线性代数（理工）

常寅山

第六章 二次型

二次型与曲线形状

问题

如何判断一个二次方程表示何种几何图形？如在二维平面上，我们知道

- $x^2+2y^2=1$ 是一个椭圆，
- 那 $7x^2+5y^2+2\sqrt{3}xy=4$ 表示什么图形呢？

作坐标变换 $\begin{cases} x=\frac{1}{2}x'+\frac{\sqrt{3}}{2}y', \\ y=-\frac{\sqrt{3}}{2}x'+\frac{1}{2}y' \end{cases}$，则

$$x'^2+2y'^2=1.$$

通过坐标变换我们知道，该方程为一个椭圆，其长轴方向为 x 轴沿着顺时针旋转 $\frac{\pi}{3}$.

二次型与曲线形状

问题

怎么用矩阵运算表示？

$$7x^2+5y^2+2\sqrt{3}xy=4,\ 即\ [x,y]\begin{bmatrix}7&\sqrt{3}\\\sqrt{3}&5\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=4,$$

$$坐标变换\ \begin{cases}x=\frac{1}{2}x'+\frac{\sqrt{3}}{2}y',\\ y=-\frac{\sqrt{3}}{2}x'+\frac{1}{2}y'\end{cases},\ 即\ \begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}\frac{1}{2}&\frac{\sqrt{3}}{2}\\-\frac{\sqrt{3}}{2}&\frac{1}{2}\end{bmatrix}\begin{bmatrix}x'\\y'\end{bmatrix},$$

$$[x',y']\begin{bmatrix}\frac{1}{2}&\frac{\sqrt{3}}{2}\\-\frac{\sqrt{3}}{2}&\frac{1}{2}\end{bmatrix}^T\begin{bmatrix}7&\sqrt{3}\\\sqrt{3}&5\end{bmatrix}\begin{bmatrix}\frac{1}{2}&\frac{\sqrt{3}}{2}\\-\frac{\sqrt{3}}{2}&\frac{1}{2}\end{bmatrix}\begin{bmatrix}x'\\y'\end{bmatrix}=4$$

$$计算得\ [x',y']\begin{bmatrix}4&0\\0&8\end{bmatrix}\begin{bmatrix}x'\\y'\end{bmatrix}=4,\ 即,\ x'^2+2y'^2=1.$$

二次型

定义 1 (二次型)

我们称数域上的一个二次齐次多项式

$$
\begin{aligned}
f(x_1,\ldots,x_n) = a_{11}x_1^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \cdots + 2a_{1n}x_1x_n \\
+a_{22}x_2^2 + 2a_{23}x_2x_3 + \cdots + 2a_{2n}x_2x_n \\
+a_{33}x_3^2 + \cdots + 2a_{3n}x_3x_n \\
+\cdots \\
+a_{nn}x_n^2
\end{aligned}
$$

为数域上的一个 n 元二次型.

注 1

当数域为实数域时, 称为实二次型; 当数域为复数域时, 称为复二次型. 我们仅讨论实二次型.

二次型

课堂练习

判断下列函数是否是二次型:

(1) $f(x,y,z)=x^2+y^2+z^2$. 是.

(2) $f(x,y)=x^4+y^4+2xy$. 否.

(3) $f(x,y)=x^2+2x+1+y^2$. 否.

二次型

课堂练习

已知 $A=(a_{ij})_{n\times n}$ 为对称矩阵, $X=(x_1,x_2,\cdots,x_n)^T$, 计算 X^TAX.

解.

$$\begin{aligned}
X^TAX = {} & a_{11}x_1^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + \cdots + a_{1n}x_1x_n \\
& a_{21}x_1x_2 + a_{22}x_2^2 + a_{23}x_2x_3 + \cdots + a_{2n}x_2x_n \\
& a_{31}x_1x_3 + a_{32}x_2x_3 + a_{33}x_3^2 + \cdots + a_{3n}x_3x_n \\
& + \cdots \\
& a_{n1}x_1x_n + a_{n2}x_2x_n + a_{n3}x_3x_n + \cdots + a_{nn}x_n^2 \\
= {} & a_{11}x_1^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \cdots + 2a_{1n}x_1x_n \\
& + a_{22}x_2^2 + 2a_{23}x_2x_3 + \cdots + 2a_{2n}x_2x_n \\
& + a_{33}x_3^2 + \cdots + 2a_{3n}x_3x_n \\
& + \cdots \\
& + a_{nn}x_n^2
\end{aligned}$$

二次型的矩阵表示

定义 2

令 $A=(a_{ij})_{n\times n}$ 为对称矩阵, $X=(x_1,\cdots,x_n)^T$, $f(x_1,x_2,\cdots,x_n)=X^TAX$ 为二次型, A 为二次型 f 的矩阵. A 的秩也称为二次型 f 的秩.

特别的, 如果二次型 $f(x_1,\cdots,x_n)$ 只含有平方项, 即

$$f(x_1,\cdots,x_n)=a_{11}x_1^2+a_{22}x_2^2+\cdots+a_{nn}x_n^2,$$

则称 f 为标准二次型, 此时 f 的矩阵为对角阵

$$\Lambda=\begin{bmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{bmatrix}.$$

二次型的矩阵表示

例 1 (P159 例 1.1.2)

设 $A=\begin{bmatrix}1 & -1 & 1\\ -1 & 2 & \frac{1}{2}\\ 1 & \frac{1}{2} & 3\end{bmatrix}$, 则以 A 为矩阵的二次型为

$$f(X)=X^TAX=x_1^2-2x_1x_2+2x_1x_3+2x_2^2+x_2x_3+3x_3^2.$$

课堂练习

写出二次型
$f(x_1,x_2,x_3)=x_1^2+2x_1x_2+4x_1x_3+x_2^2+6x_2x_3+3x_3^2$ 的矩阵，并求其秩.

