同济大学

高等代数与解析几何

教师： 蒋志洪

理学院数学系

§10.5 正定二次型与正定矩阵

定义. 实二次型$f(x_1, x_2, \cdots, x_n) = X^t AX$, 对任意一组不全为零的实数$x_1$,$x_2, \cdots, x_n$.

(1) 如果总有$f(x_1, x_2, \cdots, x_n) > 0$, 则称该二次型为正定型

(2) 如果总有$f(x_1, x_2, \cdots, x_n) \geq 0$, 则称该二次型为半正定型

(3) 如果总有$f(x_1, x_2, \cdots, x_n) < 0$, 则称该二次型为负定型

(4) 如果总有$f(x_1, x_2, \cdots, x_n) \leq 0$, 则称该二次型为半负定型

(5) 如果二次型$f(x_1, x_2, \cdots, x_n)$ 既有大于零又有小于零, 则称该二次型为不定型.

说明. 与正定二次型对应的实对称阵称为正定阵, 与半正定二次型对应的实对称阵称为半正定阵, 与负定二次型对应的实对称阵称为负定阵, 与半负定二次型对应的实对称阵称为半负定阵.

例. 设二次型f是标准形:

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

当$d_i > 0\ (i = 1, 2, \cdots, n)$时, 则$f$是正定型

当$d_i < 0\ (i = 1, 2, \cdots, n)$时, 则$f$是负定型

当$d_i \geq 0\ (i = 1, 2, \cdots, n)$时, 则$f$是半正定型

当$d_i \leq 0\ (i = 1, 2, \cdots, n)$时, 则$f$是半负定型

当$d_1, d_2, \cdots, d_n$中既有大于零又有小于零时, 则f是半负定型

所以一个对角阵是正定充要条件是对角线上的元素都大于零, 一个对角阵是半正定充要条件是对角线上的元素都大于等于零, 一个对角阵是负定充要条件是对角线上的元素都小于零, 一个对角阵是半负定充要条件是对角线上的元素都小于等于零.

命题. 可逆变量替换$X=PY$ 保持实二次型$f(x_1,x_2,\cdots,x_n)$ 的正定性不变.

证明. 设
$$f(x_1,x_2,\cdots,x_n)=X^tAX$$
则f经可逆变量替换$X=PY$ 化为
$$g(y_1,y_2,\cdots,y_n)=Y^tBY$$
这里$B=P^tAP$.

如果f是正定型但是$g(y_1,y_2,\cdots,y_n)$ 不是正定型, 则存在一组不全为零的实数$y_i=c_i\ (i=1,2,\cdots,n)$, 使得$g(c_1,\cdots,c_n)\le 0$. 设
$$\begin{pmatrix} d_1\\ d_2\\ \vdots\\ d_n\end{pmatrix}=P\begin{pmatrix} c_1\\ c_2\\ \vdots\\ c_n\end{pmatrix}\qquad \text{记 } C=\begin{pmatrix} c_1\\ c_2\\ \vdots\\ c_n\end{pmatrix}\ D=\begin{pmatrix} d_1\\ d_2\\ \vdots\\ d_n\end{pmatrix}$$
因为P可逆, 所以$d_1,d_2,\cdots,d_n$也不全为零. 而
$$0\ge g(c_1,c_2,\cdots,c_n)=C^tBC=C^tP^tAPC=D^tAD=f(d_1,d_2,\cdots,d_n)$$
其中$C=(c_1\cdots,c_n)^t$, $D=(d_1,\cdots,d_n)^t$. 这与二次型X^tAX是正定型矛盾.

推论. 矩阵合同不改变矩阵的正定性.

说明. 从上面命题的证明可以容易看出, 这个命题对于负定, 半正定, 半负定也成立.

定义. 设n阶方阵$A=(a_{ij})$的子式

$$D_i=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1i} \\ a_{21} & a_{22} & \cdots & a_{2i} \\ \cdots & & \cdots & \\ a_{i1} & a_{i2} & \cdots & a_{ii} \end{vmatrix} \quad (i=1,2,\cdots,n)$$

称为A的第i阶顺序主子式.

