§6.2 对角化

教学目的与要求 掌握矩阵 (线性变换) 可对角化的定义；理解和计算特征值的代数重数与几何重数；掌握可对角化的等价命题；能判断一个矩阵是否可对角化，在可对角化时将其对角化.

一. 可对角化定义

定义 设 φ 是数域 K 上的 n 维线性空间 V 的线性变换，若存在 V 的一组基，使得 φ 在该基下的表示矩阵为对角阵，则称 φ 是可对角化的.

定义' 设 A 是数域 K 上 n 阶方阵，若存在可逆阵 $P \in K^{n\times n}$, 使得 $P^{-1}AP$ 为对角阵，则称 A 是可对角化矩阵.

注 1 若 φ 在某组基下的表示矩阵为对角阵，则对角元在不考虑排列顺序条件下是唯一确定的，他们恰是 $f_\varphi(\lambda)$ 的所有特征值.

注 2 若 A 是可对角化矩阵，即存在可逆阵 $P \in K^{n\times n}$, 使得 $P^{-1}AP$ 为对角阵. 则 P 的列向量恰为 A 的特征向量.

注 3 设 A 在 $\mathbb{C}^{n\times n}$ 上可对角化，则 A 未必在 $K^{n\times n}$ 上可对角化，因为它可能在 K 上没有 n 个特征值. 例 $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ 在 $\mathbb{C}$ 上特征值为 $i, -i$, 但没有实特征值.

二. 特征子空间的直和

命题 设 $\lambda_1, \lambda_2 \cdots, \lambda_k$ 是数域 K 上 n 维线性空间 V 上线性变换 φ 的不同特征值， V_{λ_i} 为 φ 属于特征值 λ_i 的特征子空间，则

$$V_{\lambda_1} + V_{\lambda_2} + \cdots + V_{\lambda_k} = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k}.$$

证明 对 k 用数学归纳法. 若 $k = 1$, 则结论显然成立. 现设结论对 $k-1$ 个特征值 $\lambda_1, \lambda_2, \cdots, \lambda_{k-1}$, 它们相应的特征子空间 $V_{\lambda_1}, V_{\lambda_2}, \cdots + V_{\lambda_{k-1}}$ 之和是直和. 要证 $V_{\lambda_1} + V_{\lambda_2} + \cdots + V_{\lambda_k}$ 是直和，只需证明 $V_{\lambda_k} \cap (V_{\lambda_1} + V_{\lambda_2} + \cdots + V_{\lambda_{k-1}}) = 0$ 即

可. 设 $v\in V_{\lambda_k}\cap(V_{\lambda_1}+V_{\lambda_2}+\cdots+V_{\lambda_{k-1}})$, 则

$$v=v_1+v_2+\cdots+v_{k-1},\ \ v_i\in V_{\lambda_i},\ \ 1\le i\le k-1 \tag{1}$$

将上式两边作用 φ, 得

$$\varphi(v)=\varphi(v_1)+\varphi(v_2)+\cdots+\varphi(v_{k-1}),$$

$$\lambda_k v=\lambda_1 v_1+\lambda_2 v_2+\cdots+\lambda_{k-1}v_{k-1},$$

(1) 两边同乘以 λ_k, 减去上式，得

$$0=(\lambda_k-\lambda_1)v_1+(\lambda_k-\lambda_2)v_2+\cdots+(\lambda_k-\lambda_{k-1})v_{k-1}.$$

由归纳假设， $V_{\lambda_1}+V_{\lambda_2}+\cdots+V_{\lambda_{k-1}}$ 是直和，因此 $(\lambda_k-\lambda_i)v_i=0$, 而 $\lambda_k-\lambda_i\neq 0$, 所以 $v_i=0\,,1\le i\le k-1$, 代入 (1) 式得 $v=0$. 故命题得证. □

命题' $A\in K^{n\times n}$,$\lambda_1,\lambda_2,\cdots,\lambda_k\in K$ 是 A 的互异特征值， V_{λ_i} 为相应特征子空间，则 $V_{\lambda_1}+V_{\lambda_2}+\cdots+V_{\lambda_k}=V_{\lambda_1}\oplus V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_k}$.

推论 1 线性变换 φ 属于不同特征值的特征向量必线性无关.

