d-1}(\mathbf{v})$ 在 F 上的线性组合. 于是,

$$F[\mathcal{A}] \cdot \mathbf{v} = \langle \mathbf{v}, \mathcal{A}(\mathbf{v}), \ldots, \mathcal{A}^{d-1}(\mathbf{v}) \rangle.$$

再设 $\alpha_0, \alpha_1, \ldots, \alpha_{d-1} \in F$ 使得 $\sum_{i=0}^{d-1} \alpha_i \mathcal{A}^i(\mathbf{v}) = \mathbf{0}$. 令 $f = \alpha_{d-1}t^{d-1} + \cdots + \alpha_1 t + \alpha_0$. 则 $f(\mathcal{A})(\mathbf{v}) = \mathbf{0}$. 因为 $\deg f < d$, 所以 $f(t) = 0$, 即 $\alpha_{d-1} = \cdots = \alpha_1 = \alpha_0 = 0$. 于是, $\mathbf{v}, \mathcal{A}(\mathbf{v}), \ldots, \mathcal{A}^{d-1}(\mathbf{v})$ 线性无关. □

定义 10.3 *设 $\mathcal{A} \in \mathcal{L}(V)$, $\mathbf{v} \in V$. 如果 $V = F[\mathcal{A}] \cdot \mathbf{v}$, 则称 $\mathcal{A}$ 是 V 上的循环算子, $\mathbf{v}$ 是 V 中的循环向量, V 是关于 $\mathcal{A}$ 和 $\mathbf{v}$ 的循环空间. 简称 $\mathcal{A}$-循环空间.*

例 10.4 *考虑导数算子 $\mathcal{D} \in \mathcal{L}(\mathbb{R}[x]_n)$. 则 $\mathcal{D}^i(x^{n-1}) = \alpha_i x^{n-1}$, 其中 $\alpha_i \in (F \setminus \{0\})$, $i = 0, 1, \ldots, n-1$. 于是, $x^{n-1}, \mathcal{D}(x^{n-1}), \ldots, \mathcal{D}^{n-1}(x^{n-1})$ 是 $\mathbb{R}[x]_n$ 的一组基. 由此得出 $\mathbb{R}[x]_n$ 是关于 $\mathcal{D}$-循环空间.*

例 10.5 *设*

$$\begin{pmatrix} J_2 & O_{2\times 3} \\ O_{3\times 2} & J_3 \end{pmatrix}, \quad \text{其中 } J_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, J_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

把 A 看成 F^5 上的线性算子. 判定 F^5 是不是 A-循环的.

解. *直接计算得 $\mu_{J_2} = t^2$ 和 $\mu_{J_3} = t^3$. 则 $\mu_A = \mathrm{lcm}(t^2, t^3) = t^3$ (第二章第二讲引理 6.7). 于是, 对于任意 $\mathbf{v} \in V$, $\deg(\mu_{A,\mathbf{v}}) \le 3$. (第二章第二次习题课第四节中第一段). 根据命题 10.2 (iii) 可知, 对任意 $\mathbf{v} \in V$, $\dim(F[A] \cdot \mathbf{v}) \le 3$. 故 F^5 不是 A-循环的.* □

推论 10.6 *设 $\mathcal{A} \in \mathcal{L}(V)$. 则 $\deg \mu_{\mathcal{A}} \le \dim(V)$*

证明. 设 $n = \dim(V)$. 假设 $\deg(\mu_{\mathcal{A}}) > n$. 则存在 $\mathbf{v} \in V$ 使得 $\mu_{\mathcal{A},\mathbf{v}}$ 的次数等于 $\deg(\mu_{\mathcal{A}}) > n$ (第二章第二次习题课命题 4.1). 由命题 10.2 (iii) 可知, 子空间 $F[\mathcal{A}] \cdot \mathbf{v}$ 的维数大于 n. 矛盾. □

定理 10.7 *设 $\mathcal{A} \in \mathcal{L}(V)$ 和 $\mathbf{v} \in V$. 则 V 是 $\mathcal{A}$-循环的当且仅当 $\mu_{\mathcal{A}}$ 的次数是 $\dim(V)$.*

证明. 设 $n = \dim(V)$. 如果 V 是 $\mathcal{A}$-循环的, 则 $V = F[\mathcal{A}] \cdot \mathbf{v}$. 由命题 10.2 (iii) 可知, $\mu_{\mathcal{A},\mathbf{v}}$ 的次数等于 n. 由第二章第二次习题课第四节中第一段可知, $\mu_{\mathcal{A},\mathbf{v}} | \mu_{\mathcal{A}}$. 于是, 推论 10.6 蕴含 $\mu_{\mathcal{A}} = n$. 反之, 设 $\deg(\mu_{\mathcal{A}}) = n$. 再由第二章第二次习题课命题 4.1 可知, 存在

$\mathbf{v} \in V$ 使得 $\mu_{\mathcal{A},\mathbf{v}}$ 的次数等于 n. 由命题 10.2 (iii) 可知, $F[\mathcal{A}] \cdot \mathbf{v}$ 的维数等于 n. 我们得到 $V = F[\mathcal{A}] \cdot \mathbf{v}$. □

我们用上述命题重新考察例 10.4 和 10.5. 因为 $\mu_{\mathcal{D}} = t^n$, 所以 $\mathbb{R}[x]_n$ 是 $\mathcal{D}$-循环的. 因为 $\mu_A = t^3$, 所以 F^5 不是 A-循环的.

