第二章　线性算子

命题 7.10 设 $A \in \mathrm{M}_n(F)$,

$$\chi_A(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0, \quad a_i \in F.$$

则 $a_{n-1} = -\mathrm{tr}(A)$ 和 $a_0 = (-1)^n \det(A)$. 特别地, A 可逆当且仅当 0 不是 A 的特征根.

证明. 设 $A = (a_{i,j})_{n\times n}$ 由特征多项式的定义可知

$$\chi_A(t) = \det \begin{pmatrix} t - a_{1,1} & -a_{1,2} & \cdots & -a_{1,n} \\ -a_{2,1} & t - a_{2,2} & \cdots & -a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n,1} & -a_{n,2} & \cdots & t - a_{n,n} \end{pmatrix}.$$

由行列式的定义可知

$$\chi_A(t) = (t - a_{1,1})(t - a_{2,2}) \cdots (t - a_{n,n}) + p(t),$$

其中 $p \in F[t]$ 且 $\deg(p) < n - 1$. 故 $\deg(\chi_A) = n$ 且 $\mathrm{lc}(\chi_A) = 1$. 进而,

$$a_{n-1} - a_{1,1} - \cdots - a_{n,n} = -\mathrm{tr}(A).$$

另一方面,

$$a_0 = \chi_A(0) = \det \begin{pmatrix} -a_{1,1} & -a_{1,2} & \cdots & -a_{1,n} \\ -a_{2,1} & -a_{2,2} & \cdots & -a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n,1} & -a_{n,2} & \cdots & -a_{n,n} \end{pmatrix} = (-1)^n \det(A).$$

注意到 $\chi_A(0)=0$ 当且仅当 $\det(A)=0$. 故 0 是 A 的特征根当且仅当 A 可逆. □

注解 7.11 事实上, 我们可证上述命题中的 a_i 是 A 的所有 $n-i$ 阶主子式之和乘以 $(-1)^{n-i}$. 详见*2019-2020* 年春季学期第一章第二次习题课附录.

注解 7.12 上述命题说明*:* 如果 $\mathcal{A}\in\mathcal{L}(V)$ 且 $\dim(V)=n$, 则 $\chi_{\mathcal{A}}(t)$ 是 n 次的首一多项式.

命题 7.13 设 V 是复数域 $\mathbb{C}$ 上的线性空间 和 $\mathcal{A}\in\mathcal{L}(V)$. 则 $\mathcal{A}$ 一定有特征向量.

证明. 因为 $\chi_{\mathcal{A}}\in\mathbb{C}[t]\setminus\mathbb{C}$, 所以 $\chi_{\mathcal{A}}$ 在 $\mathbb{C}$ 中至少有一个根 λ (代数学基本定理). 即 $\mathcal{A}$ 有特征根. 于是, $\mathcal{A}$ 有特征向量. □

例 7.14 设 $A\in\mathrm{M}_n(\mathbb{C})$. 证明*:* A 相似于一个上三角矩阵.
证明. 对 n 归纳. 当 $n=1$ 时, 结论显然成立. 设 $n>1$ 且 $n-1$ 时结论成立.

考虑 $A\in\mathrm{M}_n(\mathbb{C})$. 把 A 看成 $\mathbb{C}^n$ 上在标准基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 下矩阵等于 A 的线性算子. 有上例可知, $\mathcal{A}$ 有一个 1 维 $\mathcal{A}$ 子空间 $\langle\mathbf{u}\rangle$. 根据第二章第二讲命题 *5.3*,

$$A\sim_s\begin{pmatrix}\lambda & *\\ O_{(n-1)\times1} & B\end{pmatrix},$$

其中 $\lambda \in \mathbb{C}$, $B \in \mathrm{M}_{n-1}(\mathbb{C})$. 根据归纳假设. 存在 $P \in \mathrm{GL}_{n-1}(\mathbb{C})$ 使得 $P^{-1}BP$ 是上三角的. 令

$$Q = \begin{pmatrix} 1 & O_{1\times n-1} \\ O_{(n-1)\times 1} & P \end{pmatrix}.$$

