第六章 二 次 型

在平面解析几何中,为了研究二次曲线

$$ax^2+bxy+cy^2=1 \qquad (*)$$

的几何性质,通常可以采用坐标旋转变换

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

适当地选择θ,可把方程(*)化成只含平方项的标准形式

$$mx'^2 + ny'^2 = 1$$

从而比较容易地判定曲线的类型.

在许多实际问题和数学问题(如多元函数极值)中也会遇到多元二次齐次函数. 如何寻找一个线性变换,将多元二次齐次函数化为只含平方项的标准方程? 本章我们将在更普遍的意义下解决这一问题.

§1 二次型与合同变换

一. 二次型的基本概念

定义6.1 n个变量$x_1, x_2, \ldots, x_n$的二次齐次函数

$$f(x_1, x_2, \ldots, x_n)=a_{11}x_1^2+a_{22}x_2^2+\ldots+a_{nn}x_n^2+ \\ 2a_{12}x_1x_2+2a_{13}x_1x_3+\ldots+2a_{n-1,n}x_{n-1}x_n. \quad (6.1)$$

称为一个n元二次型，简称二次型. 当系数a_{ij}均为实数时称为n元实二次型. 以下仅讨论实二次型.

以下关于(6.1)式的讨论中，对于任意的$i, j, (i\neq j)$，均取

$$a_{ij}= a_{ji},$$

则有

$$2a_{ij}x_ix_j = a_{ij}x_ix_j+a_{ji}x_jx_i \ .$$

于是(6.1)式可写成:

$f(x_1, x_2, \dots, x_n)=$

$a_{11}x_1^2+a_{12}x_1x_2+\dots+a_{1n}x_1x_n+$

$a_{21}x_2x_1+a_{22}x_2^2+\dots+a_{2n}x_2x_n+$

$\dots\dots+$

$a_{n1}x_nx_1+a_{n2}x_nx_2+\dots+a_{nn}x_n^2$

$$=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j$$

$$=\sum_{i,j=1}^{n}a_{ij}x_ix_j \qquad (6.2)$$

$f(x_1, x_2, \dots, x_n)=$

$x_1(a_{11}x_1+a_{12}x_2+\dots+a_{1n}x_n)+$

$x_2(a_{21}x_1+a_{22}x_2+\dots+a_{2n}x_n)+$

$\dots\dots+$

$x_n(a_{n1}x_1+a_{n2}x_2+\dots+a_{nn}x_n).$

利用矩阵乘法的运算符号，上式又可以表示为：

……

将(6.2)式做整理，还可写成：

$$
\begin{aligned}
f(x_1, x_2, \ldots, x_n) = \ & x_1(a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n) + \\
& x_2(a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n) + \\
& \ldots + \\
& x_n(a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n).
\end{aligned}
$$

$$
= (x_1, x_2, \ldots, x_n)\begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \ldots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{pmatrix}
$$

$$
= (x_1, x_2, \ldots, x_n)\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix}.
$$

也即:

$$f(x_1, x_2, \ldots, x_n) = (x_1, x_2, \ldots, x_n)\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}.$$

若记

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$a_{ij} = a_{ji}$

则二次型(6.1)可以写成更简洁的形式:

$$f(x_1, x_2, \ldots, x_n) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}. \tag{6.3}$$

定义6.2 称式(6.3)为**n元二次型的矩阵表示**，实对称矩阵A称为**二次型f的矩阵**，f称为**实对称矩阵A的二次型**，且称矩阵A的秩R(A)为**二次型f的秩**.

例如，二次型$f=2x_1^2+3x_2^2-x_3^2+4x_1x_2-6x_1x_3$的矩阵为

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

于是

记$\mathbf{x}=(x_1, x_2, x_3)^T$

$f=\mathbf{x}^T\mathbf{A}\mathbf{x}$

$$=(x_1,\quad x_2,\quad x_3)\begin{pmatrix} 2 & 2 & -3 \\ 2 & 3 & 0 \\ -3 & 0 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

另例，实对称矩阵

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

的二次型为

$$f=x_1^2+2x_2^2-x_3^2+4x_1x_2+2x_1x_3-2x_2x_3 .$$

定义6.3 仅含平方项的二次型

$$f(x_1, x_2,\dots, x_n)= d_1x_1^2+d_2x_2^2+\dots+d_nx_n^2 \qquad (6.4)$$

称为**标准形**.

系数$d_1, d_2,\dots, d_n$只在1, -1, 0三个数中取值的形式为

$$f(y_1, y_2,\dots, y_n)= y_1^2+\dots+y_p^2 - y_{p+1}^2 - \dots - y_r^2$$

的二次型称为**规范形**.

注1: 正项和负项的顺序可能是打乱的.

注2: 可能会有0系数项, 所以r未必等于n.

易见, 标准形的矩阵为对角矩阵. 标准形(6.4)的矩阵表示为:

$$f= \mathbf{x}^{\mathrm{T}}\mathbf{\Lambda}\mathbf{x}= (x_1,x_2,\dots,x_n)\begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \dots & \\ & & & d_n \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}.$$

设$x_1, x_2, \ldots, x_n$和$y_1, y_2, \ldots, y_n$是两组变量，则关系式：

$$\begin{cases} x_1 = c_{11}y_1 + c_{12}y_2 + \ldots + c_{1n}y_n \\ x_2 = c_{21}y_1 + c_{22}y_2 + \ldots + c_{2n}y_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n = c_{n1}y_1 + c_{n2}y_2 + \ldots + c_{nn}y_n \end{cases}$$

称为从变量$x_1, x_2, \ldots, x_n$到$y_1, y_2, \ldots, y_n$的线性变换，其作用是将二次型中$x_1, x_2, \ldots, x_n$的替换为$y_1, y_2, \ldots, y_n$.

上式中，如果记

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ y_n \end{pmatrix}, \mathbf{C} = \begin{pmatrix} c_{11} & c_{12} & \ldots & c_{1n} \\ c_{21} & c_{22} & \ldots & c_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ c_{n1} & c_{n2} & \ldots & c_{nn} \end{pmatrix},$$

则线性变换可写成矩阵形式：$\mathbf{x}=\mathbf{C}\mathbf{y}$,

则线性变换可写成矩阵形式:

$$x=Cy,$$

这里C称为**线性变换x=Cy的矩阵**.

