



合肥工业大学
HEFEI UNIVERSITY OF TECHNOLOGY

线性代数

张神星 (合肥工业大学)

办公室: 翡翠科教楼 B1810 东

Email: zhangshenxing@hfut.edu.cn

课件地址: <https://zhangshenxing.github.io>

第三章 相似和合同

- ① 方阵的相似
- ② 实对称阵的正交合同
- ③ 实对称阵的合同

第一节 方阵的相似

- 特征值与特征向量
- 特征值和特征向量的性质
- 方阵的相似
- 相似矩阵的性质
- 相似对角化

设 $Ax = \lambda x$, 则

$$(A - \lambda E)x = 0.$$

该方程有非零解当且仅当

$$|\mathbf{A} - \lambda \mathbf{E}| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0.$$

注意到该行列式是 λ 的 n 次多项式, 最高项为 $(-1)^n \lambda^n$. 称之为 A 的特征多项式.

在复数域中 n 次多项式总有 n 个根 (计算重数), 也就是说 A 的特征多项式可以写成

$$f(\lambda) = \pm(\lambda - \lambda_1) \cdots (\lambda - \lambda_n).$$

所以 A 的特征值有 n 个 (计算重数).

特征值和特征向量只对方阵存在, 且有如下性质:

- (1) 零向量不是特征向量;
- (2) 若 x 是对应 λ 的特征向量, 则它的非零倍也是;
- (3) 若 $x_1, x_2 \neq -x_1$ 是对应 λ 的特征向量, 则 $x_1 + x_2$ 也是;
- (4) $|A| = 0 \iff 0$ 是特征值; $|A| \neq 0 \iff 0$ 不是特征值;
- (5) 若 n 阶方阵 A 的各行元素之和为 k , 则 k 是 A 的一个特征值, 且特征向量为 $(1, \dots, 1)^T$.

典型例题：求特征值和特征向量

例

求 $A = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}$ 的特征值和特征向量.

解

(1) 特征多项式 $|\mathbf{A} - \lambda \mathbf{E}| = \begin{vmatrix} 1 - \lambda & 3 \\ 4 & 2 - \lambda \end{vmatrix} = \lambda^2 - 3\lambda - 10$.

(2) 由 $\lambda^2 - 3\lambda - 10 = 0$ 解得特征值 $\lambda_1 = 5, \lambda_2 = -2$.

(3) 对于 $\lambda_1 = 5$, $A - 5E = \begin{pmatrix} -4 & 3 \\ 4 & -3 \end{pmatrix} \sim \begin{pmatrix} 1 & -3/4 \\ 0 & 0 \end{pmatrix}$, 得到基础解系 $\begin{pmatrix} 3/4 \\ 1 \end{pmatrix}$. 故对应的所有特征向量为 $k(3, 4)^T, k \neq 0$.

(4) 对于 $\lambda_2 = -2$, $A + 2E = \begin{pmatrix} 3 & 3 \\ 4 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, 得到基础解系 $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. 故对应的所有特征向量为 $k(1, -1)^T, k \neq 0$.

例：求特征值和特征向量

例

求上三角阵 $A = \begin{pmatrix} a_1 & b & c \\ 0 & a_2 & d \\ 0 & 0 & a_3 \end{pmatrix}$ 的特征值.

解

特征多项式

$$|\mathbf{A} - \lambda \mathbf{E}| = \begin{vmatrix} a_1 - \lambda & b & c \\ 0 & a_2 - \lambda & d \\ 0 & 0 & a_3 - \lambda \end{vmatrix} = (a_1 - \lambda)(a_2 - \lambda)(a_3 - \lambda).$$

因此特征值为 a_1, a_2, a_3 .

上三角阵、下三角阵、对角阵的特征值就是对角元.

典型例题：求特征值和特征向量

例

求 $A = \begin{pmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}$ 的特征值和特征向量.

解

由特征多项式

$$|\mathbf{A} - \lambda \mathbf{E}| = \begin{vmatrix} -1 - \lambda & 1 & 0 \\ -4 & 3 - \lambda & 0 \\ 1 & 0 & 2 - \lambda \end{vmatrix} = (2 - \lambda) \begin{vmatrix} -1 - \lambda & 1 \\ -4 & 3 - \lambda \end{vmatrix} = (2 - \lambda)(\lambda - 1)^2 = 0$$

得到特征值 $\lambda = 1, 1, 2$.

续解

对于 $\lambda_1 = 1$,

$$\mathbf{A} - \mathbf{E} = \begin{pmatrix} -2 & 1 & 0 \\ -4 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} \stackrel{r}{\sim} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix},$$

得到基础解系 $(-1, -2, 1)^T$. 故对应的所有特征向量为 $k(-1, -2, 1)^T, k \neq 0$.

对于 $\lambda_2 = 2$,

$$\mathbf{A} - 2\mathbf{E} = \begin{pmatrix} -3 & 1 & 0 \\ -4 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

得到基础解系 $(0, 0, 1)^T$. 故对应的所有特征向量为 $k(0, 0, 1)^T, k \neq 0$.

答案

- (1) 特征值 $\lambda = -1, 2, 2$.
- (2) -1 对应的所有特征向量为 $k(1, 0, 1)^T, k \neq 0$.
- (3) 2 对应的所有特征向量为 $k_1(0, 1, -1)^T + k_2(1, 0, 4)^T, k_1, k_2$ 不全为零.

9 / 78

通过

$$|\mathbf{A} - \lambda \mathbf{E}| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix}$$

的展开可以看出, 若 $i \neq j$, 则 i 行 j 列的代数余子式中最多只会出现 λ^{n-2} 项. 所以

$$\begin{aligned} |\mathbf{A} - \lambda \mathbf{E}| &= (a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda) + \text{至多 } n-2 \text{ 次项} \\ &= (-1)^n (\lambda^n - (a_{11} + a_{22} + \cdots + a_{nn})\lambda^{n-1}) + \text{至多 } n-2 \text{ 次项}. \end{aligned}$$

根据韦达定理, 特征值之和为 $a_{11} + a_{22} + \cdots + a_{nn}$.

$\lambda = 0$ 时, 特征多项式

$$f(\lambda) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$$

的取值 $|A|$ 就是特征值的乘积.



11 / 78

- (1) 特征值之和等于迹: $\lambda_1 + \lambda_2 + \cdots + \lambda_n = \text{Tr}(\mathbf{A})$;
- (2) 特征值之积等于行列式: $\lambda_1 \lambda_2 \cdots \lambda_n = |\mathbf{A}|$.

例

$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ 的特征值为 (C).

- (A) $1, 0, 1$ (B) $1, 1, 2$ (C) $-1, 1, 2$ (D) $-1, 1, 1$

对应特征向量, 则下述矩阵有如下对应的特征值与特征向

方阵	kA	A^m	A^{-1}	A^*	$g(A)h(A)^{-1}$	A^T	$P^{-1}AP$
特征值	$k\lambda$	λ^m	λ^{-1}	$ A /\lambda$	$g(\lambda)/h(\lambda)$	λ	λ
对应特征向量	x	x	x	x	x	未必是 x	$P^{-1}x$

这里 g, h 是多项式, 且满足 $h(A)$ 可逆.

