第二章　线性算子

例 7.7 设上周最后一个例子中矩阵 B 是复矩阵. 求它的特征值和特征向量.

解. 由上例可知,

$$\chi_B(t) = t^2 + 1.$$

于是, $\mathrm{spec}_{\mathbb{C}}(B) = \{\sqrt{-1}, -\sqrt{-1}\}$. 设特征根 $\lambda_1 = \sqrt{-1}$. 它对应的特征子空间是方程组

$$(\lambda_1 E - A)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

的解空间. 即方程组

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

的解空间. 解方程组得 $V^{\lambda_1} = \langle (1, -\sqrt{-1})^t \rangle$. 类似地, 特征根 $\lambda_2 = -\sqrt{-1}$ 对应的特征子空间是 $V^{\lambda_2} = \langle (1, \sqrt{-1})^t \rangle$.

命题 7.8 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 不可逆. □

例 7.9 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,

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

命题 7.10 矩阵的特征多项式是相似不变量.

证明. 设 $A, B \in \mathrm{M}_n(F)$ 且 $A \sim_s B$. 则存在 $P \in \mathrm{GL}_n(F)$

使得 $B=P^{-1}AP$. 故

$$\begin{aligned}\chi_B(t)&=|tE-B|\\&=|tE-P^{-1}AP|\\&=|P^{-1}(tE-A)P|\\&=|P^{-1}||tE-A||P|\\&=|tE-A|=\chi_A(t).\quad\square\end{aligned}$$

注解 7.11 由上述命题可知, 矩阵的特征根都是相似不变量. 即矩阵的谱是相似不变量.

注解 7.12 矩阵的特征向量不是相似不变量. 事实上, 设 $A\in \mathrm{M}_n(F)$, $\lambda\in \mathrm{spec}_F(A)$, $\mathbf{v}\in F^n$ 是 A 关于 λ 的特征向量. 再设 $B=P^{-1}AP$, 其中 $P\in \mathrm{GL}_n(F)$. 则

$$B(P^{-1}\mathbf{v})=P^{-1}(A\mathbf{v})=P^{-1}(\lambda\mathbf{v})=\lambda\left(P^{-1}(\mathbf{v})\right).$$

故 $P^{-1}(\mathbf{v})$ 是 B 关于 λ 的特征向量. 通常 $\mathbf{v}\neq P^{-1}\mathbf{v}$.

例 7.13 设 $A\in \mathrm{M}_n(\mathbb{C})$. 则 A 有特征根.

证明. 特征多项式 $\chi_A(t)\in\mathbb{C}[t]\setminus\mathbb{C}$, 根据代数基本定理, $\mathrm{spec}_{\mathbb{C}}(A)\neq 0$. $\square$

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 $n=1$ 时显然. 设 $n>1$ 且 $n-1$ 时结论成立.

根据上例, A 有特征根有特征根. 设 λ 为 A 的特征根, $\mathbf{v} \in \mathbb{C}^n$ 是 λ 对应的一个特征向量. 则 $A\mathbf{v} = \lambda\mathbf{v}$.

令 $P \in \mathrm{GL}_n(\mathbb{C})$ 是可逆矩阵使得 $\vec{P}^{(1)} = \mathbf{v}$. 因为

$$AP = \left(A\mathbf{v}, A\vec{P}^{(2)}, \ldots, \vec{P}^{(n)}\right) = (\lambda\mathbf{v}, *, \ldots, *),$$

所以

$$\begin{aligned} P^{-1}AP &= P^{-1}\left(\lambda\vec{P}^{(1)}, *, \cdots, *\right) \\ &= \left(\lambda P^{-1}\vec{P}^{(1)}, *, \cdots, *\right) \\ &= \begin{pmatrix} \lambda & * & \cdots & * \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}, \end{aligned}$$

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

$$T = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{pmatrix}^{-1} P^{-1}AP \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{pmatrix}$$

是上三角矩阵. 故 $A \sim T$. $\square$

7.3 线性算子与矩阵的特征向量、特征值和特征多项式之间的联系

定义 7.15 设 $\mathcal{A} \in \mathcal{L}(V)$, $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 V 的一组基, $\mathcal{A}$ 在该基下的矩阵等于 A. 则 $\det(tE - A)$ 称为 $\mathcal{A}$ 的特征多项式*(characteristic polynomial)*, 记为 $\chi_{\mathcal{A}}$. 特征多项式 $\chi_{\mathcal{A}}(t)$ 在 F 中所有根的集合记为 $\mathrm{spec}_F(\mathcal{A})$, 称为 $\mathcal{A}$ 的在 F 中的谱*(spectrum)*

本节中关于矩阵特征向量和特征值的结果可以翻译成算子的语言, 表述如下设 $\mathcal{A} \in \mathcal{L}(V)$.

1. $\chi_{\mathcal{A}}$ 是良定义的 (命题 7.10).