当C可逆时, x=Cy称为**可逆线性变换**,

这时, $y=C^{-1}x$也是线性变换, 称$y=C^{-1}x$为x=Cy的**逆变换**;

当C为正交矩阵时, 称x=Cy为**正交变换**.

(正交变换当然也是一种可逆变换)

如果对n元二次型$f=\mathbf{x}^T\mathbf{A}\mathbf{x}$作变换$\mathbf{x}=\mathbf{C}\mathbf{y}$，则有

$$\begin{aligned} f&=\mathbf{x}^T\mathbf{A}\mathbf{x} \\ &=(\mathbf{C}\mathbf{y})^T\mathbf{A}(\mathbf{C}\mathbf{y}) \\ &=\mathbf{y}^T(\mathbf{C}^T\mathbf{A}\mathbf{C})\mathbf{y} \\ &=\mathbf{y}^T\mathbf{B}\mathbf{y}, \end{aligned}$$

记$\mathbf{B}=\mathbf{C}^T\mathbf{A}\mathbf{C}$

即f(也)成为$y_1, y_2, \ldots, y_n$的n元二次型，且该二次型的矩阵为

$$\mathbf{B}=\mathbf{C}^T\mathbf{A}\mathbf{C}.$$

当矩阵$\mathbf{C}$可逆时，矩阵$\mathbf{B}=\mathbf{C}^T\mathbf{A}\mathbf{C}$与矩阵$\mathbf{A}$等价. 于是有:

定理6.1 对二次型做线性变换，(其变换结果)仍为二次型；如果这个线性变换可逆，则(变换前后)二次型的秩不变.

在可逆线性变换下，二次型的矩阵由A变为$B=C^TAC$，为研究变换前后二次型矩阵的关系，我们给出如下概念.

定义6.4 设A, B是同阶方阵. 如果存在可逆矩阵C，使得

$$B=C^TAC,$$

则称A与B是合同的，或称矩阵B是A的合同矩阵.

(当矩阵C可逆时,)对方阵A做运算C^TAC，称为对A进行合同变换.

可逆矩阵C称为把A变为B的合同变换矩阵.

矩阵的合同关系具有性质:

(i) 反身性: A与A合同;

(ii) 对称性: 若A与B合同，则B与A合同;

(iii) 传递性: 若A与B合同, B与C合同，则A与C合同.

如果矩阵A与B合同，由于合同变换矩阵C可逆，所以矩阵A与B等价，于是A与B等秩(即R(A)=R(B)). 将讨论结果总结如下：

命题：如果矩阵A与B合同，则A与B等价，于是R(A)=R(B).

注：以上命题的逆命题未必成立,即

矩阵A与B等价时, A与B不一定合同.

§2 用正交变换化二次型为标准形

由定理6.1可知，要做可逆线性变换$\mathbf{x}=\mathbf{C}\mathbf{y}$，化二次型$f=\mathbf{x}^T\mathbf{A}\mathbf{x}$为标准形，只要矩阵$\mathbf{C}$满足：$\mathbf{C}^T\mathbf{A}\mathbf{C}=\mathbf{\Lambda}$为对角矩阵. 于是，利用定理5.6，即

定理5.6 设$\mathbf{A}$是实对称矩阵，则必存在正交矩阵$\mathbf{Q}$，使得$\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$ $(=\mathbf{Q}^T\mathbf{A}\mathbf{Q})$为对角矩阵.

可得如下结论.

定理6.2 任意二次型$f=\mathbf{x}^T\mathbf{A}\mathbf{x}$都可经正交变换$\mathbf{x}=\mathbf{Q}\mathbf{y}$化为标准形$f=\mathbf{y}^T\mathbf{\Lambda}\mathbf{y}$，其中$\mathbf{\Lambda}$的主对角线元素(也即标准形的n个系数)恰为矩阵$\mathbf{A}$的n个特征值.

可见，用正交变换化二次型为标准形与实对称矩阵正交相似对角化的步骤几乎是一致的.

例1 用正交变换化二次型

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

为标准形, 并给出所用的正交变换.

解 二次型的矩阵为

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

A的特征多项式为

$$\begin{vmatrix} \lambda-3 & -1 & -2 \\ -1 & \lambda-3 & 2 \\ -2 & 2 & \lambda \end{vmatrix}=\begin{vmatrix} \lambda-4 & -1 & -2 \\ \lambda-4 & \lambda-3 & 2 \\ 0 & 2 & \lambda \end{vmatrix}=\begin{vmatrix} \lambda-4 & -1 & -2 \\ 0 & \lambda-2 & 4 \\ 0 & 2 & \lambda \end{vmatrix}$$

$=(\lambda-4)(\lambda^2-2\lambda-8)=(\lambda-4)^2(\lambda+2),$

所以，矩阵$\mathbf{A}$的特征值为$\lambda_1=\lambda_2=4$, $\lambda_3=-2$. 由于

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

于是，方程组$(4E-A)\mathbf{x}=\mathbf{0}$的一个基础解系可取为：

$$\boldsymbol{\alpha}_1=(1,\ 1,\ 0)^T,\quad \boldsymbol{\alpha}_2=(2,\ 0,\ 1)^T,$$

由施密特正交化可得

$$\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1=(1,\ 1,\ 0)^T,$$

$$\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2-\frac{[\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1]}{[\boldsymbol{\beta}_1,\boldsymbol{\beta}_1]}\boldsymbol{\beta}_1=(2,\ 0,\ 1)^T-\frac{2}{2}(1,\ 1,\ 0)^T$$

$$=(1,\ -1,\ 1)^T,$$

规范化可得属于特征值$\lambda_1=\lambda_2=4$的规范正交向量组：

$$\boldsymbol{\xi}_1=\left(\frac{1}{\sqrt{2}},\ \frac{1}{\sqrt{2}},\ 0\right)^T,\ \boldsymbol{\xi}_2=\left(\frac{1}{\sqrt{3}},\ \frac{-1}{\sqrt{3}},\ \frac{1}{\sqrt{3}}\right)^T,$$

又由于

$$-2\mathbf{E}-\mathbf{A}=\begin{pmatrix}-5 & -1 & -2\\ -1 & -5 & 2\\ -2 & 2 & -2\end{pmatrix}\to\begin{pmatrix}1 & 1 & 0\\ 0 & -2 & 1\\ 0 & 0 & 0\end{pmatrix}$$

