线性代数

leip@lzu.edu.cn
兰州大学 数学与统计学院

2017 年 10 月 25 日

第六章

二次型

目录

二次型的研究起源于解析几何中化二次曲线与二次曲面的方程为标准形,
二次型理论的一个中心任务就是化二次型为标准形.
现在, 二次型的理论不仅用于几何学中, 而且也广泛的应用于数学的其它分支以及自然科学与工程技术中.
本章我们将介绍二次型的最基本的理论.

目录

在解析几何中, 当坐标原点与中心重合时, 一个有心二次曲线的一般方程可以表示为

$$ax^2 + 2bxy + cy^2 = d. \tag{6.1}$$

为便于研究这个二次曲线的几何性质, 我们可以选择适当的角度 θ, 作转轴变换

$$\begin{cases} x & = x' \cos\theta - y' \sin\theta \\ y & = x' \sin\theta + y' \cos\theta \end{cases} \tag{6.2}$$

把方程 (6.1) 化为标准方程. 在二次曲面的研究中也有类似的情况.
方程 (6.1) 的左端是一个二次齐次多项式. 从代数的观点看, 所谓化为标准方程就是用变量的线性替换 (6.2) 化简一个二次齐次多项式, 使它只含有平方项. 在许多理论问题或实际问题中, 常常需要把二次齐次多项式中变元的个数扩充到 n 个的情形. 为此, 我们引入

定义6.1.1

设 F 是一数域. 一个系数在数域 F 中含有 n 个变元 $x_1, x_2, \cdots, x_n$ 的二次齐次多项式

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) = \quad & a_{11}x_1^2 \quad +2a_{12}x_1x_2 \quad + \cdots \quad +2a_{1n}x_1x_n \\ & \qquad\quad\ +a_{22}x_2^2 \qquad\ + \cdots \quad +2a_{2n}x_2x_n \\ & \qquad\qquad\qquad\qquad\quad + \cdots \\ & \qquad\qquad\qquad\qquad\qquad\qquad +a_{nn}x_n^2, \end{aligned} \tag{6.3}$$

称为数域 F 上的一个 n 元二次型, 简称为二次型. 特别地, 当 $F = R$ 时, 称为实二次型, $F = C$ 时, 称为复二次型.

定义6.1.1

设 F 是一数域. 一个系数在数域 F 中含有 n 个变元 $x_1, x_2, \cdots, x_n$ 的二次齐次多项式

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) = \quad & a_{11}x_1^2 & +2a_{12}x_1x_2 & \quad +\cdots & +2a_{1n}x_1x_n \\ & & +a_{22}x_2^2 & \quad +\cdots & +2a_{2n}x_2x_n \\ & & & \quad +\cdots & \\ & & & & +a_{nn}x_n^2, \end{aligned} \tag{6.3}$$

称为数域 F 上的一个 n 元二次型, 简称为二次型. 特别地, 当 $F = R$ 时, 称为实二次型, $F = C$ 时, 称为复二次型.

例如

$$x_1^2 - x_1x_2 + 3x_1x_3 + 2x_2^2 - 4x_2x_3 + 3x_3^2$$

就是有理数域上的一个三元二次型, 当然也是一个实二次型. 而

$$ix_1^2 + 2(1+i)x_1x_2 - 4x_1x_3 + 2x_2^2 + (1-2i)x_2x_3 + x_3^2$$

是一个三元复二次型.

二次型是多项式, 因此两个 n 元二次型 $f(x_1, x_2, \cdots, x_n)$ 和 $g(x_1, x_2, \cdots, x_n)$ 相等当且仅当它们所有同类项的系数都相等. 从而, 若

$$f(x_1, x_2, \cdots, x_n) = g(x_1, x_2, \cdots, x_n),$$

则对任意 $c_1, c_2, \cdots, c_n \in F$, 有

$$f(c_1, c_2, \cdots, c_n) = g(c_1, c_2, \cdots, c_n).$$

与几何中化简二次曲线和二次曲面的方程一样, 在处理许多其它问题时也常常需要通过变量的线性替换来化简有关的二次型. 为此, 我们引入

定义6.1.2

设 $x_1, x_2, \cdots, x_n; y_1, y_2, \cdots, y_n$ 是两组文字. 系数在数域 F 中的一组关系式

$$\left\{\begin{array}{lcl} x_1 & = & c_{11}y_1 + c_{12}y_2 + \cdots + c_{1n}y_n \\ x_2 & = & c_{21}y_1 + c_{22}y_2 + \cdots + c_{2n}y_n \\ & & \cdots\cdots\cdots \\ x_n & = & c_{n1}y_1 + c_{n2}y_2 + \cdots + c_{nn}y_n \end{array}\right. \tag{6.4}$$

称为由 $x_1, x_2, \cdots, x_n$ 到 $y_1, y_2, \cdots, y_n$ 的一个线性替换(或线性变换), 或简称线性替换(或线性变换). 当 $F = R\ (C)$ 时, 称为实(复)线性替换. 如果 (6.4) 的系数行列式

$$\begin{vmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{vmatrix} \neq 0,$$

那么线性替换 (6.4) 称为非退化的, 否则称为退化的.

显然 (6.2) 是一个非退化的线性替换.
容易看出, 把 (6.4) 代入到 (6.3) 中, 得到的关于 $y_1, y_2, \cdots, y_n$ 的多项式仍然是一个 二次齐次多项式. 故线性替换把二次型变成二次型. 研究二次型在非退化的线性替换下的变化情况是本章的主要目的.
在讨论二次型时, 矩阵是一个强有力的工具. 因此, 我们首先把二次型与线性替换表示成矩阵的形式.

显然 (6.2) 是一个非退化的线性替换.
容易看出, 把 (6.4) 代入到 (6.3) 中, 得到的关于 $y_1, y_2, \cdots, y_n$ 的多项式仍然是一个 二次齐次多项式. 故线性替换把二次型变成二次型. 研究二次型在非退化的线性替换下的变化情况是本章的主要目的.
在讨论二次型时, 矩阵是一个强有力的工具. 因此, 我们首先把二次型与线性替换表示成矩阵的形式.
对于给定的二次型 (6.3), 如果令

$$a_{ji} = a_{ij},\ 1 \leq i < j \leq n,$$

因为 $x_i x_j = x_j x_i$, 那么二次型 (6.3) 可以写成

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) = \ & a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n \\ & + a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n \\ & + \cdots \\ & + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2 \\ = \ & \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j. \end{aligned} \tag{6.5}$$

令 $X=(x_1,x_2,\cdots,x_n)^T$,

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

则 A 是一个对称矩阵, 称为二次型 $f(x_1,x_2,\cdots,x_n)$ 的矩阵. 因此二次型的矩阵都是对称矩阵. 此时

$$f(x_1,x_2,\cdots,x_n)=X^TAX. \tag{6.6}$$

(6.6) 称为二次型 $f(x_1,x_2,\cdots,x_n)$ 的矩阵表示.
从上面的定义可以看到, 二次型 $f(x_1,x_2,\cdots,x_n)$ 的矩阵 A 的元素:
当 $i\neq j$ 时, $a_{ij}=a_{ji}$ 是它的项 x_ix_j 的系数的一半; 而 a_{ii} 是 x_i^2 项的系数. 因此二次型和它的矩阵是相互唯一决定的. 于是我们有

定理6.1.3

若 A, B 都是 n 阶对称矩阵且

$$f(x_1, x_2, \cdots, x_n) = X^T AX = X^T BX,$$

则 $A = B$.

反之, 任意一个 n 阶对称矩阵 $A = (a_{ij})$ 可以视为二次型

$$f(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

的矩阵.

定理6.1.3

若 A, B 都是 n 阶对称矩阵且

$$f(x_1, x_2, \cdots, x_n) = X^T A X = X^T B X,$$

则 $A = B$.
反之, 任意一个 n 阶对称矩阵 $A = (a_{ij})$ 可以视为二次型

$$f(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

的矩阵.

这样我们就建立了数域 F 上的 n 元二次型和 n 阶对称矩阵之间的1-1对应.

补充例题 6.1.1

求二次型 $f(x_1,x_2,x_3)=x_1^2+2x_1x_2+x_2^2-4x_2x_3-x_3^2$ 的矩阵.

补充例题 6.1.1

求二次型 $f(x_1,x_2,x_3)=x_1^2+2x_1x_2+x_2^2-4x_2x_3-x_3^2$ 的矩阵.

解： 根据二次型的矩阵的定义知,其矩阵为：

$$A=\begin{pmatrix}1 & 1 & 0\\ 1 & 1 & -2\\ 0 & -2 & -1\end{pmatrix}.$$

补充例题 6.1.2

求下列二次型的矩阵，$X=(x_1,x_2,x_3)^T$：

1. $f(X,X) \xlongequal{\text{def}} f(x_1,x_2,x_3) = X^TB_1X \xlongequal{\text{def}} (x_1,x_2,x_3)\begin{pmatrix}1&1&0\\1&1&-2\\0&-2&-1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$;
2. $f(X,X)=X^TB_2X=(x_1,x_2,x_3)\begin{pmatrix}1&2&0\\0&1&0\\0&-4&-1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$;

补充例题 6.1.2

求下列二次型的矩阵, $X=(x_1,x_2,x_3)^T$:

1. $f(X,X) \xlongequal{\text{def}} f(x_1,x_2,x_3) = X^TB_1X \xlongequal{\text{def}} (x_1,x_2,x_3)\begin{pmatrix}1 & 1 & 0\\ 1 & 1 & -2\\ 0 & -2 & -1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}$;
2. $f(X,X)=X^TB_2X=(x_1,x_2,x_3)\begin{pmatrix}1 & 2 & 0\\ 0 & 1 & 0\\ 0 & -4 & -1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}$;

解: (1)原二次型为:

$$f(X,X)=x_1^2+2x_1x_2+x_2^2-4x_2x_3-x_3^2,$$

故其矩阵为:

$$A_1=\begin{pmatrix}1 & 1 & 0\\ 1 & 1 & -2\\ 0 & -2 & -1\end{pmatrix}=B_1.$$

(2)原二次型为:

$$f(X,X) = x_1^2 + 2x_1x_2 + x_2^2 - 4x_2x_3 - x_3^2,$$

故其矩阵为：

$$A_2 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & -2 \\ 0 & -2 & -1 \end{pmatrix} \neq B_2!!$$

(2)原二次型为:

$$f(X,X) = x_1^2 + 2x_1x_2 + x_2^2 - 4x_2x_3 - x_3^2,$$

故其矩阵为：

$$A_2 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & -2 \\ 0 & -2 & -1 \end{pmatrix} \neq B_2!!$$

实际上,

$$A_2 = \frac{B_2 + B_2^T}{2}.$$

(2)原二次型为:

$$f(X,X) = x_1^2 + 2x_1x_2 + x_2^2 - 4x_2x_3 - x_3^2,$$

故其矩阵为：

$$A_2 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & -2 \\ 0 & -2 & -1 \end{pmatrix} \neq B_2!!$$

实际上,

$$A_2 = \frac{B_2 + B_2^T}{2}.$$

☞ 二次型 $f(X,X) = X^TBX$ 的矩阵为:

$$A = \frac{B + B^T}{2}.$$

如果令 $X=(x_1,x_2,\cdots,x_n)^T, Y=(y_1,y_2,\cdots,y_n)^T$,

$$C=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix},$$

那么线性替换 (6.4) 可以表示成

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

或者

$$X=CY.$$

现在我们来讨论一个二次型经过非退化的线性替换所得到的新二次型与原来的二次型之间的关系, 也就是讨论这两个二次型的矩阵之间的关系.