例 10.8 设 $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \in \mathrm{M}_n(F)$. 把 A 看成 F^n 上的线性算子, 证明: F^n 是 A-循环的当且仅当 $\lambda_1, \ldots, \lambda_n$ 两两不同. 当 F^n 是 A-循环时, 计算一个循环向量.

证明. 由第二章第二讲定理定理 *6.9* 可知, $\mu_A(t)=\mathrm{lcm}(t-\lambda_1, \ldots, t-\lambda_n)$. 于是, $\deg(\mu_A)=n$ 当且仅当 $\lambda_1, \ldots, \lambda_n$ 两两不同. 根据定理 *10.7*, F^n 是 A-循环的当且仅当 $\lambda_1, \ldots, \lambda_n$ 两两不同.

设 $\lambda_1, \ldots, \lambda_n$ 两两不同. 令 $\mathbf{v} = \mathbf{e}_1 + \cdots + \mathbf{e}_n$, 其中 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 F^n 的标准基. 因为 $A\mathbf{e}_j = \lambda_j \mathbf{e}_j$, 所以对任意 $k \in \mathbb{N}$, $A^k(\mathbf{e}_j) = \lambda^k \mathbf{e}_j$, $j = 1, 2, \ldots, n$. 于是,

$$(\mathbf{v}, A\mathbf{v}, \ldots, A^{n-1}\mathbf{v}) = (\mathbf{e}_1, \ldots, \mathbf{e}_n) \underbrace{\begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_n & \lambda_n^2 & \cdots & \lambda_n^{n-1} \end{pmatrix}}_{P}.$$

因为 P 可逆, 所以 $\mathbf{v}, A\mathbf{v}, \ldots, A^{n-1}\mathbf{v}$ 是 F^n 的一组基. 于是, $\mathbf{v}$ 是循环向量. □

注解 10.9 上面的例子表明: 如果 $\mathcal{A} \in \mathcal{L}(V)$ 可对角化, 则 V 是 $\mathcal{A}$-循环的当且仅当 $\mathcal{A}$ 的特征根两两不同.

引理 10.10 设 $\mathcal{A} \in \mathcal{L}(V)$ 且 V 是 $\mathcal{A}$-循环的. 则 $\mu_{\mathcal{A}} = \chi_A$.

证明. 设 $V = F[\mathcal{A}] \cdot \mathbf{v}$. 由定理 10.7 可知,

$$\mu_{\mathcal{A}} = t^n + f_{n-1}t^{n-1} + \cdots + f_1 t + f_0,$$

其中 $f_{n-1}, \ldots, f_1, f_0 \in F$. 由命题 10.2 (iii) 可知, $\mathcal{A}^{n-1}(\mathbf{v}), \ldots, \mathcal{A}(\mathbf{v}), \mathbf{v}$ 是 V 的一组基. 注意到 $\mu_{\mathcal{A}}(\mathcal{A})(\mathbf{v}) = 0$. 于是,

$$\mathcal{A}^n(\mathbf{v}) = -f_{n-1}\mathcal{A}^{n-1}(\mathbf{v}) - \cdots - f_1\mathcal{A}(\mathbf{v}) - f_0\mathbf{v}.$$

由此得出

$$(\mathcal{A}^n(\mathbf{v}), \mathcal{A}^{n-1}(\mathbf{v}), \ldots, \mathcal{A}^2(\mathbf{v}), \mathcal{A}(\mathbf{v})) = (\mathcal{A}^{n-1}(\mathbf{v}), \mathcal{A}^{n-2}(\mathbf{v}), \ldots, \mathcal{A}(\mathbf{v}), \mathbf{v}) \underbrace{\begin{pmatrix} -f_{n-1} & 1 & 0 & \cdots & 0 \\ -f_{n-2} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -f_1 & 0 & 0 & \cdots & 1 \\ -f_0 & 0 & 0 & \cdots & 0 \end{pmatrix}}_{A}.$$

即 A 是 $\mathcal{A}$ 在基底 $\mathcal{A}^{n-1}(\mathbf{v}),\ldots,\mathcal{A}(\mathbf{v}),\mathbf{v}$ 下的矩阵. 我们得到 $\chi_{\mathcal{A}}(t)=\det(tE-A)=\mu_{\mathcal{A}}$(见第二章第三次习题课例 2.3). □

定理 10.11 *(Hamilton-Cayley)* 设 $\mathcal{A}\in\mathcal{L}(V)$. 则 $\chi_{\mathcal{A}}(\mathcal{A})=\mathcal{O}$, 即 $\mu_{\mathcal{A}}|\chi_{\mathcal{A}}$.