于是，方程组(-2E-A)x=0的一个基础解系可取为:

$$\boldsymbol{\alpha}_3=(-1,\quad 1,\quad 2)^{\mathrm{T}},$$

所以得属于λ_3=-2的单位特征向量

$$\boldsymbol{\xi}_3=\left(\frac{-1}{\sqrt{6}},\quad \frac{1}{\sqrt{6}},\quad \frac{2}{\sqrt{6}}\right)^{\mathrm{T}}.$$

故可取正交矩阵

$$\mathbf{Q}=\begin{pmatrix}1/\sqrt{2} & 1/\sqrt{3} & -1/\sqrt{6}\\ 1/\sqrt{2} & -1/\sqrt{3} & 1/\sqrt{6}\\ 0 & 1/\sqrt{3} & 2/\sqrt{6}\end{pmatrix},$$

作正交变换: **x=Qy**, 即

$$\begin{cases} x_1 = \frac{1}{\sqrt{2}} y_1 + \frac{1}{\sqrt{3}} y_2 - \frac{1}{\sqrt{6}} y_3 \\ x_2 = \frac{1}{\sqrt{2}} y_1 - \frac{1}{\sqrt{3}} y_2 + \frac{1}{\sqrt{6}} y_3 \\ x_3 = \frac{1}{\sqrt{3}} y_2 + \frac{2}{\sqrt{6}} y_3 \end{cases}$$

二次型 $f(x_1, x_2, x_3)=3x_1^2+3x_2^2+2x_1x_2+4x_1x_3-4x_2x_3$ 变为

$$f=4y_1^2+4y_2^2-2y_3^2.$$

以上算例中,

若取正交矩阵$Q=(\xi_1, \xi_3, \xi_2)$, 做正交变换$x=Qy$, 则有

$$f=4y_1^2-2y_2^2+4y_3^2.$$

若取正交矩阵$Q=(\xi_3, \xi_1, \xi_2)$, 做正交变换$x=Qy$, 则有

$$f=-2y_1^2+4y_2^2+4y_3^2.$$

可见, 化二次型为标准形所用的正交变换以及标准形都不是唯一的. 但是, 正交变换对应的标准形中, 各项系数恰是矩阵A的所有特征值, 因此除顺序外是唯一的.

由于正交变换具有保范性, 所以正交变换具有保持几何图形不变的优点.

建议自修: 教材Pp119-122: 例6.2, 例6.3

推论1 设A, B都是n阶实对称矩阵, 如果A与B相似, 则A与B合同.

证明思路: 由于A与B相似, 因此

A与B具有全部相同的特征值: $\lambda_1, \lambda_2, \ldots, \lambda_n$.

又由于A与B都是实对称矩阵,

因此A与B都必合同于对角阵

$$\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \cdots & \\ & & & \lambda_n \end{pmatrix}.$$

根据合同关系的传递性, A与B也必然是合同的.　　证毕.

注: 推论1的逆命题一般不成立. 即: 两个合同的对称矩阵未必相似. 反例见下页.

令

$$A=\begin{pmatrix}1&&&\\&1&&\\&&\cdots&\\&&&1\end{pmatrix},\quad P=\begin{pmatrix}2&&&\\&2&&\\&&\cdots&\\&&&2\end{pmatrix},\quad B=\begin{pmatrix}4&&&\\&4&&\\&&\cdots&\\&&&4\end{pmatrix},$$

则可见：$P^T AP=B$.

因此，A与B合同．但是，显然A与B并不相似.

结论1：一般地，两个合同的对称矩阵未必相似.

提示：一般地，合同变换矩阵P只需是可逆阵.

然而，……

结论2：若A与B合同，且合同变换矩阵P为正交矩阵，则A与B必相似.

证明思路：由于 $P^T AP=B$，而P是正交阵，故 $P^{-1} AP=B$.

推论2 任意实二次型$f=\mathbf{x}^T\mathbf{A}\mathbf{x}$都可经可逆变换$\mathbf{x}=\mathbf{C}\mathbf{z}$化为规范形.

证明 由于$f=\mathbf{x}^T\mathbf{A}\mathbf{x}$可经正交变换$\mathbf{x}=\mathbf{Q}\mathbf{y}$化为标准形$f=\mathbf{y}^T\mathbf{\Lambda}\mathbf{y}$.

若$R(\mathbf{A})=r$, 则矩阵$\mathbf{A}$的n个特征值恰r有个不等于零, 不妨设

$$\lambda_i \neq 0\ (i=1, 2, \ldots, r),\ \lambda_{r+1}=\lambda_{r+2}=\ldots=\lambda_n=0,$$

则标准形的矩阵$\mathbf{\Lambda}$为:

$$\mathbf{\Lambda}=\begin{pmatrix} \lambda_1 & & & & & \\ & \cdots & & & & \\ & & \lambda_r & & & \\ & & & 0 & & \\ & & & & \cdots & \\ & & & & & 0 \end{pmatrix}.$$

记矩阵

$$\mathbf{B}=\begin{pmatrix} b_1 & & & & & \\ & \cdots & & & & \\ & & b_r & & & \\ & & & 1 & & \\ & & & & \cdots & \\ & & & & & 1 \end{pmatrix},$$

其中 $b_i=\dfrac{1}{\sqrt{|\lambda_i|}},(i=1,2,\ldots,r)$，则 $\mathbf{B}$ 是可逆矩阵，且

$$\mathbf{B}^T\boldsymbol{\Lambda}\mathbf{B}=\begin{pmatrix} \frac{\lambda_1}{|\lambda_1|} & & & & & \\ & \cdots & & & & \\ & & \frac{\lambda_r}{|\lambda_r|} & & & \\ & & & 0 & & \\ & & & & \cdots & \\ & & & & & 0 \end{pmatrix}.$$

这些数只能是1或-1

$$\mathbf{B}^{\mathrm{T}}\mathbf{\Lambda B} = \begin{pmatrix} \frac{\lambda_1}{|\lambda_1|} & & & & & \\ & \cdots & & & & \\ & & \frac{\lambda_r}{|\lambda_r|} & & & \\ & & & 0 & & \\ & & & & \cdots & \\ & & & & & 0 \end{pmatrix}.$$

所以，只要做可逆变换**y=Bz**，就有f=**y**$^{\mathrm{T}}$**Λy**=**z**$^{\mathrm{T}}$**B**$^{\mathrm{T}}$**ΛBz**.