设 $f(x_1, x_2, \cdots, x_n)$ 是一个 n 元二次型, A 是它的矩阵, 即 A 是 n 阶对称矩阵且

$$f(x_1, x_2, \cdots, x_n) = X^T AX. \tag{6.7}$$

又设 $f(x_1, x_2, \cdots, x_n)$ 经过非退化线性替换

$$X = CY$$

得到一个 $y_1, y_2, \cdots, y_n$ 的二次型

$$g(y_1, y_2, \cdots, y_n) = Y^T BY,$$

其中 B 是一个对称矩阵. 我们考察矩阵 B 与 A 的关系. 为此, 将 $X = CY$ 代入到 (6.7) 得

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) &= X^T AX = (CY)^T A(CY) = Y^T (C^T AC)Y \\ &= Y^T BY = g(y_1, y_2, \cdots, y_n). \end{aligned}$$

由于 A 是对称矩阵, 因此 C^TAC 也是对称的. 据定理 6.1.1 得

$$B = C^TAC.$$

这给出了经过非退化线性替换前后两个二次型的矩阵的关系. 为方便起见, 我们引入

定义6.1.4

数域 F 上的两个 n 阶矩阵 A, B 称为合同的, 如果存在数域 F 上的可逆矩阵 C 使得

$$B = C^TAC.$$

由于 A 是对称矩阵, 因此 C^TAC 也是对称的. 据定理 6.1.1 得

$$B = C^TAC.$$

这给出了经过非退化线性替换前后两个二次型的矩阵的关系. 为方便起见, 我们引入

定义6.1.4

数域 F 上的两个 n 阶矩阵 A, B 称为合同的, 如果存在数域 F 上的可逆矩阵 C 使得

$$B = C^TAC.$$

类似于矩阵的等价、相似关系, 矩阵的合同关系也具有
(1) 自反性: 任意方阵 A 与自身合同;
(2) 对称性: 若 A 与 B 合同, 则 B 与 A 合同;
(3) 传递性: 若 A 与 B 合同, B 与 C 合同, 则 A 与 C 合同.
此外, 容易看出合同的矩阵有相同的秩; 若 A 与 B 合同, A 是对称矩阵, 则 B 也是对称矩阵.

综上所述, 我们有

定理6.1.5

一个二次型经过非退化的线性替换所得新二次型的矩阵与原二次型的矩阵是合同的.

综上所述, 我们有

定理6.1.5

一个二次型经过非退化的线性替换所得新二次型的矩阵与原二次型的矩阵是合同的.

最后指出, 在变换二次型时, 我们总要求所作的线性替换是非退化的, 这从几何上看是自然的. 另一方面, 当线性替换 $X = CY$ 是非退化时, 线性替换 $Y = C^{-1}X$ 也是非退化的. 如果二次型 $f(x_1, x_2, \cdots, x_n)$ 经过非退化线性替换 $X = CY$ 化成二次型 $g(y_1, y_2, \cdots, y_n)$ 时, 线性替换 $Y = C^{-1}X$ 就把 $g(y_1, y_2, \cdots, y_n)$ 还原为 $f(x_1, x_2, \cdots, x_n)$. 这样就可以从所得新二次型的性质推知原二次型的一些性质.

目录

本节讨论用非退化的线性替换化简二次型的问题. 一般地, 我们可以认为最简单的二次型是只包含平方项的二次型

$$d_1x_1^2 + d_2x_2^2 + \cdots + d_nx_n^2. \tag{6.1}$$

那么一个二次型能否通过非退化的线性替换化成 (6.1) 的形式呢? 下面的定理给出了肯定的回答.

本节讨论用非退化的线性替换化简二次型的问题. 一般地, 我们可以认为最简单的二次型是只包含平方项的二次型

$$d_1x_1^2 + d_2x_2^2 + \cdots + d_nx_n^2. \tag{6.1}$$

那么一个二次型能否通过非退化的线性替换化成 (6.1) 的形式呢? 下面的定理给出了肯定的回答.

定理6.2.1

数域 F 上任意一个二次型都可以经过非退化的线性替换化成 *(6.1)* 的形式.

本节讨论用非退化的线性替换化简二次型的问题. 一般地, 我们可以认为最简单的二次型是只包含平方项的二次型

$$d_1x_1^2 + d_2x_2^2 + \cdots + d_nx_n^2. \tag{6.1}$$

那么一个二次型能否通过非退化的线性替换化成 (6.1) 的形式呢? 下面的定理给出了肯定的回答.

定理6.2.1

数域 F 上任意一个二次型都可以经过非退化的线性替换化成 *(6.1)* 的形式.

证明 设 $f(x_1, x_2, \cdots, x_n)$ 是数域 F 上的一个 n 元实二次型. 我们对变元的个数 n 归纳证明定理的结论成立.

当 $n = 1$ 时, $f(x_1) = a_{11}x_1^2$, 结论自然成立.

假设对 $n - 1$ 元的二次型, 定理结论成立. 我们考察 n 元的情况.

设

$$f(x_1,x_2,\cdots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j,$$

其中 $a_{ij}=a_{ji}$, 是一个 n 元二次型. 以下分三种情况讨论.
(1) $a_{11},a_{22},\cdots,a_{nn}$ 中至少有一个不为零, 不妨设 $a_{11}\neq 0$. 则有

$$\begin{aligned}
f(x_1,x_2,\cdots,x_n) &= a_{11}x_1^2+\sum_{j=2}^{n}a_{1j}x_1x_j+\sum_{i=2}^{n}a_{i1}x_ix_1+\sum_{i=2}^{n}\sum_{j=2}^{n}a_{ij}x_ix_j\\
&= a_{11}x_1^2+\sum_{j=2}^{n}2a_{1j}x_1x_j+\sum_{i=2}^{n}\sum_{j=2}^{n}a_{ij}x_ix_j\\
&= a_{11}(x_1+\sum_{j=2}^{n}a_{11}^{-1}a_{1j}x_j)^2-a_{11}^{-1}(\sum_{j=2}^{n}a_{1j}x_j)^2+\sum_{i=2}^{n}\sum_{j=2}^{n}a_{ij}x_ix_j\\
&= a_{11}(x_1+\sum_{j=2}^{n}a_{11}^{-1}a_{1j}x_j)^2+\sum_{i=2}^{n}\sum_{j=2}^{n}b_{ij}x_ix_j,
\end{aligned}$$

其中

$$\sum_{i=2}^{n}\sum_{j=2}^{n}b_{ij}x_ix_j=-a_{11}^{-1}(\sum_{j=2}^{n}a_{1j}x_j)^2+\sum_{i=2}^{n}\sum_{j=2}^{n}a_{ij}x_ix_j$$

是一个 $n-1$ 元二次型.

于是线性替换

$$\begin{cases} y_1 &= x_1 + \sum\limits_{j=2}^{n} a_{11}^{-1} a_{1j} x_j, \\ y_2 &= x_2, \\ &\cdots\cdots \\ y_n &= x_n. \end{cases} \quad \text{或} \quad \begin{cases} x_1 &= y_1 - \sum\limits_{j=2}^{n} a_{11}^{-1} a_{1j} y_j, \\ x_2 &= y_2, \\ &\cdots\cdots \\ x_n &= y_n. \end{cases}$$

把二次型化为

$$f(x_1, x_2, \cdots, x_n) = a_{11} y_1^2 + \sum_{i=2}^{n} \sum_{j=2}^{n} b_{ij} y_i y_j.$$

根据归纳假设, 存在非退化线性替换

$$\begin{cases} z_2 &= c_{22} y_2 + c_{23} y_3 + \cdots + c_{2n} y_n, \\ z_3 &= c_{32} y_2 + c_{33} y_3 + \cdots + c_{3n} y_n, \\ &\cdots\cdots \\ z_n &= c_{n2} y_2 + c_{n3} y_3 + \cdots + c_{nn} y_n, \end{cases}$$

把 $n-1$ 元二次型 $\sum\limits_{i=2}^{n}\sum\limits_{j=2}^{n} b_{ij}y_iy_j$ 化为

$$d_2z_2^2+d_3z_3^2+\cdots+d_nz_n^2$$

的形式. 于是非退化线性替换

$$\begin{cases} z_1 &= y_1, \\ z_2 &= c_{22}y_2+c_{23}y_3+\cdots+c_{2n}y_n, \\ & \cdots\cdots \\ z_n &= c_{n2}y_2+c_{n3}y_3+\cdots+c_{nn}y_n, \end{cases}$$

把 $f(x_1,x_2,\cdots x_n)$ 化成

$$f(x_1,x_2,\cdots x_n)=a_{11}z_1^2+d_2z_2^2+d_3z_3^2+\cdots+d_nz_n^2$$

的形式.

