二次型

黄利兵

数学科学学院

2023 年 2 月 15 日

主要内容

二次型的定义

定义

设 P 为数域, $x_1, \cdots, x_n$ 为变量. 二次齐次多项式

$$q(x_1, \cdots, x_n) = \sum_{i=1}^{n} a_{ii} x_i^2 + 2 \sum_{i<j} a_{ij} x_i x_j$$

称为 P 上的一个 n 元二次型, 简称二次型.

如果约定 $a_{ji} = a_{ij}$, $i < j$, 并引进矩阵记号 $A = (a_{ij})$, $x = (x_1, \cdots, x_n)^{\mathsf{T}}$, 则上述二次型也可写为

$$q(x) = x^{\mathsf{T}} A x.$$

称 A 为上述二次型的矩阵.

例

3 元二次型 $x_1^2 + 2x_2^2 - x_3^2 + 4x_1x_2 - 6x_2x_3 + 10x_1x_3$ 的矩阵为 $\begin{bmatrix} 1 & 2 & 5 \\ 2 & 2 & -3 \\ 5 & -3 & -1 \end{bmatrix}$.

可逆线性替换

定义

设 $x = (x_1, \cdots, x_n)^{\mathsf{T}}$ 和 $y = (y_1, \cdots, y_n)^{\mathsf{T}}$ 是两组变量. 如果 $C \in P^{n \times n}$ 是可逆矩阵, 则称关系式 $x = Cy$ 是从变量 x 到 y 的一个可逆线性替换, 也称非退化线性替换.

例

利用可逆线性替换 $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, 即 $\begin{cases} x_1 = y_1 + y_2, \\ x_2 = y_1 - y_2, \end{cases}$ 可将变量 x_1, x_2 的二次型 $x_1^2 + x_1 x_2 + x_2^2$ 化为变量 y_1, y_2 的二次型 $3y_1^2 + y_2^2$.

一般地, 关于变量 x 的二次型 $x^{\mathsf{T}} A x$ 经可逆线性替换 $x = Cy$ 可变为关于变量 y 的二次型 $y^{\mathsf{T}} B y$, 其中 $B = C^{\mathsf{T}} A C$.

矩阵的合同

定义

设 $A, B \in P^{n\times n}$, 如果存在可逆矩阵 C, 使得 $B = C^{\mathsf{T}}AC$, 则称 A 与 B 合同.

合同是一种等价关系.

- 反身性: A 与 A 合同, 因为 $A = E^{\mathsf{T}}AE$;
- 对称性: 若 A 与 B 合同, 则 B 与 A 合同;
- 传递性: 若 A 与 B 合同, B 与 C 合同, 则 A 与 C 合同.

对于一般的方阵, 寻找它的合同标准形可能是困难的. 但对于对称矩阵, 我们可以借助相关的二次型证明: 任意对称矩阵总合同于一个对角矩阵.

二次型的标准形

定理

n 元二次型 $q(x) = x^{\mathsf{T}}Ax$ 可经可逆线性替换 $x = Cy$ 变为平方和形式

$$q(x) = d_1 y_1^2 + \cdots + d_n y_n^2,$$

称为该二次型的标准形.

证明.

配方法. 如果二次型中有某个变量的平方项, 则先将所有含该变量的项配成完全平方; 这样, 剩下的表达式将不再含有这个变量, 可继续对剩下的部分配方. 如果二次型中没有平方项, 则任取一个交叉项 x_ix_j, 作可逆线性替换将它变为平方差, 再应用前面的方法配方. □

例

用可逆线性替换化二次型 $f = x_1^2 + x_2^2 + 3x_3^2 + 2x_1x_2 + 6x_2x_3 + 2x_1x_3$ 为标准形.

解答

有平方项 x_1^2, 可先将所有含 x_1 的项配成完全平方

$$f = (x_1 + x_2 + x_3)^2 + 2x_3^2 + 4x_2x_3.$$

剩下的有平方项 $2x_3^2$, 可把剩下的所有含 x_3 的项配成完全平方

$$f = (x_1 + x_2 + x_3)^2 + 2(x_3 + x_2)^2 - 2x_2^2.$$

可见, 只要令 $y_1 = x_1 + x_2 + x_3$, $y_2 = x_3 + x_2$, $y_3 = x_2$, 即

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

则原二次型化为标准形 $y_1^2 + 2y_2^2 - 2y_3^2$.

例

用可逆线性替换化二次型

$$f = 2x_1x_2 + 2x_1x_3 - 2x_1x_4 - 2x_2x_3 + 2x_2x_4 + 2x_3x_4$$

为标准形.

