第二章　线性算子

命题 5.3 设 $\mathcal{A} \in \mathcal{L}(V)$, U 是 $\mathcal{A}$ 的 d 维不变子空间, $0 < d < n$. 则存在 V 的一组基使得 $\mathcal{A}$ 在该基下的矩阵为

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

其中 $B \in \mathrm{M}_d(F)$ 是 $\mathcal{A}_U$ 的某个矩阵表示. 进而 $\mu_{\mathcal{A}_U} | \mu_{\mathcal{A}}$, $\mu_B | \mu_A$, $\mu_D | \mu_A$.

证明. 设 $\mathbf{e}_1, \ldots, \mathbf{e}_d$ 是 U 的一组基. 把它扩充为 V 的一组基 $\mathbf{e}_1, \ldots, \mathbf{e}_d, \mathbf{e}_{d+1}, \ldots, \mathbf{e}_n$. 因为 U 是 $\mathcal{A}$ 的不变子空间, 所以当 $j \in \{1, 2, \ldots, d\}$ 时, $\mathcal{A}(\mathbf{e}_j)$ 是 $\mathbf{e}_1, \ldots, \mathbf{e}_d$ 的线性组合, 即 $\mathcal{A}(\mathbf{e}_j)$ 关于 $\mathbf{e}_{d+1}, \ldots, \mathbf{e}_n$ 的坐标都等于零. 于是 $\mathcal{A}$ 在 $\mathbf{e}_1, \ldots, \mathbf{e}_d, \mathbf{e}_{d+1}, \ldots, \mathbf{e}_n$ 下的矩阵如命题所述形式, 且 B 是 $\mathcal{A}_U$ 在 $\mathbf{e}_1, \ldots, \mathbf{e}_d$ 下的矩阵.

直接计算可验证对任意 $k \in \mathbb{N}$

$$A^k = \begin{pmatrix} B^k & * \\ O & D^k \end{pmatrix},$$

其中 $*$ 是某个 $d \times (n-d)$ 阶的矩阵. 于是, 对任意 $f \in F[t]$.

$$f(A) = \begin{pmatrix} f(B) & * \\ O & f(D) \end{pmatrix}.$$

因为 $\mu_A(A)=O_{n\times n}$, 所以 $\mu_A(B)=O_{d\times d}$, $\mu_A(D)=O_{(n-d)\times(n-d)}$. 由第二章第二讲引理 4.2, $\mu_B|\mu_A$, $\mu_D|\mu_A$, 且 $\mu_{\mathcal{A}_U}|\mu_{\mathcal{A}}$. □

给定 $\mathcal{A}\in\mathcal{L}(V)$, $\{\mathbf{0}\}$ 和 V 是 $\mathcal{A}$ 的平凡的不变子空间. 下面的引理指出如何寻找 $\mathcal{A}$ 的非平凡的子空间.

引理 5.4 *设 $\mathcal{A},\mathcal{B}\in\mathcal{L}(V)$ 满足 $\mathcal{A}\mathcal{B}=\mathcal{B}\mathcal{A}$. 则 $\ker(\mathcal{B})$ 和 $\mathrm{im}(\mathcal{B})$ 是 $\mathcal{A}$ 的不变子空间.*

证明. 设 $\mathbf{x}\in\ker(\mathcal{B})$. 则

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

于是 $\mathcal{A}(\mathbf{x})\in\ker(\mathcal{B})$. 即 $\ker(\mathcal{B})$ 是 $\mathcal{A}$ 不变的. 设 $\mathbf{x}\in\mathrm{im}(\mathcal{B})$. 则存在 $\mathbf{y}\in V$ 使得 $\mathbf{x}=\mathcal{B}(\mathbf{y})$. 于是

$$\mathcal{A}(\mathbf{x})=\mathcal{A}(\mathcal{B}(\mathbf{y}))=\mathcal{B}(\mathcal{A}(\mathbf{y}))\in\mathrm{im}(\mathcal{B}).\quad\square$$