记**C=QB**，**C**必为可逆矩阵. 做线性变换**x=Cz**，则化二次型

$$f = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \mathbf{z}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}\mathbf{A}\mathbf{C}\mathbf{z} = \mathbf{z}^{\mathrm{T}}(\mathbf{Q}\mathbf{B})^{\mathrm{T}}\mathbf{A}(\mathbf{Q}\mathbf{B})\mathbf{z} = \mathbf{z}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{Q}^{\mathrm{T}}\mathbf{A}\mathbf{Q}\mathbf{B}\mathbf{z}$$

$$= \mathbf{z}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}(\mathbf{Q}^{\mathrm{T}}\mathbf{A}\mathbf{Q})\mathbf{B}\mathbf{z} = \mathbf{z}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{\Lambda}\mathbf{B}\mathbf{z} = \mathbf{z}^{\mathrm{T}}(\mathbf{B}^{\mathrm{T}}\mathbf{\Lambda}\mathbf{B})\mathbf{z}$$

$$= \frac{\lambda_1}{|\lambda_1|}z_1^2 + \frac{\lambda_2}{|\lambda_2|}z_2^2 + \cdots + \frac{\lambda_r}{|\lambda_r|}z_r^2.$$

这就是规范型.

证毕.

注：将实二次型$f=\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$化为规范形的线性变换$\mathbf{x}=\mathbf{C}\mathbf{z}$一般不再是正交变换.

§3 用配方法化二次型为标准形

如果不限于用正交变换，还可以用其他的方法寻找一个可逆线性变换$x=Cy$化二次型为标准形. 以下利用具体例子介绍常用的配方法.

例2 求一可逆线性变换化二次型

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

为标准形，并给出所用的线性变换.

解 由于这里含有x_1^2项，首先把所有含有x_1的项集中起来配成完全平方，使配方后其余项中不再含x_1. 即

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

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

$$=3(x_1^2+2/3\ x_1x_2+4/3\ x_1x_3+1/9\ x_2^2+4/9\ x_3^2+4/9\ x_2x_3)$$

$$\quad +3x_2^2-4x_2x_3-1/3\ x_2^2-4/3\ x_3^2-4/3\ x_2x_3$$

$$=3(x_1+1/3\ x_2+2/3\ x_3)^2+8/3\ x_2^2-4/3\ x_3^2-16/3\ x_2x_3$$

$$=3(x_1+1/3\ x_2+2/3\ x_3)^2+8/3\ (x_2^2-2x_2x_3+x_3^2)-4x_3^2$$

$$=3(x_1+1/3\ x_2+2/3\ x_3)^2+8/3\ (x_2-x_3)^2-4x_3^2$$

可见，只要作变换：

$$\begin{cases} y_1 = x_1 + \frac{1}{3}x_2 + \frac{2}{3}x_3 \\ y_2 = x_2 - x_3 \\ y_3 = x_3 \end{cases} \quad \text{即:} \begin{cases} x_1 = y_1 - \frac{1}{3}y_2 - y_3 \\ x_2 = y_2 + y_3 \\ x_3 = y_3 \end{cases}$$

二次型 $f(x_1, x_2, x_3)=3x_1^2+3x_2^2+2x_1x_2+4x_1x_3-4x_2x_3$ 变为

$$f=3y_1^2+\frac{8}{3}y_2^2-4y_3^2.$$

所用的线性变换 $\mathbf{x}=\mathbf{C}\mathbf{y}$ 的矩阵为

$$\mathbf{C} = \begin{pmatrix} 1 & -\frac{1}{3} & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

由于 $\mathbf{C}$ 是可逆矩阵，所以线性变换 $\mathbf{x}=\mathbf{C}\mathbf{y}$ 是可逆线性变换.

注：此时，3, 8/3, -4一般未必是二次型 f 的矩阵 **A** 的特征值.

例3 求一可逆线性变换化二次型

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

为标准形，并给出所用的线性变换.

解 由于这里不含x_1^2项，但含x_2^2项，则先对x_2配方：

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

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

$$=(x_2+x_1-2x_3)^2-(x_1^2-5x_1x_3+25/4\ x_3^2)+9/4\ x_3^2$$

$$=(x_2+x_1-2x_3)^2-(x_1-5/2\ x_3)^2+9/4\ x_3^2$$

只要作线性变换：

$$\begin{cases} y_1 = x_1 + x_2 - 2x_3 \\ y_2 = x_1 - \frac{5}{2}x_3 \\ y_3 = x_3 \end{cases} \quad \text{即:} \begin{cases} x_1 = y_2 + \frac{5}{2}y_3 \\ x_2 = y_1 - y_2 - \frac{1}{2}y_3 \\ x_3 = y_3 \end{cases}$$

二次型 $f(x_1, x_2, x_3)=x_2^2+2x_1x_2+x_1x_3-4x_2x_3$ 变为标准形

$$f=y_1^2-y_2^2+9/4\,y_3^2.$$

所用的线性变换**x=Cy**的矩阵

$$\mathbf{C}=\begin{pmatrix} 0 & 1 & \frac{5}{2} \\ 1 & -1 & -\frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix}$$

由于**C**是可逆矩阵，所以线性变换**x=Cy**是可逆线性变换.

例4 求一可逆线性变换化二次型

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

为标准形和规范形，并给出所用的线性变换.

解 由于这里不含有平方项，首先作变换

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

则有
$$f=2y_1^2-2y_2^2-4y_1y_3-6y_2y_3.$$

然后再用配方法逐项配方，则有

$$\begin{aligned} f&=2y_1^2-2y_2^2-4y_1y_3-6y_2y_3 \\ &=2(y_1^2-2y_1y_3+y_3^2)-2y_2^2-2y_3^2-6y_2y_3 \\ &=2(y_1-y_3)^2-2(y_2^2+3y_2y_3+9/4\ y_3^2)+5/2\ y_3^2 \\ &=2(y_1-y_3)^2-2(y_2+3/2\ y_3)^2+5/2\ y_3^2 \end{aligned}$$

只要作变换:

$$\begin{cases} z_1 = y_1 - y_3 \\ z_2 = y_2 + \frac{3}{2} y_3 \\ z_3 = y_3 \end{cases}$$

即: $\begin{cases} y_1 = z_1 + z_3 \\ y_2 = z_2 - \frac{3}{2} z_3 \\ y_3 = z_3 \end{cases}$ 也即: $\begin{cases} x_1 = z_1 - z_2 + \frac{5}{2} z_3 \\ x_2 = z_1 + z_2 - \frac{1}{2} z_3 \\ x_3 = z_3 \end{cases}$

原二次型化为标准形: $f=2z_1^2-2z_2^2+5/2z_3^2$.