(2) $a_{11}=a_{22}=\cdots=a_{nn}=0$, 至少有一个含 x_1 的混合项的系数不等于零. 不妨设 $a_{12}\neq 0$. 作非退化的线性替换

$$\begin{cases} x_1 & = z_1+z_2, \\ x_2 & = z_1-z_2, \\ x_3 & = z_3, \\ & \cdots\cdots \\ x_n & = z_n, \end{cases}$$

得

$$f(x_1,x_2,\cdots,x_n)=2a_{12}z_1^2-2a_{12}z_2^2+\cdots$$

由 (1) 知存在非退化的线性替换把上述二次型化为平方和的形式.

(3) $a_{11} = a_{12} = \cdots = a_{1n} = 0$. 则 $a_{21} = a_{31} = \cdots = a_{n1} = 0$. 此时

$$f(x_1, x_2, \cdots, x_n) = \sum_{i=2}^{n}\sum_{j=2}^{n} a_{ij}x_i x_j$$

是一个 $n-1$ 元二次型, 根据归纳假设, 它可以用非退化的线性替换化成平方和. 综上所述, 根据归纳法原理, 定理成立.

(3) $a_{11} = a_{12} = \cdots = a_{1n} = 0$. 则 $a_{21} = a_{31} = \cdots = a_{n1} = 0$. 此时

$$f(x_1, x_2, \cdots, x_n) = \sum_{i=2}^{n}\sum_{j=2}^{n} a_{ij}x_ix_j$$

是一个 $n-1$ 元二次型, 根据归纳假设, 它可以用非退化的线性替换化成平方和. 综上所述, 根据归纳法原理, 定理成立.
定理的证明实际上给出了把一个二次型化成平方和的具体方法, 这实质上就是中学代数中配平方的方法, 称之为 "配方法".
容易看出, 二次型 (6.1) 的矩阵是对角矩阵. 反之, 任意矩阵为对角矩阵的二次型一定具有形式 (6.1), 即平方和的形式. 这样用矩阵的语言, 定理 6.2.1 可以叙述为

定理6.2.2

数域 F 上的任意对称矩阵都合同于对角矩阵.

定理6.2.2

数域 F 上的任意对称矩阵都合同于对角矩阵.

二次型 $f(x_1, x_2, \cdots, x_n)$ 经过非退化线性替换所化成的平方和称为该二次型的标准形. 则前述定理说明,数域 F 上的任何二次型都可化为标准形.

定理6.2.2

数域 F 上的任意对称矩阵都合同于对角矩阵.

二次型 $f(x_1, x_2, \cdots, x_n)$ 经过非退化线性替换所化成的平方和称为该二次型的标准形. 则前述定理说明,数域 F 上的任何二次型都可化为标准形.

例6.2.1

化二次型

$$f(x_1, x_2, x_3) = x_1^2 + 4x_1x_2 - 4x_1x_3 + 2x_2^2 - 4x_2x_3 + 3x_3^2$$

为标准形, 并给出所用的非退化线性替换.

定理6.2.2

数域 F 上的任意对称矩阵都合同于对角矩阵.

二次型 $f(x_1, x_2, \cdots, x_n)$ 经过非退化线性替换所化成的平方和称为该二次型的**标准形**. 则前述定理说明,数域 F 上的任何二次型都可化为标准形.

例6.2.1

化二次型

$$f(x_1, x_2, x_3) = x_1^2 + 4x_1x_2 - 4x_1x_3 + 2x_2^2 - 4x_2x_3 + 3x_3^2$$

为标准形, 并给出所用的非退化线性替换.

解 对所给二次型进行配方得

$$\begin{aligned} f(x_1, x_2, x_3) &= (x_1 + 2x_2 - 2x_3)^2 - 4x_2^2 - 4x_3^2 + 8x_2x_3 + 2x_2^2 - 4x_2x_3 + 3x_3^2 \\ &= (x_1 + 2x_2 - 2x_3)^2 - 2x_2^2 + 4x_2x_3 - x_3^2 \\ &= (x_1 + 2x_2 - 2x_3)^2 - 2(x_2 - x_3)^2 + x_3^2. \end{aligned}$$

作线性替换

$$\begin{cases} y_1 &= x_1 + 2x_2 - 2x_3, \\ y_2 &= x_2 - x_3, \\ y_3 &= x_3, \end{cases}$$

即

$$\begin{cases} x_1 &= y_1 - 2y_2, \\ x_2 &= y_2 + y_3, \\ x_3 &= y_3, \end{cases} \tag{6.2}$$

把二次型化成标准形

$$f(x_1, x_2, x_3) = y_1^2 - 2y_2^2 + y_3^2.$$

因为

$$\begin{vmatrix} 1 & -2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix} = 1 \neq 0,$$

所以所作的线性替换 (6.2) 是非退化的.

上述配方过程可以用矩阵的形式表述出来.
所给二次型的矩阵

$$A=\begin{pmatrix}1&2&-2\\2&2&-2\\-2&-2&3\end{pmatrix},$$

取

$$C=\begin{pmatrix}1&-2&0\\0&1&1\\0&0&1\end{pmatrix},$$

则 C 可逆且有

$$C^TAC=\begin{pmatrix}1&0&0\\-2&1&0\\0&1&1\end{pmatrix}\begin{pmatrix}1&2&-2\\2&2&-2\\-2&-2&3\end{pmatrix}\begin{pmatrix}1&-2&0\\0&1&1\\0&0&1\end{pmatrix}=\begin{pmatrix}1&0&0\\0&-2&0\\0&0&1\end{pmatrix}.$$

例6.2.2

化二次型

$$f(x_1,x_2,x_3)=2x_1x_2+2x_1x_3-6x_2x_3$$

为标准形, 并给出所用的非退化线性替换.

例6.2.2

化二次型
$$f(x_1,x_2,x_3)=2x_1x_2+2x_1x_3-6x_2x_3$$
为标准形, 并给出所用的非退化线性替换.

解 由于所给二次型不含平方项, 首先作非退化线性替换
$$\begin{cases} x_1 &= y_1+y_2, \\ x_2 &= y_1-y_2, \\ x_3 &= y_3, \end{cases}$$
把二次型化为含平方项的形式
$$\begin{aligned} f(x_1,x_2,x_3) &= 2(y_1+y_2)(y_1-y_2)+2(y_1+y_2)y_3-6(y_1-y_2)y_3 \\ &= 2y_1^2-4y_1y_3-2y_2^2+8y_2y_3. \end{aligned}$$
然后再对新的二次型进行配方得
$$\begin{aligned} f(x_1,x_2,x_3) &= 2(y_1-y_3)^2-2y_2^2+8y_2y_3-2y_3^2 \\ &= 2(y_1-y_3)^2-2(y_2-2y_3)^2+6y_3^2. \end{aligned}$$

作线性替换

$$\begin{cases} z_1 &= y_1 - y_3, \\ z_2 &= y_2 - 2y_3, \\ z_3 &= y_3, \end{cases} \quad \text{或} \quad \begin{cases} y_1 &= z_1 + z_3, \\ y_2 &= z_2 + 2z_3, \\ y_3 &= z_3, \end{cases}$$

把二次型化为标准形

$$f(x_1, x_2, x_3) = 2z_1^2 - 2z_2^2 + 6z_3^2.$$

所用的线性替换是

$$\begin{aligned} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 & 3 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}, \end{aligned}$$

由于

$$\begin{vmatrix} 1 & 1 & 3 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{vmatrix} = -2 \neq 0,$$

这是一个非退化线性替换.
因为原二次型的矩阵

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -3 \\ 1 & -3 & 0 \end{pmatrix},$$

所以

$$\begin{pmatrix} 1 & 1 & 3 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix}^T \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -3 \\ 1 & -3 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 3 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 6 \end{pmatrix}.$$

需要注意的是, 二次型的标准形是不唯一的. 如在例 6.2.2 中, 进一步作非退化线性替换

$$\begin{cases} z_1 &= \frac{1}{\sqrt{2}} w_1, \\ z_2 &= \frac{1}{\sqrt{2}} w_2, \\ z_3 &= \frac{1}{\sqrt{6}} w_3, \end{cases}$$

就把原二次型化成标准形

$$f(x_1, x_2, x_3) = w_1^2 - w_2^2 + w_3^2.$$

利用矩阵的工具, 我们可以给出用非退化线性替换化二次型为标准形的另一种方法——同步初等变换法, 等价地, 这也是把对称矩阵合同变换为对角矩阵的方法.

根据定理 6.2.2, 对任意对称矩阵 A, 总存在可逆矩阵 C 使 $D = C^T AC$ 是对角矩阵. 由 C 可逆知 C 可以表示成初等矩阵的乘积, 即存在初等矩阵 $P_1, P_2, \cdots, P_s$ 使

$$C = P_1 P_2 \cdots P_s. \tag{6.3}$$

于是

$$\begin{aligned} D = C^T AC &= (P_1 P_2 \cdots P_s)^T A P_1 P_2 \cdots P_s \\ &= P_s^T (\cdots (P_2^T (P_1^T A P_1) P_2) \cdots) P_s. \end{aligned} \tag{6.4}$$

为此需要考察矩阵 $P_i^T B P_i$ 与矩阵 B 的关系, 其中 P_i 是初等矩阵.

初等矩阵有三类: $P(i,j)$, $P(i(c))$, $P(i,j(k))$. 显然

$$P(i,j)^T = P(i,j), \quad P(i(c))^T = P(i(c)), \quad P(i,j(k))^T = P(j,i(k)).$$

因此

$$P(i,j)^T BP(i,j) = P(i,j)BP(i,j),$$

这相当于先交换 B 的第 i 行与第 j 行, 然后交换所得矩阵的第 i 列与第 j 列.

$$P(i(c))^T BP(i(c)) = P(i(c))BP(i(c)),$$

这相当于先用非零数 c 乘以 B 的第 i 行, 然后用 c 乘以所得矩阵的第 i 列.

$$P(i,j(k))^T BP(i,j(k)) = P(j,i(k))BP(i,j(k)),$$

这相当于把 B 的第 i 行的 k 倍加到第 j 行, 然后把所得矩阵的第 i 列的 k 倍加到第 j 列.