由此可知, 若 $g(\mathbf{A}) = \mathbf{O}$, 则 \mathbf{A} 的所有特征值 λ 均满足 $g(\lambda) = 0$.

例

设 3 阶方阵 A 的特征值为 2, 2, 3, 则

- (1) A^2 的特征值为 4, 4, 9;
- (2) $A^2 - 2A + E$ 的特征值为 1, 1, 4;
- (3) $|A^2 - 2A + E| =$ 4;
- (4) A^{-1} 的特征值为 1/2, 1/2, 1/3;
- (5) A^* 的特征值为 4, 6, 6;
- (6) $A_{11} + A_{22} + A_{33} =$ 16, 其中 A_{ij} 表示 A 的代数余子式.

例：特征值和特征向量的性质

例

设 3 阶方阵 A 的特征值为 $1, -1, 2$. 求 $|A^* + 3A - 2E|$.

解

由于 $|A| = -2$, $A^* = -2A^{-1}$. 于是

$$A^* + 3A - 2E = -2A^{-1} + 3A - 2E$$

的特征值为 $-1, -3, 3$, $|A^* + 3A - 2E| = (-1) \times (-3) \times 3 = 9$.

练习

若 4 阶方阵 A 的特征值为 $1, 2, -2, 0$, 则下列矩阵可逆的是 (C).

- (A) A (B) $A - 2E$ (C) $A + E$ (D) $A - E$

例：特征值和特征向量的性质

例

若 $\alpha = (1, k, 1)^T$ 为 $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ 的逆 A^{-1} 的特征向量, 求 k .

解

α 也是 A 的特征向量.

$$\mathbf{A}\boldsymbol{\alpha} = \begin{pmatrix} k+3 \\ 2k+2 \\ k+3 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ k \\ 1 \end{pmatrix},$$

因此 $\lambda = k + 3, 2k + 2 = (k + 3)k, k^2 + k - 2 = 0, k = -2$ 或 1 .

特征值和特征向量的性质

定理

若 $\lambda_1, \dots, \lambda_m$ 是 A 的 m 个两两不同的特征值, 则其对应的特征向量 $\alpha_1, \dots, \alpha_m$ 线性无关.

即对应于不同特征值的特征向量线性无关.

证明

设 $k_1\alpha_1 + \cdots + k_m\alpha_m = \mathbf{0}$. 左乘 A^k 得到 $k_1\lambda_1^k\alpha_1 + \cdots + k_m\lambda_m^k\alpha_m = \mathbf{0}$. 令 $k = 1, 2, \dots, m-1$, 我们得到

$$(k_1 \alpha_1, \dots, k_m \alpha_m) \begin{pmatrix} 1 & \lambda_1 & \dots & \lambda_1^{m-1} \\ 1 & \lambda_2 & \dots & \lambda_2^{m-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_m & \dots & \lambda_m^{m-1} \end{pmatrix} = \mathbf{O}.$$

注意到第二个方阵的行列式是范德蒙行列式, 当 $\lambda_1, \dots, \lambda_m$ 两两不同时它非零, 从而

$$(k_1 \alpha_1, \dots, k_m \alpha_m) = \mathbf{O}, \quad k_1 = \dots = k_m = 0. \quad \square$$

设 λ_1 对应线性无关的特征向量 α_1, α_2 , λ_2 对应线性无关的特征向量 β_1, β_2 . 若 $\lambda_1 \neq \lambda_2$, 则 $\alpha_1, \alpha_2, \beta_1, \beta_2$ 也是线性无关的. 这是因为

$$\mathbf{A}(k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2) = \lambda_1(k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2),$$

$$\mathbf{A}(\ell_1\boldsymbol{\beta}_1 + \ell_2\boldsymbol{\beta}_2) = \lambda_2(\ell_1\boldsymbol{\beta}_1 + \ell_2\boldsymbol{\beta}_2).$$

若 $k_1\alpha_1 + k_2\alpha_2 + \ell_1\beta_1 + \ell_2\beta_2 = 0$, 同理可证明这些向量都是零向量.

由此也可以知道, 不同特征值的特征向量的线性组合不可能还是特征向量.

设 $f: \mathbb{C}^n \rightarrow \mathbb{C}^n$ 是一个线性映射. 对于任意 $x \in \mathbb{C}^n$, 它可以唯一表达为

$$\mathbf{x} = x_1 \mathbf{e}_1 + \cdots + x_n \mathbf{e}_n.$$

通过将 $f(e_1), \dots, f(e_n)$ 表达为这组**标准基**的线性组合, 我们建立了线性映射 f 与一个 n 阶方阵 A 的联系.

一般的线性空间并没有这样的标准正交基, 或者 \mathbb{C}^n 本身我们也可以选择其它基. 则这种对应会有什么变化呢? 设 $\alpha_1, \dots, \alpha_n$ 是 \mathbb{C}^n 中一组线性无关的向量,

$$P = (\alpha_1, \dots, \alpha_n).$$

则

$$f(\alpha_i) = A\alpha_i = (\alpha_1, \dots, \alpha_n)P^{-1}A\alpha_i,$$

所以 $f(\alpha_i)$ 表达为 $\alpha_1, \dots, \alpha_n$ 线性组合的系数形成的向量是 $P^{-1}A\alpha_i$. 它们构成矩阵

$$(P^{-1}A\alpha_1, \dots, P^{-1}A\alpha_n) = P^{-1}AP.$$

方阵的相似

也就是说, 若我们将线性空间 \mathbb{C}^n 换一组基表达, 线性映射对应的矩阵就会变成 $P^{-1}AP$. 我们称 f 在不同基下矩阵为相似的.

定义

若存在可逆矩阵 P 使得 $B = P^{-1}AP$, 则称方阵 A 与 B 相似.

注意, 相似只对方阵有定义.

命题

方阵的相似满足

- (1) 自反性: A 与自身相似;
- (2) 对称性: 若 A 相似于 B , 则 B 相似于 A ;
- (3) 传递性: 若 A 相似于 B , B 相似于 C , 则 A 相似于 C .

若 A, B 相似 ($B = P^{-1}AP$), 则 A, B 等价 ($B = PAQ$). 反之未必成立, 例如 $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. 这是因为和 $A = E$ 相似的只有它自己.

例

若 3 阶方阵 $A \xrightarrow{r_1 \leftrightarrow r_3} B \xrightarrow{c_1 \leftrightarrow c_3} C$, 则 A 与 C (**C**).

- (A) 等价但不相似 (B) 相似但不等价
(C) 等价而且相似 (D) 既不等价也不相似

定理

若 A 与 B 相似, 则二者的特征多项式相同, 从而特征值也相同.