命题. 设实二次型$f(x_1,x_2,\cdots,x_n)=X^tAX$, 如果$A$的所有顺序主子式都不等于零, 则存在可逆变换$X=PY$, 使得$f$经可逆变换后变成

$$D_1y_1^2+\frac{D_2}{D_1}y_2^2+\cdots+\frac{D_i}{D_{i-1}}y_i^2+\cdots+\frac{D_n}{D_{n-1}}y_n^2$$

证明. 等价证明A合同于对角阵$\mathrm{diag}\{D_1, \frac{D_2}{D_1}, \cdots, \frac{D_n}{D_{n-1}}\}$, 对$A$的阶数进行数学归纳, 当$A$的阶数是1时, 结论显然成立, 归纳假设$n-1$阶成立, 当$A$的阶数为$n$时, 记

$$A_i = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1i} \\ a_{21} & a_{22} & \cdots & a_{2i} \\ \cdots & & \cdots & \\ a_{i1} & a_{i2} & \cdots & a_{ii} \end{pmatrix} \quad A = \begin{pmatrix} A_{n-1} & \beta^t \\ \beta & a_{nn} \end{pmatrix} \quad P_1 = \begin{pmatrix} E_{n-1} & \mathbf{0} \\ -\beta A_{n-1}^{-1} & 1 \end{pmatrix}$$

则

$$P_1 A P_1^T = \begin{pmatrix} E_{n-1} & \mathbf{0} \\ -\beta A_{n-1}^{-1} & 1 \end{pmatrix} \begin{pmatrix} A_{n-1} & \beta^t \\ \beta & a_{nn} \end{pmatrix} \begin{pmatrix} E_{n-1} & -A_{n-1}^{-1}\beta^T \\ \mathbf{0} & 1 \end{pmatrix} = \begin{pmatrix} A_{n-1} & 0 \\ 0 & a_{nn} - \beta A_{n-1}\beta^T \end{pmatrix}$$

两边取行列式得

$$D_n = |A| = |P_1 A P_1^T| = |A_{n-1}|(a_{nn} - \beta A_{n-1}\beta^T) = D_{n-1}(a_{nn} - \beta A_{n-1}\beta^T)$$

由此推出

$$a_{nn} - \beta A_{n-1}\beta^T = \frac{D_n}{D_{n-1}}$$

所以

$$P_1AP_1^T=\begin{pmatrix} A_{n-1} & 0 \\ 0 & \frac{D_n}{D_{n-1}} \end{pmatrix}$$

由于A_{n-1}是$n-1$阶矩阵, 由归纳假设结论成立, 再注意A_{n-1}的顺序主子式就是A的顺序主子式, 所以存在可逆阵P_2使得

$$P_2A_{n-1}P_2^T=\mathrm{diag}\{D_1,\frac{D_2}{D_1},\cdots,\frac{D_{n-2}}{D_{n-1}}\}$$

令

$$P=\begin{pmatrix} P_2 & 0 \\ 0 & 1 \end{pmatrix}P_1$$

则

$$PAP^T=\mathrm{diag}\{D_1,\frac{D_2}{D_1},\cdots,\frac{D_n}{D_{n-1}}\}$$

定理. 实二次型$f(x_1, x_2, \cdots, x_n) = X^tAX$满足如下等价条件之一时, 它一定是正定型.

(1) 对所有不为零的$X = (x_1, x_2, \cdots, x_n)^t \in \mathbb{R}^n$都有$X^tAX > 0$,

(2) 对称阵A的所有特征值$\lambda_1, \lambda_2, \cdots, \lambda_n$全大于零

(3) 对称阵A的正惯性指数$p = R(A) = n$.

(4) A与单位阵E合同

(5) A的所有顺序主子式皆大于零.

证明. (1) $\Rightarrow$ (2) 因为A是实对称阵, 所以存在正交阵O, 使$O^tAO = O^{-1}AO = \mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$, 所以经变量替换$X = OY$, X^tAX可化为标准形

$$X^tAX = Y^t(O^tAO)Y = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$$

由于可逆变量替换不改变正定性, 所以$\lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$ 是正定型, 于是$\lambda_i > 0$ $(i = 1, 2, \cdots, n)$.