则 P 可逆且

$$\begin{aligned}
& Q^{-1} \begin{pmatrix} \lambda & * \\ O_{(n-1)\times 1} & B \end{pmatrix} Q \\
= & \begin{pmatrix} 1 & O_{1\times (n-1)} \\ O_{(n-1)\times 1} & P^{-1} \end{pmatrix} \begin{pmatrix} \lambda & * \\ O_{(n-1)\times 1} & B \end{pmatrix} \begin{pmatrix} 1 & O_{1\times (n-1)} \\ O_{(n-1)\times 1} & P \end{pmatrix} \\
= & \underbrace{\begin{pmatrix} \lambda & * \\ O_{(n-1)\times 1} & P^{-1}BP \end{pmatrix}}_{T}.
\end{aligned}$$

因为 $P^{-1}BP$ 已经是上三角矩阵, 所以 T 也是是上三角矩阵. 显然, $A \sim_s T$. $\square$

注解 7.15 上述例子说明: 算子 $\mathcal{A} \in \mathcal{L}(V)$ 在 V 的某组基下的矩阵是上三角的.

例 7.16 设 $A \in \mathrm{M}_n(F)$ 是如下分块上三角形

$$\begin{pmatrix} A_1 & * & * & \cdots & * \\ O & A_2 & * & \cdots & * \\ O & O & A_3 & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & * \\ O & O & O & \cdots & A_k \end{pmatrix}.$$

证明: $\chi_A = \chi_{A_1} \cdots \chi_{A_k}$.

证明.

$$\begin{aligned} \chi_A(t) &= \det(tE - A) \\ &= \begin{pmatrix} tE_{n_1} - A_1 & * & * & \cdots & * \\ O & tE_{n_2} - A_2 & * & \cdots & * \\ O & O & tE_{n_3} - A_3 & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & * \\ O & O & O & \cdots & tE_{n_k} - A_k \end{pmatrix}, \end{aligned}$$

其中 A_i 是 n_i 阶方阵, $i = 1, 2, \ldots, k$. 于是,

$$\chi_A(t) = \prod_{i=1}^{k} \det(tE_{n_i} - A_i) = \prod_{i=1}^{k} \chi_{A_i}(t). \quad \square$$

例 7.17 设 V 是 $\mathbb{C}$ 上的 n 维线性空间, $\mathcal{A} \in \mathcal{L}(V)$. 证明: $\mathcal{A}$ 有 $n-1$ 维的 $\mathcal{A}$-子空间.

证明. 注意到 $\mathcal{A}$ 的对偶算子 $\mathcal{A}^* \in \mathcal{L}(V^*)$ 且 $\mathcal{A}^*$ 的维数也是 n. 有上述命题 $\mathcal{A}^*$ 有特征向量 f. 再设 λ 是 f 对应的特征根.

设 $U=\ker(f)$. 则 $\dim(U)=n-1$. 设 $\mathbf{u}\in U$. 我们有

$$f(\mathcal{A}(\mathbf{u}))=(f\circ\mathcal{A})(\mathbf{u})=(\mathcal{A}^*(f))(\mathbf{u})=(\lambda f)(\mathbf{u})=\lambda(f(\mathbf{u}))=0.$$

故 $\mathcal{A}(\mathbf{u})\in U$. 于是, U 是 $\mathcal{A}$-子空间. $\square$

8 对角化

定义 8.1 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 $A\in \mathrm{M}_n(F)$ 相似于一个对角矩阵, 则称 A 是可对角化的.

定理 8.2 *(可对角化判别法I)* 设 $n=\dim(V)$ 和 $\mathcal{A}\in\mathcal{L}(V)$. 则 $\mathcal{A}$ 可对角化当且仅当 $\mathcal{A}$ 有 n 个线性无关的特征向量.