这是因为若 $P^{-1}AP = B$, 则

$$P^{-1}(A - \lambda E)P = P^{-1}AP - \lambda P^{-1}EP = B - \lambda E.$$

两边取行列式并利用行列式的可乘性得到 $|A - \lambda E| = |B - \lambda E|$.

注意反过来未必成立, 例如 $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. 的特征多项式相同, 但它们不相似.

相似矩阵具有相同的特征值, 但对应的特征向量未必相同.

推论

若 A 与对角阵 $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ 相似, 则 $\lambda_1, \dots, \lambda_n$ 是 A 的 n 个特征值.

- 线性代数 ► 第三章 相似和合同 ► 1 方阵的相似 ► D 相似矩阵的性质

例：相似矩阵的性质

例

若 4 阶方阵 A 与 B 相似, A 的特征值为 $1, 2, 3, 4$, 则 $|B + E| = 120$.

例

设 3 阶可逆阵 A, B 相似, A^{-1} 的特征值为 $1/2, 1/3, 1/4$, 则 $|E - B| = -6$.

例

设矩阵 $A = \begin{pmatrix} 2 & -1 & 4 \\ 0 & a & 7 \\ 0 & 0 & 3 \end{pmatrix}$ 与 $\Lambda = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & b \end{pmatrix}$ 相似, 求 a 与 b 的值.

解

二者特征值相同, 从而 $a = 1, b = 3$.

练习

- (1) 若 3 阶矩阵 A 与 B 相似, 且 $|A - E| = 0, |A + 2E| = 0, |2A - E| = 0$, 则 $|B^{-1} - E| = \underline{3/4}$.
- (2) 若 3 阶矩阵 A 与 B 相似, 且存在非零矩阵 C 使得 $AC = 2C, |A + 2E| = 0, |2A - E| = 0$, 则 $|B^{-1} - E| = \underline{3/4}$.

若 $AB = kB$, 则 B 的每个非零列向量均为 A 的属于特征值 k 的特征向量.

相似对角化的定义

定义

若 n 阶方阵 A 相似于某个对角阵

$$\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_n),$$

则称 A 可(相似)对角化.

若 A 可对角化, 则存在 P 使得 $A = P^{-1}\Lambda P$. 于是 $A^k = P^{-1}\Lambda^k P$.

设 $P = (p_1, \dots, p_n)$, 则

$$P^{-1}AP = \Lambda \iff AP = P\Lambda \iff (Ap_1, \dots, Ap_n) = (\lambda_1 p_1, \dots, \lambda_n p_n),$$

即 $Ap_i = \lambda_i p_i$. 由于 P 可逆, A 拥有 n 个线性无关的特征向量 p_1, \dots, p_n . 反之, 若 A 拥有 n 个线性无关的特征向量, 则选择 P 以它们为列向量即可使 A 对角化.

定理

n 阶矩阵 A 可对角化 $\iff A$ 有 n 个线性无关的特征向量.

推论

若 A 的特征值两两不同, 则 A 可对角化.

反之未必成立.

例

设 3 阶方阵 B 的特征值为 $1, 2, -1$, $A = B^3 - 2B$, 求 A 的特征值及相似对角阵.

解

A 的特征值为 $-1, 4, 1$, 相似对角阵为 $\text{diag}(-1, 4, 1)$.

回忆 k 重特征值对应的线性无关的特征向量最多 k 个.

定理

若 λ 是 A 的 k 重特征值, 则 A 可对角化 $\iff \forall \lambda, R(A - \lambda E) = n - k$, 即 λ 对应的线性无关的特征向量恰有 k 个.

例：可对角化的刻画

例

判断下列矩阵 A 能否相似对角化: (1) A 是二阶实矩阵且 $|A| < 0$;

$$(2) \mathbf{A} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix}; (3) \mathbf{A} = \begin{pmatrix} -2 & 1 & 1 \\ 0 & 2 & 0 \\ -4 & 1 & 3 \end{pmatrix}; (4) \mathbf{A} = \begin{pmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}.$$

解

(1) 特征值一正一负, 能对角化; (2) 特征值为 $1, 3, 0$, 能对角化;

(3) $(2 - \lambda)(\lambda^2 - \lambda - 2) \implies \lambda = -1, 2, 2$. $\mathbf{A} - 2\mathbf{E} = \begin{pmatrix} -4 & 1 & 1 \\ 0 & 0 & 0 \\ -4 & 1 & 1 \end{pmatrix}$ 秩 1, 能对角化;

(4) $(2 - \lambda)(\lambda^2 - 2\lambda + 2) \implies \lambda = 2, 1, 1$. $\mathbf{A} - \mathbf{E} = \begin{pmatrix} -2 & 1 & 0 \\ -4 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$ 秩 2, 不能对角化.

例：方阵的相似的应用

例

设 $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$, 计算 A^k .

解

A 特征值为 $1, 2$, 对应的特征向量可以取 $p_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, p_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. 设 $P = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, 则

$P^{-1}AP = \begin{pmatrix} 1 & \\ & 2 \end{pmatrix}$. 因此

$$\mathbf{A}^k = \mathbf{P} \begin{pmatrix} 1 & \\ & 2 \end{pmatrix}^k \mathbf{P}^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \\ & 2^k \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2^k - 1 \\ 0 & 2^k \end{pmatrix}.$$

例：对角化的计算

例

$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ 能否对角化? 若能, 求 P 使得 $P^{-1}AP$ 是对角阵.

解

(1) 上三角阵 A 特征值为 $1, 0, 0$.

(2) 对于 $\lambda_1 = 1$, $\mathbf{A} - \mathbf{E} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \sim \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, 取特征向量 $\mathbf{p}_1 = (1, 0, 0)^T$.

(3) 对于 $\lambda_2 = \lambda_3 = 0$, A 对应的基础解系可以取 $p_2 = (-1, 1, 0)^T, p_3 = (-1, 0, 1)^T$.

(4) 因此 A 可对角化, 取 $P = \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, 则 $P^{-1}AP = \text{diag}(1, 0, 0)$.

例：对角化的性质

练习

已知矩阵 $A = \begin{pmatrix} 0 & 0 & 1 \\ a & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ 可对角化, 则 -1.

练习

设 $P^{-1}AP = \text{diag}(1, 2, 3)$. 若 $P = (\alpha_1, \alpha_2, \alpha_3), Q = (\alpha_1, \alpha_3, \alpha_2)$, 则 $Q^{-1}AQ = \text{diag}(1, 3, 2)$.

练习

若 $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, 则 (C).

- (A) A, C 相似, B, C 相似
(C) A, C 相似, B, C 不相似

- (B) A, C 不相似, B, C 相似
(D) A, C 不相似, B, C 不相似

任何方阵都相似于约当标准形

$$\text{diag}(\mathbf{J}_{k_1}(\lambda_1), \dots, \mathbf{J}_{k_t}(\lambda_t)),$$

其中 $\mathbf{J}_k(\lambda) = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$ 是 k 阶方阵. 当 $k_1 = \cdots = k_t = 1$ 时就是对角阵.