命题 5.5 *设 $\mathcal{A}\in\mathcal{L}(V)$, $f\in F[t]$. 则 $\ker(f(\mathcal{A}))$ 和 $\mathrm{im}(f(\mathcal{A}))$ 都是 $\mathcal{A}$ 的不变子空间.*

证明. 因为 $\mathcal{A}f(\mathcal{A})=f(\mathcal{A})\mathcal{A}$, 所以 $\ker(f(\mathcal{A}))$ 和 $\mathrm{im}(f(\mathcal{A}))$ 都是 $\mathcal{A}$ 的不变子空间(引理 5.4). □

为了简单起见, 当 U 是 $\mathcal{A}$ 的不变子空间时, 我们说 U 是 $\mathcal{A}$-不变的或许 $\mathcal{A}$-子空间.

命题 5.6 *设 $\mathcal{A}\in\mathcal{L}(V)$, U_1,U_2 是 $\mathcal{A}$-子空间. 则 U_1+U_2 和 $U_1\cap U_2$ 都是 $\mathcal{A}$-子空间.*

证明. 设 $\mathbf{x} \in U_1 + U_2$. 则存在 $\mathbf{x}_1 \in U_1$, $\mathbf{x}_2 \in U_2$ 使得 $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$. 于是,

$$\mathcal{A}(\mathbf{x}) = \mathcal{A}(\mathbf{x}_1) + \mathcal{A}(\mathbf{x}_2) \in U_1 + U_2.$$

设 $\mathbf{x} \in U_1 \cap U_2$, 则 $\mathcal{A}(\mathbf{x}) \in U_1$ 且 $\mathcal{A}(\mathbf{x}) \in U_2$. 由此可知, $\mathcal{A}(\mathbf{x}) \in U_1 \cap U_2$. $\square$

引理 5.7 设 $\mathcal{A} \in \mathcal{L}(V)$, U_1, U_2 是非平凡 $\mathcal{A}$-子空间, 且 $V = U_1 \oplus U_2$. 设 $\epsilon_1, \ldots, \epsilon_{d_1}$ 是 U_1 的基, $\delta_1, \ldots, \delta_{d_2}$ 是 U_2 的基. 则在 V 的基底 $\epsilon_1, \ldots, \epsilon_{d_1}, \delta_1, \ldots, \delta_{d_2}$ 下 $\mathcal{A}$ 的矩阵是

$$A = \begin{pmatrix} A_1 & O \\ O & A_2 \end{pmatrix},$$

其中 $A_i \in \mathrm{M}_{d_i}(F)$ 是 $\mathcal{A}_{U_i}$ 在对应基下的矩阵, $i = 1, 2$. 进而 $\mu_{\mathcal{A}} = \mathrm{lcm}(\mu_{\mathcal{A}_{U_1}}, \mu_{\mathcal{A}_{U_2}})$ *(取首一的最小公倍式)*.

证明. 注意到 $V = U_1 \oplus U_2$ 蕴含 $d_1 + d_2 = n (= \dim(V))$ 且 $\epsilon_1, \ldots, \epsilon_{d_1}, \delta_1, \ldots, \delta_{d_2}$ 线性无关. 所以 $\epsilon_1, \ldots, \epsilon_{d_1}, \delta_1, \ldots, \delta_{d_2}$ 是 V 的一组基. 对 $i \in \{1, 2, \ldots, d_1\}$, $\mathcal{A}(\epsilon_i) \in U_1$, $\mathcal{A}(\epsilon_i)$ 是 $\epsilon_1, \ldots, \epsilon_{d_1}$ 的线性组合, 它关于 $\delta_1, \ldots, \delta_{d_2}$ 的坐标都是零. 于是, 存在 $A_1 \in \mathrm{M}_{d_1}(F)$ 使得

$$(\mathcal{A}(\epsilon_1), \ldots, \mathcal{A}(\epsilon_{d_1})) = (\epsilon_1, \ldots, \epsilon_{d_1}) A_1.$$

类似地, 存在 $A_2 \in \mathrm{M}_{d_1}(F)$ 使得

$$(\mathcal{A}(\delta_1), \ldots, \mathcal{A}(\delta_{d_1})) = (\delta_1, \ldots, \delta_{d_2}) A_2.$$

于是 A_i 是 $\mathcal{A}_{U_i}$ 在对应基底下的矩阵, $i=1,2$. 进而, $\mathcal{A}$ 在 V 的基底 $\epsilon_1,\dots,\epsilon_{d_1},\delta_1,\dots,\delta_{d_2}$ 下的矩阵等于 A.

