

0.1 对角化

定义 0.1 (可对角化线性变换)

若 n 维线性空间 V 上的线性变换 φ 在某组基 $\{e_1, e_2, \dots, e_n\}$ 下的表示矩阵为对角阵:

$$\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

则称 φ 为**可对角化线性变换**.

定理 0.1 (线性变换可对角化的充要条件)

设 φ 是 n 维线性空间 V 上的线性变换, 则 φ 可对角化的充分必要条件是 φ 有 n 个线性无关的特征向量。

证明 若 φ 是 V 上可对角化线性变换, 则可设 φ 在某组基 $\{e_1, e_2, \dots, e_n\}$ 下的表示矩阵为对角阵:

$$\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

此时 $\varphi(e_i) = \lambda_i e_i$, 即 e_1, e_2, \dots, e_n 是 φ 的特征向量, 于是 φ 有 n 个线性无关的特征向量。

反过来, 若 n 维线性空间 V 上的线性变换 φ 有 n 个线性无关的特征向量 e_1, e_2, \dots, e_n , 则这组向量构成了 V 的一组基, 且 φ 在这组基下的表示矩阵显然是一个对角阵。□

定义 0.2 (可对角化矩阵)

设 A 是 n 阶矩阵, 若 A 相似于对角阵, 即存在可逆阵 P , 使 $P^{-1}AP$ 为对角阵, 则称 A 为**可对角化矩阵**.

引理 0.1

设 A 是 n 阶矩阵, φ 是线性空间 V 上由矩阵 A 乘法诱导的线性变换, 即 $\varphi(\alpha) = A\alpha, \forall \alpha \in V$ 。设 $\{e_1, e_2, \dots, e_n\}$ 是 V 的一组标准基, 则 φ 在这组基下的矩阵就是 A 。证明:

- (1) 矩阵 A 与线性变换 φ 的特征值相同;
- (2) 矩阵 A 可对角化等价于线性变换 φ 可对角化。

证明

- (1) 若 λ 为矩阵 A 的特征值, 则存在 $\xi \in V$, 使得 $\varphi(\xi) = A\xi = \lambda\xi$, 因此矩阵 A 的特征值也是线性变换 φ 的特征值。

若 λ 为线性变换 φ 的特征值, 则存在 $\eta \in V$, 使得 $\varphi(\eta) = A\eta = \lambda\eta$, 因此线性变换 φ 的特征值也是矩阵 A 的特征值。

故矩阵 A 与线性变换 φ 的特征值相同。

- (2) 若矩阵 A 可对角化, 则存在可逆矩阵 P , 使得 $P^{-1}AP$ 为对角矩阵。

从而 $(e_1, e_2, \dots, e_n)P$ 的列向量也是 V 的一组基, 于是由命题?? 可知 φ 在这组基下的矩阵为 $P^{-1}AP$ 是对角矩阵, 故 φ 也可对角化。

若线性变换 φ 可对角化, 则存在 V 的一组基 $\{f_1, f_2, \dots, f_n\}$, 使得 φ 在这组基下的矩阵 B 为对角矩阵。设基 $\{e_1, e_2, \dots, e_n\}$ 到基 $\{f_1, f_2, \dots, f_n\}$ 的过渡矩阵为 G , 则由命题?? 可知 $B = G^{-1}AG$ 。因此矩阵 A 也可对角化。

故矩阵 A 可对角化等价于线性变换 φ 可对角化。

□

定理 0.2 (矩阵可对角化的充要条件)

设 A 是 n 阶矩阵, 则 A 可对角化的充分必要条件是 A 有 n 个线性无关的特征向量.

♡

证明 设 φ 是线性空间 V 上由矩阵 A 乘法诱导的线性变换.

若矩阵 A 有 n 个线性无关的特征值, 则由引理??(1) 可知线性变换 φ 也有相同的 n 个线性无关的特征值, 于是由定理??可知线性变换 φ 可对角化, 从而再由引理??(2) 可知矩阵 A 也可对角化.