推论 1' 矩阵 A 属于不同特征值的特征向量必线性无关.

推论 2 线性变换 φ 有 n 个不同特征值，则必存在 V 的某个基，使 φ 在该基下的表示矩阵为对角阵 (φ 必可对角化), 反之未必.

推论 2' 矩阵 A 在 K 上有 n 个不同特征值，则必存在可逆矩阵 P, 使 $P^{-1}AP$ 为对角阵 (A 必可对角化), 反之未必.

三. 线性变换的代数重数与几何重数

命题 设 φ 是 n 维线性空间 V 的线性变换， λ_0 是 φ 的特征值. 设 λ_0 是 φ 的特征多项式的 m 重根， λ_0 的特征子空间 V_{λ_0} 的维数为 t, 则 $t\le m$.

证明 设 $\xi_1,\xi_2,\cdots,\xi_t$ 是 V_{λ_0} 的一组基，由定义知 $\varphi(\xi_i)=\lambda_0\xi_i,1\le i\le t$. 将 $\xi_1,\xi_2,\cdots,\xi_t$ 扩为 V 的一组基 $\xi_1,\xi_2,\cdots,\xi_t,\xi_{t+1},\cdots,\xi_n$, 则 φ 在该基下的表示矩阵为

$$A=\begin{pmatrix}\lambda_0 I_t & A_{12}\\ 0 & A_{22}\end{pmatrix}_{n\times n},$$

从而 $f_\varphi(\lambda)=(\lambda-\lambda_0)^t g(\lambda)$, 即 λ_0 至少是 $f_\varphi(\lambda)$ 的 t 重根. 即 $t\le m$. $\square$

定义 设 φ 是数域 K 上 n 维线性空间 V 上的线性变换， λ_0 是 $f_\varphi(\lambda)$ 的 m 重根， V_0 为 φ 的属于 λ_0 的特征子空间. 则称 m 为 λ_0 的代数重数， $t=\dim V_0$ 称为 λ_0 的几何重数.

四. 可对角化的充分必要条件

定理 设 φ 是数域 K 上 n 维线性空间 V 的线性变换，则下列命题等价：

(1) φ 可对角化；

(2) φ 有 n 个线性无关的特征向量；

(3) $V=V_{\lambda_1}\oplus V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_s}$, 这里 $\lambda_1,\lambda_2\cdots,\lambda_s$ 是 φ 的全部互异特征值；

(4) $\displaystyle\sum_{i=1}^{s}\dim V_{\lambda_i}=n$, 这里 $\lambda_1,\lambda_2\cdots,\lambda_s$ 是 φ 的全部特征值.

(5) φ 的特征多项式的所有根全部在 K 上，且特征值的代数重数等于几何重数.

证明 (1) $\Rightarrow$ (2) 设 φ 可对角化，则 φ 在 V 的某组基 $\xi_1,\xi_2,\cdots,\xi_n$ 下的表示矩阵为对角阵，即

$$\varphi(\xi_1,\xi_2,\cdots,\xi_n)=(\xi_1,\xi_2,\cdots,\xi_n)\begin{pmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_n\end{pmatrix},$$

即 $\varphi(\xi_i)=\lambda_i\xi_i$, $1\le i\le n$. 所以 $\xi_1,\xi_2,\cdots,\xi_n$ 即为 φ 的特征向量且线性无关.

(2) $\Rightarrow$ (3) 由上面的命题知 $V_{\lambda_1}+V_{\lambda_2}+\cdots+V_{\lambda_s}=V_{\lambda_1}\oplus V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_s}$. 因为 (2) 成立，故 $\dim(V_{\lambda_1}\oplus V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_s})=n$, 故 $V=V_{\lambda_1}\oplus V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_s}$.

(3) $\Rightarrow$ (4) 显然.

(4) $\Rightarrow$ (5) (反证法) 设 m_i 为 λ_i 的代数重数， t_i 为 λ_i 的几何重数，则 $m_i\ge t_i$, $1\le i\le s$. $f_\varphi(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_s)^{m_s}g(\lambda)$, 这里 $g(\lambda_i)\ne 0$, $1\le i\le s$. 根据条件有 $\displaystyle\sum_{i=1}^{s}m_i\le n=\sum_{i=1}^{s}\dim V_{\lambda_i}=\sum_{i=1}^{s}t_i\le\sum_{i=1}^{s}m_i$, 所以 $t_i=m_i$ 且 $\displaystyle\sum_{i=1}^{s}m_i=n$. 这样 $f_\varphi(x)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_s)^{m_s}$, 即 φ 的特征多项式的所有根全部在 K 上，且特征值的代数重数等于几何重数.