2. $\chi_{\mathcal{A}}$ 是次数为 $\dim(V)$ 的首一多项式, $\mathcal{A}$ 可逆当且仅当 $\chi_{\mathcal{A}}(0) \neq 0$. （命题 7.8)

下面我们描述定义 7.15 中 $\mathcal{A}$ 与 A 特征向量之间的关系. 设 $\mathbf{v} \in V \setminus \mathbf{0}$ 且

$$\mathbf{v} = (\mathbf{e}_1, \ldots, \mathbf{e}_n) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

则 $\mathbf{v}$ 是关于特征值 λ 的特征向量当且仅当

$$\mathcal{A}(\mathbf{v})=\lambda\mathbf{v} \Longleftrightarrow (\mathbf{e}_1,\ldots,\mathbf{e}_n)A\begin{pmatrix}v_1\\ \vdots\\ v_n\end{pmatrix}=\lambda(\mathbf{e}_1,\ldots,\mathbf{e}_n)\begin{pmatrix}v_1\\ \vdots\\ v_n\end{pmatrix}$$

$$\Longleftrightarrow A\begin{pmatrix}v_1\\ \vdots\\ v_n\end{pmatrix}=\lambda\begin{pmatrix}v_1\\ \vdots\\ v_n\end{pmatrix}.$$

故 $\mathbf{v}$ 是 $\mathcal{A}$ 关于 λ 的特征向量当且仅当其坐标是 A 关于 λ 的特征向量. 由此可知, 计算线性算子的特征值和特征向量的问题可以转化为计算其矩阵表示的特征值和特征向量的问题.

例 7.16 设 V 的一组基是 $\mathbf{e}_1,\ldots,\mathbf{e}_n$, 算子 $\mathcal{J}$ 由

$$\mathcal{J}(\mathbf{e}_1)=\mathbf{0},\mathcal{J}(\mathbf{e}_2)=\mathbf{e}_1,\mathcal{J}(\mathbf{e}_3)=\mathbf{e}_2,\ldots,\mathcal{J}(\mathbf{e}_n)=\mathbf{e}_{n-1}$$

确定. 求 $\mathcal{J}$ 的所有特征向量和特征值.

解. 算子 $\mathcal{J}$ 在 $\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n$ 的矩阵是

$$J=\begin{pmatrix}0&1&0&\cdots&0&0\\ 0&0&1&\cdots&0&0\\ \vdots&\vdots&\vdots&\ddots&\vdots&\vdots\\ 0&0&0&\cdots&0&1\\ 0&0&0&\cdots&0&0\end{pmatrix}.$$

故 $\chi_{\mathcal{A}}(t)=t^n$. 故 $V^0=\ker(\mathcal{J})=\langle\mathbf{e}_1\rangle$.

例 7.17 设 V 是 $\mathbb{C}$ 上的 n 维线性空间, $\mathcal{A} \in \mathcal{L}(V)$. 则 $\mathcal{A}$ 有特征值, 且 $\mathcal{A}$ 在某组基下的矩阵的上三角的.

证明. 根据例 *7.13, 7.14* 和线性算子与矩阵表示之间的关系直接可得. □

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 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 $P^{-1}AP$ 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 由上周讲义例 *7.3* 可知, $\mathbb{R}[x]^{(n)}$ 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 $n>1$ 时 $\mathcal{D}$ 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 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 $\mathbf{e}_{i,1}, \dots \mathbf{e}_{i,d_i}$ 是 V^{λ_i} 的一组基, $i = 1, 2, \dots, k$. 基底中的元素都是特征向量. 由直和分解可知,

$$\mathbf{e}_{1,1}, \dots, \mathbf{e}_{1,d_1}, \dots, \mathbf{e}_{k,1}, \dots, \mathbf{e}_{k,d_k}$$

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

$$\mathbf{e}_{1,1}, \dots, \mathbf{e}_{1,d_1}, \dots, \mathbf{e}_{k,1}, \dots, \mathbf{e}_{k,d_k}.$$

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

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

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

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 $d\leq m$. □

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) $\forall\, \lambda \in \operatorname{spec}_F(\mathcal{A})$, λ 的几何重数等于它的代数重数.

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

设 $\mathcal{A}$ 可对角化. 由例 8.11, $\mathcal{A}$ 在某组基下的矩阵是

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

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. □

定理 8.20 *(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 $A=\mathrm{diag}(\lambda_1,\ldots,\lambda_n)$. 则

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

(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 5.11 (扩展的核核分解定理之极小多项式版), 我们有

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

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 5.5, 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 $d_i = \dim(U_i)$. 故 $\mathcal{A}$ 在 V 的某组基下的矩阵是

$$\begin{pmatrix} \beta_1 E_{d_1} & O & \cdots & O \\ O & \beta_2 E_{d_2} & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & \beta_k E_{d_k} \end{pmatrix}. \quad \square$$

推论 8.21 *(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 8.22 *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. $\square$

注解 8.23 *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 8.24 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.