若矩阵 A 可对角化, 则由引理??(2) 可知线性变换 φ 也可对角化, 从而由定理??可知 φ 有 n 个线性无关的特征值, 于是由引理??(1) 可知矩阵 A 也有相同的 n 个线性无关的特征值. □

定理 0.3

若 $\lambda_1, \lambda_2, \dots, \lambda_k$ 为 n 维线性空间 V 上的线性变换 φ 的不同的特征值, 记 λ_i 的特征子空间为 $V_i (1 \leq i \leq k)$, 则

$$V_1 + V_2 + \dots + V_k = V_1 \oplus V_2 \oplus \dots \oplus V_k.$$

♡

证明 对 k 用数学归纳法. 若 $k = 1$, 结论显然成立. 现设对 $k - 1$ 个不同的特征值 $\lambda_1, \lambda_2, \dots, \lambda_{k-1}$, 它们相应的特征子空间 V_1, V_2, \dots, V_{k-1} 之和是直和. 我们要证明 $V_1, V_2, \dots, V_{k-1}, V_k$ 之和为直和, 这只需证明:

$$V_k \cap (V_1 + V_2 + \dots + V_{k-1}) = 0. \quad (1)$$

即可. 设 $v \in V_k \cap (V_1 + V_2 + \dots + V_{k-1})$, 则

$$v = v_1 + v_2 + \dots + v_{k-1}, \quad (2)$$

其中 $v_i \in V_i (i = 1, 2, \dots, k - 1)$. 在(??)式两边作用 φ , 得

$$\varphi(v) = \varphi(v_1) + \varphi(v_2) + \dots + \varphi(v_{k-1}). \quad (3)$$

但 $v, v_1, v_2, \dots, v_{k-1}$ 都是 φ 的特征向量或零向量, 因此

$$\lambda_k v = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_{k-1} v_{k-1}. \quad (4)$$

在(??)式两边乘以 λ_k 减去(??)式得

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

由于 v_1, v_2, \dots, v_{k-1} 是直和, 因此 $(\lambda_k - \lambda_i)v_i = 0$, 而 $\lambda_k - \lambda_i \neq 0$, 从而 $v_i = 0 (i = 1, 2, \dots, k - 1)$. 这就证明了(??)式. □

推论 0.1

线性变换 φ 属于不同特征值的特征向量必线性无关.

♡

证明 设 $\lambda_1, \lambda_2, \dots, \lambda_k$ 是线性变换 φ 的 k 个不同特征值, 由定理??可知 $V_{\lambda_1} \oplus V_{\lambda_2} \oplus \dots \oplus V_{\lambda_k}$. 于是任取 $\alpha_i \in V_{\lambda_i} (1 \leq i \leq k)$ 且 $\alpha_i \neq 0$, 假设 $\alpha_1, \alpha_2, \dots, \alpha_k$ 线性无关, 则存在一组不全为零的数 b_1, b_2, \dots, b_k , 使得

$$b_1 \alpha_1 + b_2 \alpha_2 + \dots + b_k \alpha_k = 0.$$

不妨设 $b_1 \neq 0$, 则

$$\alpha_1 = -\frac{b_2}{b_1} \alpha_2 - \frac{b_3}{b_1} \alpha_3 - \dots - \frac{b_k}{b_1} \alpha_k \in V_{\lambda_1} \cap (V_{\lambda_2} \oplus \dots \oplus V_{\lambda_k}).$$


又由 $V_{\lambda_1} \oplus V_{\lambda_2} \oplus \dots \oplus V_{\lambda_k}$ 及直和的等价条件可知,

$$V_{\lambda_1} \cap (V_{\lambda_2} \oplus \dots \oplus V_{\lambda_k}) = \{0\},$$

从而 $\alpha_1 = 0$, 这与 $\alpha_i \neq 0 (1 \leq i \leq k)$ 矛盾! □

推论 0.2

若 n 维线性空间 V 上的线性变换 φ 有 n 个不同的特征值, 则 φ 必可对角化.