再作变换:

$$\begin{cases} z_1 = \frac{1}{\sqrt{2}} w_1 \\ z_2 = \frac{1}{\sqrt{2}} w_2 \\ z_3 = \sqrt{\frac{2}{5}} w_3 \end{cases}$$

也即: $\begin{cases} x_1 = \frac{1}{\sqrt{2}}\omega_1 - \frac{1}{\sqrt{2}}\omega_2 + \sqrt{\frac{5}{2}}\omega_3 \\ x_2 = \frac{1}{\sqrt{2}}\omega_1 + \frac{1}{\sqrt{2}}\omega_2 - \frac{1}{\sqrt{10}}\omega_3 \\ x_3 = \sqrt{\frac{2}{5}}\omega_3 \end{cases}$

原二次型化为规范形: $f=w_1^2-w_2^2+w_3^2$.

由前三例可见，用配方法化二次型为标准形步骤为：

情形I：二次型中含有平方项$a_{ii}x_i^2$，则把所有含x_i的项集中配方，要求配方以后其余项中不再含有x_i；

情形II：二次型中不含平方项，但含$2a_{ij}x_i x_j$项，则先作可逆线性变换：

$$\begin{cases} x_i = y_i - y_j \\ x_j = y_i + y_j \\ x_k = y_k \quad , \quad k \neq i,j \end{cases}$$

把二次型化为含有平方项，再进行配方．最后把两次可逆线性变换复合成一个可逆线性变换．

可见，任意二次型都能找到可逆变换**x=Cy**，把二次型化成标准形或规范形．但是，随着配方的方法不同，二次型的标准形和所用的可逆变换可能是不同(不唯一)的．

寻找一个可逆线性变换化二次型为标准形也可以用矩阵的合同对角化来实现.

由于可逆变换$\mathbf{x}=\mathbf{C}\mathbf{y}$化二次型$f=\mathbf{x}^T\mathbf{A}\mathbf{x}$为标准形，就是使$\mathbf{C}^T\mathbf{A}\mathbf{C}=\mathbf{\Lambda}$为对角矩阵.

由于矩阵$\mathbf{C}$可逆，记$\mathbf{C}=\mathbf{P}_1\mathbf{P}_2\dots\mathbf{P}_s$(这里$\mathbf{P}_1$, $\mathbf{P}_2,\dots,$ $\mathbf{P}_s$均为初等矩阵)，则有:

$$\mathbf{\Lambda}= \mathbf{P}_s^T\mathbf{P}_{s-1}^T\dots\mathbf{P}_2^T\mathbf{P}_1^T\mathbf{A}\mathbf{P}_1\mathbf{P}_2\dots\mathbf{P}_{s-1}\mathbf{P}_s.$$

由初等矩阵的性质可知，矩阵$\mathbf{A}$可以经过有限次初等行变换和完全相同的初等列变换变为对角矩阵$\mathbf{\Lambda}$.

例如，例2中的二次型，由于

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

$$\mathbf{A}=\begin{pmatrix}3&1&2\\1&3&-2\\2&-2&0\end{pmatrix}\overset{\substack{r_2-1/3\ r_1\\ r_3-2/3\ r_1}}{\sim}\begin{pmatrix}3&1&2\\0&8/3&-8/3\\0&-8/3&-4/3\end{pmatrix}$$

$$\overset{\substack{c_2-1/3\ c_1\\ c_3-2/3\ c_1}}{\sim}\begin{pmatrix}3&0&0\\0&8/3&-8/3\\0&-8/3&-4/3\end{pmatrix}\overset{r_3+\ r_2}{\sim}\begin{pmatrix}3&0&0\\0&8/3&-8/3\\0&0&-4\end{pmatrix}$$

$$\overset{c_3+\ c_2}{\sim}\begin{pmatrix}3&0&0\\0&8/3&-0\\0&0&-4\end{pmatrix}=\mathbf{\Lambda}.$$

上述初等变换写成矩阵形式就是:

$$\begin{pmatrix}1&0&0\\0&1&0\\0&1&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\\frac{-2}{3}&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\\frac{-1}{3}&1&0\\0&0&1\end{pmatrix}\mathbf{A}\begin{pmatrix}1&\frac{-1}{3}&0\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&\frac{-2}{3}\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&1\\0&0&1\end{pmatrix}$$

$$=\mathbf{\Lambda}.$$

也即
$$C^TAC=\Lambda,$$
其中

$$\mathbf{C}=\begin{pmatrix}1 & \frac{-1}{3} & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & \frac{-2}{3}\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 1\\ 0 & 0 & 1\end{pmatrix}$$

$$=\mathbf{E}(1+2(-1/3))\ \mathbf{E}(1+3(-2/3))\ \mathbf{E}(2+3(1))$$

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

只要做变换**x=Cy**, 二次型变成标准形:

$$f=3y_1^2+8/3\ y_2^2-4y_3^2.$$

注: 以上做法中的变换**x=Cy**一般**未必是正交变换**.

§4 正定二次型

一. 惯性定理与正定二次型

将一个二次型化成标准形，所用的线性变换是不唯一的，所得的标准形也不是唯一的，但标准形中所含平方项的个数却是相同的，它等于二次型的秩，也等于二次型矩阵非零特征值的个数. 不同的标准形之间还有什么关系呢?

定理6.3(惯性定理) 设实二次型$f=x^TAx$，其秩为r，在不同的可逆线性变换x=Cy和x=Dz下，f的标准形分别为

$$f=\lambda_1y_1^2+\lambda_2y_2^2+\dots+\lambda_ry_r^2\ (\lambda_i\neq0,\ i=1,2,\dots,r.),$$

$$f=\mu_1z_1^2+\mu_2z_2^2+\dots+\mu_rz_r^2\ (\mu_i\neq0,\ i=1,2,\dots,r.),$$

则$\lambda_1, \lambda_2, \dots, \lambda_r$中正数的个数与$\mu_1, \mu_2, \dots, \mu_r$中正数的个数相同.(负系数的个数当然也相同)

此定理的证明思路参见下页

关于定理6.3的证明大意(建议自修)

只需证明：对于给定的二次型，其规范型中的正项数和负项数都是确定不变的.