综上所述, 若 P_i 是一个初等矩阵, 则 $P_i^T BP_i$ 相当于先对 B 作一次初等行变换, 然后对所得矩阵作一次同样的初等列变换. 为方便起见, 我们称之为对矩阵实施同步初等变换.

定理 6.2.2 及 (6.4) 表明, 任意对称矩阵都可以经过适当同步初等变换化为对角矩阵.

为了求出上述的可逆矩阵 C, 比较 (6.3) 和 (6.4) 便有: 当我们对 A 施行同步初等变换把 A 化成对角矩阵时, 对 E 只作其中的列变换, 则 E 就变成了 C. 于是我们就得到了用非退化线性替换化二次型为标准形的第二种方法-初等变换法: 设二次型 $f(x_1, x_2, \cdots, x_n)$ 的矩阵为 A. 作初等变换

$$\begin{pmatrix} A \\ E \end{pmatrix} \xrightarrow[\text{对}E\text{只作其中的初等列变换}]{\text{对}A\text{作同步初等变换}} \begin{pmatrix} D \\ C \end{pmatrix},$$

则

$$D = C^T AC.$$

若 $D = diag\{d_1, d_2, \cdots, d_n\}$ 是对角矩阵, 则作非退化线性替换 $X = CY$ 就可以把 $f(x_1, x_2, \cdots, x_n)$ 化成

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) &= X^T AX = Y^T (C^T AC)Y = Y^T DY \\ &= d_1 y_1^2 + d_2 y_2^2 + \cdots + d_n y_n^2. \end{aligned}$$

例6.2.3

化二次型

$$f(x_1, x_2, x_3) = 2x_1x_2 + 2x_1x_3 - 6x_2x_3$$

为标准形, 并给出所用的非退化线性替换.

例6.2.3

化二次型

$$f(x_1,x_2,x_3)=2x_1x_2+2x_1x_3-6x_2x_3$$

为标准形, 并给出所用的非退化线性替换.

解 $f(x_1,x_2,x_3)$ 的矩阵为

$$A=\begin{pmatrix}0&1&1\\1&0&-3\\1&-3&0\end{pmatrix}.$$

作初等变换

$$\begin{pmatrix}A\\E\end{pmatrix}=\begin{pmatrix}0&1&1\\1&0&-3\\1&-3&0\\1&0&0\\0&1&0\\0&0&1\end{pmatrix}\to\begin{pmatrix}1&1&-2\\1&0&-3\\1&-3&0\\1&0&0\\0&1&0\\0&0&1\end{pmatrix}\to\begin{pmatrix}2&1&-2\\1&0&-3\\-2&-3&0\\1&0&0\\1&1&0\\0&0&1\end{pmatrix}$$

$$\to \begin{pmatrix} 2 & 1 & -2 \\ 0 & -1/2 & -2 \\ -2 & -3 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 2 & 0 & -2 \\ 0 & -1/2 & -2 \\ -2 & -2 & 0 \\ 1 & -1/2 & 0 \\ 1 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\to \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1/2 & -2 \\ 0 & -2 & -2 \\ 1 & -1/2 & 1 \\ 1 & 1/2 & 1 \\ 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & 6 \\ 1 & -1/2 & 3 \\ 1 & 1/2 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

于是

$$C^T AC = diag\{2, -1/2, 6\}.$$

所作线性替换为 $X = CY$, 即

$$\begin{cases} x_1 &= y_1 - 1/2y_2 + 3y_3, \\ x_2 &= y_1 + 1/2y_2 - y_3, \\ x_3 &= y_3. \end{cases}$$

所得的标准形为

$$f(x_1, x_2, x_3) = 2y_1^2 - \frac{1}{2}y_2^2 + 6y_3^2.$$

所作线性替换为 $X = CY$, 即

$$\begin{cases} x_1 & = y_1 - 1/2y_2 + 3y_3, \\ x_2 & = y_1 + 1/2y_2 - y_3, \\ x_3 & = y_3. \end{cases}$$

所得的标准形为

$$f(x_1, x_2, x_3) = 2y_1^2 - \frac{1}{2}y_2^2 + 6y_3^2.$$

对于实二次型, 除了用配方法和初等变换法把它化成标准形外, 还有一种重要的方法-正交变换法.

设 $f(x_1, x_2, \cdots, x_n) = X^T AX$ 是实二次型, 则它的矩阵 A 是实对称矩阵. 据定理 5.3.3 知, 存在正交矩阵 P, 使得

$$P^{-1}AP = P^T AP = diag\{\lambda_1, \lambda_2, \cdots, \lambda_n\},$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 A 的全部特征值.

用二次型的语言表述就是

定理6.2.3

任意一个实二次型 $f(x_1, x_2, \cdots, x_n) = X^T AX$ 都可以经过非退化的实线性替换 $X = PY$ 化成

$$f(x_1, x_2, \cdots, x_n) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2,$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 A 的全部特征值, P 是正交矩阵.

用二次型的语言表述就是

定理6.2.3

任意一个实二次型 $f(x_1, x_2, \cdots, x_n) = X^TAX$ 都可以经过非退化的实线性替换 $X = PY$ 化成

$$f(x_1, x_2, \cdots, x_n) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2,$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 A 的全部特征值, P 是正交矩阵.

如果 P 是正交矩阵, 那么称线性替换 $X = PY$ 为正交(线性)替换.

例6.2.4

用正交线性替换化实二次型

$$f(x_1, x_2, x_3) = x_1^2 + 4x_1x_2 + 4x_1x_3 + x_2^2 + 4x_2x_3 + x_3^2$$

为标准形, 并给出所用的正交线性替换.

解 二次型的矩阵为

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

据例 5.3.1 知, 正交矩阵

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}$$

使得 $P^{-1}AP = P^TAP = diag\{-1, -1, 5\}$. 因此作正交线性替换 $X = PY$ 就可以把二次型化为标准形 $-y_1^2 - y_2^2 + 5y_3^2$.

例6.2.5

设平面上一条有心二次曲线的方程为

$$x^2 + 4xy + y^2 = 3.$$

问这是什么曲线？

例6.2.5

设平面上一条有心二次曲线的方程为

$$x^2 + 4xy + y^2 = 3.$$

问这是什么曲线？

解 考察二次型 $f(x, y) = x^2 + 4xy + y^2$. 它的矩阵

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

A 的特征多项式为

$$|\lambda E - A| = \begin{vmatrix} \lambda - 1 & -2 \\ -2 & \lambda - 1 \end{vmatrix} = (\lambda + 1)(\lambda - 3).$$

A 的特征值为 $\lambda_1 = -1$, $\lambda_2 = 3$.
对应特征值 -1, 解得 $(-E - A)X = 0$ 的一个基础解系为 $(1, -1)^T$, 单位化得 $\eta_1 = \frac{1}{\sqrt{2}}(1, -1)^T$.

对应特征值 3, 解得 $(3E-A)X=0$ 的一个基础解系 $(1,1)^T$, 单位化得 $\eta_2=\frac{1}{\sqrt{2}}(1,1)^T$.
令

$$P=(\eta_1\ \eta_2)=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}.$$

则 P 是正交矩阵, 且 $P^TAP=diag\{-1,3\}$.
作正交线性替换

$$\begin{pmatrix}x\\ y\end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}\begin{pmatrix}x'\\ y'\end{pmatrix}=\begin{pmatrix}cos\frac{-\pi}{4} & -sin\frac{-\pi}{4}\\ sin\frac{-\pi}{4} & cos\frac{-\pi}{4}\end{pmatrix}\begin{pmatrix}x'\\ y'\end{pmatrix},$$

则二次曲线的方程变为

$$-x'^2+3y'^2=3,$$

即

$$-\frac{x'^2}{3}+y'^2=1.$$

所以这是一条双曲线.

补充例题 6.2.1

化下列二次型为标准形，并给出所用的可逆线性替换。

1. $f(x_1, x_2, x_3) = 2x_1^2 + 2x_2^2 + 2x_3^2 + 2x_1x_2 + 2x_1x_3 + 2x_2x_3$；
2. $f(x_1, x_2, x_3) = 2x_1^2 + 2x_2^2 + 2x_3^2 - 2x_1x_2 - 2x_1x_3 - 2x_2x_3$。

补充例题 6.2.1

化下列二次型为标准形，并给出所用的可逆线性替换。

1. $f(x_1,x_2,x_3)=2x_1^2+2x_2^2+2x_3^2+2x_1x_2+2x_1x_3+2x_2x_3$；
2. $f(x_1,x_2,x_3)=2x_1^2+2x_2^2+2x_3^2-2x_1x_2-2x_1x_3-2x_2x_3$。

解 (1)因为

$$\begin{aligned} f(x_1,x_2,x_3) &= 2x_1^2+2x_2^2+2x_3^2+2x_1x_2+2x_1x_3+2x_2x_3 \quad (6.5)\\ &= (x_2+x_3)^2+(x_1+x_3)^2+(x_1+x_2)^2 \quad (6.6)\end{aligned}$$

故只需作线性替换

$$\begin{cases} y_1 = x_2+x_3, \\ y_2 = x_1+x_3, \\ y_3 = x_1+x_2, \end{cases}$$

把二次型化成标准形

$$f(x_1,x_2,x_3)=y_1^2+y_2^2+y_3^2.$$

所用的可逆线性替换为

$$\begin{cases} x_1 &= \frac{1}{2}(y_2 + y_3 - y_1), \\ x_2 &= \frac{1}{2}(y_1 + y_3 - y_2), \\ x_3 &= \frac{1}{2}(y_1 + y_2 - y_3), \end{cases} \tag{6.7}$$

因为

$$\begin{vmatrix} -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = \frac{1}{8} \begin{vmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{vmatrix} = \frac{1}{2} \neq 0,$$

所以所作的线性替换 (6.7) 是非退化的.