 **笔记** 注意这个推论只是可对角化的充分条件而非必要条件, 比如说纯量变换 $\varphi = cI_V$ 当然可对角化, 但 φ 的 n 个特征值都是 c .

证明 设 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是线性变换 φ 的 n 个不同特征值, 则任取 $\alpha_i \in V_{\lambda_i}$ ($1 \leq i \leq n$), 由推论??可知 $\alpha_1, \alpha_2, \dots, \alpha_n$ 一定线性无关. 从而由定理??可知, φ 一定可对角化. \square

定理 0.4 (线性变换可对角化的充要条件)

设 φ 是 n 维线性空间 V 上的线性变换, $\lambda_1, \lambda_2, \dots, \lambda_k$ 是 φ 的全部不同的特征值, V_i ($i = 1, 2, \dots, k$) 是特征值 λ_i 的特征子空间, 则 φ 可对角化的充要条件是

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k.$$

证明 先证充分性. 设

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k,$$


分别取 V_i 的一组基 $\{e_{i1}, e_{i2}, \dots, e_{it_i}\}$ ($i = 1, 2, \dots, k$), 则由直和的等价条件 (4) 知这些向量拼成了 V 的一组基, 并且它们都是 φ 的特征向量. 因此 φ 有 n 个线性无关的特征向量, 从而定理??可知 φ 可对角化.

再证必要性. 设 φ 可对角化, 则由定理??可知 φ 有 n 个线性无关的特征向量 $\{e_1, e_2, \dots, e_n\}$, 它们构成了 V 的一组基. 不失一般性, 可设这组基中前 t_1 个是关于特征值 λ_1 的特征向量; 接下去的 t_2 个是关于特征值 λ_2 的特征向量; \dots ; 最后 t_k 个是关于特征值 λ_k 的特征向量. 对任一 $\alpha \in V$, 设 $\alpha = a_1 e_1 + a_2 e_2 + \cdots + a_n e_n$, 则 α 可写成 V_1, V_2, \dots, V_k 中向量之和, 因此由定理??可知

$$V = V_1 + V_2 + \cdots + V_k = V_1 \oplus V_2 \oplus \cdots \oplus V_k.$$


 \square **定义 0.3 (线性变换的几何重数与代数重数)**

设 λ_0 是 n 维线性空间 V 上的线性变换 φ 的一个特征值, V_0 是属于 λ_0 的特征子空间, 称 $\dim V_0$ 为 λ_0 的**度数或几何重数**. λ_0 作为 φ 的特征多项式根的重数称为 λ_0 的**重数或代数重数**.

 **笔记** 由线性映射的维数公式可知, 特征值 λ_0 的度数 $\dim V_0 = \dim \text{Ker}(\lambda_0 I_V - \varphi) = n - r(\lambda_0 I_V - \varphi)$, 而特征值 λ_0 的重数则由特征多项式 $|\lambda I_V - \varphi|$ 的因式分解决定.

定义 0.4 (矩阵的几何重数与代数重数)

设 λ_0 是 n 阶方阵的 A 的一个特征值, V_0 是属于 λ_0 的特征子空间, 称 $\dim V_0$ 为 λ_0 的**度数或几何重数**. λ_0 作为 A 的特征多项式根的重数称为 λ_0 的**重数或代数重数**.

 **笔记** 由线性方程组的理论可知, 特征值 λ_0 的度数 $\dim V_0 = n - r(\lambda_0 I_n - A)$, 若将 A 看作由矩阵 A 乘法诱导的 V 上的线性变换, 则由线性变换的维数公式可知 $\dim V_0 = \dim \text{Ker}(\lambda_0 I_V - A) = n - r(\lambda_0 I_V - A)$. 而特征值 λ_0 的重数则由特征多项式 $|\lambda I_n - A|$ 的因式分解决定.

引理 0.2 (特征值的几何重数总小于代数重数)

设 φ 是 n 维线性空间 V 上的线性变换, λ_0 是 φ 的一个特征值, 则 λ_0 的度数总是小于等于 λ_0 的重数.

证明 设特征值 λ_0 的重数为 m , 度数为 t , 又 V_0 是属于 λ_0 的特征子空间, 则 $\dim V_0 = t$. 设 $\{e_1, \dots, e_t\}$ 是 V_0 的一组基. 由于 V_0 中的非零向量都是 φ 关于 λ_0 的特征向量, 故

$$\varphi(e_i) = \lambda_0 e_i, \quad i = 1, \dots, t.$$

将 $\{e_1, \dots, e_t\}$ 扩充为 V 的一组基, 记为 $\{e_1, \dots, e_t, e_{t+1}, \dots, e_n\}$, 则 φ 在这组基下的表示矩阵为

$$A = \begin{pmatrix} \lambda_0 I_t & * \\ O & B \end{pmatrix},$$

其中 B 是一个 $n-t$ 阶方阵. 因此, 线性变换 φ 的特征多项式具有如下形式:

$$|\lambda I_V - \varphi| = |\lambda I_n - A| = (\lambda - \lambda_0)^t |\lambda I_{n-t} - B|,$$