答案: 矩阵为 $A=\begin{bmatrix}1 & 1 & 2\\ 1 & 1 & 3\\ 2 & 3 & 3\end{bmatrix}$, 秩为 $r(A)=3$.

二次型的矩阵表示

注 2

二次型与对称矩阵是一一对应的.

注 3

给定一个不对称的方阵 B, 列向量 $X=(x_1,\cdots,x_n)^T$, 则 $f(x_1,\cdots,x_n)=X^TBX$ 也为二次型, 我们可以找到一个对称阵 $A=\frac{1}{2}(B+B^T)$, 使得 $f=X^TAX$.

事实上, $X^TBX=\sum_{i=1}^{n}\sum_{j=1}^{n}b_{ij}x_ix_j$,

x_i^2 项的系数为 b_{ii},

x_ix_j $(i\neq j)$ 项的系数为 $b_{ij}+b_{ji}$,

令 $a_{ij}=\dfrac{b_{ij}+b_{ji}}{2}$, 即 $A=\frac{1}{2}(B+B^T)$, 则 $X^TAX=X^TBX$.

二次型的矩阵表示

例 2 (P159 例 1.1.2)

设 $A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 2 & 3 & 1 \end{bmatrix}$, $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, 写出二次型 $f(x_1, x_2, x_3) = X^T A X$ 的矩阵.

解. $B = \frac{1}{2}(A + A^T) = \begin{bmatrix} 1 & \frac{5}{2} & \frac{5}{2} \\ \frac{5}{2} & 2 & 2 \\ \frac{5}{2} & 2 & 1 \end{bmatrix}$ 即为 $f(x_1, x_2, x_3)$ 的矩阵.

二次型与曲线形状

回到最初的问题:

问题

给定二次型 $f(x_1, x_2, \cdots, x_n) = X^TAX$, 如何判断 $f(x_1, \cdots, x_n) = c$ 的形状?

由于 A 是对称矩阵, 一个简单的想法是将 A 正交对角化: $Q^{-1}AQ = Q^TAQ = \Lambda$,
此时, 令 $X = QY$, 则

$$X^TAX = (QY)^TA(QY) = Y^TQ^TAQY = Y^T\Lambda Y,$$

$Y^T\Lambda Y$ 为标准二次型, 我们较容易判断它的形状.
而 $Y^T\Lambda Y$ 与 X^TAX 为同一个二次型在不同直角坐标系下的表示.

可逆线性变换与二次型

课堂练习

已知 $A=\begin{bmatrix} 2 & -2 & 0 \\ -2 & 1 & -2 \\ 0 & -2 & 0 \end{bmatrix}$, $X=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

- 写出二次型 $f(x_1,x_2,x_3)=X^TAX$;
- 令 $P=\begin{bmatrix} 1 & 1 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$, $Y=\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$, 作可逆线性变换 $X=PY$, 将 f 写成关于 y_1,y_2,y_3 的二次型.
- 计算 P^TAP.

$f(x_1,x_2,x_3)=2x_1^2+x_2^2-4x_1x_2-4x_2x_3=2y_1^2-y_2^2+4y_3^2$.
$P^TAP=\mathrm{diag}\{2,-1,4\}$.

合同

定义 3

设 A, B 都是 n 阶方阵, 若存在可逆矩阵 P 使得 $B = P^T A P$, 则称 A 与 B合同(或 A 合同于 B), 记为 $A \simeq B$.

命题 1

合同关系满足以下性质:

(1) 反身性: $A \simeq A$

(2) 对称性: $A \simeq B$, 则 $B \simeq A$.

(3) 传递性: 当 $A \simeq B, B \simeq C$ 时, 有 $A \simeq C$.

即, 合同关系是一种等价关系.

合同

定理 1

任一实对称矩阵 A 都合同于一对角阵.

证明.

任一实对称矩阵 A, 存在正交矩阵 Q, 使得 $Q^TAQ = Q^{-1}AQ = \Lambda$ 为对角阵. $\square$

第六章 二次型

二次型化为标准形

我们知道, 对二次型 X^TAX, 若有可逆矩阵 P, 使得 $P^TAP = \Lambda$ 为对角阵, 则令 $X = PY$, $Y^T\Lambda Y$ 为标准二次型. 即, 要想将二次型化为标准形, 我们需要找可逆矩阵 P, 使得 P^TAP 为对角阵.

我们知道, 一个将实对称阵 A 对角化的办法是找正交矩阵 Q, 使得 $Q^TAQ = Q^{-1}AQ = \Lambda$. 具体步骤如下:

- 求 A 的特征值 $\lambda_1, \lambda_2, \cdots, \lambda_s$;
- 对每个特征值, 解方程 $(\lambda_i E - A)X = 0$, 取其基础解系, 并施密特 (Schmidt) 正交规范化;
- 将 n 个正交规范的特征向量组作为列向量, 即为 Q.