解答

由于 f 中不含平方项, 可先作替换

$$x_1 = y_1 + y_2, \quad x_2 = y_1 - y_2, \quad x_3 = y_3, \quad x_4 = y_4,$$

将它化为

$$f = 2y_1^2 - 2y_2^2 + 4y_2y_3 - 4y_2y_4 + 2y_3y_4.$$

再逐步配方, 得到

$$f = 2y_1^2 - 2(y_2 - y_3 + y_4)^2 + 2\left(y_3 - \frac{1}{2}y_4\right)^2 + \frac{3}{2}y_4^2.$$

解答 (续)

因此, 令 $z_1 = y_1$, $z_2 = y_2 - y_3 + y_4$, $z_3 = y_3 - y_4/2$, $z_4 = y_4$, 则原二次型化为标准形

$$2z_1^2 - 2z_2^2 + 2z_3^2 + \frac{3}{2}z_4^2,$$

相应的可逆线性替换为

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & -1/2 \\ 1 & -1 & -1 & 1/2 \\ 0 & 0 & 1 & 1/2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix}.$$

从上面例子可以看出, 用配方法化二次型为标准形, 虽然操作起来比较简单, 但有两个明显的缺陷

- 每一次配方时都需要重新整理剩下的项, 一旦没有平方项, 就需要多做一次变量替换;
- 为求出最终的可逆线性替换, 需要把多次替换的结果复合起来, 还要求出逆向的替换.

初等变换法

求二次型标准形的第二种方法是初等变换法. 做法是: 取二次型的矩阵 A, 对矩阵 $\begin{pmatrix} A \\ E_n \end{pmatrix}$ 作一次初等列变换, 再作一次相应的初等行变换 (一次初等列变换与一次初等行变换称为相应的, 是指对应的初等矩阵互为转置); 作一次初等列变换, 再作一次相应的初等行变换; $\cdots$; 直到将 A 变为对角矩阵 D. 这时 E_n 变为可逆矩阵 C.

我们断言: 作可逆线性替换 $x = Cy$, 就可将原二次型 $x^\mathsf{T} Ax$ 化为标准形 $y^\mathsf{T} Dy$.

事实上, 对 A 作一次初等变换, 相当于乘以相应的初等矩阵. 假设初等列变换对应的矩阵分别为 $P_1, P_2, \cdots, P_s$, 则有

$$P_s^\mathsf{T} \cdots P_2^\mathsf{T} P_1^\mathsf{T} A P_1 P_2 \cdots P_s = D, \quad E_n P_1 P_s \cdots P_s = C.$$

可见 $C^\mathsf{T} AC = D$.

例

用可逆线性替换化二次型 $f = 2x_1x_2 - 6x_2x_3 + 2x_1x_3$ 为标准形.

解答

原二次型的矩阵为 $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & -3 \\ 1 & -3 & 0 \end{bmatrix}$. 我们对矩阵 $\begin{bmatrix} A \\ E_3 \end{bmatrix}$ 作以下初等变换: 第 2 列的 1/2 倍加到第 1 列, 第 2 行的 1/2 倍加到第 1 行; 第 1 列的 (-1) 倍加到第 2 列, 第 1 行的 (-1) 倍加到第 2 行; 第 1 列的 1/2 倍加到第 3 列, 第 1 行的 1/2 倍加到第 3 行; 第 2 列的 5/2 倍加到第 3 列, 第 2 行的 5/2 倍加到第 3 行, 则 A 变为对角矩阵 $D = \mathrm{diag}(1, -1, 6)$, 而 E_3 变为

$$C = \begin{bmatrix} 1 & -1 & 3 \\ \frac{1}{2} & \frac{1}{2} & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

因此, 作可逆线性替换 $x = Cy$, 则原二次型化为标准形 $y_1^2 - y_2^2 + 6y_3^2$.

标准形的唯一性

从前面的讨论中可以看到, 将二次型化为标准形的过程并不唯一, 所得的结果往往也不唯一. 但从矩阵合同的角度可以看出, 不同标准形中非零平方项的个数是相同的, 都等于该二次型的矩阵的秩, 也称它为该二次型的秩.
现在讨论复数域上的二次型.

定理

复数域上的 n 元二次型总可经非退化线性替换化为

$$z_1^2 + z_2^2 + \cdots + z_r^2,$$

其中 r 为该二次型的秩. 这个标准形也称为该二次型的规范形.

证明.