设 $p=\operatorname{lcm}(\mu_{\mathcal{A}_{U_1}},\mu_{\mathcal{A}_{U_2}})$. 由引理 5.3, $\mu_{\mathcal{A}_{U_1}}|\mu_{\mathcal{A}},\mu_{\mathcal{A}_{U_2}}|\mu_{\mathcal{A}}$. 于是 $p|\mu_{\mathcal{A}}$. 又因为 $\mu_{\mathcal{A}_{U_1}}|p,\mu_{\mathcal{A}_{U_2}}|p$, 所以 $p(\mathcal{A}_{U_1})=\mathcal{O}$ 和 $p(\mathcal{A}_{U_2})=\mathcal{O}$ (引理 5.4). 于是

$$p(A)=\begin{pmatrix}p(A_1) & O\\ O & p(A_2)\end{pmatrix}=\begin{pmatrix}O & O\\ O & O\end{pmatrix}.$$

由此和引理 5.4, $\mu_{\mathcal{A}}|p$. 再利用首一性得出 $p=\mu_{\mathcal{A}}$. □

例 5.8 设

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

计算 μ_A.

解. 由上述引理 $\mu_A=\operatorname{lcm}(\mu_{(1)},\mu_{(0)})=\operatorname{lcm}(t-1,t)=(t-1)t$. □

定理 5.9 设 $\mathcal{A}\in\mathcal{L}(V)$, $U_1,\dots,U_k$ 是非平凡 $\mathcal{A}$-子空间满足 $V=U_1\oplus\cdots\oplus U_k$. 设 Z_i 是 U_i 的一组基, $i=1,\dots,k$. 则 $\mathcal{A}$ 在 V 的基底 $Z_1\cup\cdots\cup Z_k$ 下的矩阵

$$A=\begin{pmatrix}A_1 & O & \cdots & O\\ O & A_2 & \cdots & O\\ \vdots & \vdots & \ddots & \vdots\\ O & O & \cdots & A_k\end{pmatrix},$$

其中 A_i 是 $\mathcal{A}_{U_i}$ 在 Z_i 下的矩阵, $i = 1, 2, \ldots, k$. 进而, $\mu_{\mathcal{A}} = \mathrm{lcm}(\mu_{\mathcal{A}_{U_1}}, \ldots, \mu_{\mathcal{A}_{U_k}})$.

证明. 对 k 归纳. 当 $k = 1$ 时, 定理显然成立. 设 $k > 1$ 且 $k-1$ 时定理成立. 设 $W = U_1 \oplus \cdots \oplus U_{k-1}$. 则 $V = W \oplus U_k$, $Y = Z_1 \cup \cdots \cup Z_{k-1}$ 是 W 的基. 由引理 5.7, $\mathcal{A}$ 在基底 $W \cup Z_k$ 下的矩阵是

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

其中 B 是 $\mathcal{A}_W$ 在 Y 下的矩阵, A_k 是 $\mathcal{A}_{U_k}$ 在 Z_k 下的矩阵, 且 $\mu_{\mathcal{A}} = \mathrm{lcm}(\mu_{\mathcal{A}_W}, \mu_{A_{U_k}})$.