由此可知, A^* 的所有特征值的就是 A 的 n 个特征值中 $n-1$ 个相乘得到的.

- (1) 当 $|A| \neq 0$ 时, A^* 的所有特征值为 $|A|/\lambda$.
- (2) 当 $|A| = 0$ 且 $\lambda_1 = 0$ 是一重特征值, 则 A^* 唯一的非零特征值为 A 非零特征值之乘积.
- (3) 当 $|A| = 0$ 且 $\lambda_1 = 0$ 是 ≥ 2 重特征值, 则 $A^* = O$.

例：特征值的性质

练习

若 4 阶实矩阵 A^* 的特征值为 $-1, 1, 2, 4$, 则下列矩阵可逆的是 (D).

- (A) $A + 2E$ (B) $A - 2E$ (C) $A + \frac{1}{2}E$ (D) $A - E$

练习

若 n 阶实对称矩阵 A 满足 $A^2 - A = O$, 则下列矩阵不可逆的是 (C).

- (A) $A + 2E$ (B) $A + E$ (C) $A - E$ (D) $A - 2E$

练习

若 3 阶方阵 A 的特征值互不相同且 $|A| = 0$, 则 $R(A) = 2$.

例：对角化的计算

例

设 $A = \begin{pmatrix} 1 & -1 & 1 \\ x & 4 & y \\ -3 & -3 & 5 \end{pmatrix}$ 有 3 个线性无关特征向量, $\lambda = 2$ 是 A 的二重特征值. 求可逆阵 P , 使得 $P^{-1}AP$ 为对角阵.

解

由 $\text{Tr}(\mathbf{A}) = 10$ 可知特征值为 $2, 2, 6$. 由 $R(\mathbf{A} - 2\mathbf{E}) = 1$ 可知 $x = 2, y = -2$.

$$\mathbf{A} - 2\mathbf{E} = \begin{pmatrix} -1 & -1 & 1 \\ x & 2 & y \\ -3 & -3 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \implies \mathbf{p}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{p}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

$$\mathbf{A} - 6\mathbf{E} = \begin{pmatrix} -5 & -1 & 1 \\ 2 & -2 & -2 \\ -3 & -3 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -1/3 \\ 0 & 1 & 2/3 \\ 0 & 0 & 0 \end{pmatrix} \implies \mathbf{p}_3 = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}, \mathbf{P} = \begin{pmatrix} -1 & 1 & 1 \\ 1 & 0 & -2 \\ 0 & 1 & 3 \end{pmatrix}.$$

例：对角化的计算

例

已知 n 阶方阵 A 的各行元素之和为 2, $A \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$, 则 A 相似于 (C).

(A) $\text{diag}(1, 1, 2)$ (B) $\text{diag}(2, 1, 1)$ (C) $\text{diag}(2, 1, -1)$ (D) $\text{diag}(2, -1, -1)$

练习

设 $A = \begin{pmatrix} 2 & a & 2 \\ 5 & b & 3 \\ -1 & 1 & -1 \end{pmatrix}$ 有特征值 ± 1 , 问 A 能否相似对角化?

答案

$|\mathbf{A} \pm \mathbf{E}| = 0 \implies a = -1, b = -3.$ $\text{Tr}(\mathbf{A}) = -2 \implies \lambda_3 = -2,$ 可对角化.

例：对角化的计算

例

设 A 为三阶方阵, $\alpha_1, \alpha_2, \alpha_3$ 是线性无关的三维列向量且 $A\alpha_1 = 2\alpha_1, A\alpha_2 = 3\alpha_2 + 2\alpha_3, A\alpha_3 = 2\alpha_2 + 3\alpha_3$. 求 A 并证明 A 可对角化.

解

$A(\alpha_1, \alpha_2, \alpha_3) = (\alpha_1, \alpha_2, \alpha_3)B, B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 2 \\ 0 & 2 & 3 \end{pmatrix}$. 由于 B 是实对称矩阵, 因此 B 能对角化, 从而 A 也可以.

为什么实对称矩阵一定能对角化? 在下一节中我们将回答这个问题.

设第 n 个月的兔子数量为 a_n , 那么

设 $x_n = \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix}$, 则 $x_{n+1} = Ax_n$, $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$. 通过计算可知

因此 $x_n = A^{n-1}x_1 = P\Lambda^{n-1}P^{-1}x_1$, $a_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right)$,

$$a_{12} \approx \frac{1.618^n}{\sqrt{5}} = 143.9 \approx 144.$$

对于一般的线性递推数列

$$a_{n+k} - c_{k-1}a_{n+k-1} - \cdots - c_1a_{n+1} - c_0a_n = 0,$$

$\mathbf{x}_n = (a_{n+k-1}, \dots, a_{n+1}, a_n)^T$ 满足 $\mathbf{x}_{n+1} = \mathbf{A}\mathbf{x}_n$, 其中 $\mathbf{A} = \begin{pmatrix} c_{k-1} & \cdots & c_1 & c_0 \\ 1 & & & \\ & \ddots & & \\ & & 1 & \end{pmatrix}$ 的特

征多项式为

$$(-1)^k(\lambda^k - c_{k-1}\lambda^{k-1} - \cdots - c_1\lambda - c_0) = 0.$$

若特征值两两不同, 可知 a_n 的通项为 $\lambda_1^n, \dots, \lambda_k^n$ 的线性组合; 若特征值有重数, 由约当块的幂次形式可知

$$a_n = \sum_{\lambda} f_{\lambda}(n)\lambda^n,$$

其中 f_{λ} 是多项式, 且次数不超过 λ 的重数 -1 .

$$y^{(k)}(t) -$$

$x = (y^{(k-1)}, \dots, y', y)^T$ 满足 $x' = Ax$. 于是 “ $x = \exp(At)$ ”.

$$\exp(\mathbf{A}) = \sum_{i=0}^{\infty} \frac{\mathbf{A}^i}{i!}.$$
$$y(t) = \sum_{\lambda} f_{\lambda}(t) e^{\lambda t},$$

线性代数 ▶ 第三章 相似和合同 ▶ 1 方阵的相似 ▶ E 相似对角化

第二节 实对称阵的正交合同

- 实二次型
- 实对称阵和实二次型的合同

定义

若 n 元多项式 $f(x_1, \dots, x_n)$ 满足

$$f(\lambda x_1, \dots, \lambda x_n) = \lambda^2 f(x_1, \dots, x_n), \quad \forall \lambda \in \mathbb{R},$$

则称 f 为二次齐次多项式或实二次型. 它可以写成

$$f(x) = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2 + 2a_{12}x_1x_2 + \cdots + 2a_{n-1,n}x_{n-1}x_n.$$

本课程仅讨论实二次型. 根据定义, f 不能包含一次项和常数项. 若 f 的交叉项 $x_i x_j (i < j)$ 系数均为零, 则称 f 为实二次型的标准形.

