

例 8.2.16. 用初等变换的方法把二次型 $Q(x_1, x_2, x_3) = 2x_1x_2 + 4x_1x_3$ 化为标准形.

解. 二次型的矩阵为 $\mathbf{A} = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix}$. 与例 8.2.15 中的情形不同, 这儿的 \mathbf{A} 的对角线上的元素全为零. 我们的策略是用 \mathbf{A} 的非零的 $(1, 2)$ 与 $(2, 1)$ 元素造出一个对角线上的非零元. 于是

$$\begin{pmatrix} \mathbf{A} \\ \mathbf{I} \end{pmatrix} = \begin{pmatrix} \boxed{0} & \boxed{1} & 2 \\ \boxed{1} & 0 & 0 \\ 2 & 0 & 0 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{c_2 \rightarrow c_1} \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 0 \\ 2 & 0 & 0 \\ \hline 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{r_2 \rightarrow r_1} \begin{pmatrix} \boxed{2} & \boxed{1} & \boxed{2} \\ 1 & 0 & 0 \\ 2 & 0 & 0 \\ \hline 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{\begin{matrix} -\frac{1}{2}c_1 \rightarrow c_2 \\ -c_1 \rightarrow c_3 \end{matrix}} \begin{pmatrix} 2 & 0 & 0 \\ 1 & -\frac{1}{2} & -1 \\ 2 & -1 & -2 \\ \hline 1 & -\frac{1}{2} & -1 \\ 1 & \frac{1}{2} & -1 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{\begin{matrix} -\frac{1}{2}r_1 \rightarrow r_2 \\ -r_1 \rightarrow r_3 \end{matrix}} \begin{pmatrix} 2 & 0 & 0 \\ 0 & \boxed{-\frac{1}{2}} & \boxed{-1} \\ 0 & -1 & -2 \\ \hline 1 & -\frac{1}{2} & -1 \\ 1 & \frac{1}{2} & -1 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{-2c_2 \rightarrow c_3} \begin{pmatrix} 2 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & -1 & 0 \\ \hline 1 & -\frac{1}{2} & 0 \\ 1 & \frac{1}{2} & -2 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{-2r_2 \rightarrow r_3} \begin{pmatrix} 2 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 \\ \hline 1 & -\frac{1}{2} & 0 \\ 1 & \frac{1}{2} & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

最后的分块矩阵中, 上面的方阵为对角阵, 故下面的矩阵 $\mathbf{P} = \begin{pmatrix} 1 & -\frac{1}{2} & 0 \\ 1 & \frac{1}{2} & -2 \\ 0 & 0 & 1 \end{pmatrix}$ 为所求,

满足 $\mathbf{P}^T \mathbf{A} \mathbf{P} = \text{diag}(2, -1/2, 0)$. 此时的坐标变换为

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{2} & 0 \\ 1 & \frac{1}{2} & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix},$$

而二次型为

$$Q(x_1, x_2, x_3)|_{\mathbf{X}=\mathbf{P}\mathbf{Y}} = 2y_1^2 - \frac{1}{2}y_2^2. \quad \square$$

注 8.2.17. (1) 由于初等变换的选取的不同, 最终结果中的标准形可能不同.

(2) 我们也可以把矩阵写成 (\mathbf{A}, \mathbf{I}) 的形式, 先同时对 \mathbf{A}, \mathbf{I} 做行变换, 然后只对 \mathbf{A} 做相应列变换. 变换的最终目标是得到 $(\mathbf{D}, \mathbf{P}^T)^*$, 其中 \mathbf{D} 为对角阵. 做法与我们上面的操作是等价的. 学生只需要熟悉一种方法即可.

*注意这儿是 \mathbf{P}^T , 而不再是 \mathbf{P}

注 8.2.18. 设 a_1, \dots, a_n 是欧氏空间 V 的一组基, 它的 Gram 矩阵为 G . 若 b_1, \dots, b_n 为 V 的另外一组基, 使得 a_1, \dots, a_n 到 b_1, \dots, b_n 的过渡矩阵为 P . 此时, 不难验证 b_1, \dots, b_n 的 Gram 矩阵必为 $P^T G P$. 若进一步地, 有 $P^T G P = I$, 则 b_1, \dots, b_n 为 V 的一组标准正交基.