(5) $\Rightarrow$ (1) 因 $n=\sum_{i=1}^{s} m_i=\sum_{i=1}^{s} t_i$, 由上面的命题知 $V=V_{\lambda_1}\oplus V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_s}$. 取 V_{λ_i} 的一组基 $\alpha_{i1},\alpha_{i2},\cdots,\alpha_{im_i}$,$1\le i\le s$, 则 $\alpha_{11},\cdots,\alpha_{1m_1},\alpha_{21},\cdots,\alpha_{2m_2},\cdots,\alpha_{s1},\cdots,$ α_{sm_s} 恰是 V 的一组基.

$$\varphi(\alpha_{11},\cdots,\alpha_{1m_1},\alpha_{21},\cdots,\alpha_{2m_2},\cdots,\alpha_{s1},\cdots,\alpha_{sm_s})$$

$$=(\alpha_{11},\cdots,\alpha_{1m_1},\alpha_{21},\cdots,\alpha_{2m_2},\cdots,\alpha_{s1},\cdots,\alpha_{sm_s})\begin{pmatrix}\lambda_1 I_{m_1} & & & \\ & \lambda_2 I_{m_2} & & \\ & & \ddots & \\ & & & \lambda_s I_{m_s}\end{pmatrix},$$

即 φ 可对角化. $\square$

注 对 n 阶方阵 A, 有对应的结论.

五. 判定 A 是否可对角化和求可逆阵 P 的方法

步骤 1. 求 $f_\varphi(\lambda)$ 所有互异特征值 λ_i, 重数 m_i, $1\le i\le s$;
步骤 2. 求 V_{λ_i} 的一组基 $\xi_{i1},\cdots,\xi_{it_i}$, $1\le i\le s$;
步骤 3. 若有某 $t_i<m_i$, 则 A 不可对角化，若所有 $t_i=m_i$, 则 A 可对角化;
步骤 4. 令 $P=(\xi_{i1},\cdots,\xi_{it_i},\cdots,\xi_{s1},\cdots,\xi_{st_s})$, 则

$$P^{-1}AP=\begin{pmatrix}\lambda_1 I_{t_1} & & & \\ & \lambda_2 I_{t_2} & & \\ & & \ddots & \\ & & & \lambda_s I_{t_s}\end{pmatrix}.$$

例 1 判断 A 是否相似对角阵，若是，求出可逆阵 P, 使 $P^{-1}AP$ 为对角阵.

$$(1)\quad A=\begin{pmatrix}0&1&0\\0&0&1\\0&0&0\end{pmatrix}\qquad (2)\quad B=\begin{pmatrix}1&0&0\\-2&5&-2\\-2&4&-1\end{pmatrix}$$

解 (1) 因为 $f_A(\lambda)=\begin{vmatrix}\lambda&-1&0\\0&\lambda&-1\\0&0&\lambda\end{vmatrix}=\lambda^3$, 所以 $\lambda_1=0$ 是三重根. 而 $r\begin{pmatrix}0&-1&0\\0&0&-1\\0&0&0\end{pmatrix}=$ 2, 所以特征值 λ_1 的几何重数 (1) 不等于其代数重数 (3), 故 A 不可对角化.

(2) $f_B(\lambda)=\begin{pmatrix}\lambda-1 & 0 & 0\\ 2 & \lambda-5 & 2\\ 2 & -4 & \lambda+1\end{pmatrix}=(\lambda-1)^2(\lambda-3)$, 故有 $\lambda_1=1$ (2 重); $\lambda_2=3$ (1 重).

解 $(\lambda_1-A)X=0$, 即 $\begin{pmatrix}0 & 0 & 0\\ 2 & -4 & 2\\ 2 & -4 & 2\end{pmatrix}X=0$, 得到基础解系 $\xi_{11}=\begin{pmatrix}2\\1\\0\end{pmatrix}$, $\xi_{12}=\begin{pmatrix}-1\\0\\1\end{pmatrix}$.