实二次型的矩阵形式

设实二次型 f 的 x_i^2 项的系数为 a_{ii} , $x_i x_j (i < j)$ 项的系数为 $2a_{ij}$. 设 $a_{ji} = a_{ij}$, 对称阵 $A = (a_{ij})_n$, 则

$$\begin{aligned}(x_1, \dots, x_n) \mathbf{A} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} &= \left(\sum_{i=1}^n a_{i1} x_i, \dots, \sum_{i=1}^n a_{in} x_i \right) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= \sum_{j=1}^n \left(\sum_{i=1}^n a_{ij} x_i \right) x_j = \sum_{i,j=1}^n a_{ij} x_i x_j = f(x_1, \dots, x_n),\end{aligned}$$

即 $f = \mathbf{x}^T \mathbf{A} \mathbf{x}$, 其中 $\mathbf{x} = (x_1, \dots, x_n)^T$, \mathbf{A} 为实对称阵.

反过来, 任给一个实对称阵 A , 多项式 $f(x) = x^T A x$ 显然满足

$$f(\lambda \mathbf{x}) = (\lambda \mathbf{x})^\top \mathbf{A}(\lambda \mathbf{x}) = \lambda^2 f(\mathbf{x}),$$

故 f 是实二次型. 因此实二次型 f 与对称阵 A 之间存在一一对应的关系.

例：实二次型的矩阵形式

例

写出实二次型 $f(x_1, x_2, x_3) = x_1^2 + 4x_1x_2 + 4x_2^2 + 2x_1x_3 + x_3^2 + 4x_2x_3$ 对应的矩阵.

解

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}.$$

若 f 是标准形, 则 f 对应矩阵 A 是对角阵.

定义

若实方阵 A 满足 $A^T A = E$, 则称 A 为正交阵.

正交阵满足如下性质:

- (1) $A = (\alpha_1, \dots, \alpha_n)$ 是正交阵 $\iff \alpha_1, \dots, \alpha_n$ 是标准正交向量组.
- (2) A 是正交阵 $\iff A^T = A^{-1}$.
- (3) A 是正交阵 $\implies |A| = \pm 1$ 且 A^T, A^{-1}, A^* 都是正交阵.

定义

若 P 为正交阵, 称线性变换 $y = Px$ 为正交变换.

例如, \mathbb{R}^2 上的正交变换就是绕原点的旋转、反射, 以及它们的复合. 由于

$$[Px, Py] = x^T P^T P y = x^T y = [x, y],$$

因此正交变换保持向量的长度和夹角.

定义

- (1) 若存在可逆线性变换 $x = Py$ 使得实二次型 f 在变量 x, y 下的矩阵分别为 A, B , 则称矩阵 A **合同** 或 **相合** 于 B .
- (2) 若 P 是正交阵, 则称矩阵 A **正交合同** 或 **正交相合** 于 B .

若 A 是对称阵, P 可逆, 则 $P^T A P$ 也是对称阵. 由

$$f = x^T A x = y^T P^T A P y = y^T (P^T A P) y$$

可知 A (正交) 合同于 B 当且仅当存在可逆 (正交) 方阵 P 使得 $B = P^T A P$.

命题

对称方阵的 (正交) 合同满足

- (1) 自反性: A 与自身正交合同;
- (2) 对称性: 若 A (正交) 合同于 B , 则 B (正交) 合同于 A ;
- (3) 传递性: 若 A (正交) 合同于 B , B (正交) 合同于 C , 则 A (正交) 合同于 C .

合同、等价、相似有如下关系:

- (1) 若 A, B 合同, 则 A, B 等价, $R(A) = R(B)$. 反之未必.
 (2) 若 A, B 正交合同, 则 A, B 相似. 反之, 若实对称阵 A, B 相似, 则二者正交合同.

实二次型的对角化

定理

对于实对称阵 A , 存在正交阵 P 使得 $P^T A P$ 是对角阵. 从而 A 对应的实二次型在线性变换 $y = Px$ 下变为标准形.

命题

实对称阵的特征值一定是实数, 从而其特征向量均可取实向量.

证明

设 A 是实对称阵, 非零向量 x 满足 $Ax = \lambda x$. 两边取转置和共轭并右乘 x 得到

$$\bar{\lambda} \bar{x}^T x = \bar{x}^T \bar{A}^T x = \bar{x}^T A x = \lambda \bar{x}^T x.$$

显然 $\bar{x}^T x = |x_1|^2 + \cdots + |x_n|^2 > 0$, 因此 λ 是实数. 由于特征向量是方程 $(A - \lambda E)x = 0$ 的解, 系数矩阵是实方阵, 因此特征向量可取实向量. \square

实二次型的对角化的证明

证明 (定理的证明)

归纳证明 A 存在 n 个两两正交的单位特征向量. 假设我们已找到 $k < n$ 个两两正交的单位特征向量 e_1, \dots, e_k , 分别对应特征值 $\lambda_1, \dots, \lambda_k$. 设 V 是与这些向量均正交的实向量全体, 即方程 $(e_1, \dots, e_k)^T x = 0$ 的解空间. 由于系数秩为 k , 存在基础解系 $v_1, \dots, v_{n-k} \in V$. 对任意 i, j ,

$$[e_i, \mathbf{A}v_j] = e_i^T \mathbf{A}v_j = (\mathbf{A}e_i)^T v_j = \lambda_i e_i^T v_j = 0 \implies \mathbf{A}v_j \in V.$$

设 $(n-k)$ 阶矩阵 B 的第 j 列是 Av_i 表示为 v_1, \dots, v_{n-k} 线性组合的系数, 即

$$A(v_1, \dots, v_{n-k}) = (v_1, \dots, v_{n-k})B.$$

设非零向量 x 满足 $Bx = \lambda x$, 则

$$A(v_1, \dots, v_{n-k})x = \lambda(v_1, \dots, v_{n-k})x,$$

即非零向量 $(v_1, \dots, v_{n-k})x$ 是 A 关于 λ 的特征向量, 于是 $\lambda \in \mathbb{R}$ 且可选择 x 使得它是实向量, 它和 e_1, \dots, e_k 均正交. 令 e_{k+1} 为该向量的标准化. 归纳可知 A 存在 n 个两两正交的特征向量, 它们构成的正交阵 $P = (e_1, \dots, e_n)$ 满足题述要求. \square

由于特征值 λ 对应的实特征向量就是 P 中 λ 对应的那些列向量的线性组合, 因此:

推论

实对称阵的不同特征值对应的实特征向量正交.

练习

设 $\alpha_1 = (1, -3, 1)^T$, $\alpha_2 = (1, a, 2)^T$ 是实实对称阵 A 对应特征值 $\lambda_1 = 1$ 和 $\lambda_2 = 2$ 的特征向量, 则 $a = \underline{1}$.