证明. 任取非零向量 $\mathbf{u}\in V$. 设 $U=F[\mathcal{A}]\cdot\mathbf{u}$ 和 $d=\dim(U)$. 因为 U 是 $\mathcal{A}$ 子空间(命题 10.2 (ii)), 所以存在 V 的一组基使得 $\mathcal{A}$ 在该基下的矩阵是

$$A=\begin{pmatrix} B & C \\ O & D\end{pmatrix},$$

其中 $B\in\mathrm{M}_d(F)$ 是 $\mathcal{A}_U$ 的某个矩阵表示. 因为 U 是 A_U-循环的, 所以 $\mu_{A_U}=\chi_{A_U}$ (引理 10.10). 于是 $\chi_{\mathcal{A}_U}(\mathcal{A}_U)=\mathcal{O}_d$. 特别地, $\chi_{\mathcal{A}_U}(\mathcal{A}_U)(\mathbf{u})=\mathbf{0}$. 从而, $\chi_{\mathcal{A}_U}(\mathcal{A})(\mathbf{u})=\mathbf{0}$.

由第二章第三次讲义例 8.15, 我们有 $\chi_A(t)=\chi_B(t)\chi_D(t)$, 即 $\chi_{\mathcal{A}}(t)=\chi_{\mathcal{A}_U}(t)\chi_D(t)$. 于是,

$$\chi_{\mathcal{A}}(\mathcal{A})(\mathbf{u})=(\chi_{\mathcal{A}_U}(\mathcal{A})\chi_D(\mathcal{A}))(\mathbf{u})=\chi_D(\mathcal{A})(\chi_{\mathcal{A}_U}(\mathcal{A})(\mathbf{u}))=\chi_D(\mathcal{A})(\underbrace{\mu_{\mathcal{A}_U}(\mathcal{A})(\mathbf{u})}_{\mathbf{0}})=\mathbf{0}.$$

由 $\mathbf{u}$ 的任意性可知 $\chi_{\mathcal{A}}(\mathcal{A})=\mathcal{O}$. □

推论 10.12 设 $\mathcal{A}\in\mathcal{L}(V)$. 证明*:* $\mathrm{spec}_F(\mathcal{A})$ 与 $\mu_{\mathcal{A}}$ 在 F 中的根的集合相同.

证明. 设 S 是 $\mu_{\mathcal{A}}$ 在 F 中的根的集合. 由 Hamilton-Cayley 定理, $S\subset\mathrm{spec}_F(\mathcal{A})$. 设 $\lambda\in\mathrm{spec}_F(\mathcal{A})$. 则 $\mu_{\mathcal{A}}(\lambda)$ 是 $\mu_{\mathcal{A}}(\mathcal{A})$=$\mathcal{O}$ 的特征根(第二章第三讲作业题 4). 于是, $\mu_{\mathcal{A}}(\lambda)$=0, 即 $\mathrm{spec}_F(\mathcal{A})\subset S$. □

例 10.13 利用上述推论, 我们有以下结论.

(i) 对合算子在某组基下的矩阵是

$$\begin{pmatrix} E_k & O \\ O & -E_\ell\end{pmatrix}.$$

(ii) 幂零算子的特征根都等于零, 但非零的幂零算子不可能对角化.

例 10.14 设

$$A=\begin{pmatrix} 7 & -4 \\ 14 & -8\end{pmatrix}.$$

计算 A^{66}.

解. 计算得 $\chi_{\mathcal{A}}(t)=t^2+t$. 由 *Hamilton-Cayley* 定理, $A^2=-A$. 于是,

$$A^{66}=(A^2)^{33}=-A^{33}=-AA^{32}=-AA^{16}=-AA^8=-AA^4=-AA^2=(-A)(-A)=A^2=-A.$$

我们得到

$$A^{66} = \begin{pmatrix} -7 & 4 \\ -14 & 8 \end{pmatrix}. \quad \square$$

另解. 由对角化可知,

$$\begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} = P^{-1}AP,$$

其中

$$P = \begin{pmatrix} 4 & 1 \\ 7 & 2 \end{pmatrix}.$$

于是

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = P^{-1}A^{66}P \implies A^{66} = P\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}P^{-1} = \begin{pmatrix} -7 & 4 \\ -14 & 8 \end{pmatrix}. \quad \square$$