这表明 λ_0 的重数至少为 t , 即 $t \leq m$. □

定义 0.5 (完全的特征向量系)

设 λ_0 是 φ (或 A) 的 m 重特征值, 即它是 φ (或 A) 的特征多项式的 m 重根. 此时若有 $m = \dim V_{\lambda_0}$, 即 λ_0 的代数重数和几何重数相等, 则称 λ_0 有**完全的特征向量系**. 若对 φ (或 A) 的任一特征值, 其代数重数和几何重数都相等, 则称 φ (或 A) 有**完全的特征向量系**. ♣

定理 0.5 (线性变换可对角化的充要条件)

设 φ 是 n 维线性空间 V 上的线性变换, 则 φ 可对角化的充分必要条件是 φ 有完全的特征向量系. ♡

证明 设 $\lambda_1, \lambda_2, \dots, \lambda_k$ 是 φ 的全部不同的特征值, 它们对应的特征子空间、重数和度数分别记为 $V_i, m_i, t_i (i = 1, 2, \dots, k)$. 由重数的定义以及引理??可知 $m_1 + m_2 + \dots + m_k = n, t_i \leq m_i, i = 1, 2, \dots, k$.

由定理??可知, 我们只要证明 φ 有完全的特征向量系当且仅当 $V = V_1 \oplus V_2 \oplus \dots \oplus V_k$.

若 $V = V_1 \oplus V_2 \oplus \dots \oplus V_k$, 则

$$\begin{aligned} n = \dim V &= \dim(V_1 \oplus V_2 \oplus \dots \oplus V_k) \\ &= \dim V_1 + \dim V_2 + \dots + \dim V_k \\ &= \sum_{i=1}^k t_i \leq \sum_{i=1}^k m_i = n, \end{aligned}$$

因此 $t_i = m_i (i = 1, 2, \dots, k)$, 即 φ 有完全的特征向量系. 反过来, 若 φ 有完全的特征向量系, 则

$$\dim(V_1 \oplus V_2 \oplus \dots \oplus V_k) = \sum_{i=1}^k t_i = \sum_{i=1}^k m_i = n = \dim V,$$

又 $V_1 \oplus V_2 \oplus \dots \oplus V_k \subset V$, 故 $V = V_1 \oplus V_2 \oplus \dots \oplus V_k$ 成立. □

定理 0.6 (线性变换可对角化的充要条件汇总)

几何形式:

- (1) 设 φ 是 n 维线性空间 V 上的线性变换, 则 φ 可对角化的充分必要条件是 φ 有 n 个线性无关的特征向量.
- (2) 设 φ 是 n 维线性空间 V 上的线性变换, $\lambda_1, \lambda_2, \dots, \lambda_k$ 是 φ 的全部不同的特征值, $V_i (i = 1, 2, \dots, k)$ 是特征值 λ_i 的特征子空间, 则 φ 可对角化的充要条件是

$$V = V_1 \oplus V_2 \oplus \dots \oplus V_k.$$

- (3) 设 φ 是 n 维线性空间 V 上的线性变换, 则 φ 可对角化的充分必要条件是 φ 有完全的特征向量系.

代数形式:

- (1) 设 A 是 n 阶矩阵, 则 A 可对角化的充分必要条件是 A 有 n 个线性无关的特征向量.
- (2) 设 A 是 n 阶矩阵, $\lambda_1, \lambda_2, \dots, \lambda_k$ 是 A 的全部不同的特征值, $V_i (i = 1, 2, \dots, k)$ 是特征值 λ_i 的特征子空间, 则 A 可对角化的充要条件是

$$V = V_1 \oplus V_2 \oplus \dots \oplus V_k.$$

- (3) 设 A 是 n 阶矩阵, 则 A 可对角化的充分必要条件是 A 有完全的特征向量系. ♡

证明 几何形式:

- (1) 证明见定理??.
- (2) 证明见定理??.
- (3) 证明见定理??.

代数形式: 由几何形式的结论及引理??立即得到证明.

□