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

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

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 $\epsilon_1, \ldots, \epsilon_n$ 线性无关. □

推论 8.3 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 8.4 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 8.5 *(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 $P^{-1}\mathrm{M}_nP$ 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 $(1,\lambda_2)^t$. 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
$$P^{-1}AP=\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix}.$$
例 8.6 如例 **??** 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.

引理 8.7 设 $\mathcal{A}\in\mathcal{L}(V)$, $\lambda_1,\ldots,\lambda_k\in\operatorname{spec}_F(\mathcal{A})$ 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

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

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,

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

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 1.11 (ii), $V^{\lambda_1}+\cdots+V^{\lambda_k}$ 是直和. □

推论 8.8 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 8.7), 所以 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 线性无关(第一章第一讲定理 1.11 (ii)). 于是, 特征向量 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 是 V 的一组基. 由定理 8.2, $\mathcal{A}$ 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. $\square$

例 8.9 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 $\mathbb{C}$ 中有三个互不相同的根. 由上述推论, A 可以对角化. $\square$

定理 8.10 *(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 8.2). 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 8.7). 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 8.2, $\mathcal{A}$ 可对角化. □

例 8.11 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

$$\begin{pmatrix} \lambda_1 E_{d_1} & O & \cdots & O \\ O & \lambda_2 E_{d_2} & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & \lambda_k E_{d_k} \end{pmatrix}.$$

推论 8.12 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 8.13 *(可对角化判别法III)* 设 $\mathcal{A} \in \mathcal{L}(V)$ 且

$$\operatorname{spec}_F(\mathcal{A}) = \{\lambda_1, \cdots, \lambda_k\}.$$

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 8.7 可知,

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

于是

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

由定理 8.10, $\mathcal{A}$ 可对角化当且仅当

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

推论 8.14 设 $A \in \mathrm{M}_n(F)$ 且 $\operatorname{spec}_F(A) = \{\lambda_1, \cdots, \lambda_k\}$. 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 8.15 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

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

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 8.13)*. 于是, $\mathrm{rank}(A) = 0$, 即 $A = O$. 矛盾. □

定义 8.16 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 8.17 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(第二章第三讲第十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 5.3). 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 的代数重数是最大的整数 m 使得 $(t-\lambda)^m|\chi_{\mathcal{A}}(t)$. 故 $d\le m$. $\square$

定理 8.18 *(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) $\forall\,\lambda\in\mathrm{spec}_F(\mathcal{A})$, λ 的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 8.11, $\mathcal{A}$ 在某组基下的矩阵是

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

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 8.13, $\mathcal{A}$ 可对角化. □

推论 8.19 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 $\chi_A(t)$ 的所有根都在 F 中;

(ii) 对任意 $\lambda\in\mathrm{spec}_F(A)$, λ 的几何重数等于其代数重数.

证明. 把 A 看成 F^n 上在标准基下矩阵为 A 的算子. □ 最后一个可对角化判别法本质上与特征向量和特征值无关.

引理 8.20 设 $\mathcal{A}\in\mathcal{L}(V)$ 和 $f\in F[t]$ 零化 $\mathcal{A}$. 设 $f=pq$, 其中 $p,q\in F[t]$ 满足 $\gcd(p,q)=1$. 则对任意 $\mathbf{x}\in\ker(p(\mathcal{A}))\backslash\{\mathbf{0}\}$, $q(\mathcal{A})(\mathbf{x})\neq\mathbf{0}$.

证明. 根据核核分解定理, $V=\ker(p(\mathcal{A}))\oplus\ker(q(\mathcal{A}))$. 于是, $\ker(p(\mathcal{A}))\cap\ker(q(\mathcal{A}))=\{\mathbf{0}\}$. 故 $\mathbf{x}\notin\ker(q(\mathcal{A}))$. □.