正交变换法

例 3 (P160 例 6.2.1)

将二次型 $f(x_1,x_2,x_3)=x_1x_2+x_1x_3-3x_2x_3$ 化为标准形.

解. 二次型 f 的矩阵为 $A=\begin{bmatrix}0 & \frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & 0 & -\frac{3}{2}\\ \frac{1}{2} & -\frac{3}{2} & 0\end{bmatrix}$.

A 的特征多项式为

$$|\lambda E-A|=\begin{vmatrix}\lambda & -\frac{1}{2} & -\frac{1}{2}\\ -\frac{1}{2} & \lambda & \frac{3}{2}\\ -\frac{1}{2} & \frac{3}{2} & \lambda\end{vmatrix}=(\lambda-\frac{3}{2})(\lambda-\frac{\sqrt{17}-3}{4})(\lambda+\frac{\sqrt{17}+3}{4}),$$

A 的特征值为 $\lambda_1=\frac{3}{2}$,$\lambda_2=\frac{\sqrt{17}-3}{4}$, $\lambda_3=-\frac{\sqrt{17}+3}{4}$

正交变换法

解续. 对 $\lambda_1=\frac{3}{2}$, 解齐次方程组 $(\frac{3}{2}E-A)X=0$,

$$(\frac{3}{2}E-A)=\begin{bmatrix}\frac{3}{2} & -\frac{1}{2} & -\frac{1}{2}\\ -\frac{1}{2} & \frac{3}{2} & \frac{3}{2}\\ -\frac{1}{2} & \frac{3}{2} & \frac{3}{2}\end{bmatrix}\to\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 1\\ 0 & 0 & 0\end{bmatrix}$$

其基础解系为 $\xi_1=[0,1,-1]^T$, 将其单位化得

$$\eta_1=[0,\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2}]^T.$$

对 $\lambda_2=\frac{\sqrt{17}-3}{4}$, 解齐次方程组 $(\frac{\sqrt{17}-3}{4}E-A)X=0$, 其基础解系为 $\xi_2=[\frac{\sqrt{17}+3}{2},1,1]^T$, 将其单位化得

$$\eta_2=[\sqrt{\frac{3+\sqrt{17}}{2\sqrt{17}}},\sqrt{\frac{2}{17+3\sqrt{17}}},\sqrt{\frac{2}{17+3\sqrt{17}}}]^T.$$

正交变换法

解续. 对 $\lambda_3=-\frac{\sqrt{17}+3}{4}$, 解齐次方程组 $(-\frac{\sqrt{17}+3}{4}E-A)X=0$, 其基础解系为 $\xi_3=[\frac{3-\sqrt{17}}{2},1,1]^T$, 将其单位化得

$$\eta_3=[-\sqrt{\frac{\sqrt{17}-3}{2\sqrt{17}}},\sqrt{\frac{2}{17-3\sqrt{17}}},\sqrt{\frac{2}{17-3\sqrt{17}}}]^T.$$

令 $P=[\eta_1,\eta_2,\eta_3]=\begin{bmatrix} 0 & \sqrt{\frac{3+\sqrt{17}}{2\sqrt{17}}} & -\sqrt{\frac{\sqrt{17}-3}{2\sqrt{17}}} \\ \frac{2}{2} & \sqrt{\frac{2}{17+3\sqrt{17}}} & \sqrt{\frac{2}{17-3\sqrt{17}}} \\ -\frac{2}{2} & \sqrt{\frac{2}{17+3\sqrt{17}}} & \sqrt{\frac{2}{17-3\sqrt{17}}} \end{bmatrix}$, 则 P 为正交矩阵, 且 $P^TAP=\Lambda=\begin{bmatrix} \frac{3}{2} & & \\ & \frac{\sqrt{17}-3}{4} & \\ & & -\frac{\sqrt{17}+3}{4} \end{bmatrix}$

正交变换法

解续. 作可逆线性变换 $X = PY$ 可得,

$$
\begin{aligned}
f(X) &= X^T A X = Y^T P^T A P Y = Y^T \Lambda Y \\
&= \frac{3}{2} y_1^2 + \frac{\sqrt{17} - 3}{4} y_2^2 - \frac{\sqrt{17} + 3}{4} y_3^2.
\end{aligned}
$$

此即 $f(X)$ 的标准形.

正交变换法

例 4 (P162 例 6.2.2)

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

解. 二次型 f 的矩阵为 $A=\begin{bmatrix}2&2&-2\\2&5&-4\\-2&-4&5\end{bmatrix}$.

A 的特征多项式为

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

A 有特征值 $\lambda_1=\lambda_2=1,\ \lambda_3=10.$

正交变换法

解续. 对 $\lambda_1 = \lambda_2 = 1$, 解齐次方程组 $(E - A)X = 0$,

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

得基础解系为 $\xi_1 = (0,1,1)^T$, $\xi_2 = (2,0,1)^T$. 将 ξ_1, ξ_2 正交化得

$\beta_1 = \xi_1 = (0,1,1)^T$,

$\beta_2 = \xi_2 - \frac{(\xi_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1 = (2, -\frac{1}{2}, \frac{1}{2})^T$

再将 β_1, β_2 单位化得

$\eta_1 = \frac{1}{\|\beta_1\|}\beta_1 = (0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})^T$,

$\eta_2 = \frac{1}{\|\beta_2\|}\beta_2 = (\frac{2\sqrt{2}}{3}, -\frac{\sqrt{2}}{6}, \frac{\sqrt{2}}{6})^T$