(2) $\Rightarrow$ (3) 实二次型$f(x_1, x_2, \cdots, x_n) = X^tAX$可以经过正交变量替换$X = OY$化为标准形

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$

其中$\lambda_1, \lambda_2, \cdots, \lambda_n$是矩阵$A$的所有特征值皆大于零, 于是实二次型的标准形中, 非零平方项个数等于二次型的秩; 正平方项的个数等于二次型的正惯性指数. 因此, $R(A) =$正惯性指数$= n$.

(3) $\Rightarrow$ (4) 由于$R(A) =$正惯性指数$= n$, 所以A与单位阵E合同.

(4) $\Rightarrow$ (5) 由于A与单位阵E合同, 实二次型X^tAX经可逆变量替换$X=PZ$可化为规范形

$$z_1^2+z_2^2+\cdots+z_n^2$$

它当然是正定型. 由于可逆变换不改变正定性, 原二次型X^tAX也必为正定型.

如果A的某个顺序主子式$D_i=0$, 设A_i为D_i对应的A的子矩阵, 则A_i不可逆, 所以存在一组不全为零的数$c_1,c_2,\cdots,c_i$使得

$$A_i\begin{pmatrix}c_1\\c_2\\\vdots\\c_i\end{pmatrix}=0$$

记

$$A=\begin{pmatrix}A_i & B^T\\B & C\end{pmatrix}\quad X=\begin{pmatrix}c_1\\c_2\\\vdots\\c_i\\0\\\vdots\\0\end{pmatrix}$$

则

$$X^TAX = 0$$

这与f是正定矛盾. 所以A的所有顺序主子式都不等于零, 因此f可以经可逆变换后变成

$$D_1y_1^2 + \frac{D_2}{D_1}y_2^2 + \cdots + \frac{D_i}{D_{i-1}}y_i^2 + \cdots + \frac{D_n}{D_{n-1}}y_n^2$$

由于f是正定的所以

$$D_1 > 0, \frac{D_2}{D_1} > 0, \cdots, \frac{D_n}{D_{n-1}} > 0$$

由此推出

$$D_1 > 0, D_2 > 0, \cdots, D_n > 0$$

(5) $\Rightarrow$ (1) 由于A的所有顺序主子式都大于零, 所以f可以经可逆变换后变成

$$D_1y_1^2 + \frac{D_2}{D_1}y_2^2 + \cdots + \frac{D_i}{D_{i-1}}y_i^2 + \cdots + \frac{D_n}{D_{n-1}}y_n^2$$

由于系数都大于零, 所以是正定的, 由于可逆变换不改变正定性, 所以f是正定的.

推论. 对称阵A为正定阵的充分必要条件是存在上三角阵

$$T=\begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ & t_{22} & \cdots & t_{2n} \\ & & \ddots & \\ & & & t_{nn} \end{pmatrix},\quad \text{其中} t_{ii}>0,\quad i=1,2,\cdots,n$$

使得$A=T^tT$.

证明. (必要性) 正定对称阵A合同于单位阵E, 即存在可逆阵P, 使$A=P^tEP=P^tP$. P是实可逆阵, 所以正交阵Q和可逆上三角阵R, 使$P=QR$. 于是, $A=P^tP=R^tQ^tQR=R^tR$. 由P的QR分解的实际做法可知, 上三角阵R中主对角元素皆大于零, 所以$T=R$ 即为所求.

(充分性) 如果$A=T^tT$, 则A与单位阵合同. 所以A正定.

例. 设A是正定对称阵, 则对任意正整数m, 必存在唯一的正定对称阵B, 使得$B^m = A$.

证明: 因为A是正定对称阵, 所以存在正交阵O, 使

$$A = O\mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}O^t$$

其中$\lambda_1, \lambda_2, \cdots, \lambda_n$是$A$ 的特征值, 它们都是大于零的实数, 存在m次方根 $\sqrt[m]{\lambda_i} > 0$, 且

$$A = O\mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}O^{-1} = (O\mathrm{diag}\{\sqrt[m]{\lambda_1}, \cdots, \sqrt[m]{\lambda_n}\}O^{-1})^m$$

令$X = O\mathrm{diag}\{\sqrt[m]{\lambda_1}, \cdots, \sqrt[m]{\lambda_n}\}O^{-1}$, 它是正定对称阵, 使$X^m = A$.