(2)因为

$$f(x_1,x_2,x_3) = 2x_1^2+2x_2^2+2x_3^2-2x_1x_2-2x_1x_3-2x_2x_3 \tag{6.8}$$

$$= (x_2-x_3)^2+(x_1-x_3)^2+(x_1-x_2)^2 \tag{6.9}$$

故只需作线性替换

$$\begin{cases} y_1 = x_2-x_3, \\ y_2 = x_1-x_3, \\ y_3 = x_1-x_2, \end{cases} \tag{6.10}$$

把二次型化成标准形

$$f(x_1,x_2,x_3) = y_1^2+y_2^2+y_3^2.$$

所用的可逆线性替换为

$$\begin{cases} x_1 & =?!, \\ x_2 & =?!, \\ x_3 & =?!, \end{cases} \tag{6.11}$$

因为

$$\begin{vmatrix} 0 & 1 & -1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{vmatrix} = 0,$$

所以所作的线性替换 (6.10) 是退化的.

☞ 错！

☞ 正确解法：

$$
\begin{aligned}
f(x_1,x_2,x_3) &= 2x_1^2+2x_2^2+2x_3^2-2x_1x_2-2x_1x_3-2x_2x_3 && (6.12)\\
&= 2(x_1-\frac{1}{2}x_2-\frac{1}{2}x_3)^2+\frac{3}{2}(x_2^2+x_3^2)-3x_2x_3 && (6.13)\\
&= 2(x_1-\frac{1}{2}x_2-\frac{1}{2}x_3)^2+\frac{3}{2}(x_2-x_3)^2 && (6.14)
\end{aligned}
$$

故只需作线性替换

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

☞ 正确解法：

$$
\begin{aligned}
f(x_1,x_2,x_3) &= 2x_1^2+2x_2^2+2x_3^2-2x_1x_2-2x_1x_3-2x_2x_3 && (6.12)\\
&= 2(x_1-\frac{1}{2}x_2-\frac{1}{2}x_3)^2+\frac{3}{2}(x_2^2+x_3^2)-3x_2x_3 && (6.13)\\
&= 2(x_1-\frac{1}{2}x_2-\frac{1}{2}x_3)^2+\frac{3}{2}(x_2-x_3)^2 && (6.14)
\end{aligned}
$$

故只需作线性替换

$$
\begin{cases}
y_1 &= x_1-\frac{1}{2}x_2-\frac{1}{2}x_3,\\
y_2 &= x_2-x_3,\\
y_3 &= x_3,
\end{cases} \tag{6.15}
$$

即

$$
\begin{cases}
x_1 &= y_1+y_2+2y_3,\\
x_2 &= y_2+y_3,\\
x_3 &= y_3,
\end{cases} \tag{6.16}
$$

把二次型化成标准形

$$f(x_1,x_2,x_3)=2y_1^2+\frac{3}{2}y_2^2.$$

目录

由上一节的讨论我们知道二次型的标准形不是唯一的. 但是经过非退化线性替换, 新二次型的矩阵与原二次型的矩阵是合同的, 而合同的矩阵有相同的秩. 由于标准形的矩阵是对角矩阵, 对角矩阵的秩等于它对角线上非零元素的个数. 因此, 在一个二次型的标准形中, 系数不为零的平方项的个数等于二次型的矩阵的秩, 它是由二次型唯一确定的, 与所作的非退化线性替换无关. 于是我们有

定理6.3.1

一个二次型的标准形中, 系数不为零的平方项的个数与所作非退化线性替换无关, 它就等于该二次型的矩阵的秩.

由上一节的讨论我们知道二次型的标准形不是唯一的. 但是经过非退化线性替换, 新二次型的矩阵与原二次型的矩阵是合同的, 而合同的矩阵有相同的秩. 由于标准形的矩阵是对角矩阵, 对角矩阵的秩等于它对角线上非零元素的个数. 因此, 在一个二次型的标准形中, 系数不为零的平方项的个数等于二次型的矩阵的秩, 它是由二次型唯一确定的, 与所作的非退化线性替换无关. 于是我们有

定理6.3.1

一个二次型的标准形中, 系数不为零的平方项的个数与所作非退化线性替换无关, 它就等于该二次型的矩阵的秩.

我们把二次型矩阵的秩也称为二次型的秩. 这是二次型的一个在非退化线性替换下的不变量.

我们看到, 在一般数域内, 二次型的标准型中系数不为零的平方项的个数(等于二次型的秩)是唯一的, 但非零平方项的系数是可以任意的. 如果我们限定在特殊的数域内, 标准形还可以进一步化简. 以下我们分别就复数域和实数域分别进行讨论. 首先考察复数域的情形.
设 $f(x_1, x_2, \cdots, x_n)$ 是一个复二次型. 则由定理 6.2.1 知, 经过一适当的(复系数)非退化线性替换, $f(x_1, x_2, \cdots, x_n)$ 可以化成标准形, 不妨设它的标准型为

$$d_1y_1^2 + d_2y_2^2 + \cdots + d_ry_r^2, \tag{6.1}$$

其中 r 是二次型的秩, $d_i \neq 0,\ i = 1, 2, \cdots, r$.
注意到复数总可以开平方.

进一步作非退化线性替换

$$\begin{cases} y_1 & = (1/\sqrt{d_1})z_1, \\ & \cdots\cdots, \\ y_r & = (1/\sqrt{d_r})z_r, \\ y_{r+1} & = z_{r+1}, \\ & \cdots\cdots, \\ y_n & = z_n, \end{cases}$$

把 $f(x_1, x_2, \cdots, x_n)$ 变成

$$z_1^2 + z_2^2 + \cdots + z_r^2. \tag{6.2}$$

称 (6.2) 为复二次型 $f(x_1, x_2, \cdots, x_n)$ 的规范形. 显然, 规范形完全由原二次型的秩唯一确定. 因此我们有

定理6.3.2

任意复二次型总可以经过适当的非退化线性替换可化成规范形, 规范形是唯一的, 它由二次型的秩唯一确定.

定理6.3.2

任意复二次型总可以经过适当的非退化线性替换可化成规范形, 规范形是唯一的, 它由二次型的秩唯一确定.

用矩阵的语言, 定理 6.3.2 可以表述为

定理6.3.3

复数域上的任意对称矩阵 A 都合同于一个形如

$$D = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

的对角矩阵, 其中 $r = r(A)$. D 称为 A 的合同标准形.

定理6.3.2

任意复二次型总可以经过适当的非退化线性替换可化成规范形, 规范形是唯一的, 它由二次型的秩唯一确定.

用矩阵的语言, 定理 6.3.2 可以表述为

定理6.3.3

复数域上的任意对称矩阵 A 都合同于一个形如

$$D=\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

的对角矩阵, 其中 $r=r(A)$. D 称为 A 的合同标准形.

作为定理 6.3.3 的推论, 我们有

推论6.3.4

复数域上的两个 n 阶对称矩阵 A 与 B 合同的充分必要条件是 $r(A)=r(B)$.

接下来我们考察实数域的情形.

设 $f(x_1, x_2, \cdots, x_n)$ 是一实二次型. 则 $f(x_1, x_2, \cdots, x_n)$ 可经过适当的非退化实线性替换化成标准形, 如果需要可以适当排列文字的次序(这相当于作一个非退化线性替换), 不妨设标准形为

$$d_1 y_1^2 + \cdots + d_p y_p^2 - d_{p+1} y_{p+1}^2 - \cdots - d_r y_r^2,$$

其中 r 等于二次型的秩, $d_i > 0$, $i = 1, 2, \cdots, r$. 进一步作实非退化线性替换

$$\begin{cases} y_1 &= (1/\sqrt{d_1}) z_1, \\ & \cdots\cdots, \\ y_r &= (1/\sqrt{d_r}) z_r, \\ y_{r+1} &= z_{r+1}, \\ & \cdots\cdots, \\ y_n &= z_n, \end{cases}$$

把 $f(x_1, x_2, \cdots, x_n)$ 化成

$$z_1^2 + \cdots + z_p^2 - z_{p+1}^2 - \cdots - z_r^2. \tag{6.3}$$

称 (6.3) 为实二次型 $f(x_1, x_2, \cdots, x_n)$ 的**规范形**.
关于实二次型的规范形, 我们有

定理6.3.5

任意实二次型总可以经过适当的实非退化线性替换化成规范形, 且规范形是唯一的.

把 $f(x_1,x_2,\cdots,x_n)$ 化成

$$z_1^2+\cdots+z_p^2-z_{p+1}^2-\cdots-z_r^2. \tag{6.3}$$

称 (6.3) 为实二次型 $f(x_1,x_2,\cdots,x_n)$ 的规范形.
关于实二次型的规范形, 我们有

定理6.3.5

任意实二次型总可以经过适当的实非退化线性替换化成规范形, 且规范形是唯一的.

定理 6.3.5 通常称为惯性定理.

证明 只需要证明唯一性. 设实二次型 $f(x_1, x_2, \cdots, x_n)$ 经过非退化实线性替换 $X = BY$ 化成规范形

$$f(x_1, x_2, \cdots, x_n) = y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2. \tag{6.4}$$

又设它经过另一非退化实线性替换 $X = CZ$ 化成规范形

$$f(x_1, x_2, \cdots, x_n) = z_1^2 + \cdots + z_q^2 - z_{q+1}^2 - \cdots - z_r^2. \tag{6.5}$$

为证明唯一性, 只需证明 $p = q$.
用反证法. 假设 $p > q$. 由 (6.4) 和 (6.5) 得

$$y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2 = z_1^2 + \cdots + z_q^2 - z_{q+1}^2 - \cdots - z_r^2, \tag{6.6}$$

其中 $Z = C^{-1}BY$. 令 $C^{-1}B = (g_{ij})_{n\times n}$. 则

$$\begin{cases} z_1 & = g_{11}y_1 + g_{12}y_2 + \cdots + g_{1n}y_n, \\ z_2 & = g_{21}y_1 + g_{22}y_2 + \cdots + g_{2n}y_n, \\ & \cdots\cdots \\ z_n & = g_{n1}y_1 + g_{n2}y_2 + \cdots + g_{nn}y_n. \end{cases} \tag{6.7}$$

考虑齐次线性方程组

$$\begin{cases} g_{11}y_1+g_{12}y_2+\cdots+g_{1n}y_n &= 0, \\ \cdots\cdots & \\ g_{q1}y_1+g_{q2}y_2+\cdots+g_{qn}y_n &= 0, \\ y_{p+1} &= 0, \\ \cdots\cdots & \\ y_n &= 0. \end{cases} \tag{6.8}$$

由于方程组含有 n 个未知量, 而只有 $q+(n-p)=n-(p-q)<n$ 个方程, 因此它有非零解. 设

$$Y_0=(k_1,\cdots,k_p,k_{p+1},\cdots,k_n)^T$$

是 (6.8) 的一个非零解. 则 $k_{p+1}=\cdots=k_n=0$, $k_1,\cdots,k_p$ 不全为零. 把 Y_0 代入到 (6.6) 的左端得

$$k_1^2+\cdots+k_p^2>0. \tag{6.9}$$

把 Y_0 代入到 (6.7) 中解得

$$Z_0 = (t_1, \cdots, t_q, t_{q+1}, \cdots, t_n)^T.$$

由于 Y_0 是 (6.8) 的解, 故有 $t_1 = \cdots = t_q = 0$. 把 Z_0 代入到 (6.6) 的右端得

$$-z_{q+1}^2 - \cdots - z_r^2 \leq 0.$$

这与 (6.9) 矛盾. 因此 $p \leq q$.
同理可证 $q \leq p$. 从而 $p = q$.