练习

若 3 阶实实对称阵 A 满足 $A^2 = A$, $R(A) = 1$, 则 A 的特征值为 $0, 0, 1$.

例：对阵矩阵的性质

例

设 3 阶实对称阵 A 的特征值为 $6, 3, 3$, 与特征值 6 对应的特征向量为 $\alpha_1 = (1, 1, 1)^T$, 求 A .

解

由于有两个与特征值 3 对应的线性无关特征向量, 因此与 α_1 正交的向量都是与特征值 3 对应的特征向量. 由 $\alpha_1^T x = 0$ 得 $a_2 = (-1, 1, 0)^T, a_3 = (-1, 0, 1)^T$. 故

$$\mathbf{A} = \mathbf{P} \text{diag}(6, 3, 3) \mathbf{P}^{-1} = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 6 & & \\ & 3 & \\ & & 3 \end{pmatrix} \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{pmatrix}.$$

另解

注意到 $A - 3E$ 行和为 3, 迹为 3, 可设 $A - 3E = \begin{pmatrix} c & a & b \\ a & b & c \\ b & c & a \end{pmatrix}$, $a + b + c = 3$. 再由 $R(A - 3E) = 1$ 可知 $a = b = c = 1$.

实二次型对角化的步骤

对称阵正交实二次型的对角化, 或求正交变换 $x = Py$ 将实二次型 f 化为标准形的步骤:

- (1) 写出 f 对应的对称阵 A .
- (2) 求出 A 的特征值.
- (3) 若特征值是 $k \geq 1$ 重的, 求出 k 个特征向量后, 用格拉姆-施密特方法将其正交单位化.
- (4) 这些特征向量构成正交阵 P , $P^T A P$ 为这些特征向量对应的特征值构成的对角阵.
- (5) 写出正交变换 $x = P y$ 以及对应的实二次型

$$f = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2.$$

典型例题：实二次型的对角化

例

设 $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & x \end{pmatrix}$, $B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & y & 0 \\ 0 & 0 & -1 \end{pmatrix}$ 相似, 求 x, y 以及正交阵 P 使得 $P^{-1}AP = B$.

解

由 \mathbf{A}, \mathbf{B} 相似得 $|\mathbf{A}| = -2 = |\mathbf{B}| = -2y$, $\text{Tr}(\mathbf{A}) = 2 + x = \text{Tr}(\mathbf{B}) = 1 + y$, 故 $x = 0, y = 1$.

- 对于 $\lambda_1 = 2$, $\mathbf{A} - 2\mathbf{E} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 1 & -2 \end{pmatrix} \sim \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow \boldsymbol{\alpha}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$

续解

- 对于 $\lambda_2 = 1$, $\mathbf{A} - \mathbf{E} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow \boldsymbol{\alpha}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$.

- 对于 $\lambda_3 = -1$, $\mathbf{A} - \mathbf{E} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow \boldsymbol{\alpha}_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$

- 将特征向量单位化得到 $P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$.

例

求正交变换 $x = Py$ 化 $f = 4x_1^2 + 4x_2^2 + 4x_3^2 + 4x_1x_2 + 4x_2x_3 + 4x_3x_1$ 为标准形.

解

- f 对应 $\mathbf{A} = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$. 由 $|\mathbf{A} - \lambda \mathbf{E}| = -(\lambda - 2)^2(\lambda - 8)$ 得到特征值 $8, 2, 2$.
- 对于 $\lambda_1 = 8$, $\mathbf{A} - 8\mathbf{E} = \begin{pmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{pmatrix} \xrightarrow{r} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow \boldsymbol{\alpha}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. 将其单位化得到 $\mathbf{e}_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

典型例题: 实二次型的对角化

续解

- 对于 $\lambda_2, \lambda_3 = 2$, $A - 2E = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow \alpha_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \alpha_3 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$

将其正交单位化得到 $\beta_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, e_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix},$

$$\beta_3 = \alpha_3 - \frac{[\alpha_3, \beta_2]}{[\beta_2, \beta_2]} \beta_2 = \alpha_3 - \frac{1}{2} \beta_2 = \begin{pmatrix} -1/2 \\ -1/2 \\ 1 \end{pmatrix}, \quad e_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}.$$

- 因此经过正交变换 $x = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix} y$, f 化为标准形 $f = 8y_1^2 + 2y_2^2 + 2y_3^2.$

典型例题：实二次型的对角化

例

设实二次型 $f = ax_1^2 + x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3 + 4x_3x_1$ 经过正交变换 $x = Py$ 化为 $f = 5y_1^2 - y_2^2 - y_3^2$. 求常数 a 和正交阵 P .

解

$$f \text{ 对应 } \mathbf{A} = \begin{pmatrix} a & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}, \text{Tr}(\mathbf{A}) = a + 2 = 5 - 1 - 1, a = 1.$$

同上例可得 $\mathbf{P} = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix}$.

典型例题：实二次型的对角化

练习

设实二次型 $f = x_1^2 + 2x_2^2 + ax_3^2 - 4x_1x_2 - 4x_2x_3$ 经过正交变换 $x = Py$ 化为 $f = 2y_1^2 + 5y_2^2 + by_3^2$. 求常数 a, b 和正交阵 P .

答案

$$a = 3, b = -1, \mathbf{P} = \frac{1}{3} \begin{pmatrix} -2 & 1 & 2 \\ 1 & -2 & 2 \\ 2 & 2 & 1 \end{pmatrix}.$$

例

设实二次型 $f = a(x_1^2 + x_2^2 + x_3^2) + 4x_1x_2 + 4x_2x_3 + 4x_3x_1$ 经正交变换化成标准形 $f(y_1, y_2, y_3) = 6y_1^2$, 则 $a = 2$.

第三节 实对称阵的合同

- 惯性指数
- 实二次型规范形
- 正定二次型

引例：二次曲线的分类

设 A, B, C 是不全为零的实数. 二次曲线 $Ax^2 + Bxy + Cy^2 = 1$ 左侧的实二次型对应方阵 $A = \begin{pmatrix} A & B/2 \\ B/2 & C \end{pmatrix}$. 由

$$|\mathbf{A} - \lambda \mathbf{E}| = \begin{vmatrix} A - \lambda & B/2 \\ B/2 & C - \lambda \end{vmatrix} = \lambda^2 - (A + C)\lambda + (AC - B^2/4)$$

可知,

(1) 当 $B^2 - 4AC > 0$ 时, A 特征值一正一负, 从而通过正交变换 $\begin{pmatrix} x \\ y \end{pmatrix} = P \begin{pmatrix} s \\ t \end{pmatrix}$ 可知该曲线为双曲线.

(2) 同理, $B^2 - 4AC < 0$ 时该曲线为椭圆 (或空集);

(3) $B^2 - 4AC = 0$ 时该曲线为两条直线 (若有一次项则为抛物线).