再证唯一性. 如果有正定对称阵X,Y满足$X^m=Y^m=A$, 因为X,Y都是正定对称阵, 可写

$$\begin{aligned} X &= O_1\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}O_1^{-1} \\ Y &= O_2\mathrm{diag}\{\mu_1,\mu_2,\cdots,\mu_n\}O_2^{-1} \end{aligned}$$

其中O_1,O_2是正交阵, 且

$$0<\lambda_1\le\lambda_2\le\cdots\le\lambda_n,\quad 0<\mu_1\le\mu_2\le\cdots\le\mu_n$$

分别是X,Y特征值, 于是X^m的特征值$0<\lambda_1^m\le\lambda_2^m\le\cdots\le\lambda_n^n$, Y^m的特征值$0<\mu_1^m\le\mu_2^m\le\cdots\le\mu_n^m$. 由$X^m=Y^m$, 得$\lambda_i^m=\mu_i^m(i=1,2,\cdots,n)$. 但$\lambda_i,\mu_i$皆为正实数, 因此$\lambda_i=\mu_i$ $(i=1,2,\cdots,n)$. 记

$$D=\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}=\mathrm{diag}\{\mu_1,\mu_2,\cdots,\mu_n\}.$$

由$X^m=Y^m$, 得

$$O_1D^mO_1^{-1}=O_2D^mO_2^{-1}$$

设$O=O_2^{-1}O_1$, 有$OD^m=D^mO$. 现在可把D改写成

$$D^m=\begin{pmatrix}\lambda_1^mE_1 & & & \\ & \lambda_2^mE_2 & & \\ & & \ddots & \\ & & & \lambda_t^mE_t\end{pmatrix}$$

其中$\lambda_1,\cdots,\lambda_t$是$D$ 中全部互不相同的特征值. 比较等式$OD^m=D^mO$两边, 得

$$O=\begin{pmatrix}Q_1 & & & \\ & Q_2 & & \\ & & \ddots & \\ & & & Q_t\end{pmatrix}$$

其中Q_i 是与E_i有相同阶数的正交阵. 这样$OD=DO$, 即$O_1DO_1^{-1}=O_2DO_2^{-1}$. 从而$X=Y$.

- **多元函数的极值**

多元函数的泰劳(Talor)展开式

$$f(X_0+\Delta X)=f(X_0)+(\Delta X)^t\frac{\partial f}{\partial X}+\frac{1}{2!}(\Delta X)^tA(\Delta X)+o(\|\Delta X\|^2)$$

其中$f(X_0)=f(x_1^0,\cdots,x_n^0)$,

$$\Delta X=\begin{pmatrix}\Delta x_1\\ \vdots\\ \Delta x_n\end{pmatrix},\quad \frac{\partial f}{\partial X}=\begin{pmatrix}\frac{\partial f}{\partial x_1}\\ \vdots\\ \frac{\partial f}{\partial x_n}\end{pmatrix}_{X=X_0}\quad A=\begin{pmatrix}\frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1\partial x_n}\\ \frac{\partial^2 f}{\partial x_2\partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2\partial x_n}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial^2 f}{\partial x_n\partial x_1} & \frac{\partial^2 f}{\partial x_n\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2}\end{pmatrix}_{X=X_0}$$

如果$X=X_0$是极值点, 则必有$\frac{\partial f}{\partial X}|_{X_0}=0$. 这时有

$$f(X_0+\Delta X)-f(X_0)=\frac{1}{2!}(\Delta X)^tA(\Delta X)+o(\|\Delta X\|^2)$$

于是在X_0的足够小邻域里, 上式右端的正负号由二次型

$$(\Delta X)^tA(\Delta X)$$

确定. 特别当A为正定对称阵时, $f(X)$在X_0处取极小值; 当A为负定对称阵时, $f(X)$在X_0处取极大值.

注意. (1) 上面讨论的正定型的这些结果都可直接转移负定型对称阵. 因为当$f(x_1, x_2, \cdots, x_n)$是负定型时, $-f(x_1, x_2, \cdots, x_n)$就是正定型.

(2) 对于半正(负)定型有类似判别条件.

练习. 习题10.4:1(1),3,4,5,6,7.