定义6.3.6

在实二次型 $f(x_1, x_2, \cdots, x_n)$ 的规范形中，正平方项的个数 p 称为 $f(x_1, x_2, \cdots, x_n)$ 的正惯性指数；负平方项的个数 $r-p$ 称为 $f(x_1, x_2, \cdots, x_n)$ 的负惯性指数；它们的差 $p-(r-p)=2p-r$ 称为 $f(x_1, x_2, \cdots, x_n)$ 的符号差. 设 A 是一个实对称矩阵. 称二次型 X^TAX 的正惯性指数(负惯性指数, 符号差)为 A 的正惯性指数(负惯性指数, 符号差).

定义6.3.6

在实二次型 $f(x_1, x_2, \cdots, x_n)$ 的规范形中，正平方项的个数 p 称为 $f(x_1, x_2, \cdots, x_n)$ 的正惯性指数；负平方项的个数 $r-p$ 称为 $f(x_1, x_2, \cdots, x_n)$ 的负惯性指数；它们的差 $p-(r-p) = 2p-r$ 称为 $f(x_1, x_2, \cdots, x_n)$ 的符号差. 设 A 是一个实对称矩阵. 称二次型 $X^T AX$ 的正惯性指数(负惯性指数, 符号差)为 A 的正惯性指数(负惯性指数, 符号差).

用矩阵语言惯性定理可以表述为

定理6.3.7

任意 n 阶实对称矩阵 A 都合同于一个形如

$$D = \begin{pmatrix} E_p & 0 & 0 \\ 0 & -E_{r-p} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

的对角矩阵, 其中 $r = r(A)$, p 是 A 的正惯性指数.

作为定理 6.3.7 的推论, 我们有

推论6.3.8

两个 n 阶实对称矩阵 A 与 B 合同的充分必要条件是 A, B 有相同的秩和相同的正(负)惯性指数.

作为定理 6.3.7 的推论, 我们有

推论6.3.8

两个 n 阶实对称矩阵 A 与 B 合同的充分必要条件是 A, B 有相同的秩和相同的正(负)惯性指数.

根据惯性定理可以看出, 虽然实二次型的标准形不是唯一的, 但是标准形中系数为正的平方项的个数是唯一的. 因此, 惯性定理也可以叙述为: 实二次型的标准形中系数为正的平方项的个数是唯一确定的, 它等于二次型的正惯性指数, 系数为负的平方项的个数等于二次型的负惯性指数.

例 6.2.3 中的二次型的正惯性指数等于 2, 负惯性指数等于 1, 因此它的符号差等于 1.

☞ 设 A 是实对称矩阵. 据定理 6.2.3 知,

- A 的秩等于 A 的非零特征值的个数,
- A 的正惯性指数等于 A 的正特征值的个数,
- A 的负惯性指数等于 A 的负特征值的个数.

目录

正定二次型与正定矩阵在数学、物理以及工程技术等领域都有广泛的应用. 本节我们将重点讨论正定二次型与正定矩阵.

定义6.4.1

实二次型 $f(x_1, x_2, \cdots, x_n)$ 称为正定(负定, 半正定, 半负定)的, 如果对于任意非零实向量 $X_0 = (c_1, c_2, \cdots, c_n)^T \in R^n$, 都有

$$f(c_1, c_2, \cdots, c_n) > 0 \ (< 0, \ \geq 0, \ \leq 0).$$

否则, 称 $f(x_1, x_2, \cdots, x_n)$ 是不定的.

正定二次型与正定矩阵在数学、物理以及工程技术等领域都有广泛的应用. 本节我们将重点讨论正定二次型与正定矩阵.

定义6.4.1

实二次型 $f(x_1, x_2, \cdots, x_n)$ 称为正定(负定, 半正定, 半负定)的, 如果对于任意非零实向量 $X_0 = (c_1, c_2, \cdots, c_n)^T \in R^n$, 都有

$$f(c_1, c_2, \cdots, c_n) > 0 \ (< 0, \ \geq 0, \ \leq 0).$$

否则, 称 $f(x_1, x_2, \cdots, x_n)$ 是不定的.

由定义容易看出, 实二次型 $f(x_1, x_2, \cdots, x_n)$ 负定 (半负定) 当且仅当 $-f(x_1, x_2, \cdots, x_n)$ 正定 (半正定).

显然, 实二次型 $f(x_1, x_2, \cdots, x_n) = x_1^2 + x_2^2 + \cdots + x_n^2$ 是正定的. 一般地, 容易证明: 实二次型

$$f(x_1, x_2, \cdots, x_n) = d_1x_1^2 + d_2x_2^2 + \cdots + d_nx_n^2$$

是正定的当且仅当 $d_i > 0, \ i = 1, 2, \cdots, n$.

这样标准形式的实二次型的正定性就很容易判定.

我们知道实二次型都可以经过非退化实线性替换化成标准形, 那么一个自然的问题是: 非退化实线性替换是否改变实二次型的正定性?
对此, 我们有

我们知道实二次型都可以经过非退化实线性替换化成标准形, 那么一个自然的问题是: 非退化实线性替换是否改变实二次型的正定性?
对此, 我们有

定理6.4.2

非退化实线性替换保持实二次型的正定 (负定, 半正定, 半负定, 不定)性.

我们知道实二次型都可以经过非退化实线性替换化成标准形, 那么一个自然的问题是: 非退化实线性替换是否改变实二次型的正定性?
对此, 我们有

定理6.4.2

非退化实线性替换保持实二次型的正定 (负定, 半正定, 半负定, 不定)性.

据定理 6.4.2, 要判断一个实二次型是否正定, 可以先把它化成标准形, 根据标准形正定与否就可以判定原二次型是否正定. 由此可知, 例 6.2.1, 6.2.2, 6.2.4 中的二次型都不是正定的. 事实上, 它们都是不定的.

证明 我们只需证明正定的情形, 其它情形类似可证.
设正定实二次型 $f(x_1, x_2, \cdots, x_n) = X^T AX$, 其中 A 是二次型的矩阵, 经过非退化实线性替换 $X = CY$ 变成二次型 $g(y_1, y_2, \cdots, y_n) = Y^T BY$, 其中 B 是它的矩阵. 则 $B = C^T AC$, 且

$$g(y_1, y_2, \cdots, y_n) = Y^T BY = X^T AX = f(x_1, x_2, \cdots, x_n).$$

对于任意的非零 n 维实列向量 $Y_0 = (k_1, k_2, \cdots, k_n)^T$,
$X_0 = CY_0 = (c_1, c_2, \cdots, c_n)^T$ 也是 n 维非零实列向量. 由于 $f(x_1, x_2, \cdots, x_n)$ 是正定的, 因此

$$X_0^T AX_0 = f(c_1, c_2, \cdots, c_n) > 0.$$

于是

$$g(k_1, k_2, \cdots, k_n) = f(c_1, c_2, \cdots, c_n) > 0.$$

所以 $g(y_1, y_2, \cdots, y_n)$ 是正定二次型.
反之, 显然 $g(y_1, y_2, \cdots, y_n)$ 可经过非退化实线性替换 $Y = C^{-1}X$ 化成 $f(x_1, x_2, \cdots, x_n)$. 如果 $g(y_1, y_2, \cdots, y_n)$ 是正定的, 那么 $f(x_1, x_2, \cdots, x_n)$ 也是正定的. 这就证明了定理成立.

下面的定理给出了正定二次型的一些判定条件.

定理6.4.3

设 n 元实二次型 $f(x_1, x_2, \cdots, x_n) = X^TAX$, A 是它的矩阵（实对称）. 则下列条件等价:

1. $f(x_1, x_2, \cdots, x_n)$ 是正定的;
2. $f(x_1, x_2, \cdots, x_n)$ 的正惯性指数等于 n;
3. A 的特征值都大于零;
4. $f(x_1, x_2, \cdots, x_n)$ 的规范形为: $y_1^2 + y_2^2 + \cdots + y_n^2$;
5. 存在 n 阶实可逆矩阵 P 使得 $A = P^TP$.

下面的定理给出了正定二次型的一些判定条件.

定理6.4.3

设 n 元实二次型 $f(x_1, x_2, \cdots, x_n) = X^T AX$, A 是它的矩阵（实对称）. 则下列条件等价:

1. $f(x_1, x_2, \cdots, x_n)$ 是正定的;
2. $f(x_1, x_2, \cdots, x_n)$ 的正惯性指数等于 n;
3. A 的特征值都大于零;
4. $f(x_1, x_2, \cdots, x_n)$ 的规范形为: $y_1^2 + y_2^2 + \cdots + y_n^2$;
5. 存在 n 阶实可逆矩阵 P 使得 $A = P^T P$.