例 8.2.19. 设 V 是次数不超过 2 的实系数多项式构成的 \mathbb{R} 上的线性空间, 配有内积 $(f, g) := \int_0^1 f(x)g(x) dx$. 求 V 的一组标准正交基.

解. **(思路一)** 对于基 $1, x, x^2$ 作 Schmidt 正交化, 我们可以得到

$$\varepsilon_1 = 1, \quad \varepsilon_2 = \sqrt{12} \left(x - \frac{1}{2} \right), \quad \varepsilon_3 = \sqrt{180} \left(x^2 - x + \frac{1}{6} \right).$$

(思路二) 对基 $1, x, x^2$ 的度量矩阵 $\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix}$ 作相合变换:

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & 0 & 1 & 0 \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & 0 & 0 & 1 \end{pmatrix} \xrightarrow{\begin{matrix} -\frac{1}{2}r_1 \rightarrow r_2 \\ -\frac{1}{3}r_1 \rightarrow r_3 \\ -\frac{1}{2}c_1 \rightarrow c_2 \\ -\frac{1}{3}c_1 \rightarrow c_3 \end{matrix}} \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & \frac{1}{12} & \frac{1}{12} & -\frac{1}{2} & 1 & 0 \\ 0 & \frac{1}{12} & \frac{4}{45} & -\frac{1}{3} & 0 & 1 \end{pmatrix} \xrightarrow{\begin{matrix} -r_2 \rightarrow r_3 \\ -c_2 \rightarrow c_3 \end{matrix}} \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & \frac{1}{12} & 0 & -\frac{1}{2} & 1 & 0 \\ 0 & 0 & \frac{1}{180} & \frac{1}{6} & -1 & 1 \end{pmatrix}.$$

一般而言, 这一方法并不能保证得到与 Schmidt 正交化相同的计算结果, 除非操作时遵从指定的顺序

这说明向量组 $(1, x, x^2) \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{6} & -1 & 1 \end{pmatrix}^T = (1, x - \frac{1}{2}, x^2 - x + \frac{1}{6})$ 的 Gram 矩阵为 $\text{diag}(1, \frac{1}{12}, \frac{1}{180})$. 由此, 我们仍然可以得到标准正交基

$$\varepsilon_1 = 1, \quad \varepsilon_2 = \sqrt{12} \left(x - \frac{1}{2} \right), \quad \varepsilon_3 = \sqrt{180} \left(x^2 - x + \frac{1}{6} \right). \quad \square$$

注 8.2.20. 本节介绍了三种方法用于将二次型化为标准形. 其中的主轴化的方法 (坐标的正交变换) 是最重要的, 要求大家熟练掌握. 另一方面, 在下一节中, 我们会谈到二次型的正负惯性指数, 利用配方法容易帮助大家快速找到这些重要的指标.

8.3 相合不变量与分类

之前我们已经证明了任意实二次型都可以化成标准形, 但是标准形不唯一. 例如, 二次型 $Q(x_1, x_2, x_3) = 4x_1^2 - 3x_2^2$ 已是 Q 的标准形. 但是若采用坐标变换

$$\begin{cases} y_1 = 2x_1, \\ y_2 = \sqrt{3}x_2, \\ y_3 = x_3, \end{cases}$$

则 $Q(y_1, y_2, y_3) = y_1^2 - y_2^2$ 是 Q 的另外一个标准形. 那么, 什么是二次型“最终极”的标准形呢? 等价地, 实对称阵可以通过相合变换得到一个什么形式的最简单的对角阵呢?

定理 8.3.1 (惯性定理). 设 A 是一个 n 阶实对称矩阵, 则存在可逆矩阵 P , 使得

$$P^T A P = \begin{pmatrix} I_r & & \\ & -I_s & \\ & & O \end{pmatrix}. \quad (8.1)$$

并且这里的 r 与 s 是由 A 唯一决定的.

显然, 在等式 (8.1) 中, $r + s = \text{rank}(A)$. 我们将把等式右边的矩阵称为 A 的规范形 (normal form), 而相应的二次型称为对应于 A 的规范二次型. 上面的定理说明 A 的规范形是存在且唯一的.