首先可经可逆线性替换化为 $d_1 y_1^2 + d_2 y_2^2 + \cdots + d_r y_r^2$. 我们取非零复数 c_i 使得 $c_i^2 = d_i$, $1 \le i \le r$, 再作可逆线性替换 $y_i = z_i / c_i$, $1 \le i \le r$; $y_j = z_j$, $r < j \le n$, 则原二次型化为 $z_1^2 + z_2^2 + \cdots + z_r^2$. □

这个结论也可用矩阵语言来叙述: 任意复对称方阵 A 总合同于 $\mathrm{diag}(E_r, 0)$, 其中 $r = \mathrm{rank}(A)$.

再来讨论实数域上的二次型, 也称实二次型. 我们将证明, 不同标准形中, 正平方项的个数是相同的, 负平方项的个数也是相同的.

定理 (惯性定理)

实数域上的 n 元二次型 f 总可经非退化线性替换化为

$$f = z_1^2 + \cdots + z_p^2 - z_{p+1}^2 - \cdots - z_{p+q}^2,$$

右式称为 f 的规范形, 其中 p, q 不依赖于非退化线性替换的选取, 分别称为 f 的正惯性指数和负惯性指数.

证明.

首先将 f 化为标准形

$$f = d_1 y_1^2 + d_2 y_2^2 + \cdots + d_r y_r^2,$$

在非零实数 $d_1, \cdots, d_r$ 中, 不妨设 $d_1, \cdots, d_p$ 为正数, $d_{p+1}, \cdots, d_r$ 为负数. 再令

$$y_i = z_i / \sqrt{d_i}, \quad 1 \le i \le p; \qquad y_j = z_j / \sqrt{-d_j}, \quad p < j \le r,$$

则有 $f = z_1^2 + \cdots + z_p^2 - z_{p+1}^2 - \cdots - z_r^2$.

证明 (续).

下面再证明 p 不依赖于可逆线性替换的选取, 从而 $q = r - p$ 也不依赖于可逆线性替换的选取.
假设有另一可逆线性替换可将 f 化为另一规范形

$$f = w_1^2 + \cdots + w_s^2 - w_{s+1}^2 - \cdots - w_r^2.$$

如果 $s \neq p$, 不妨设 $s < p$, 那么, 关于变量 $z_1, \cdots, z_n$ 的线性方程组

$$w_1 = \cdots = w_s = 0, z_{p+1} = \cdots = z_n = 0$$

的变量个数 n 大于方程个数 $s + (n - p)$, 从而它必有非零解. 将非零解代入第一个规范形, 所得的值 > 0; 代入第二个规范形, 所得的值 ≤ 0, 矛盾! 因此 $s = p$. 这就证明了 p 与可逆线性替换的选取无关. □

这个结论也可用矩阵语言来叙述: 任意实对称方阵 A 总合同于 $\mathrm{diag}(E_p, -E_q, 0)$, 其中 p, q 分别称为 A 的正、负惯性指数.

思考题

(**) 证明实对称方阵 A 的正、负惯性指数分别等于 A 的正、负特征值的个数.

正定二次型

定义

如果一个 n 元实二次型 f 在任意非零列向量 $x \in \mathbb{R}^{n\times 1}$ 处的值 $f(x) > 0$, 则称 f 是正定的.

例

二元二次型 $x^2 + y^2$ 是正定的. 三元二次型 $x^2 + y^2 + z^2 - xy - yz - zx$ 不是正定的, 因为它在 $(1,1,1)^\mathsf{T}$ 处的值不是正数.

命题

n 元实二次型 f 是正定的, 当且仅当它的正惯性指数为 n.

证明.

取 f 的规范形 $z_1^2 + \cdots + z_p^2 - z_{p+1}^2 - \cdots - z_r^2$.
如果正惯性指数 $p < n$, 则当 $z_1 = \cdots = z_p = 0$, $z_{p+1} = \cdots = z_n = 1$ 时, f 的值 ≤ 0, 与 f 正定矛盾.
反之, 若正惯性指数 $p = n$, 即 $f = z_1^2 + z_2^2 + \cdots + z_n^2$, 易知 f 是正定的. □

例

考虑 3 元实二次型 $f = x^2 + y^2 + z^2 + xy + yz + zx$. 由

$$f = \frac{1}{2}\left((x+y)^2 + (y+z)^2 + (z+x)^2\right)$$

可知 f 是正定的. 将其中的变量 z 替换为 0, 得到 2 元二次型 $g = x^2 + y^2 + xy$, 它仍是正定的.