正交变换法

解续. 对 $\lambda_3 = 10$, 解齐次线性方程组 $(10E - A)X = 0$,

$$10E - A = \begin{bmatrix} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{bmatrix} \to \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

基础解系为 $\xi_3 = (\frac{1}{2}, 1, -1)^T$, 单位化得 $\eta_3 = (\frac{1}{3}, \frac{2}{3}, -\frac{2}{3})^T$.
令 $P = [\eta_1, \eta_2, \eta_3]$, 则 P 为正交矩阵, 且

$$P^TAP = \Lambda = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 10 \end{bmatrix}$$

作线性变换 $X = PY$, 则

$$f(X) = X^TAX = Y^T\Lambda Y = y_1^2 + y_2^2 + 10y_3^2.$$

正交变换法

注 4

我们注意到，如果只关心 f 的标准形，而不关心正交矩阵 P，我们只需要知道 A 的特征值即可.

定理 2

任一二次型 $f(X) = X^TAX$，其中 A 为 n 阶对称矩阵，则存在正交变换 $X = PY$，其中 P 为正交矩阵，使得

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

其中，$\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 A 的 n 个特征值.

正交变换法

例 5 (P164 例 6.2.3)

求多元函数 $f(x,y,z)=x^2+3y^2+4z^2+4xy+2xz-2yz$ 在约束条件 $x^2+y^2+z^2=1$ 下的最大值与最小值.

解. 令 $X=(x,y,z)^T$, $A=\begin{bmatrix}1 & 2 & 1\\ 2 & 3 & -1\\ 1 & -1 & 4\end{bmatrix}$, 则

$$f(x,y,z)=f(X)=X^TAX,$$

约束条件为 $\|X\|=1$.

分析: 对正交变换: $X=PY$, $\|X\|=\|Y\|$, 故可考虑将 A 正交相似对角化, 从而将二次型化为标准形, 此时约束条件等价于 $\|Y\|=1$.

正交变换法

解续. A 的特征值为 $\lambda_1 = 4$, $\lambda_2 = 2+\sqrt{7}$, $\lambda_3 = 2-\sqrt{7}$, 则存在正交矩阵 P 使得

$$P^TAP = \Lambda = \begin{bmatrix} 4 & & \\ & 2+\sqrt{7} & \\ & & 2-\sqrt{7} \end{bmatrix}.$$

令 $X = PY$, 记 $Y = (x_1, y_1, z_1)^T$, 则

$\|X\| = \sqrt{(PY, PY)} = \sqrt{Y^TP^TPY} = \sqrt{Y^TY} = \|Y\|$,

且 $f(X) = Y^T\Lambda Y = 4x_1^2 + (2+\sqrt{7})y_1^2 + (2-\sqrt{7})z_1^2$,

故当 $Y = (0,0,1)^T$ 时, $f(X)$ 有最小值 $(2-\sqrt{7})$; 当 $Y = (0,1,0)^T$ 时, $f(X)$ 有最大值 $2+\sqrt{7}$.

正交变换法

课堂练习

用正交变换法将

$$f(X)=f(x_1,x_2,x_3,x_4)=2x_1^2+2x_2^2+2x_3^2+2x_4^2+2x_1x_2+2x_3x_4$$

化为标准形, 并求其在限制 $\|X\|=1$ 下的最大值.

提示: $A=\begin{bmatrix}2&1&0&0\\1&2&0&0\\0&0&2&1\\0&0&1&2\end{bmatrix}$ 的特征值 $1,1,3,3$, 对应的特征向量 $\xi_1=\begin{bmatrix}0\\0\\-1\\1\end{bmatrix}$,$\xi_2=\begin{bmatrix}-1\\1\\0\\0\end{bmatrix}$, $\xi_3=\begin{bmatrix}0\\0\\1\\1\end{bmatrix}$, $\xi_4=\begin{bmatrix}1\\1\\0\\0\end{bmatrix}$

配方法

正交变换法有时计算较繁琐, 如果所作的可逆线性变换不要求正交, 我们还可以考虑配方法. 我们需要用公式

$$a^2+2ab+b^2=(a+b)^2,\quad (a+b)(a-b)=a^2-b^2,$$

消去交叉项, 将二次型化为标准形.

例 6 (P165 例 6.2.4)

用配方法将 $f(x_1,x_2,x_3)=x_1^2+x_3^2+x_1x_2$ 化为标准形.

分析: 可以考虑配方
$x_1^2+x_1x_2=x_1^2+x_1x_2+\frac{1}{4}x_2^2-\frac{1}{4}x_2^2=(x_1+\frac{1}{2}x_2)^2-\frac{1}{4}x_2^2$,
以去掉交叉项 x_1x_2.

配方法

解.

$$\begin{aligned} f(x_1,x_2,x_3) &= x_1^2+x_3^2+x_1x_2 \\ &= x_1^2+x_1x_2+\frac{1}{4}x_2^2-\frac{1}{4}x_2^2+x_3^2 \\ &= (x_1+\frac{1}{2}x_2)^2-\frac{1}{4}x_2^2+x_3^2. \end{aligned}$$

令 $\begin{cases} y_1 = x_1+\frac{1}{2}x_2 \\ y_2 = x_2 \\ y_3 = x_3 \end{cases}$，即 $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$，则

$$f(X) = y_1^2-\frac{1}{4}y_2^2+y_3^2.$$

配方法

例 7 (P165 例 6.2.5)

用配方法将 $f(x_1, x_2, x_3) = x_1x_2 + x_1x_3 - 3x_2x_3$ 化为标准形.