对 $\mathcal{A}_W$, W, $U_1, \ldots, U_{k-1}$ 用归纳假设得

$$B = \begin{pmatrix} A_1 & O & \cdots & O \\ O & A_2 & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & A_{k-1} \end{pmatrix},$$

其中 A_i 是 $\mathcal{A}_{U_i}$ 在 Z_i 下的矩阵, $i = 1, 2, \ldots, k-1$. 进而, $\mu_{\mathcal{A}_W} = \mathrm{lcm}(\mu_{\mathcal{A}_{U_1}}, \ldots, \mu_{\mathcal{A}_{U_{k-1}}})$. 于是, A 是所要求的形式. 注意到

$$\begin{aligned} &\mathrm{lcm}\left(\mu_{\mathcal{A}_{U_1}}, \ldots, \mu_{\mathcal{A}_{U_k}}\right) \\ &= \mathrm{lcm}\left(\mathrm{lcm}\left(\mu_{\mathcal{A}_{U_1}}, \ldots, \mu_{\mathcal{A}_{U_{k-1}}}\right), \mu_{A_{U_k}}\right) \\ &= \mathrm{lcm}(\mu_{\mathcal{A}_W}, \mu_{A_{U_k}}) = \mu_{\mathcal{A}}. \quad \square \end{aligned}$$

定理 5.10 *(核像分解 II)*[1] 设 $\mathcal{A} \in \mathcal{L}(V)$. 则

$$V = \ker(\mathcal{A}) \oplus \mathrm{im}(\mathcal{A}) \iff t^2 \nmid \mu_{\mathcal{A}}.$$

证明. 设 $K = \ker(\mathcal{A})$ 和 $I = \mathrm{im}(\mathcal{A})$.

($\Longrightarrow$) 设 $V = K \oplus I$. 如果 $K = \{\mathbf{0}\}$, 则 $\mathcal{A}$ 是单射(第一章第一讲命题 2.2). 于是, $\mathcal{A}$ 可逆(第二章第一讲推论 1.19). 根据第二章第二讲命题 4.12, $\mu_{\mathcal{A}}(0) \neq 0$, 从而 $t^2 \nmid \mu_{\mathcal{A}}$. 如果 $K_{\mathcal{A}} = V$. 则 $\mathcal{A} = \mathcal{O}$. 此时 $\mu_{\mathcal{A}} = t$. 于是 $t^2 \nmid \mu_{\mathcal{A}}$.

设 K 和 I 都是非平凡的. 则 $\mathcal{A}_K$ 是零算子. 于是 $\mu_{\mathcal{A}_K} = t$. 设 $\mathbf{v} \in I$. 如果 $\mathcal{A}_I(\mathbf{v}) = \mathbf{0}$, 则 $\mathcal{A}(\mathbf{v}) = \mathbf{0}$. 于是, $\mathbf{v} \in K \cap I$. 有直和条件可知, $\mathbf{v} = \mathbf{0}$. 即 $\mathcal{A}_I$ 是双射(第二章第一讲推论 1.19). 根据第二章第二讲命题 4.12, $t \nmid \mu_{\mathcal{A}_I}(0) \neq 0$. 我们得到

$$\mu_{\mathcal{A}} = \mathrm{lcm}(\mu_{\mathcal{A}_K}, \mu_{\mathcal{A}_I}) = t\mu_{\mathcal{A}_I}.$$

于是, $t^2 \nmid \mu_{\mathcal{A}}$.

($\Longleftarrow$) 如果 $t \nmid \mu_{\mathcal{A}}$, 则 $\mathcal{A}$ 可逆(第二章第二讲命题 4.12). 如果 $\mu_{\mathcal{A}} = t$, 则 $\mathcal{A} = \mathcal{O}$. 在这两种情形下, $V = K \oplus I$ 显然成立.

设 $\mu_{\mathcal{A}} = tp$ 满足 $\gcd(t, p) = 1$. 由核核分解定理

$$V = K \oplus \ker(p(\mathcal{A})).$$

[1]袁力, 沈洁. 常州工学院学报 27 卷第二期, 2014 年 4 月.