一般地, 我们有

命题

设 $f = f(X_1, X_2)$ 是正定的 n 元实二次型, 其中我们把 n 个变量分为两组 $X_1 = (x_1, \cdots, x_s)^{\mathsf{T}}$, $X_2 = (x_{s+1}, \cdots, x_n)^{\mathsf{T}}$. 令 $g(X_1) = f(X_1, \mathbf{0})$ 为 r 元实二次型, 则 g 仍是正定的.

证明.

由于 f 正定, 所以当 $X_1 \neq \mathbf{0}$ 时, $f(X_1, \mathbf{0}) > 0$. 这就证明了 g 是正定的. □

正定矩阵

定义

如果 $A \in \mathbb{R}^{n\times n}$ 是对称的, 且二次型 $x^{\mathsf{T}}Ax$ 正定, 则称 A 为正定矩阵.

由前面关于正定二次型的定理, 我们有

定理

n 阶实对称方阵是正定矩阵, 当且仅当它的正惯性指数为 n, 也当且仅当它合同于 E_n

推论

正定矩阵的行列式大于 0.

证明.

若 A 正定, 则它合同于 E_n, 即有可逆矩阵 P 使得 $A = P^{\mathsf{T}}E_nP$. 取行列式得 $|A| = |P|^2 > 0$. $\square$

主子式

设 A 为对称矩阵, 考虑二次型 $f(x) = x^{\mathsf{T}} Ax$. 如果将 n 个变量分为两组 X_1 和 X_2, 其中

$$X_1 = (x_{i_1}, \cdots, x_{i_k})^{\mathsf{T}}, X_2 = (x_{i_{k+1}}, \cdots, x_{i_n})^{\mathsf{T}},$$

在 f 中, 令 X_2 这组变量都取 0, 就得到 k 元实二次型 $g(X_1)$. 不难看出, 二次型 g 的矩阵, 恰好是 A 中第 $i_1, \cdots, i_k$ 行和第 $i_1, \cdots, i_k$ 列交叉位置的 $k \times k$ 子矩阵. 我们把这样的子矩阵称为 A 的 $(i_1, i_2, \cdots, i_k)$-主子矩阵, 简称主子矩阵.

定义

A 的主子矩阵的行列式称为 A 的主子式. 特别地, 前 k 行、前 k 列所构成的子矩阵称为 A 的 k 阶顺序主子矩阵, 它的行列式称为 A 的 k 阶顺序主子式.

前面提到, 如果 f 是正定的实二次型, 那么, 将其中部分变量取为 0 所得的 (变量个数较少的) 二次型仍是正定的. 这就告诉我们,

命题

如果 A 是正定矩阵, 则它的主子矩阵仍是正定的; 因而 A 的主子式都 > 0.

我们能证明, 如果一个实对称矩阵的所有主子式都 > 0, 则它是正定的. 事实上, 有如下更强的结论

定理

设 A 是 n 阶实对称矩阵, 则 A 是正定矩阵, 当且仅当 A 的所有顺序主子式都 > 0.

证明.

若 A 是正定矩阵, 则 A 的所有主子式都 > 0. 特别地, 顺序主子式都 > 0.
反之, 如果 A 的顺序主子式都大于 0, 我们对 n 用数学归纳法证明 A 合同于 E_n.

证明 (续).

当 $n=1$ 时, 结论是显然的.
假设结论对 $n-1$ 阶矩阵成立, 那么, 对于 n 阶矩阵 A, 可作分块

$$A=\begin{bmatrix} A_1 & \beta \\ \beta^{\mathsf{T}} & a \end{bmatrix}, \quad A_1\in\mathbb{R}^{(n-1)\times(n-1)}, \quad \beta\in\mathbb{R}^{(n-1)\times 1}, \quad a\in\mathbb{R}.$$

注意 A_1 的顺序主子式都大于零, 由归纳假设知 A_1 正定, 即存在可逆矩阵 P_1 使得

$$P_1^{\mathsf{T}}A_1P_1=E_{n-1}.$$

令 $\delta=-A_1^{-1}\beta$, 并令 $P=\begin{bmatrix} P_1 & \delta \\ 0 & 1 \end{bmatrix}$, 则有

$$P^{\mathsf{T}}AP=\begin{bmatrix} E_{n-1} & \\ & c \end{bmatrix}$$

其中 $c=a+\beta^{\mathsf{T}}\delta$. 上式两端取行列式可得 $c=|A|\cdot|P|^2>0$, 因此 A 的正惯性指数为 n, A 是正定的. □

例

判断 3 元二次型 $5x_1^2 + x_2^2 + 5x_3^2 + 4x_1x_2 - 8x_1x_3 - 4x_2x_3$ 是否正定.