分析: 此二次型中没有平方项, 我们用平方差公式构造平方项, 令 $\begin{cases} x_1 = y_1 + y_2 \\ x_2 = y_1 - y_2 \end{cases}$, 则 $x_1x_2 = y_1^2 - y_2^2$.

解.

令 $\begin{cases} x_1 = y_1 + y_2 \\ x_2 = y_1 - y_2 \\ x_3 = y_3 \end{cases}$, 即 $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$, 则

$f(x_1, x_2, x_3) = y_1^2 - y_2^2 + (y_1 + y_2)y_3 - 3(y_1 - y_2)y_3$

$= y_1^2 - 2y_1y_3 - y_2^2 + 4y_2y_3$

配方法

解续. $$\begin{aligned} f(x_1,x_2,x_3) &= y_1^2 - 2y_1y_3 - y_2^2 + 4y_2y_3 \\ &= (y_1-y_3)^2 - y_3^2 - y_2^2 + 4y_2y_3 \\ &= (y_1-y_3)^2 - (y_2-2y_3)^2 + 3y_3^2 \end{aligned}$$

令 $\begin{cases} z_1 = y_1 - y_3 \\ z_2 = y_2 - 2y_3 \\ z_3 = y_3 \end{cases}$，即 $\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$，则

$$f = z_1^2 - z_2^2 + 3z_3^2.$$

配方法

在上面的例子中, 我们所作的变换为:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}, \quad \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix},$$

故 $$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 & 3 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}.$$

注意

在使用配方法时, 要注意, 所用的线性变换必须是可逆的.

配方法

课堂练习

将二次型 $f(X)=(x_1-x_2)^2+(x_2-x_3)^2+(x_1-x_3)^2$ 化为标准形, 并指出其秩为多少.

典型错误!!!

令 $\begin{cases} y_1=x_1-x_2 \\ y_2=x_2-x_3 \\ y_3=x_1-x_3 \end{cases}$, 则 $f=y_1^2+y_2^2+y_3^2$, 秩为 3.

因为使用的线性变换 $Y=PX$, 其中 $P=\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & -1 \end{bmatrix}$ 不可逆!

配方法

解. 令 $\begin{cases} y_1 = x_1 - x_2 \\ y_2 = x_2 - x_3 \\ y_3 = x_3 \end{cases}$，即 $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$，则

$$f = y_1^2 + y_2^2 + (y_1 + y_2)^2 = 2y_1^2 + 2y_1y_2 + 2y_2^2$$
$$= 2(y_1 + \frac{1}{2}y_2)^2 + \frac{3}{2}y_2^2$$

再令 $\begin{cases} z_1 = y_1 + \frac{1}{2}y_2 \\ z_2 = y_2 \\ z_3 = y_3 \end{cases}$，即 $\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$，

则 f 的标准形为 $f = 2z_1^2 + \dfrac{3}{2}z_2^2$，秩为 2.

配方法小结

配方法小结

- 若二次型有平方项, 通过配方, 使用 $a^2 + 2ab + b^2 = (a+b)^2$ 消去交叉项;
- 若无平方项只有交叉项时, 选一个交叉项, 如 x_1x_2, 令

$$\begin{cases} x_1 = y_1 + y_2 \\ x_2 = y_1 - y_2 \\ x_3 = y_3 \\ \quad \vdots \\ x_n = y_n \end{cases},$$

从而, 构造出平方项.

配方法小结

定理 3

对任一二次型 $f(X)=X^TAX$, A 为 n 阶对称矩阵, 通过配平处理, 都可以找到一个可逆线性替换 $X=CY$, 使得

$$f(X)=g(Y)=d_1y_1^2+d_2y_2^2+\cdots+d_ny_n^2.$$

推论 4

任意一个 n 阶对称矩阵都合同于对角阵.

配方法与正交变换法

回顾: $f(x_1, x_2, x_3) = x_1x_2 + x_1x_3 - 3x_2x_3$ 的标准形:

- 用正交变换法计算得到标准形:
  $f = \frac{3}{2}y_1^2 + \frac{\sqrt{17}-3}{4}y_2^2 - \frac{\sqrt{17}+3}{4}y_3^2$.
- 用配方法计算得到标准形: $f = z_1^2 - z_2^2 + 3z_3^2$.

注 5

二次型的标准形不唯一. (因为用的是相合变换.)

配方法与正交变换法

配方法与正交变换法比较:

(1) 用配方法得到的标准形的系数 $d_1, d_2, \ldots, d_n$ 不一定是二次型对应的矩阵 A 的特征值.
用正交变换法得到的标准形的系数 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 一定是二次型对应的实对称阵 A 的特征值.

(2) 配方法时, 可逆线性变换 $X = PY$, P 不一定是正交矩阵.

(3) 配方法适用于所有二次型 $f(X) = X^T AX$, 即 A 是复矩阵时也可用.
正交变换法只适用于实对称矩阵 A 对应的二次型 $f(X) = X^T AX$.

合同变换法

定理 5

对任意实对称矩阵 A, 存在初等矩阵 $P_1, \cdots, P_s$ 使得

$$P_s^T \cdots P_1^T A P_1 \cdots P_s \text{为对角阵}.$$

证明.