下面我们验证 $I=\ker(p(\mathcal{A}))$. 设 $\mathbf{x}\in I$. 则存在 $\mathbf{y}\in V$ 使得 $\mathbf{x}=\mathcal{A}(\mathbf{y})$. 于是,

$$p(\mathcal{A})(\mathbf{x})=p(\mathcal{A})(\mathcal{A}(\mathbf{y}))=(p(\mathcal{A})\mathcal{A})(\mathbf{y})=(tp)(\mathcal{A})(\mathbf{y})=\mu_{\mathcal{A}}(\mathcal{A})(\mathbf{y})=\mathbf{0}.$$

由此得出 $I\subset\ker(p(\mathcal{A}))$. 另一方面, 由直和分解的性质和核像维数公式可知

$$\dim(\ker(p(\mathcal{A})))=\dim(V)-\dim(K)=\dim(I).$$

我们推出 $I=\ker(p(\mathcal{A}))$. $\square$

例 5.11 设 $\mathcal{A}\in\mathcal{L}(V)$ 满足 $\mathcal{A}^3-\mathcal{A}^2-\mathcal{A}=\mathcal{O}$. 证明: $\ker(\mathcal{A})\oplus\mathrm{im}(\mathcal{A})=V$.

证明. 设 $f(t)=t^3-t^3-t$. 则 $f(\mathcal{A})=\mathcal{O}$. 由引理 **??**, $\mu_{\mathcal{A}}(t)|f(t)$. 因为 $t^2\nmid f$, 所以 $t^2\nmid\mu_{\mathcal{A}}(t)$. 由上述定理, $\ker(\mathcal{A})\oplus\mathrm{im}(\mathcal{A})=V$. $\square$

例 5.12 设 $\mathcal{A}\in\mathcal{L}(V)$ 满足 $\ker(\mathcal{A})\oplus\mathrm{im}(\mathcal{A})=V$. 设 $\mathbf{e}_1,\ldots,\mathbf{e}_r$ 是 $\mathrm{im}(\mathcal{A})$ 的一组基, $\mathbf{e}_{r+1},\ldots,\mathbf{e}_n$ 是 $\ker(\mathcal{A})$ 的一组基. 则 $\mathbf{e}_1,\ldots,\mathbf{e}_r,\mathbf{e}_{r+1},\ldots,\mathbf{e}_n$ 是 V 的一组基. 因为 $\mathrm{im}(\mathcal{A})$ 和 $\ker(\mathcal{A})$ 都是 $\mathcal{A}$-子空间, 且 $\mathcal{A}(\mathbf{e}_j)=\mathbf{0}$, $j=r+1,r+2,\ldots,n$, 所以 $\mathcal{A}$ 在该基底下的矩阵是

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

其中 $B\in\mathrm{M}_r(F)$ 满秩. 当 $r=n$ 时, $B=A$. 否则, $\mu_A=\mathrm{lcm}(\mu_B,t)$.

6 不可分子空间

定义 6.1 设 $\mathcal{A} \in \mathcal{L}(V)$, U 是 $\mathcal{A}$-子空间. 如果 U 不能写成两个非零的 $\mathcal{A}$-子空间的直和, 则称 U 是 $\mathcal{A}$-不可分的.

命题 6.2 设 $\mathcal{A} \in \mathcal{L}(V)$. 则 V 是有限个 $\mathcal{A}$-不可分子空间的直和.

证明. 设 $n = \dim(V)$. 我们对 n 归纳. 当 $n = 1$ 时, V 本身是 $\mathcal{A}$ 不可分的. 定理显然成立. 设 $n > 1$ 且当空间维数小于 n 时定理成立. 如果 V 是 $\mathcal{A}$ 不可分的, 则定理成立. 否则存在两个非零 $\mathcal{A}$-子空间 U, W 使得 $V = U \oplus W$. 则 $\dim(U)$ 和 $\dim(W)$ 的维数都小于 n. 由归纳假设, $U = U_1 \oplus \cdots \oplus U_k$, 其中 U_i 是 A_U 不可分的, 从而也是 $\mathcal{A}$ 不可分的. 同样, $W = W_1 \oplus \cdots \oplus W_\ell$, 其中 W_j 是 $\mathcal{A}_W$ 不可分的, 从而也是 $\mathcal{A}$ 不可分的. 于是

$$V = U \oplus W = U_1 \oplus \cdots \oplus U_k \oplus W_1 \oplus \cdots \oplus W_\ell. \quad \square$$

7 特征向量和特征多项式

在本节中, V 是域 F 上的有限维线性空间且 $\dim(V)>0$.

