第二章　线性算子

例 5.5 设

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

计算 μ_A.

解. 由上一讲引理 *5.4*,

$$\mu_A=\mathrm{lcm}(\mu_{(1)},\mu_{(0)})=\mathrm{lcm}(t-1,t)=(t-1)t. \ \square$$

定理 5.6 设 $\mathcal{A}\in\mathcal{L}(V)$, $U_1,\ldots,U_k$ 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, $i=1,\ldots,k$. 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. 由上一讲引理 5.4, $\mathcal{A}$ 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 $\mu_{\mathcal{A}} = \mathrm{lcm}(\mu_{\mathcal{A}_W}, \mu_{\mathcal{A}_{U_k}})$.

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

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

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 5.7 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 $\mathbf{x} \in \operatorname{im}(\mathcal{B})$. 则存在 $\mathbf{y} \in V$ 使得 $\mathbf{x} = \mathcal{B}(\mathbf{y})$. 于是

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

命题 5.8 设 $\mathcal{A} \in \mathcal{L}(V)$, $f \in F[t]$. 则 $\ker(f(\mathcal{A}))$ 和 $\operatorname{im}(f(\mathcal{A}))$ 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 $\operatorname{im}(f(\mathcal{A}))$ 都是 $\mathcal{A}$ 的不变子空间(引理 5.7). $\square$

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 $f = f_{n-1}x^{n-1} + \cdots + f_1 x + f_0$, 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

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

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 $\mathcal{D}$ 得特征向量当且仅当 $f \in \mathbb{R} \setminus \{0\}$, 这些特征向量对应得特征值是 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, $\mathbf{x} \in \ker(\mathcal{A})$. 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

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

由非零解 $(\alpha_1, \ldots, \alpha_n)$. 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

$$(\lambda_1E-A)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

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, $\mathrm{spec}_{\mathbb{C}}(\mathcal{B}) = \{\sqrt{-1}, -\sqrt{-1}\}$. 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 $V^{\lambda_2} = \langle (1, \sqrt{-1})^t \rangle$.