对于任意的标准形

$$f = d_1y_1^2+\dots+d_py_p^2 - d_{p+1}y_{p+1}^2 - \dots - d_ry_r^2 \ (d_i\neq 0,\ i= 1, 2,\dots, r.),$$

均做可逆线性变换

$$\begin{cases} y_1 = \dfrac{1}{\sqrt{d_1}} z_1, \\ \quad\cdots\ \cdots \\ y_r = \dfrac{1}{\sqrt{d_r}} z_r, \\ y_{r+1} = z_{r+1}, \\ \quad\cdots\ \cdots \\ y_n = z_n. \end{cases}$$

使以上的二次型化为规范形：

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

因此，以下假设二次型$f(x_1, x_2, \ldots, x_n)$经过可逆线性变换

$$\mathbf{x} = \mathbf{B}\mathbf{y}$$

化成规范形：

$$f = y_1^2 + \ldots + y_p^2 - y_{p+1}^2 - \ldots - y_r^2, \tag{*1}$$

又假设二次型$f(x_1, x_2, \ldots, x_n)$还可经过可逆线性变换

$$\mathbf{x} = \mathbf{C}\mathbf{z}$$

化成规范形：

$$f = z_1^2 + \ldots + z_q^2 - z_{q+1}^2 - \ldots - z_r^2. \tag{*2}$$

以下往证：p=q.

采用反证法. 假设p>q, 则由以上假设(*1)与(*2)可知

$$y_1^2 + \ldots + y_p^2 - y_{p+1}^2 - \ldots - y_r^2 = z_1^2 + \ldots + z_q^2 - z_{q+1}^2 - \ldots - z_r^2, \tag{*3}$$

其中

$$\mathbf{z} = \mathbf{C}^{-1}\mathbf{B}\mathbf{y}\ . \tag{*4}$$

令:

$$\mathbf{G}=\mathbf{C}^{-1}\mathbf{B}=\begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{pmatrix},$$

于是(*4)化为

$$\begin{cases} z_1 = g_{11}y_1 + g_{12}y_2 + \cdots + g_{1n}y_n, \\ z_2 = g_{21}y_1 + g_{22}y_2 + \cdots + g_{2n}y_n, \\ \cdots\cdots, \\ z_n = g_{n1}y_1 + g_{n2}y_2 + \cdots + g_{nn}y_n. \end{cases} \tag{*5}$$

然后，考虑方程组

$$\begin{cases} g_{11}y_1 + g_{12}y_2 + \cdots + g_{1n}y_n = 0, \\ \cdots\cdots \\ g_{q1}y_1 + g_{q2}y_2 + \cdots + g_{qn}y_n = 0, \\ y_{p+1} = 0, \\ \cdots\cdots, \\ y_n = 0,. \end{cases} \tag{*6}$$

可见, 方程组(*6)存在n个变量以及

$$q+(n-p)= n-(p-q) \le n$$

个方程. 因此, 方程组(*6)必有非零解.

设方程组(*6)的一个非零解为

$$(y_1,\dots, y_p, y_{p+1},\dots, y_n)= (k_1,\dots k_p, k_{p+1},\dots, k_n),$$

则其中必有 $k_{p+1}=\dots= k_n= 0$, 且存在某$k_i \ne 0$ $(1\le i\le p)$成立$k_i \ne 0$,

将此解(的后半部分$k_{p+1}=\dots= k_n= 0$)代入(*3)左边, 可得

$$k_1^2+\dots+k_p^2>0, \qquad (*7)$$

(*6)的解(的前半部分)结合(*5), 意味着

$$z_1=\dots=z_q= 0,$$

以上结果代入(*3)的右边, 可得

$$-z_{q+1}^2-\dots-z_r^2\le 0. \qquad (*8)$$

比较(*7)与(*8), 得出矛盾. 因此假设p>q是不成立的.

从而可得$p\le q$.

同理可证: $q\le p$.

从而得到: $p=q$.

证毕.

回定理6.3

定义6.5 二次型f的标准形中的正系数的个数称为f的**正惯性指数**，负系数的个数称为f的**负惯性指数**.

惯性定理指出，对二次型做任意实可逆变换，其(正、负)惯性指数都不变.

推论1 n阶实对称矩阵A与B合同的充分必要条件是：A与B对应的二次型的惯性指数相同.

思考，或教师讲解.

推论2 n阶实对称矩阵A与B合同的充分必要条件是：

$$\begin{cases} \text{A的正特征值个数与B的正特征值个数相同,} \\ \text{A的负特征值个数与B的负特征值个数相同,} \\ \text{A的取值为0的特征值个数与B的取值为0的特征值个数相同.} \end{cases}$$

推论3 若n阶实对称矩阵A与B相似，则A与B合同.

定义6.6 设$f=\mathbf{x}^T\mathbf{A}\mathbf{x}$为实二次型.

如果对于任意$\mathbf{x}\neq\mathbf{0}$, 都有$f>0$, 则称f为正定二次型, 并称实对称矩阵A为(对称)正定矩阵;

如果对于任意$\mathbf{x}\neq\mathbf{0}$, 都有$f<0$, 则称f为负定二次型, 并称实对称矩阵A称为(对称)负定矩阵.

定义6.7 设$f=\mathbf{x}^T\mathbf{A}\mathbf{x}$为实二次型.

如果对于任意$\mathbf{x}\neq\mathbf{0}$, 都有$f\geq 0$, 则称f为半正定二次型, 并称实对称矩阵A为(对称)半正定矩阵;

如果对于任意$\mathbf{x}\neq\mathbf{0}$, 都有$f\leq 0$, 则称f为半负定二次型, 并称实对称矩阵A称为(对称)半负定矩阵.

二次型的正定性在很多技术领域中有着广泛的应用.

注: 既不是半正定也不是半负定的方阵称为不定矩阵.

二. 正定二次型(正定矩阵)的判定

定理6.4 对于n元实二次型$f=\mathbf{x}^T\mathbf{A}\mathbf{x}$,
f为正定二次型的充分必要条件是: f的正惯性指数等于n;
f为负定二次型的充分必要条件是: f的负惯性指数等于n.

证明 设二次型$f=\mathbf{x}^T\mathbf{A}\mathbf{x}$经可逆变换$\mathbf{x}=\mathbf{C}\mathbf{y}$化为标准形:

$$f=\mathbf{x}^T\mathbf{A}\mathbf{x}=\mathbf{y}^T\mathbf{C}^T\mathbf{A}\mathbf{C}\mathbf{y}=d_1y_1^2+d_2y_2^2+\dots+d_ny_n^2.$$

以下主要证明正定时的结论.