7.1 特征向量

定义 7.1 设 $\mathcal{A} \in \mathcal{L}(V)$, $\mathbf{v} \in V \setminus \{\mathbf{0}\}$. 如果 $\langle \mathbf{v} \rangle$ 是 $\mathcal{A}$ -子空

间, 则称 $\mathbf{v}$ 是 $\mathcal{A}$ 的一个特征向量 *(eigenvector)*.

命题 7.2 设 $\mathcal{A} \in \mathcal{L}(V)$, $\mathbf{v} \in V \setminus \{\mathbf{0}\}$. 则下列结论等价:

(i) $\mathbf{v}$ 是 $\mathcal{A}$ 的特征向量;

(ii) $\mathcal{A}(\mathbf{v}) \in \langle \mathbf{v} \rangle$;

(iii) 存在 $\lambda \in F$ 使得 $\mathcal{A}(\mathbf{v}) = \lambda \mathbf{v}$.

证明. (i) $\Longrightarrow$ (ii) 显然.

(ii) $\Longrightarrow$ (iii) 因为 $\mathcal{A}(\mathbf{v}) \in \langle \mathbf{v} \rangle$, 所以存在 $\lambda \in F$ 使得 $\mathcal{A}(\mathbf{v}) = \lambda \mathbf{v}$.

(iii) $\Longrightarrow$ (i) 设 $\mathbf{x} \in \langle \mathbf{v} \rangle$. 则存在 $\alpha \in F$ 使得 $\mathbf{x} = \alpha \mathbf{v}$. 于是,

$$\mathcal{A}(\mathbf{x}) = \mathcal{A}(\alpha \mathbf{v}) = \alpha \mathcal{A}(\mathbf{v}) = \alpha \lambda \mathbf{v} \in \langle \mathbf{v} \rangle \Longrightarrow \langle \mathbf{v} \rangle \text{ 是 } \mathcal{A} \text{ 不变的}. \quad \square$$

从上述命题可知 $\mathbf{v}$ 是 $\mathcal{A}$ 特征向量当且仅当存在 $\lambda \in F$ 使得 $\mathcal{A}(\mathbf{v}) = \lambda \mathbf{v}$. 我们称 λ 是关于特征向量 $\mathbf{v}$ 的特征值 (eigenvalue). 简称 $\mathcal{A}$ 的特征根. 反之, 设 $\lambda \in F$ 是 $\mathcal{A}$ 的特征值. 令

$$V^{\lambda} = \{\mathbf{x} \in V \mid \mathcal{A}(\mathbf{x}) = \lambda \mathbf{x}\}$$

称为 $\mathcal{A}$ 关于 λ 的特征子空间(eigenspace). 下面我们来验证 V^{λ} 是 $\mathcal{A}$-子空间.

设 $\alpha, \beta \in F$, $\mathbf{x}, \mathbf{y} \in V^\lambda$. 则

$$\mathcal{A}(\alpha\mathbf{x}+\beta\mathbf{y}) = \alpha\mathcal{A}(\mathbf{x})+\beta\mathcal{A}(\mathbf{y}) = \alpha\lambda\mathbf{x}+\beta\lambda\mathbf{y} = \lambda(\alpha\mathbf{x}+\beta\mathbf{y}).$$

由此可知 $\alpha\mathbf{x}+\beta\mathbf{y} \in V^\lambda$. 即 V^λ 是子空间. 因为

$$\mathcal{A}(\mathbf{x}) = \lambda\mathbf{x} \in V^\lambda,$$

所以 V^λ 是 $\mathcal{A}$ 不变的.

例 7.3 设 $\mathcal{D}$ 是 $\mathbb{R}[x]_n$ 中的导数算子. 求 $\mathcal{D}$ 所有特征值和特征向量.