证明. (存在性的证明) 在前面一节中, 我们已经见到, 对于实对称阵 A , 存在实矩阵 P_0 , 使得 $P_0^T A P_0 = \text{diag}(\mu_1, \dots, \mu_n)$ 为对角阵. 通过交换 P_0 的列向量组的排列顺序, 我们不妨假设这些对角线上的元素满足

$$\mu_1, \dots, \mu_r > 0, \quad \mu_{r+1}, \dots, \mu_{r+s} < 0, \quad \mu_{r+s+1} = \dots = \mu_n = 0.$$

若令

$$P = P_0 \text{diag} \left(\frac{1}{\sqrt{\mu_1}}, \dots, \frac{1}{\sqrt{\mu_r}}, \frac{1}{\sqrt{-\mu_{r+1}}}, \dots, \frac{1}{\sqrt{-\mu_{r+s}}}, 1, \dots, 1 \right),$$

$$\text{则有 } P^T A P = \begin{pmatrix} I_r & & \\ & -I_s & \\ & & O \end{pmatrix}.$$

(唯一性的证明) 假设 $\begin{pmatrix} I_r & & \\ & -I_s & \\ & & O \end{pmatrix}$ 与 $\begin{pmatrix} I_p & & \\ & -I_q & \\ & & O \end{pmatrix}$ 都是 \mathbf{A} 的规范形, 则 $r + s = \text{rank}(\mathbf{A}) = p + q$. 考虑二次型 $Q(x_1, \dots, x_n) := \mathbf{X}^\top \mathbf{A} \mathbf{X}$. 此时, 存在两个不同的坐标的可逆变换 $\mathbf{X} = \mathbf{P}_1 \mathbf{Y}$ 和 $\mathbf{X} = \mathbf{P}_2 \mathbf{Z}$ 分别使得

$$\begin{aligned} Q(\mathbf{X})|_{\mathbf{X}=\mathbf{P}_1 \mathbf{Y}} &= y_1^2 + \dots + y_r^2 - y_{r+1}^2 - \dots - y_{r+s}^2, \\ Q(\mathbf{X})|_{\mathbf{X}=\mathbf{P}_2 \mathbf{Z}} &= z_1^2 + \dots + z_p^2 - z_{p+1}^2 - \dots - z_{p+q}^2. \end{aligned}$$

若 $r \neq p$, 不妨假设 $r < p$. 此时, $\mathbf{Y} = \mathbf{P}_1^{-1} \mathbf{X}$ 与 $\mathbf{Z} = \mathbf{P}_2^{-1} \mathbf{X}$. 不妨设 \mathbf{P}_1^{-1} 的行向量依次为 $\mathbf{B}_1, \dots, \mathbf{B}_n$, 而 \mathbf{P}_2^{-1} 的行向量依次为 $\mathbf{C}_1, \dots, \mathbf{C}_n$. 则对任意 i , 我们有 $y_i = \mathbf{B}_i \mathbf{X}$, $z_i = \mathbf{C}_i \mathbf{X}$. 接下来, 我们来考察如下特别挑选的齐次线性方程组

$$\begin{cases} y_1 = 0, \\ \vdots \\ y_r = 0, \\ z_{p+1} = 0, \\ \vdots \\ z_n = 0, \end{cases} \Leftrightarrow \begin{pmatrix} \mathbf{B}_1 \\ \vdots \\ \mathbf{B}_r \\ \mathbf{C}_{p+1} \\ \vdots \\ \mathbf{C}_n \end{pmatrix} \mathbf{X} = \mathbf{0},$$

其中系数矩阵为 $(r + (n - p)) \times n$ 维的. 由于 $r < p$, 这导致 $r + (n - p) < n$. 这意味着上面的齐次线性方程组的系数矩阵不是列满秩的, 从而存在非零解 \mathbf{X}_0 . 此时,

$$\mathbf{Y}_0 = \mathbf{P}_1^{-1} \mathbf{X}_0 = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ y_{r+1} \\ \vdots \\ y_n \end{pmatrix} \neq \mathbf{0}, \quad \mathbf{Z}_0 = \mathbf{P}_2^{-1} \mathbf{X}_0 = \begin{pmatrix} z_1 \\ \vdots \\ z_p \\ 0 \\ \vdots \\ 0 \end{pmatrix} \neq \mathbf{0}.$$