解答

注意该二次型的矩阵为

$$\begin{bmatrix} 5 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{bmatrix},$$

它的顺序主子式为 5, 1, 1, 都是正数, 所以该二次型是正定的.

思考题

- (****) 若 $A = (a_{ij})$, B 为实方阵, 定义 $A \otimes B$ 为分块矩阵 $(a_{ij}B)$. 现在, 设 A, B 都是正定矩阵, 证明 $A \otimes B$ 也是正定矩阵.
- (****) 若 $A = (a_{ij})$ 和 $B = (b_{ij})$ 都是 n 阶正定矩阵, 证明 $A \circ B = (a_{ij}b_{ij})$ 也是正定矩阵.

定义

设 f 是 n 元实二次型. 如果 f 的负惯性指数为 n, 则称为负定的; 如果 f 的负惯性指数为零, 则称为半正定的; 如果 f 的正惯性指数为零, 则称为半负定的. 如果 f 的正、负惯性指数都大于零, 则称为不定的. 相应地, 我们也称该二次型的矩阵为负定、半正定、半负定或不定的.

类似于正定矩阵, 我们对半正定矩阵有以下刻画.

定理

设 A 为 n 阶实对称矩阵. 则以下条件等价:

(1) A 是半正定矩阵;

(2) 存在 $P \in \mathbb{R}^{n\times n}$, 使得 $A = P^{\mathsf{T}}P$;

(3) A 的所有主子式都是非负数;

证明.

(1)$\Longrightarrow$(2). 若 A 是半正定矩阵, 则 A 合同于 $D=\operatorname{diag}(E_r,0)$, 即存在可逆矩阵 Q 使得 $A=Q^{\mathsf T}DQ$. 令 $P=DQ$, 就有 $A=P^{\mathsf T}P$.

(2)$\Longrightarrow$(1). 二次型 $f(x)=x^{\mathsf T}Ax=x^{\mathsf T}P^{\mathsf T}Px\geq 0$, 所以 f 的负惯性指数为 0, 即 f 半正定, 从而 A 半正定.

(1)$\Longrightarrow$(3). 当 A 为半正定矩阵时, 二次型 $f(x)=x^{\mathsf T}Ax$ 恒取非负值, 从而任意把其中一部分变量取为零时, 所得的二次型仍非负. 这就说明 A 的任意主子矩阵仍是半正定的. 又由 (2) 可知半正定矩阵的行列式非负, 所以 A 的所有主子式都非负.

(3)$\Longrightarrow$(1). 当 A 的所有主子式都非负时, 我们对 n 用数学归纳法证明 A 是半正定的.
当 $n=1$ 时, 结论是平凡的.
假设结论对 $n-1$ 阶矩阵成立, 那么, 对于 n 阶矩阵 A, 可作分块

$$A=\begin{bmatrix} A_1 & \beta \\ \beta^{\mathsf T} & a \end{bmatrix},\quad A_1\in\mathbb{R}^{(n-1)\times(n-1)},\quad \beta\in\mathbb{R}^{(n-1)\times 1},\quad a\in\mathbb{R}.$$

证明 (续).

注意 a 是一个 1 阶的主子式, 所以 $a \geq 0$.
如果 $a = 0$, 则考虑包含 a 的那些 2 阶主子式即可得到 $\beta = \mathbf{0}$, 这时由归纳假设可知结论成立.
如果 $a > 0$, 我们可对 A 施行成对的初等行列变换 (将最后一列/行的适当倍数加到其他各列/行), 将它变为

$$B = \begin{bmatrix} B_1 & 0 \\ 0 & a \end{bmatrix}, \quad B_1 = A_1 - \beta\beta^{\mathsf{T}}/a.$$

容易看到, 在变换前后, $(i_1, \cdots, i_k, n)$-主子矩阵的行列式是不变的. 也就是说, A 的 $(i_1, \cdots, i_k, n)$-主子式, 等于 B 的 $(i_1, \cdots, i_k, n)$-主子式, 等于 B_1 的 $(i_1, \cdots, i_k)$-主子式乘以 a. 由条件可知, B_1 的所有主子式都非负, 因而 B_1 是半正定的, 它合同于 $\mathrm{diag}(E_r, 0)$. 进而可知 B 合同于 $\mathrm{diag}(E_r, 0, a)$. 可见 B 是半正定的, 从而 A 也是半正定的. □