解. 设 $f = f_{n-1}x^{n-1}+\cdots+f_1x+f_0$, 其中 $f_{n-1}, \ldots, f_1, f_0 \in \mathbb{R}$. 如果 $\mathcal{D}(f) = \lambda f$, 其中 $\lambda \in \mathbb{R}$. 则

$$(n-1)f_{n-1}x^{n-2}+\cdots+f_1 = \lambda(f_{n-1}x^{n-1}+\cdots+f_1x+f_0).$$

上式成立当且仅当 $\lambda = 0$ 且 $f_{n-1} = \cdots = f_1 = 0$. 注意到对任意 $r \in \mathbb{R}$, $\mathcal{D}(r) = 0 = 0r$. 于是, f 是 $\mathcal{D}$ 的特征向量当且仅当 $f \in \mathbb{R}\setminus\{0\}$, 这些特征向量对应的特征值是 0. 而 $V^0 = \mathbb{R}$. □

当我们把矩阵 $A \in \mathrm{M}_n(F)$ 看成 $\mathcal{L}(F^n)$ 中由 $A(\mathbf{x}) = A\mathbf{x}$ 定义得线性算子时, 我们同样有矩阵 A 的特征向量, 特征值和特征子空间的概念.

例 7.4 求数乘矩阵的特征向量和特征值.

解. 设 $A = \lambda E$, 其中 $\lambda \in F$. 则对任意 $\mathbf{x} \in F^n$,

$$A\mathbf{x} = \lambda E\mathbf{x} = \lambda\mathbf{x}.$$

于是, 任何 F^n 中的非零向量都是 A 的特征向量, 它们对应的特征值都是 λ. 进而, $V^\lambda = F^n$. □

例 7.5 设 $\mathcal{A} \in \mathcal{L}(V)$ 且 $\ker(\mathcal{A}) \neq \{\mathbf{0}\}$. 证明: 0 是 $\mathcal{A}$ 的特征值且 $V^0 = \ker(\mathcal{A})$.

证明. 设 $\mathbf{v} \in \ker(\mathcal{A}) \setminus \{\mathbf{0}\}$. 则 $\mathcal{A}(\mathbf{v}) = \mathbf{0} = 0\mathbf{v}$. 于是, 0 是 $\mathcal{A}$ 的特征值. 设 $\mathbf{x} \in \ker(\mathcal{A})$. 则 $\mathcal{A}(\mathbf{x}) = 0\mathbf{x}$. 于是, $\mathbf{x} \in V^0$. 反之, 设 $\mathbf{y} \in V^0$. 则 $\mathcal{A}(\mathbf{y}) = 0\mathbf{y} = \mathbf{0}$. 于是 $\mathbf{y} \in \ker(\mathcal{A})$. 由此得出 $V^0 = \ker(\mathcal{A})$. □

7.2 特征多项式

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. 设 $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$, 其中 $x_1, \ldots, x_n \in F$ 不全等于零. 则 $\mathbf{x}$ 是 $\mathcal{A}$ 的特征向量当且仅当存在 $\lambda \in F$ 使得 $\mathcal{A}(\mathbf{x}) = \lambda\mathbf{x}$, 即

$$\lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \iff (\lambda E - A) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

由此推出 $\mathbf{x}$ 是 $\mathcal{A}$ 的特征向量蕴含 $\det(\lambda E - A) = 0$.

设 $\chi_A(t)=\det(tE-A)\in F[t]$. 则 $\mathbf{x}$ 是 $\mathcal{A}$ 的特征向量蕴含着它对应的特征值 λ 是 $\chi_A(t)$ 的根. 反之, 设 $\lambda\in F$ 是 $\chi_A(t)$ 的根. 则方程组

$$(\lambda E-A)\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix}$$

有非零解 $(\alpha_1,\ldots,\alpha_n)$. 于是, $\mathbf{v}=\alpha_1\mathbf{e}_1+\cdots+\alpha_n\mathbf{e}_n$ 满足 $\mathcal{A}(\mathbf{v})=\lambda\mathbf{v}$. 由此推出 $\lambda\in F$ 是 $\chi_A(t)$ 的根当且仅当 λ 是 $\mathcal{A}$ 的特征值.