则我们同时有

$$\begin{aligned} Q(\mathbf{X}_0) &= Q|_{\mathbf{X}=\mathbf{P}_1 \mathbf{Y}}(\mathbf{Y}_0) = -y_{r+1}^2 - \dots - y_{r+s}^2 \leq 0, \quad (r + s \text{ 可能严格小于 } n, \text{ 从而 } Q \text{ 可能为 } 0) \\ Q(\mathbf{X}_0) &= Q|_{\mathbf{X}=\mathbf{P}_2 \mathbf{Z}}(\mathbf{Z}_0) = z_1^2 + \dots + z_p^2 > 0, \quad (p > r \geq 0, \text{ 从而 } Q \text{ 严格大于 } 0) \end{aligned}$$

矛盾. 故 $r = p$, $s = q$, 由矩阵 \mathbf{A} 唯一确定. □

定义 8.3.2. 惯性定理 8.3.1 中的 r 称为 A 的正惯性指数 (positive index of inertia), s 称为 A 的负惯性指数 (negative index of inertia)*. 由上面定理的证明可知, 这两个指数分别是 A 的任意一个标准形中正项的项数与负项的项数, 从而由主轴定理 (定理 8.2.3) 可知, 它们也分别是 A 的特征值的正项与负项的个数. 相应地, $r-s$ 称为 A 的符号差 (signature). 这儿的 r, s 和 $r-s$ 也分别被称为 A 所对应的二次型 Q 的正惯性指数、负惯性指数和符号差.

请想清楚
原因

综合前面的结果, 我们不难得到如下的定理.

定理 8.3.3. 任何实对称阵都可以 (实) 相合等价于一个实对角阵, 并且两个相同维数的实对称阵相合等价, 当且仅当它们的正、负惯性指数相等, 当且仅当它们的秩和符号差相等.

请想清楚
原因

例 8.3.4. 设 A 为 n 阶实可逆矩阵. 若 A 与 $-A$ 相合, 则 n 为偶数. 进一步地, 若 A 为实对称矩阵, 则 A 的正惯性指数为 $\frac{n}{2}$.

证明. 由于 A 与 $-A$ 相合, 存在可逆矩阵 P 使得 $-A = P^T A P$, 故 $|-A| = |P^T A P|$, 即 $(-1)^n |A| = |A| \cdot |P|^2$. 又 $|A| \neq 0$, 故 $(-1)^n = |P|^2 > 0$, 从而 n 为偶数.

若进一步地, A 为实对称阵, 关于惯性指数, 我们显然有 $r_A = s_{-A}$, $s_A = r_{-A}$. 而若 A 与 $-A$ 相合, 则 $r_A = r_{-A}$, $s_A = s_{-A}$, 从而皆等于 $\frac{n}{2}$. \square

例 8.3.5. 求 n 元实二次型 $\sum_{i=1}^n \sum_{j=1}^n a_i a_j x_i x_j$ 的规范形, 其中 a_1, a_2, \dots, a_n 不全为零.

解. 记 $\mathbf{a} = (a_1, \dots, a_n)^T$, 于是,

$$\sum_{i=1}^n \sum_{j=1}^n a_i a_j x_i x_j = \left(\sum_i a_i x_i \right) \left(\sum_j a_j x_j \right) = \mathbf{x}^T \mathbf{a} \mathbf{a}^T \mathbf{x}.$$

(解法一) 不妨设 $a_1 \neq 0$. 此时, 可以考虑坐标变换如下: 令 $y_1 = \sum_i a_i x_i$, 而对 $j = 2, 3, \dots, n$, 令 $y_j = x_j$. 那么二次型就写成了 y_1^2 , 这显然为相应二次型的规范形.

(解法二) 于是, 二次型的矩阵为 $A = \mathbf{a} \mathbf{a}^T$. 由于 $\text{rank}(A) = 1$, A 的特征值为 $\text{tr}(A) = \text{tr}(\mathbf{a} \mathbf{a}^T) = \text{tr}(\mathbf{a}^T \mathbf{a}) = \sum_i a_i^2$ (1 重), 和 0 ($n-1$ 重). 从而, 二次型的规范形为 z_1^2 . \square

例 8.3.6. 求如下 n 元实二次型的规范形, 其中 $n \geq 2$.