我们已经知道, 存在可逆矩阵 P 使得 $P^T AP$ 为对角阵, 而 P 可以分解为若干初等矩阵的乘积 $P = P_1 \cdots P_s$, 故

$$P^T AP = (P_1 \cdots P_s)^T A(P_1 \cdots P_s) = P_s^T \cdots P_1^T A P_1 \cdots P_s. \qquad \square$$

定义 4

设 P 为初等矩阵, 我们称变换 $A \to P^T AP$ 为对 A 作一次合同变换.

合同变换法

- 对初等矩阵 $P = P(i,j)$, $P^T = P$, $A \to P^TAP$ 为先将 A 的第 i, j 两行交换得 P^TA, 再将 P^TA 的 i, j 两列交换;
- 对初等矩阵 $P = P(i(\lambda))$, $P^T = P$, $A \to P^TAP$ 为先将 A 的第 i 行乘以 λ 得 P^TA, 再将 P^TA 的第 i 列乘以 λ;
- 对初等矩阵 $P = P(i,j(k))$, $P^T = P(j,i(k))$, $A \to P^TAP$ 为先将 A 的第 i 行的 k 倍加到第 j 行得 P^TA, 再将 P^TA 的第 i 列的 k 倍加到第 j 列.

总结: 对初等矩阵 P, $A \overset{\text{初等行变换}}{\longrightarrow} P^TA \overset{\text{相同的初等列变换}}{\longrightarrow} P^TAP$.

合同变换法

作合同变换:

- $\begin{bmatrix} A \\ E \end{bmatrix} \to \begin{bmatrix} P_s^T \cdots P_1^T A P_1 \cdots P_s \\ E P_1 \cdots P_s \end{bmatrix} = \begin{bmatrix} \Lambda \\ P \end{bmatrix}$
- $[A, E] \to [P_s^T \cdots P_1^T A P_1 \cdots P_s, P_s^T \cdots P_1^T E] = [\Lambda, P^T]$

合同变换法

例 8 (P168 例 6.2.6)

用合同变换法化 $f(X)=x_1x_2+x_1x_3-3x_2x_3$ 为标准形, 并写出所用可逆变换.

解. 二次型对应矩阵 $A=\begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{3}{2} \\ \frac{1}{2} & -\frac{3}{2} & 0 \end{bmatrix}$,

对 $\begin{bmatrix} A \\ E \end{bmatrix}$ 作合同变换, 得 $\left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & -\frac{1}{4} & 0 \\ 0 & 0 & 3 \\ \hline 1 & -\frac{1}{2} & 3 \\ 1 & \frac{1}{2} & -1 \\ 0 & 0 & 1 \end{array}\right]=\begin{bmatrix} P^TAP \\ P \end{bmatrix}$

合同变换法

解续. 作可逆线性变换 $X=PY=\begin{bmatrix}1 & -\frac{1}{2} & 3\\ 1 & \frac{1}{2} & -1\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}y_1\\ y_2\\ y_3\end{bmatrix}$, 则 $f(X)=y_1^2-\frac{1}{4}y_2^2+3y_3^2$.

注: 还可以继续作合同变换

$$\left[\begin{array}{ccc}1 & 0 & 0\\ 0 & -\frac{1}{4} & 0\\ 0 & 0 & 3\\ \hline 1 & -\frac{1}{2} & 3\\ 1 & \frac{1}{2} & -1\\ 0 & 0 & 1\end{array}\right] \xrightarrow{\text{第二行 }\times 2,\text{ 第二列 }\times 2} \left[\begin{array}{ccc}1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 3\\ \hline 1 & -1 & 3\\ 1 & 1 & -1\\ 0 & 0 & 1\end{array}\right] = \begin{bmatrix}P'^TAP'\\ P'\end{bmatrix}$$

作可逆线性变换 $X=P'Z$, 其中 $Z=(z_1,z_2,z_3)^T$, 可得标准形 $f(X)=z_1^2-z_2^2+3z_3^2$.

合同变换法

课堂练习

用合同变换法将 $f(X) = x_1^2 + x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3 + 3x_2x_3$ 化为标准形.

答案见学习指导 P105 例 5

第六章 二次型

惯性定理

我们知道, 二次型的标准形不唯一, 所以, 我们想要在这些不同的标准形中找到不变量.

我们发现, 标准形中所含的非零项数是相同的 (= 秩), 且正 (负) 项的个数不变.

本节我们只考虑实二次型.

惯性定理

对 n 元实二次型 $f(X) = X^TAX$, 经过可逆线性变换可化为标准形, 即存在可逆线性变换 $X = PY$, 使得

$$
\begin{aligned}
f(X) &= X^TAX = Y^T\Lambda Y \\
&= c_1y_1^2 + \cdots + c_sy_s^2 - d_1y_{s+1}^2 - \cdots - d_ty_{s+t}^2 + 0y_{s+t+1}^2 + \cdots + 0y_n^2 \\
&= c_1y_1^2 + \cdots + c_sy_s^2 - d_1y_{s+1}^2 - \cdots - d_ty_{s+t}^2
\end{aligned}
$$