定义 7.6 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)*

注意到矩阵 A 与基底选取有关. 为了验证上述定义的合理性, 我们设 B 是 $\mathcal{A}$ 在基底 $(\epsilon_1,\ldots,\epsilon_n)=(\mathbf{e}_1,\ldots,\mathbf{e}_n)P$ 下的矩阵, 其中 $P\in\mathrm{GL}_n(F)$. 则 $B=P^{-1}AP$. 我们有

$$\det(tE-B)=\det(tE-P^{-1}AP)=\det(P^{-1}(tE-A)P)=\det(tE-A).$$

这样就验证了上述定义的合理性.

类似地, 设 $A \in \mathrm{M}_n(F)$. 我们称 $\chi_A = \det(tE - A)$ 为矩阵 A 的特征多项式. 上述合理性验证也说明 χ_A 是相似不变量. 从而, spec_A 中的元素都是相似不变量.

注解 7.7 根据上述讨论, $\mathcal{A}$ 的特征值也称为 $\mathcal{A}$ 的特征根*(eigenroot)*.

下面我们演示通过特征多项式求特征向量的方法.

例 7.8 设 $\mathbf{e}_1, \mathbf{e}_2$ 是 $\mathbb{R}^2$ 的标准基, $\mathcal{A}, \mathcal{B} \in \mathcal{L}(\mathbb{R}^2)$ 分别由公式 $\mathcal{A}(\mathbf{e}_1) = \mathbf{e}_2, \mathcal{A}(\mathbf{e}_2) = \mathbf{e}_1$ 和 $\mathcal{B}(\mathbf{e}_1) = \mathbf{e}_2$, $\mathcal{B}(\mathbf{e}_2) = -\mathbf{e}_1$ 给出. 计算 $\mathcal{A}$ 和 $\mathcal{B}$ 的特征子空间.

解. 算子 $\mathcal{A}$ 和 $\mathcal{B}$ 在 $\mathbf{e}_1, \mathbf{e}_2$ 下的矩阵分别是

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

于是

$$\chi_{\mathcal{A}}(t) = \det(tE - A) = \det \begin{pmatrix} t & -1 \\ -1 & t \end{pmatrix} = t^2 - 1,$$

和

$$\chi_{\mathcal{B}}(t) = \det(tE - B) = \det \begin{pmatrix} t & 1 \\ -1 & t \end{pmatrix} = t^2 + 1.$$

于是, $\mathrm{spec}_{\mathbb{R}}(\mathcal{A}) = \{1, -1\}$, $\mathrm{spec}_{\mathcal{B}} = \emptyset$. 从而 $\mathcal{B}$ 没有特征根, 从而没有特征向量和特征子空间.

特征根 $\lambda_1 = 1$, 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

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

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

例 7.9 设 $\mathbf{e}_1, \mathbf{e}_2$ 是 $\mathbb{C}^2$ 的标准基, $\mathcal{B} \in \mathcal{L}(\mathbb{C}^2)$ 由公式 $\mathcal{B}(\mathbf{e}_1) = \mathbf{e}_2$, $\mathcal{B}(\mathbf{e}_2) = -\mathbf{e}_1$ 给出. 计算 $\mathcal{B}$ 的特征子空间.

解. 算子 $\mathcal{B}$ 在 $\mathbf{e}_1, \mathbf{e}_2$ 下的矩阵是

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

于是

$$\chi_{\mathcal{B}}(t) = \det(tE - B) = \det\begin{pmatrix} t & 1 \\ -1 & t \end{pmatrix} = t^2 + 1.$$

于是, $\operatorname{spec}_{\mathbb{C}}(\mathcal{B}) = \{\sqrt{-1}, -\sqrt{-1}\}$. 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 7.10 设 V 是复数域 $\mathbb{C}$ 上的线性空间, $\mathcal{A} \in \mathcal{L}(V)$. 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. 于是有特征向量. □

例 7.11 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. 由上例可知, 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,*

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

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 也是上三角矩阵. 显然, $A \sim_s T$. $\square$

命题 7.12 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 不是 A 的特征根当且仅当 A 可逆. □

注解 7.13 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.

例 7.14 设 $A \in \mathrm{M}_n(F)$ 是如下分块上三角形矩阵

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

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,

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