定理 8.21 *(扩展的核核分解)* 设 $\mathcal{A}\in\mathcal{L}(V)$, $f(t)\in F[t]\setminus\{0\}$ 且 $f(\mathcal{A})=\mathcal{O}$. 设 $f=q_1q_2\cdots q_s$, 其中 $q_1,\ldots,q_s\in F[t]$ 两两互素. 对 $i\in\{1,2,\ldots,s\}$, 令 $K_i=\ker(q_i(\mathcal{A}))$ 和 $\mathcal{A}_i=\mathcal{A}_{K_i}$. 则

$$V=K_1\oplus\cdots\oplus K_s.$$

证明. 首先, 我们来证明断言: $K_1+\cdots+K_s$ 是直和. 设

$$\mathbf{x}_1+\cdots+\mathbf{x}_s=\mathbf{0},$$

其中 $\mathbf{x}_1\in K_1,\ldots,\mathbf{x}_s\in K_s$. 设 $p_i=q_1\cdots q_{i-1}q_{i+1}\cdots q_s$, $i=1,2,\ldots,s$. 把 $p_i(\mathcal{A})$ 作用于上式两侧得

$$p_i(\mathcal{A})(\mathbf{x}_1)+\cdots+p_s(\mathcal{A})(\mathbf{x}_s)=\mathbf{0}.$$

由 p_i 和 $\mathbf{x}_j$ 得定义可知, 当 $i\neq j$ 时, 我们有 $p_i(\mathcal{A})(\mathbf{x}_j)=\mathbf{0}$. 故上式可化简为

$$p_i(\mathcal{A})(\mathbf{x}_i)=\mathbf{0}.$$

因为 $\gcd(q_i,p_i)=1$, 所以 $\mathbf{x}_i=\mathbf{0}$ (引理 8.20). 根据第一章第一讲定理 1.11 (ii), 断言成立.

对 s 归纳. 当 $s=1$ 时, $V=K_1$. 故结论是平凡的.

设 $s>1$ 且 $s-1$ 时定理成立. 令 $q=q_2\cdots q_s$. 则 $\gcd(q_1,q)=1$. 根据核核分解定理,

$$V=K_1\oplus\ker(q(\mathcal{A})). \tag{1}$$

设 $\mathcal{B}$ 是 $\mathcal{A}$ 在 $\ker(q(\mathcal{A}))$ 上的限制映射. 因为 $\ker(q(\mathcal{A}))$ 是 $\mathcal{A}$-子空间(第二章第三次讲义命题 5.8), 所以 $\mathcal{B}$ 是 $\ker(q(\mathcal{A}))$

上的线性算子. 因为 $q(\mathcal{A})$ 在 $\ker(q(\mathcal{A}))$ 上是把所有向量都映成零向量, 所以 $q(\mathcal{B})$ 是 $\ker(q(\mathcal{A}))$ 上的零算子. 对 $\mathcal{B}$, $\ker(q(\mathcal{A}))$, 和 $q = q_2 \cdots q_s$ 用归纳假设得

$$\ker(q(\mathcal{A})) = \ker(q_2(\mathcal{B})) \oplus \cdots \oplus \ker(q_s(\mathcal{B})).$$

再根据 (1) 得

$$V = K_1 \oplus \ker(q_2(\mathcal{B})) \oplus \cdots \oplus \ker(q_s(\mathcal{B})). \tag{2}$$

因为 $\mathcal{B}$ 是 $\mathcal{A}$ 的限制算子, 所以 $\ker(q_i(\mathcal{B}))\subset K_i$, $i = 2, \ldots, s$. 故

$$\ker(q_2(\mathcal{B})) \oplus \cdots \oplus \ker(q_s(\mathcal{B})) \subset K_2 + \cdots + K_s.$$