其中, $c_i > 0, d_j > 0$, $i = 1, \cdots, s$, $j = 1, \cdots, t$, $s + t = r(A) \leq n$.
我们再作可逆线性变换

$$
\begin{cases}
z_i = \sqrt{c_i}y_i, & i = 1, \cdots, s, \\
z_{s+j} = \sqrt{d_j}y_{s+j}, & j = 1, \cdots, t \\
z_{s+t+m} = y_{s+t+m}, & m = 1, \cdots, n-r
\end{cases}
\quad 则
$$

$$
f = z_1^2 + \cdots + z_s^2 - z_{s+1}^2 - \cdots - z_{s+t}^2, \quad (s + t = r(A) \leq n).
$$

惯性定理

定义 5

对 n 元实二次型 $f(X) = X^T AX$, 经过可逆线性变换 $X = PZ$, 有

$$f(X) = z_1^2 + \cdots + z_s^2 - z_{s+1}^2 - \cdots - z_{s+t}^2, \quad (s + t = r(A) \leq n).$$

我们称其为 f 的规范形（normal form），称 s 为正惯性指标, t 为负惯性指标.

定理 6

任一实二次型都可以通过可逆线性变换化为规范形, 且规范形唯一. 即正, 负惯性指标是两个不变量.

规范形唯一的证明

规范形唯一的证明

不妨设二次型 $f(X)=X^TAX$, A 为 n 阶实对称阵, 经过可逆线性变换 $X=PY$ 与 $X=QZ$ 有规范形

$$\begin{aligned} f(X) &= y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_{p+q}^2 \\ &= z_1^2+\cdots+z_s^2-z_{s+1}^2-\cdots-z_{s+t}^2, \end{aligned}$$

容易知道, $p+q=s+t=r(A)$, 我们将证明 $p=s$.
我们将假设 $p>s$ 并得出矛盾.
记 $G=Q^{-1}P=(g_{ij})_{n\times n}$, 则其可逆且 $Z=GY$.

规范形唯一的证明

规范形唯一的证明续

考虑齐次线性方程组
$$\begin{cases} g_{11}y_1 + g_{12}y_2 + \cdots + g_{1n}y_n = 0 \\ \vdots \\ g_{s1}y_1 + g_{s2}y_2 + \cdots + g_{sn}y_n = 0 \\ y_{p+1} = 0 \\ \vdots \\ y_n = 0 \end{cases}$$
其有 n 个未知数, $s+n-p<n$ 个方程, 故有非零解. 取非零解 $Y_0=(a_1,a_2,\cdots,a_n)^T$, 显然
$a_{p+1}=a_{p+2}=\cdots=a_n=0$, 故
$a_1^2+\cdots+a_p^2-a_{p+1}^2-\cdots-a_{p+q}^2=a_1^2+\cdots+a_p^2>0.$

规范形唯一的证明

规范形唯一的证明续.

令 $Z_0 = GY_0 = (k_1, k_2, \cdots, k_n)^T$, 则

$$k_1^2+\cdots+k_s^2-k_{s+1}^2-\cdots-k_{s+t}^2 = a_1^2+\cdots+a_p^2-a_{p+1}^2-\cdots-a_{p+q}^2 > 0$$

但是, 由 Y_0 的选取可知, $k_1 = \cdots = k_s = 0$, 故

$$k_1^2 + \cdots + k_s^2 - k_{s+1}^2 - \cdots - k_{s+t}^2 = -k_{s+1}^2 - \cdots - k_{s+t}^2 \leq 0$$

我们得到矛盾.
故假设 $p > s$ 错误, 我们有 $p \leq s$. 类似可得, $s \leq p$. 从而 $p = s$. □

惯性定理

推论 7

任意实对称矩阵 A 合同于对角阵 $\begin{bmatrix} E_s & & \\ & -E_t & \\ & & 0 \end{bmatrix}$, 其中 $s+t=R(A)$.

注 6

设 A 为对称方阵. 则二次型 $f(X)=X^TAX$ 的正惯性指标 s 等于 A 的正特征值个数, 负惯性指标等于 A 的负特征值个数.

惯性定理

例 9

设 A 为 n 阶实对称阵, 二次型 $f(X) = X^TAX$.

- 若 $f(X) = x_1^2 + x_2^2 + \cdots + x_n^2$, 则 $f(X) \geq 0$ 恒成立.
- 若 $f(X)$ 经过可逆线性变换 $X = PY$ 可化为规范型 $f(X) = y_1^2 + y_2^2 + \cdots + y_n^2$, 则 $f(X) = X^TAX \geq 0$ 恒成立.
- 若 A 的特征值均为正, 则 $f(X) = X^TAX \geq 0$ 恒成立.
- 若 $f(X)$ 经过可逆线性变换 $X = PY$ 可化为规范型 $f(X) = -y_1^2 - y_2^2 - \cdots - y_n^2$, 则 $f(X) = X^TAX \leq 0$ 恒成立.
- 若 A 的特征值均为负, 则 $f(X) = X^TAX \leq 0$ 恒成立.

惯性定理

例 10

设 n 元二次型 $f(X)=x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_{s+t}^2$.

- 若 $X=(x_1,\cdots,x_s,0,\cdots,0)^T$, 则 $f(X)\geq 0$;
- 若 $X=(0,\cdots,0,x_{s+1},\cdots,x_{s+t},0,\cdots,0)^T$, 则 $f(X)\leq 0$;
- 若 $X=(0,\cdots,0,x_{s+t+1},\cdots,x_n)^T$, 则 $f(X)=0$.