证明 (1) $\Rightarrow$ (2) 设正定二次型 $f(x_1, x_2, \cdots, x_n)$ 经过非退化实线性替换化成标准形

$$d_1 y_1^2 + d_2 y_2^2 + \cdots + d_n y_n^2. \tag{6.1}$$

据定理 6.4.2 知二次型 (6.1) 仍是正定的, 因此 $d_i > 0$, $i = 1, 2, \cdots, n$, 从而 $f(x_1, x_2, \cdots, x_n)$ 的正惯性指数等于 n.

(2) $\Rightarrow$ (3) 因为实二次型得正惯性指数就等于它的矩阵的正特征值的个数, 因此 (3) 成立.

(3) $\Rightarrow$ (4) 设 A 的全部特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 并且 $\lambda_i > 0$, $i = 1, 2, \cdots, n$. 据定理 6.2.3, 实二次型 $f(x_1, x_2, \cdots, x_n) = X^T AX$ 可以经过正交线性替换化成标准形

$$\lambda_1 z_1^2 + \lambda_2 z_2^2 + \cdots + \lambda_n z_n^2.$$

于是 $f(x_1, x_2, \cdots, x_n)$ 的规范形为

$$y_1^2 + y_2^2 + \cdots + y_n^2.$$

(4) $\Rightarrow$ (5) 设 $f(x_1, x_2, \cdots, x_n)$ 的规范形为: $y_1^2 + y_2^2 + \cdots + y_n^2$. 则存在非退化实线性替换 $X = CY$ 使得

$$f(x_1, x_2, \cdots, x_n) = X^T AX = Y^T (C^T AC) Y = y_1^2 + y_2^2 + \cdots + y_n^2.$$

因此 $C^T AC = E$. 令 $P = C^{-1}$. 则 $A = P^T P$.

(5) $\Rightarrow$ (1) 若 (5) 成立, 则有

$$f(x_1, x_2, \cdots, x_n) = X^T AX = (PX)^T (PX).$$

对任意 $0 \neq X \in R^n$, 由 P 是可逆实矩阵得 $PX \neq 0$ 且 $PX \in R^n$.
设 $PX = (c_1, c_2, \cdots, c_n)^T$. 则有

$$f(x_1, x_2, \cdots, x_n) = (PX)^T (PX) = c_1^2 + c_2^2 + \cdots + c_n^2 > 0.$$

所以 $f(x_1, x_2, \cdots, x_n)$ 是正定的.

定义6.4.4

实对称矩阵 A 称为正定 *(负定, 半正定, 半负定, 不定)* 的, 如果二次型 X^TAX 是正定 *(负定, 半正定, 半负定, 不定)* 的.

定义6.4.4

实对称矩阵 A 称为正定 (负定, 半正定, 半负定, 不定) 的, 如果二次型 X^TAX 是正定 (负定, 半正定, 半负定, 不定) 的.

据定理 6.4.3, 我们有

定理6.4.5

设 A 是 n 阶实对称矩阵. 则下列条件等价:

1. A 是正定的;
2. A 的特征值都大于零;
3. A 合同于单位矩阵;
4. 存在 n 阶可逆实矩阵 P 使得 $A = P^TP$.

定义6.4.4

实对称矩阵 A 称为正定 (负定, 半正定, 半负定, 不定) 的, 如果二次型 X^TAX 是正定 (负定, 半正定, 半负定, 不定) 的.

据定理 6.4.3, 我们有

定理6.4.5

设 A 是 n 阶实对称矩阵. 则下列条件等价:

1. A 是正定的;
2. A 的特征值都大于零;
3. A 合同于单位矩阵;
4. 存在 n 阶可逆实矩阵 P 使得 $A = P^TP$.

由定理 6.4.5 , 立即可得

推论6.4.6

正定矩阵的行列式大于零.

有时我们需要直接从二次型的矩阵来判断这个二次型是不是正定的, 而不是通过它的标准形或规范性. 为此我们引入

定义6.4.7

在 n 阶矩阵 $A=(a_{ij})$ 中, 由它的第 i_1 行, 第 i_2 行, $\cdots$, 第 i_k 行和第 i_1 列, 第 i_2 列, $\cdots$, 第 i_k 列组成的 k 阶子式称为 A 的一个 k 阶主子式. 特别地, 由 A 的前 k 行和前 k 列组成的 k 阶子式

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix}$$

称为 A 的 k 阶顺序主子式, 记为 $\Delta_k(A)$.

有时我们需要直接从二次型的矩阵来判断这个二次型是不是正定的, 而不是通过它的标准形或规范性. 为此我们引入

定义6.4.7

在 n 阶矩阵 $A=(a_{ij})$ 中, 由它的第 i_1 行, 第 i_2 行, $\cdots$, 第 i_k 行和第 i_1 列, 第 i_2 列, $\cdots$, 第 i_k 列组成的 k 阶子式称为 A 的一个 k 阶主子式. 特别地, 由 A 的前 k 行和前 k 列组成的 k 阶子式

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix}$$

称为 A 的 k 阶顺序主子式, 记为 $\Delta_k(A)$.

显然 n 阶矩阵 A 有且仅有 n 个顺序主子式 $\Delta_1(A), \Delta_2(A), \cdots, \Delta_n(A)$.

例6.4.1

矩阵

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

的 4 个顺序主子式分别为:

$$\Delta_1(A)=1$$

$$\Delta_2(A)=\begin{vmatrix}1 & -1\\ 0 & 2\end{vmatrix}=2,$$

$$\Delta_3(A)=\begin{vmatrix}1 & -1 & 2\\ 0 & 2 & 2\\ 1 & 0 & 3\end{vmatrix}=0,$$

$$\Delta_4(A)=|A|=-3.$$

定理6.4.8

设 A 为 n 阶实对称阵,则实二次型

$$f(x_1, x_2, \cdots, x_n) = X^T AX = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j$$

是正定的充分必要条件是矩阵 A 的顺序主子式全大于零.

定理6.4.8

设 A 为 n 阶实对称阵,则实二次型

$$f(x_1, x_2, \cdots, x_n) = X^T AX = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

是正定的充分必要条件是矩阵 A 的顺序主子式全大于零.

证明 必要性. 设二次型

$$f(x_1, x_2, \cdots, x_n) = X^T AX = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

是正定的. 对于 $k = 1, 2, \cdots, n$, 令

$$f_k(x_1, x_2, \cdots, x_k) = \sum_{i=1}^{k} \sum_{j=1}^{k} a_{ij} x_i x_j.$$

则 $f_k(x_1, x_2, \cdots, x_k)$ 是一个 k 元正定二次型. 事实上, 对任意一组不全为零的实数 $c_1, c_2, \cdots, c_k$, 由于 $f(x_1, x_2, \cdots, x_n)$ 是正定的,

因此

$$f_k(c_1, c_2, \cdots, c_k) = \sum_{i=1}^{k} \sum_{j=1}^{k} a_{ij} c_i c_j = f(c_1, \cdots, c_k, 0, \cdots, 0) > 0.$$

所以 $f_k(c_1, c_2, \cdots, c_k)$ 是正定的. 于是 $f_k(c_1, c_2, \cdots, c_k)$ 的矩阵

$$A_k = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$$

是正定矩阵. 据推论 6.4.6 得 $\Delta_k(A) = |A_k| > 0$. 这就证明了 A 的顺序主子式全大于零.

因此

$$f_k(c_1, c_2, \cdots, c_k) = \sum_{i=1}^{k}\sum_{j=1}^{k} a_{ij}c_ic_j = f(c_1, \cdots, c_k, 0, \cdots, 0) > 0.$$

所以 $f_k(c_1, c_2, \cdots, c_k)$ 是正定的. 于是 $f_k(c_1, c_2, \cdots, c_k)$ 的矩阵

$$A_k = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$$

是正定矩阵. 据推论 6.4.6 得 $\Delta_k(A) = |A_k| > 0$. 这就证明了 A 的顺序主子式全大于零.

充分性. 对二次型的元数进行归纳证明.

当 $n = 1$ 时, $f(x_1) = a_{11}x_1^2$. 由于 $a_{11} = \Delta_1(A) > 0$, 因此 $f(x_1)$ 是正定的.

假设结论对于 $n-1$ 元实二次型成立, 考察 n 元的情形.

令

$$A_{n-1}=\begin{pmatrix} a_{11} & \cdots & a_{1,n-1} \\ \cdots & \cdots & \cdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} \end{pmatrix}, \quad \alpha=\begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{n-1,n} \end{pmatrix}.$$

则

$$A=\begin{pmatrix} A_{n-1} & \alpha \\ \alpha' & a_{nn} \end{pmatrix}.$$

由于 A 的顺序主子式全大于零, 从而 A_{n-1} 的顺序主子式全大于零. 据归纳法假设, A_{n-1} 是正定矩阵. 因此存在 $n-1$ 阶可逆实矩阵 Q 使得 $Q^TA_{n-1}Q=E_{n-1}$. 令

$$C_1=\begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix}.$$

则

$$C_1^T A C_1 = \begin{pmatrix} Q^T & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A_{n-1} & \alpha \\ \alpha^T & a_{nn} \end{pmatrix} \begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} E_{n-1} & Q^T\alpha \\ \alpha^T Q & a_{nn} \end{pmatrix}.$$

再令

$$C_2 = \begin{pmatrix} E_{n-1} & -Q^T\alpha \\ 0 & 1 \end{pmatrix}.$$

则有

$$\begin{aligned} C_2^T C_1^T A C_1 C_2 &= \begin{pmatrix} E_{n-1} & 0 \\ -\alpha^T Q & 1 \end{pmatrix} \begin{pmatrix} E_{n-1} & Q^T\alpha \\ \alpha^T Q & a_{nn} \end{pmatrix} \begin{pmatrix} E_{n-1} & -Q^T\alpha \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} E_{n-1} & 0 \\ 0 & a_{nn} - \alpha^T Q Q^T \alpha \end{pmatrix}. \end{aligned}$$

记 $a = a_{nn} - \alpha^T QQ^T \alpha$. 由已知条件 $|A| > 0$, 因此

$$a = |C_2|^2 |C_1|^2 |A| > 0.$$

令

$$C_3 = \begin{pmatrix} E_{n-1} & 0 \\ 0 & 1/\sqrt{a} \end{pmatrix}$$

且 $C = C_3 C_2 C_1$. 则 C 是可逆实矩阵, 并且

$$C^T AC = \begin{pmatrix} E_{n-1} & 0 \\ 0 & 1/\sqrt{a} \end{pmatrix} \begin{pmatrix} E_{n-1} & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} E_{n-1} & 0 \\ 0 & 1/\sqrt{a} \end{pmatrix} = E,$$

即 A 合同于单位矩阵. 据定理 6.4.5 知 A 是正定矩阵, 也就是说, $f(x_1, x_2, \cdots, x_n)$ 是正定二次型.
根据归纳法原理, 充分性得证.