由断言和 (2), $V = K_1 \oplus K_2 \oplus \cdots \oplus K_s$. □

定理 8.22 *(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=diag$(\lambda_1, \ldots, \lambda_n)$. 则

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

(见第二章第三讲引理 5.7). 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, $i\neq j$. 由定理 **??**

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

其中 $U_i=\ker(\mathcal{A}-\beta_i\mathcal{E})$, $i=1,\ldots,k$. 根据第二章第三讲引理 5.4, 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 $d_i=\dim(U_i)$. 由第二章第三讲定理 5.9, $\mathcal{A}$ 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 8.23 *(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 8.15. 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 8.24 *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 8.25 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 8.26 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.

9 循环子空间

定义 9.1 设 $\mathcal{A}\in\mathcal{L}(V)$ 和 $\mathbf{v}\in V$. 设 $f(t)\in F[t]$. 如果 $f(\mathcal{A})(\mathbf{v})=\mathbf{0}$, 则称 f 是关于 $\mathcal{A}$ 和 $\mathbf{v}$ 的零化多项式. 关于 $\mathcal{A}$ 和 $\mathbf{v}$ 的零化多项式中次数最小的非零多项式称为关于 $\mathcal{A}$ 和 $\mathbf{v}$ 的极小多项式. 我们用 $\mu_{\mathcal{A},\mathbf{v}}$ 记首一的关于 $\mathcal{A}$ 和 $\mathbf{v}$ 的极小多项式.

因为 $\mu_{\mathcal{A}}(\mathcal{A})=\mathcal{O}$, 所以对任意 $\mathbf{v}\in V$, $\mu_{\mathcal{A}}(\mathcal{A})(\mathbf{v})=\mathbf{0}$. 故对于任何 $\mathbf{v}\in V$, 关于 $\mathcal{A}$ 零化 $\mathbf{v}$ 的非零多项式存在. 于是, 关于 $\mathcal{A}$ 和 $\mathbf{v}$ 的极小多项式的多项式存在.

设 $f(\mathcal{A})(\mathbf{v})=\mathbf{0}$. 由多项式带余除法可知

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

其中 $q,r\in F[t]$, $\deg(r)<\deg(\mu_{\mathcal{A},\mathbf{v}})$. 带入 $\mathcal{A}$ 得

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

两侧同时作用在 $\mathbf{v}$ 上得到

$$\mathbf{0} = q(\mathcal{A})\mu_{\mathcal{A},\mathbf{v}}(\mathcal{A})(\mathbf{v}) + r(\mathcal{A})(\mathbf{v}) = \mathbf{0} + r(\mathcal{A})(\mathbf{v}).$$

于是, $r(\mathcal{A})(\mathbf{v}) = \mathbf{0}$. 因为 $\deg(r) < \deg(\mu_{\mathcal{A},\mathbf{v}})$, 所以 $r(t) = 0$. 由此得出 $\mu_{\mathcal{A},\mathbf{v}}|f$. 特别地, $\mu_{\mathcal{A},\mathbf{v}}|\mu_{\mathcal{A}}$. 于是我们有

引理 9.2 设 $\mathcal{A} \in \mathcal{L}(V)$ 和 $\mathbf{v} \in V$. 设 $f(t) \in F[t]$. 则 $f(t)$ 关于 $\mathcal{A}$ 零化 $\mathbf{v}$ 当且仅当 $\mu_{\mathcal{A},\mathbf{v}}(t)|f(t)$.

定义 9.3 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 9.4 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 $F[\mathcal{A}] \cdot \mathbf{v}$ 的一组基. 特别地, $d = \dim(F[\mathcal{A}] \cdot \mathbf{v})$.

证明. (i) 设

$$p = p_k t^k + p_{k-1} t^{k-1} + \cdots + p_0,$$

其中 $p_k, p_{k-1}, \ldots, p_0 \in F$. 则

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

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 $q(t) = \sum_{i=0}^{\ell} \alpha_i t^i$. 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. $\square$