充分性: 设f的正惯性指数为n, 则由惯性定理可知

$$d_i>0 \quad (i=1, 2,\dots, n).$$

对任意$\mathbf{x}\neq\mathbf{0}$, 都有$\mathbf{y}=\mathbf{C}^{-1}\mathbf{x}\neq\mathbf{0}$, 故

$$f=d_1y_1^2+d_2y_2^2+\dots+d_ny_n^2>0,$$

即f是正定二次型.

必要性：设f是正定二次型，以下往证f的正惯性指数等于n.

用反证法.

假设f的正惯性指数不等于n，则有某个$d_i \le 0$，$i \in \{1, 2, \ldots, n\}$，则取

$$\mathbf{y}=\mathbf{e}_i=(0,\ldots 0, 1, 0,\ldots 0)^T,$$ (第i个标准单位向量)

第i个分量

注意$\mathbf{x}=\mathbf{C}\mathbf{y}\neq\mathbf{0}$，且有

$$f=\mathbf{y}^T\mathbf{C}^T\mathbf{A}\mathbf{C}\mathbf{y} = d_1 0^2+\ldots+d_{i-1}0^2+d_i 1^2+d_{i+1}0^2+\ldots+d_n 0^2=d_i \le 0,$$

这与“f是正定二次型”的假设矛盾.

所以必有$d_i>0$ $(i=1, 2,\ldots, n)$，也即f的正惯性指数必然等于n.

关于负定二次型的相应结论也可以类似地给出证明.

证毕.

推论1 对于n元实二次型$f=\mathbf{x}^T\mathbf{A}\mathbf{x}$,

f为半正定二次型的充分必要条件是:

f的负惯性指数等于0;

f为半负定二次型的充分必要条件是:

f的正惯性指数等于0.

推论2 对于实对称矩阵**A**,

A正定的充要条件是: **A**的n个特征值全是正数;

A负定的充要条件是: **A**的n个特征值全是负数;

A半正定的充要条件是: **A**没有负特征值;

A半负定的充要条件是: **A**没有正特征值.

定义6.8 设$\mathbf{A}=(a_{ij})_{n\times n}$，形如

$$\mathbf{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2k} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} & \cdots & a_{kn} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} & \cdots & a_{nn} \end{pmatrix},$$

则$\mathbf{A}$的左上角的k阶子式，即

$$D_k=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix}, \quad k\in\{1, 2, \ldots, n\}.$$

称为矩阵$\mathbf{A}$的k阶顺序主子式.

这是行列式

注：显然，$D_1=a_{11}$，$D_n=\det(\mathbf{A})$.

进入例题

定理6.5 对于n阶实对称矩阵A，

A正定的充分必要条件是：

A的所有顺序主子式都大于0；

A负定的充分必要条件是：

$$\begin{cases} \text{A的所有奇数阶顺序主子式都小于0,} \\ \text{A的所有偶数阶顺序主子式都大于0.} \end{cases}$$

关于定理6.5的证明思路，参见下页(可自修).

进入证明

关于定理6.5的证明思路大意(建议自修)

以下主要对矩阵A为对称正定的情形给出证明

1. 必要性

首先往证：若A是对称正定矩阵，则A的任意一个顺序主子阵A_k也必然是对称正定的.

采用反证法. 假设某个顺序主子阵A_k不是正定的(但A_k显然是对称的)，则存在某个非零向量$x \in R^k$，使

$$x^T A_k x \le 0.$$

因此，可取非零向量

$$y = \begin{pmatrix} x \\ 0 \end{pmatrix} \in R^n,\ 其中 0 \in R^{n-k}, 为零向量.$$

使

$$y^T A y = (x^T, 0^T)\begin{pmatrix} A_k & B \\ B^T & C \end{pmatrix}\begin{pmatrix} x \\ 0 \end{pmatrix} = x^T A_k x \le 0.$$

这与A是正定矩阵的条件矛盾，故假设错误，A_k必为正定的.

由于A_k对称正定，所以A_k的全体特征值必全为正实数，因此

$$\det(A_k)=A_k\text{的全体特征值之乘积}>0.$$

必要性得证.

2. 充分性

对矩阵的阶数采用归纳法.

对于1阶矩阵，定理结论显然成立.

现在假设对所有阶数小于n-1的对称矩阵，当其各阶顺序主子式都大于0时，此矩阵必为正定矩阵.

以下讨论中，对于n阶对称矩阵

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

设**A**的各阶顺序主子式都大于0. 往证**A**为正定矩阵.

由于A的1阶顺序主子式大于0，即$a_{11}>0$，故必存在可逆矩阵

$$\mathbf{P}=\begin{pmatrix}1&&&\\p_{21}&1&&\\&&\cdots&\\&&&1\end{pmatrix}\begin{pmatrix}1&&&\\&1&&\\p_{31}&&\cdots&\\&&&1\end{pmatrix}\cdots\begin{pmatrix}1&&&\\&1&&\\&&\cdots&\\p_{n1}&&&1\end{pmatrix}$$

$$=\begin{pmatrix}1&&&\\p_{21}&1&&\\\cdots&&\cdots&\\p_{n1}&&&1\end{pmatrix}，$$ 其中 $p_{i1}=\dfrac{-a_{i1}}{a_{11}}$， $i=2,\ldots,n$. 使

$$\mathbf{PAP}^{\mathrm{T}}=\begin{pmatrix}1&&&\\p_{21}&1&&\\\cdots&&\cdots&\\p_{n1}&&&1\end{pmatrix}\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&\cdots&\cdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}\begin{pmatrix}1&p_{21}&\cdots&p_{n1}\\&1&&\\&&\cdots&\\&&&1\end{pmatrix}$$

$$=\begin{pmatrix}a_{11}&0&\cdots&0\\0&a'_{22}&\cdots&a'_{2n}\\\cdots&\cdots&\cdots&\cdots\\0&a'_{n2}&\cdots&a'_{nn}\end{pmatrix}=\begin{pmatrix}a_{11}&\mathbf{0}^{\mathrm{T}}\\\mathbf{0}&\mathbf{A}'_{n-1}\end{pmatrix}.$$

其中$\mathbf{A}'_{n-1}$为$n-1$阶对称矩阵.

因上述合同变换PAP^T相当于只对A做一系列的(第三类)初等变换,故不改变A的所有2阶至n阶顺序主子式的值.

因此, A的2阶顺序主子式= $a_{11}\times A'_{n-1}$的1阶顺序主子式,

A的3阶顺序主子式= $a_{11}\times A'_{n-1}$的2阶顺序主子式,

……

A的n阶顺序主子式= $a_{11}\times A'_{n-1}$的n-1阶顺序主子式.