记 $a = a_{nn} - \alpha^T QQ^T \alpha$. 由已知条件 $|A| > 0$, 因此

$$a = |C_2|^2 |C_1|^2 |A| > 0.$$

令

$$C_3 = \begin{pmatrix} E_{n-1} & 0 \\ 0 & 1/\sqrt{a} \end{pmatrix}$$

且 $C = C_3 C_2 C_1$. 则 C 是可逆实矩阵, 并且

$$C^T AC = \begin{pmatrix} E_{n-1} & 0 \\ 0 & 1/\sqrt{a} \end{pmatrix} \begin{pmatrix} E_{n-1} & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} E_{n-1} & 0 \\ 0 & 1/\sqrt{a} \end{pmatrix} = E,$$

即 A 合同于单位矩阵. 据定理 6.4.5 知 A 是正定矩阵, 也就是说, $f(x_1, x_2, \cdots, x_n)$ 是正定二次型.
根据归纳法原理, 充分性得证.
用矩阵语言, 定理 6.4.8 可以表述为: 实对称矩阵正定的充分必要条件是它的顺序主子式全大于零.

例6.4.2

判断下列实二次型

(1) $f(x_1, x_2, x_3) = x_1^2 + 4x_1x_2 - 4x_1x_3 + 2x_2^2 - 4x_2x_3 + 3x_3^2$.

(2) $f(x_1, x_2, x_3) = 5x_1^2 + x_2^2 + 5x_3^2 + 4x_1x_2 - 8x_1x_3 - 4x_2x_3$.

是否正定.

例6.4.2

判断下列实二次型

(1) $f(x_1,x_2,x_3)=x_1^2+4x_1x_2-4x_1x_3+2x_2^2-4x_2x_3+3x_3^2$.

(2) $f(x_1,x_2,x_3)=5x_1^2+x_2^2+5x_3^2+4x_1x_2-8x_1x_3-4x_2x_3$.

是否正定.

解 (1) **方法 1** 因为 $f(1,0,1)=0$, 所以 $f(x_1,x_2,x_3)$ 不是正定的.

方法 2 据例 6.2.1 知 $f(x_1,x_2,x_3)$ 可经过适当的非退化实线性替换化成标准形

$$f(x_1,x_2,x_3)=y_1^2-2y_2^2+y_3^2.$$

所以 $f(x_1,x_2,x_3)$ 不是正定的.

或者由规范形

$$z_1^2+z_2^2-z_3^2.$$

知 $f(x_1,x_2,x_3)$ 的正惯性指数等于 2, 所以它不是正定的.

方法 3 $f(x_1,x_2,x_3)$ 的矩阵为

$$A=\begin{pmatrix}1&2&-2\\2&2&-2\\-2&-2&3\end{pmatrix}.$$

A 的 2 阶顺序主子式 $\Delta_2(A)=\begin{vmatrix}1&2\\2&2\end{vmatrix}=-2<0$, 所以 A不是正定矩阵, 从而二次型 $f(x_1,x_2,x_3)$ 不是正定的. 或者, 因为

$$|A|=\begin{vmatrix}1&2&-2\\2&2&-2\\-2&-2&3\end{vmatrix}=-2<0,$$

所以 A 不是正定矩阵, 从而二次型 $f(x_1,x_2,x_3)$ 不是正定的.

(2) **方法 1** $f(x_1, x_2, x_3)$ 的矩阵为

$$A = \begin{pmatrix} 5 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{pmatrix},$$

它的顺序主子式为 $\Delta_1(A) = 5 > 0$,

$$\Delta_2(A) = \begin{vmatrix} 5 & 2 \\ 2 & 1 \end{vmatrix} = 1 > 0, \quad \Delta_3(A) = \begin{vmatrix} 5 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{vmatrix} = 1 > 0,$$

因此 A 是正定矩阵, 从而二次型 $f(x_1, x_2, x_3)$ 正定.

方法 2 根据方法 1, A 的特征多项式为

$$|\lambda E - A| = \begin{vmatrix} \lambda-5 & -2 & 4 \\ -2 & \lambda-1 & 2 \\ 4 & 2 & \lambda-5 \end{vmatrix} = (\lambda-1)(\lambda^2-10\lambda+1),$$

它的特征值为

$$\begin{aligned} \lambda_1 &= 1 > 0, \\ \lambda_2 &= 5-2\sqrt{6} > 0, \\ \lambda_3 &= 5+2\sqrt{6} > 0. \end{aligned}$$

所以 A 是正定矩阵, 从而二次型 $f(x_1, x_2, x_3)$ 正定.

例6.4.3

t 取何值时, 实对称矩阵

$$A = \begin{pmatrix} 1 & t & -1 \\ t & 4 & 2 \\ -1 & 2 & 2 \end{pmatrix}$$

是正定的？

例6.4.3

t 取何值时, 实对称矩阵

$$A=\begin{pmatrix}1 & t & -1\\ t & 4 & 2\\ -1 & 2 & 2\end{pmatrix}$$

是正定的?

解 A 的顺序主子式为

$$\Delta_1(A)=1>0,\quad \Delta_2(A)=\begin{vmatrix}1 & t\\ t & 4\end{vmatrix}=4-t^2,\quad \Delta_3(A)=|A|=-2t^2-4t.$$

因此 A 正定的充分必要条件为

$$\begin{cases}4-t^2 & >0,\\ -2t^2-4t & >0.\end{cases}$$

解不等式组得 $-2<t<0$. 所以当 $-2<t<0$ 时, A 正定.

利用定理 4.6.5 和 4.6.8, 可以得到正定矩阵的下列性质.

定理6.4.9

设 A, B 都是 n 阶正定矩阵. 则

1. 对于任意正实数 s, sA 是正定矩阵.
2. A^{-1}, A^* 都是正定矩阵.
3. 对于任意可逆实矩阵 C, C^TAC 是正定矩阵.
4. $A+B$ 是正定矩阵.
5. $D = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ 是正定矩阵.

利用定理 4.6.5 和 4.6.8, 可以得到正定矩阵的下列性质.

定理6.4.9

设 A, B 都是 n 阶正定矩阵. 则

1. 对于任意正实数 s, sA 是正定矩阵.
2. A^{-1}, A^* 都是正定矩阵.
3. 对于任意可逆实矩阵 C, C^TAC 是正定矩阵.
4. $A+B$ 是正定矩阵.
5. $D = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ 是正定矩阵.

证明 (1) 因为 A 是正定矩阵, 所以

$$\Delta_k(A) > 0, \quad k = 1, 2, \cdots, n,$$

从而

$$\Delta_k(sA) = s^k \Delta_k(A) > 0, \quad k = 1, 2, \cdots, n,$$

故 rA 是正定的.

(2) 设 A 的全部特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$. 因为 A 正定, 所以 $\lambda_i > 0$, $i = 1, 2, \cdots, n$. 于是 A^{-1} 的全部特征值为 $\lambda_1^{-1}, \lambda_2^{-1}, \cdots, \lambda_n^{-1}$, 并且 $\lambda_i^{-1} > 0$, $i = 1, 2, \cdots, n$, 因而 A^{-1} 是正定矩阵.
因为 A 是正定的, 所以 $|A| > 0$. 由 $A^* = |A|A^{-1}$ 及 (1) 得 A^* 是正定的.
(3) 这就是定理 6.4.2 关于正定的矩阵表示.
(4) 显然 $A + B$ 是实对称矩阵. 因为 A, B 都是 n 阶正定矩阵, 所以二次型 X^TAX 和 X^TBX 都是正定二次型. 于是对于任意非零实向量 $\alpha \in R^n$,

$$\alpha^T(A + B)\alpha = \alpha^TA\alpha + \alpha^TB\alpha > 0,$$

因此二次型 $X^T(A + B)X$ 是正定的, 从而 $A + B$ 是正定矩阵.
(5) 因为

$$|\lambda E_{2n} - D| = \begin{vmatrix} \lambda E_n - A & 0 \\ 0 & \lambda E_n - B \end{vmatrix} = |\lambda E_n - A||\lambda E_n - B|,$$

所以 A 的全部特征值与 B 的全部特征值合起来就是 D 的全部特征值. 由 A, B 的特征值全大于零得 D 的特征值全大于零, 因此 D 是正定矩阵.

基于正定与负定的关系, 不难得出负定二次型与负定矩阵的判别条件及性质. 下面的定理给出了实二次型(实对称矩阵)是半正定的判定条件, 略去其证明, 见习题6.4之第6题. 关于半负定的结果可以类似的给出.

定理6.4.10

设实二次型 $f(x_1, x_2, \cdots, x_n) = X^T AX$, 其中 A 是它的矩阵, $r = r(A)$. 则下列条件等价:

1. $f(x_1, x_2, \cdots, x_n)$ 是半正定的;
2. $f(x_1, x_2, \cdots, x_n)$ 的正惯性指数等于 r;
3. $f(x_1, x_2, \cdots, x_n)$ 规范形为 $y_1^2 + y_2^2 + \cdots + y_r^2$;
4. A 的特征值都大于或等于零;
5. 存在 n 阶实矩阵 C 使得 $A = C^T C$;
6. A 与 $\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$ 合同;
7. A 的所有主子式都大于或等于零.

值得注意的是, 与正定性相比较, 这里 (7) 中, 仅有顺序主子式都大于或等于零不能保证半正定性. 例如, 实二次型 $f(x_1, x_2) = -x_2^2$ 显然是半负定, 而它的矩阵

$$A = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$$

的顺序主子式都等于零.