可以看出我们有时候只关心实二次型对应的矩阵的特征值的符号.

定义

称实二次型 f 对应矩阵 A 的秩为 f 的秩.

对于可逆阵 P , $P^T A P$ 的秩和 A 相同. 因此若可逆线性变换 $x = Py$ 将 f 变为标准形, 则 y_i^2 系数非零为恰有 r 个.

定理 (慣性定理)

设实二次型 $f = x^T A x$ 的秩为 r . 若可逆线性变换 $x = P y = Q z$ 分别将 f 变为

$$f = k_1 y_1^2 + \cdots + k_r y_r^2$$

$$f = \ell_1 z_1^2 + \cdots + \ell_r z_r^2,$$

则 k_1, \dots, k_r 中正的个数和 ℓ_1, \dots, ℓ_r 中正的个数相同.

证明

由题设可知

$$B := \text{diag}(k_1, \dots, k_r, 0, \dots, 0) = P^T A P,$$

$$C := \text{diag}(\ell_1, \dots, \ell_r, 0, \dots, 0) = Q^T A Q.$$

设 $R = P^{-1}Q = (\alpha_1, \dots, \alpha_n)$, 则

$$C = R^T B R = \text{diag}(k_1 \alpha_1^T \alpha_1, \dots, k_r \alpha_r^T \alpha_r, 0, \dots, 0).$$

从而 $\ell_i = k_i \|\alpha_i\|^2$, 二者符号相同.



定义

把 f 标准型系数中为正数的个数称为实二次型 f 的正惯性指数 p , 为负数的个数称为实二次型 f 的负惯性指数 $q = r - p$.

推论

实二次型 $f = x^T A x$ 的正 (负) 惯性指数等于实对称阵 A 的正 (负) 特征值的个数.

推论

设 A, B 均为 n 阶实对称阵. A 与 B 合同当且仅当 A 的正特征值的个数等于 B 的正特征值的个数, 且 A 的负特征值的个数等于 B 的负特征值的个数.

例：惯性指数的应用

例

设 $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, 则 A 合同于 (D).

$$(A) \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$$

$$(B) \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

(C) $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$

(D) $\begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}$

例

矩阵 $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ 与 $\begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix}$ (B).

(A) 合同且相似

(B) 合同但不相似

(C) 不合同但相似

(D) 既不合同也不相似

确定了二次曲面的类别.

$$\frac{z^2}{c^2} = 1.$$

$$\frac{z^2}{a^2} - \frac{z^2}{c^2} = 1.$$

$$\frac{z^2}{a^2} - \frac{z^2}{c^2} = 1.$$

$$= 1.$$

$$= 1.$$

- 确定了二次曲面的类别.
- $$\frac{z^2}{c^2} = 1.$$
- $$\frac{z^2}{a^2} - \frac{z^2}{c^2} = 1.$$
- $$\frac{z^2}{a^2} - \frac{z^2}{c^2} = 1.$$
- $$= 1.$$
- $$= 1.$$

定义

若实二次型的标准形的系数只在 $-1, 0, 1$ 三个数中取值, 则称此实二次型为**规范形**.

定理

任意一个实二次型总可经过适当的可逆线性变换化为规范形, 且规范形是唯一的 (可任意交换变量顺序):

$$f = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2,$$

其中 p, q 分别为正负惯性指数.

推论

任意 n 阶实对称矩阵 A 合同于对角矩阵

$$\begin{pmatrix} \mathbf{E}_p & & \\ & -\mathbf{E}_q & \\ & & \mathbf{O}_{n-p-q} \end{pmatrix},$$

其中 p, q 分别为正负特征值个数 (计算重数), $R(\mathbf{A}) = p + q$.

例

若实对称矩阵 A 合同于 $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 2 & 0 \end{pmatrix}$, 则通过可逆线性变换 $x = Py$ 可将二次型 $x^T Ax$ 化为规范形 $y_1^2 + y_2^2 - y_3^2$.

定义

设 $f = x^T A x$ 是二次型.

- (1) 若对任意 $\boldsymbol{x} \neq \mathbf{0}$, 都有 $f(\boldsymbol{x}) > 0$, 则称 f 为**正定二次型**, 并称实对称阵 \boldsymbol{A} 为**正定矩阵**, 也记作 $\boldsymbol{A} > 0$.
- (2) 若对任意 \boldsymbol{x} , 都有 $f(\boldsymbol{x}) \geq 0$, 则称 f 为**半正定二次型**, 并称实对称阵 \boldsymbol{A} 为**半正定矩阵**, 也记作 $\boldsymbol{A} \geq 0$.
- (3) 若 $-f$ 正定, 则称 f 为**负定二次型**, 并称实对称阵 \boldsymbol{A} 为**负定矩阵**, 也记作 $\boldsymbol{A} < 0$.
- (4) 若 $-f$ 半正定, 则称 f 为**半负定二次型**, 并称实对称阵 \boldsymbol{A} 为**半负定矩阵**, 也记作 $\boldsymbol{A} \leq 0$.
- (5) 除此之外, 称 f **不定**.

可逆线性变换 $x = Py$ 不会影响正定性. A 正定 $\iff P^T A P$ 正定.

例

- (1) $f(x_1, x_2, x_3) = x_1^2 + x_2^2$ 半正定.
- (2) $f(x_1, x_2, x_3) = x_1^2 - 2x_2^2 + x_3^2$ 不定.
- (3) $f(x_1, x_2, x_3) = (x_1 + x_2)^2 + (x_2 + x_3)^2 + (x_3 + x_1)^2$ 正定.
- (4) $f(x_1, x_2, x_3) = (x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_3 - x_1)^2$ 半正定.
- (5) 椭球面 $f(x, y, z) = 1$ 对应的二次型正定.
- (6) 单叶/双叶双曲面 $f(x, y, z) = 1$ 对应的二次型不定.

定理

设 A 是 n 阶实对称阵, $f = x^T A x$. 如下命题等价:

- (1) $A > 0$ 正定, 即 f 正定.
- (2) f 的正惯性指数为 n , 即 A 特征值全为正.
- (3) 存在正交阵 P 使得 $A = P^T P$.

Hurwitz

- (4) (赫尔维茨 定理) A 的各阶顺序主子式都为正, 即

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0, \quad \dots, \quad |\mathbf{A}| > 0.$$

将(4)中 $>$ 换成 \geq 即可判断半正定, 这也等价于 f 的负惯性指数为 0, 即 A 特征值全非负.

例：正定的性质和判定

推论

若 A 正定, 则 $|A| > 0$ 且对角元全为正.

例

若 $f(x_1, x_2, x_3) = 2x_1^2 + x_2^2 + x_3^2 + 2x_1x_2 + tx_2x_3$ 正定, 求 t 的取值范围.