解 $(\lambda_2-A)X=0$, 即 $\begin{pmatrix}2 & 0 & 0\\ 2 & -2 & 2\\ 2 & -4 & 4\end{pmatrix}X=0$, 得到基础解系 $\xi_{21}=\begin{pmatrix}0\\1\\1\end{pmatrix}$.

故 A 可对角化. 令 $P=(\,\xi_{11}\quad\xi_{12}\quad\xi_{21}\,)$, 则 $P^{-1}AP=\begin{pmatrix}1 & & \\ & 1 & \\ & & 3\end{pmatrix}$.

亦可令 $Q=\begin{pmatrix}1 & -1 & 0\\ 1 & 0 & 1\\ 1 & 1 & 1\end{pmatrix}$, 则 $Q^{-1}AQ=\begin{pmatrix}1 & & \\ & 1 & \\ & & 3\end{pmatrix}$.

例 2 计算 A^{10}, 其中 $A=\begin{pmatrix}1 & 0\\ 1 & -2\end{pmatrix}$.

解 用上例方法求得存在可逆阵 $P=\begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}$, 使得 $P^{-1}AP=\begin{pmatrix}1 & 0\\ 0 & -2\end{pmatrix}$, 故 $A^{10}=\left(P\begin{pmatrix}1 & 0\\ 0 & -2\end{pmatrix}P^{-1}\right)^{10}=P\begin{pmatrix}1 & 0\\ 0 & -2\end{pmatrix}^{10}P^{-1}=\begin{pmatrix}1 & 0\\ 1-2^{10} & 2^{10}\end{pmatrix}$.

例 3 设 3 阶矩阵 A 的特征值为 1,1,3, 对应的特征向量依次为 $(2,1,0)'$, $(-1,0,1)'$, $(0,1,1)'$. 求矩阵 A.

解 令 $P=\begin{pmatrix}2 & -1 & 0\\ 1 & 0 & 1\\ 0 & 1 & 1\end{pmatrix}$, $B=\begin{pmatrix}1 & & \\ & 1 & \\ & & 3\end{pmatrix}$. 则 $P^{-1}AP=B$. 从而 $A=PBP^{-1}=\begin{pmatrix}1 & 0 & 0\\ -2 & 5 & -2\\ -2 & 4 & -1\end{pmatrix}$.

例 4 求证:

(1) 若矩阵 A 适合 $A^2=I_n$, 则 A 必可对角化;

(2) 若矩阵 A 适合 $A^2=A$, 则 A 必可对角化;

(3) 若 A 是非零矩阵且 $A^k=0$, 则 A 必不能对角化;

(4) 若实矩阵 A 适合 $A^2+A+I_n=0$, 则 A 在实数域上不可对角化.

证明 (1) 因为 $A^2=I_n$, 即 $(I+A)(I-A)=0$, 所以 $r(I+A)+r(I-A)\le n$. 另一方面， $r(I-A)+r(I+A)\ge r(2I)=n$. 因此 $r(I+A)+r(I-A)=n$. 注意到 $\mathrm{dim}V_1=n-r(I-A)$, $\mathrm{dim}V_{-1}=n-r(I+A)$, 故 $\mathrm{dim}V_1+\mathrm{dim}V_{-1}=n$. 所以 A 可对角化.

(2) 同 (1) 类似讨论.

(3) 因为 $A^k=0$, 所以若 $AX=\lambda X, X\neq 0$, 则 $\lambda=0$. 若 A 可对角化，则 A 相似与 0, 从而 $r(A)=0$, 与题设 $A\neq 0$ 矛盾.

(4) 因 A 适合 $A^2+A+I_n=0$, 若 $X\neq 0, AX=\lambda X$, 则 $\lambda^2+\lambda+1=0$, 而 $\lambda^2+\lambda+1=0$ 无实根，故 A 无实根，从而在实数域上不可对角化. □

作业 P_{241} 4, 5, 9, 10(2-4), 11, 12; P_{249} 7, 8.

思考题 P_{241} 7，8.

选做题 P_{251} 17，18.

挑战题 P_{229} 15.