(1) $\sum_{i=1}^n x_i^2 + \sum_{1 \leq i < j \leq n} x_i x_j.$

(2) $\sum_{1 \leq i < j \leq n} x_i x_j.$

*因此, 有些教材里分别选用 p 和 n 来表示实对称阵的正、负惯性指数

解. 设 \mathbf{J} 是 n 阶的元素全为 1 的矩阵.

(1) 二次型的矩阵为 $\mathbf{A} = \frac{1}{2}(\mathbf{I} + \mathbf{J})$. 由于 \mathbf{J} 的特征值为 n (1 重) 和 0 ($n-1$ 重), \mathbf{A} 的特征值为 $\frac{n+1}{2}$ (1 重) 和 $\frac{1}{2}$ ($n-1$ 重). 由于 \mathbf{A} 的特征值都是正的, 相应二次型有规范形 $z_1^2 + z_2^2 + \cdots + z_n^2$.

(2) 二次型的矩阵为 $\mathbf{B} = \frac{1}{2}(\mathbf{J} - \mathbf{I})$. 类似可以看出, \mathbf{B} 的特征值为 $\frac{n-1}{2}$ (1 重) 和 $\frac{-1}{2}$ ($n-1$ 重), 从而 \mathbf{B} 的正负惯性指数分别为 1 和 $n-1$, 从而相应二次型有规范形 $z_1^2 - z_2^2 - \cdots - z_n^2$. \square

例 8.3.7. 设 n 阶实对称阵 \mathbf{A} , \mathbf{B} 和 $\mathbf{A} + \mathbf{B}$ 的正惯性指数分别为 r_A , r_B 和 r_{A+B} . 证明: $r_A + r_B \geq r_{A+B}$.

证明. 记 \mathbf{A} , \mathbf{B} 和 $\mathbf{A} + \mathbf{B}$ 的正惯性指数分别为 r_A , r_B 和 r_{A+B} , 负惯性指数分别为 s_A , s_B 和 s_{A+B} . 于是, 存在可逆矩阵 \mathbf{P}_1 , \mathbf{P}_2 和 \mathbf{P}_3 使得

$$\begin{aligned} \mathbf{X}^\top \mathbf{A} \mathbf{X} |_{\mathbf{X}=\mathbf{P}_1 \mathbf{Y}} &= y_1^2 + \cdots + y_{r_A}^2 - y_{r_A+1}^2 - \cdots - y_{r_A+s_A}^2, \\ \mathbf{X}^\top \mathbf{B} \mathbf{X} |_{\mathbf{X}=\mathbf{P}_2 \mathbf{Z}} &= z_1^2 + \cdots + z_{r_B}^2 - z_{r_B+1}^2 - \cdots - z_{r_B+s_B}^2, \\ \mathbf{X}^\top (\mathbf{A} + \mathbf{B}) \mathbf{X} |_{\mathbf{X}=\mathbf{P}_3 \mathbf{W}} &= w_1^2 + \cdots + w_{r_{A+B}}^2 - w_{r_{A+B}+1}^2 - \cdots - w_{r_{A+B}+s_{A+B}}^2. \end{aligned}$$

用反证法, 假设 $r_A + r_B < r_{A+B}$. 类似于惯性定理的证明中的处理手法, 我们解关于 \mathbf{X} 的如下方程组:

$$\begin{cases} y_i = 0, & \text{若 } 1 \leq i \leq r_A, \\ z_j = 0, & \text{若 } 1 \leq j \leq r_B, \\ w_k = 0, & \text{若 } r_{A+B} + 1 \leq k \leq n. \end{cases}$$

由于此时 $r_A + r_B + (n - r_{A+B}) < n$, 该齐次方程组存在非零解 \mathbf{X}_0 . 对于它, 我们同时有

$$\mathbf{X}_0^\top \mathbf{A} \mathbf{X}_0 \leq 0, \quad \mathbf{X}_0^\top \mathbf{B} \mathbf{X}_0 \leq 0 \quad \text{以及} \quad \mathbf{X}_0^\top (\mathbf{A} + \mathbf{B}) \mathbf{X}_0 > 0.$$

这显然有矛盾. \square