解

f 对应 $A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & t/2 \\ 0 & t/2 & 1 \end{pmatrix}$, 顺序主子式 $2 > 0$, $\begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = 1 > 0$, $|A| = 1 - \frac{t^2}{2} > 0$ 得
到 $-\frac{\sqrt{2}}{2} < t < \frac{\sqrt{2}}{2}$.

例：正定的性质和判定

例

判别二次型 $f = -5x^2 - 6y^2 - 4z^2 + 4xy + 4xz$ 是否正定.

解

由于 $A = \begin{pmatrix} -5 & 2 & 2 \\ 2 & -6 & 0 \\ 2 & 0 & -4 \end{pmatrix}$. 顺序主子式 $-5 < 0$, 因此不是正定的.
或由 $f(1, 0, 0) < 0$ 得到.

例

若实对称阵 A 正定, 证明 $|A + E| > 1$.

证明

由 A 正定可知其特征值均为正, 从而 $A+E$ 特征值都大于 1. 从而 $|A+E| > 1$. \square

例：正定的性质和判定

例

设 3 阶实对称阵 A 满足 $A^2 + 2A = O, R(A) = 2$.

- (1) 求 A 的全部特征值.
- (2) 当 k 为何值时, 矩阵 $A + kE$ 为正定矩阵.

解

- (1) 由 $A^2 + 2A = O$ 可知 A 的特征值满足 $\lambda^2 + 2\lambda = 0, \lambda = 0, -2$. 由 $R(A) = 2$ 可知 A 特征值为 $0, -2, -2$.
- (2) $A + kE$ 特征值为 $k, k - 2, k - 2$. 因此 $k > 2$.

例

设 A 是 $m \times n$ 实矩阵且 $R(A) = n$. 证明 $A^T A$ 正定.

证明

显然 $A^T A$ 是对称的. 注意到

$$x^T(A^T A)x = (Ax)^T Ax = \|Ax\|.$$

由于 $R(A) = n$, $Ax = 0$ 只有零解. 因此当 $x \neq 0$ 时, $Ax \neq 0$, 从而

$$x^T(A^T A)x = \|Ax\|^2 > 0.$$



例：正定的性质和判定

例

设 A 为 n 阶实对称矩阵. 证明 $R(A) = n \iff$ 存在一个 n 阶实方阵 B 使得 $AB + B^T A$ 正定.

证明

显然 $AB + B^T A$ 是对称的, 且

$$x^T(AB + B^T A)x = (Ax)^T Bx + (Bx)^T (Ax) = 2[Ax, Bx].$$

若 $R(A) = n$, 令 $B = A$, 则当 $x \neq 0$ 时, $Ax \neq 0$, 从而

$$[Ax, Bx] = \|Ax\| > 0.$$

若 $R(\mathbf{A}) < n$, 则存在 $\mathbf{x} \neq \mathbf{0}$ 使得 $\mathbf{Ax} = \mathbf{0}$, 从而 $[\mathbf{Ax}, \mathbf{Bx}] = 0$, $\mathbf{AB} + \mathbf{B}^T \mathbf{A}$ 不正定. □

例: 正定的性质和判定

例

设二次型 $f = \mathbf{x}^T \mathbf{A} \mathbf{x}$. 证明: 当 $\|\mathbf{x}\| = 1$ 时, $f(\mathbf{x})$ 的最大 (小) 值为实对称阵 \mathbf{A} 的最大 (小) 特征值.

证明

将 \mathbf{A} 的特征值排序为 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. 存在正交变换 $\mathbf{x} = \mathbf{P} \mathbf{y}$ 使得

$$\mathbf{\Lambda} = \mathbf{P}^T \mathbf{A} \mathbf{P} = \text{diag}(\lambda_1, \dots, \lambda_n), \quad f = \mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{y}^T \mathbf{\Lambda} \mathbf{y} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2.$$

由于正交变换保持长度, 因此 $\|\mathbf{x}\| = 1 \iff \|\mathbf{y}\| = 1$.

$$f = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 \leq \lambda_1 (y_1^2 + \cdots + y_n^2) = \lambda_1,$$

且等式可在 $\mathbf{y} = (1, 0, \dots, 0)$ 处取得.

$$f = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 \geq \lambda_n (y_1^2 + \cdots + y_n^2) = \lambda_n,$$

且等式可在 $\mathbf{y} = (0, \dots, 0, 1)$ 处取得. 故 f 的最大值为 λ_1 , 最小值为 λ_n . □

实对称阵可用于判断多元函数的极值. 设 $f(\boldsymbol{x}) = f(x_1, \dots, x_n)$ 是一个 n 元实函数, \boldsymbol{a} 是其定义域内一点, 且 f 在 \boldsymbol{a} 附近具有连续的二阶偏导数. 记 $f''_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$, 则 $f''_{ij} = f''_{ji}$. 于是 $A = (f''_{ij}(\boldsymbol{a}))$ 是 n 阶实对称阵.

定理

设 f 在 a 处各阶偏导均为零.

- (1) 若 A 正定, 则 f 在 a 处取极小值;
- (2) 若 A 负定, 则 f 在 a 处取极大值.

若 A 不定, 则无法判断 a 是否是极值点.

$$\mathbf{A} = \mathbf{U}_{m \times m} \mathbf{\Sigma}_{m \times n} \mathbf{V}_{n \times n}^T,$$

$$\mathbf{A} = \mathbf{U}_{m \times m} \mathbf{\Sigma}_{m \times n} \mathbf{V}_{n \times n}^T,$$

首先对 $A^T A$ 这一半正定对称阵做正交合同对角化

$$\mathbf{A}^T \mathbf{A} = \mathbf{V}^T \mathbf{\Lambda} \mathbf{V}, \quad \mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_n), \quad \mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_n), \lambda_1 \geq \dots \geq \lambda_n \geq 0.$$

对于 $1 \leq j \leq r = R(A)$, 令 $u_j = \frac{1}{\sigma_j} A v_j$, 设 u_{r+1}, \dots, u_m 是 $A^T x = 0$ 的一组标

准正交基础解系, 则 $U = (u_1, \dots, u_m)$ 是正交阵, 且 $A = U\Sigma V^T$.

$$A = U'_{m \times r} \Sigma'_{r \times r} V'_{r \times n},$$

$$A = U'_{m \times r} \Sigma'_{r \times r} V'_{r \times n},$$

这意味着当 r 相比 m, n 较小时, 只需存储 $(m + n + 1)r$ 个元素即可还原 A . 这是一种**无损压缩**算法. 如果我们只截取前 $k < r$ 个奇异值以及对应的 U, V 部分, 则可以对 A 进行**有损压缩**到 A' . 例如 A 表示一张图像的像素信息, 保留它较大的奇异值往往对它的信息影响很小. 有时候, 我们甚至需要主动舍弃较小的奇异值, 只保留较大的奇异值来实现**信号降噪**.

当 $m = n = 3$ 时, 可以看出线性变换可以分解为旋转、放缩、旋转的复合.