于是由$a_{11}>0$以及A的各阶顺序主子式都大于0的条件,立即得到: A'_{n-1}的各阶顺序主子式必大于0.

再由假设条件可得: A'_{n-1}必为对称正定矩阵.

因此, 矩阵 $\mathbf{PAP}^{\mathrm{T}}=\begin{pmatrix} a_{11} & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & A'_{n-1} \end{pmatrix}$必然也是对称正定矩阵.

故矩阵$\mathbf{A}=\mathbf{P}^{-1}\begin{pmatrix} a_{11} & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & A'_{n-1} \end{pmatrix}(\mathbf{P}^{-1})^{\mathrm{T}}$必为正定矩阵.

当A是负定矩阵时, -A为正定矩阵, 类似也可证得对应结论.

证毕.

回到定理6.5

例5 判断二次型$f(x_1, x_2, x_3)=2x_1^2+3x_2^2+3x_3^2-4x_2x_3$的正定性.

解

法1 由于$f(x_1, x_2, x_3)= 2x_1^2+3(x_2-2/3\ x_3)^2+5/3\ x_3^2$,

所以，二次型$f(x_1, x_2, x_3)$是正定二次型.

法2 对二次型f的矩阵 $\mathbf{A}=\begin{pmatrix}2 & 0 & 0\\ 0 & 3 & -2\\ 0 & -2 & 3\end{pmatrix}$作合同变换，有

$$\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 2/3 & 1\end{pmatrix}\begin{pmatrix}2 & 0 & 0\\ 0 & 3 & -2\\ 0 & -2 & 3\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 2/3\\ 0 & 0 & 1\end{pmatrix}=\begin{pmatrix}2 & 0 & 0\\ 0 & 3 & 0\\ 0 & 0 & 5/3\end{pmatrix},$$

所以矩阵**A**与对角矩阵diag(2, 3, 5/3) 合同.

因此，二次型$f(x_1, x_2, x_3)$是正定二次型.

法3 计算矩阵$\mathbf{A}$的特征多项式，由于

$$|\lambda\mathbf{E}-\mathbf{A}|=\begin{vmatrix}\lambda-2 & 0 & 0\\ 0 & \lambda-3 & 2\\ 0 & 2 & \lambda-3\end{vmatrix}$$

$$=(\lambda-2)(\lambda-1)(\lambda-5),$$

所以，$\mathbf{A}$的全部特征值为2, 1, 5.

因此$\mathbf{A}$是正定矩阵，二次型$f(x_1, x_2, x_3)$是正定二次型.

法4 求矩阵$\mathbf{A}$的各阶顺序主子式，由于

$$D_1=2>0,$$

$$D_2=6>0,$$

$$D_3=10>0,$$

所以，$\mathbf{A}$是正定矩阵，二次型$f(x_1, x_2, x_3)$是正定二次型.

例6 确定参数t, 使二次型

$f(x_1, x_2, x_3)=x_1^2+2x_2^2+3x_3^2+2tx_1x_2-4tx_2x_3$

为正定二次型.

解 二次型的矩阵为

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

由

$$D_1=1>0,$$

$$D_2=2-t^2>0,$$

$$D_3=3(2-t^2)+2t(-2t)=6-7t^2>0,$$

得$t^2<6/7$，即：

$-\sqrt{6/7}<t<\sqrt{6/7}$ 时, $f(x_1, x_2, x_3)$正定.

例7 设**A**是对称正定矩阵，请证明: det(**A**+**E**)>1.

证明： 由于**A**是对称正定矩阵，所以**A**的特征值λ_i (i = 1, 2,…, n)都大于0.

又由于$\lambda_i + 1$ (i = 1, 2,…, n)是**A**+**E**的全体n个特征值，且对于每个$\lambda_i + 1$均有

$$\lambda_i + 1 > 1, \ (i = 1, 2, \ldots, n).$$

因此可得

$$\det(\mathbf{A}+\mathbf{E}) = \prod_{i=1}^{n}(\lambda_i + 1) > 1.$$

证毕.

例8 设实对称矩阵$\mathbf{A}$满足$\mathbf{A}^2-5\mathbf{A}+6\mathbf{E}=\mathbf{O}$，请证明矩阵$\mathbf{A}$是正定的.

证明： 设λ是矩阵$\mathbf{A}$的任一特征值，ξ是属于λ的特征向量，则有

$$\mathbf{A}\xi=\lambda\xi,$$

于是有

$$\mathbf{0}=\mathbf{O}\xi=(\mathbf{A}^2-5\mathbf{A}+6\mathbf{E})\xi=\mathbf{A}^2\xi-5\mathbf{A}\xi+6\xi=\lambda^2\xi-5\lambda\xi+6\xi,$$

也即

$$(\lambda^2-5\lambda+6)\xi=\mathbf{0}.$$

由于$\xi\neq\mathbf{0}$，所以$\mathbf{A}$的任意特征值λ必满足方程

$$\lambda^2-5\lambda+6=0,$$

可得$\lambda=2$或$\lambda=3$.

因此，矩阵$\mathbf{A}$的特征值都大于0，所以$\mathbf{A}$是正定矩阵. 证毕.

教材作业题目，强烈建议自修

其中(以下题目中的正定矩阵，均可理解为对称正定矩阵)

经典且难度中等：P132: 6, 7, 8, 9, 15, 16.

经典且略有难度：P132: 17, 19. 其中17题可推广为以下命题：

命题：若$\boldsymbol{A}$为n阶对称矩阵，则

许多习题(如19题)可用此命题结论来做

$$\max_{\boldsymbol{x}\in \mathrm{R}^n,\text{且}|\boldsymbol{x}|=1} \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \lambda_{\max}(\boldsymbol{A}),$$

此即17题

$$\min_{\boldsymbol{x}\in \mathrm{R}^n,\text{且}|\boldsymbol{x}|=1} \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \lambda_{\min}(\boldsymbol{A}).$$

其中，

$\lambda_{\max}(\boldsymbol{A})$和$\lambda_{\min}(\boldsymbol{A})$分别代表$\boldsymbol{A}$的最大特征值和最小特征值，

$|\boldsymbol{x}|$代表向量$\boldsymbol{x}=(x_1, x_2, \ldots, x_n)^{\mathrm{T}}$的长度，也即

$$|\boldsymbol{x}| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$