即

- 当 $X\in\mathrm{span}\{e_1,\cdots,e_s\}$ 时, $f(X)\geq 0$;
- 当 $X\in\mathrm{span}\{e_{s+1},\cdots,e_{s+t}\}$ 时, $f(X)\leq 0$;
- 当 $X\in\mathrm{span}\{e_{s+t+1},\cdots,e_n\}$ 时, $f(X)=0$.

当 X 不属于以上三种情况时, 需要根据具体情况判断 $f(X)$ 的正负.

惯性定理

例 11

设 n 元二次型 $f(X)$ 通过可逆线性变换 $X=PY$ 有规范型 $f(X)=y_1^2+\cdots+y_s^2-y_{s+1}^2-\cdots-y_{s+t}^2$. 记 $P=[\alpha_1,\cdots,\alpha_n]$.

- 若 $Y=(y_1,\cdots,y_s,0,\cdots,0)^T$, $X=PY$, 则 $f(X)\geq 0$. 即当 $X\in\operatorname{span}\{\alpha_1,\cdots,\alpha_s\}$ 时, $f(X)\geq 0$.
- 若 $Y=(0,\cdots,0,y_{s+1},\cdots,y_{s+t},0,\cdots,0)^T$, $X=PY$, 则 $f(X)\leq 0$. 即 $X\in\operatorname{span}\{\alpha_{s+1},\cdots,\alpha_{s+t}\}$ 时, $f(X)\leq 0$.
- 若 $Y=(0,\cdots,0,y_{s+t+1},\cdots,y_n)^T$, $X=PY$, 则 $f(X)=0$. 即 $X\in\operatorname{span}\{\alpha_{s+t+1},\cdots,\alpha_n\}$ 时, $f(X)=0$.

当 X 不属于以上三种情况时, 需要根据具体情况判断 $f(X)$ 的正负.

惯性定理的几何直观

设 n 阶实对称阵 A 有 s 个正特征值 $\lambda_1, \cdots, \lambda_s$, t 个负特征值 $\lambda_{s+1}, \cdots, \lambda_{s+t}$, 及 $n-s-t$ 个零特征值. 对应的线性无关特征向量 $\alpha_1, \cdots, \alpha_n$. 则

- 当 $X \in \operatorname{span}\{\alpha_1, \cdots, \alpha_s\}$ 时, $f(X) \geq 0$.
- 当 $X \in \operatorname{span}\{\alpha_{s+1}, \cdots, \alpha_{s+t}\}$ 时, $f(X) \leq 0$.
- 当 $X \in \operatorname{span}\{\alpha_{s+t+1}, \cdots, \alpha_n\}$ 时, $f(X) = 0$.

通过可逆线性变换 $X = PY$, 我们将原来的标准基下的坐标 X, 变成了新基 P^{-1} 下的坐标 Y, 二次型的形式改变了, 但只是换了坐标, 本质并没有改变.

正定二次型

定义 6

设 n 元实二次型 $f(X)=X^TAX$, 如果对任意 $X\neq 0$,

- $f(X)>0$, 则称 $f(X)$ 为正定二次型;
- $f(X)\geq 0$, 则称 $f(X)$ 为半正定二次型;
- $f(X)<0$, 则称 $f(X)$ 为负定二次型;
- $f(X)\leq 0$, 则称 $f(X)$ 为半负定二次型;
- 如果 $f(X)$ 可正可负, 则称 $f(X)$ 为不定二次型.

正定二次型

例 12 (P170 例 6.3.1)

记 $X=(x_1,\cdots,x_n)^T\in\mathrm{R}^n$, 则 $f(X)=x_1^2+\cdots+x_n^2$ 是正定二次型; $g(X)=x_1^2+\cdots+x_r^2\ (r<n)$ 是半正定二次型; $h(X)=x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_{s+t}^2\ (s+t\leq n)$ 是不定二次型.

注 7

容易看出, 对 n 元实二次型 $f(X)$,

- f 负定当且仅当 $-f$ 正定.
- f 半负定当且仅当 $-f$ 半正定.

正定二次型

注 8

容易看出, 对 n 元二次型的规范形, 若 s, t 分别为正, 负惯性指标, 则

- 当 $s = n$ 时, 正定;
- 当 $s < n$, $t = 0$ 时, 半正定;
- 当 $t = n$ 时, 负定;
- 当 $s = 0$, $t < n$ 时, 半负定;
- 当 $s \geq 1$, $t \geq 1$ 时, 不定.

正定矩阵

定义 7

设 A 为 n 阶实矩阵, $f(X) = X^T AX$ （半）正定, 则称 A 为（半）正定矩阵.

可逆线性变换不改变二次型类型

定理 8

二次型经过可逆线性变换，其类型不变

证明.

设 $f(X) = X^TAX$, $X = PY$ 为可逆线性变换. X^TAX 与 $Y^T(P^TAP)Y$ 值域相同.
且 $X = PY = 0$ 当且仅当 $Y = 0$.
故二次型经过可逆线性变换，其类型不变. □

注 9

二次型的可逆线性变换对应着实对称矩阵的合同变换，故可通过对矩阵进行合同变换，化为标准形或规范形后，再判断二次型的类型.

正定矩阵性质

定理 9

对 n 阶实对称矩阵 A, $f(X) = X^TAX$, 以下命题等价:

1. A 正定;
2. A 的特征值全为正数;
3. $f(X)$ 的正惯性指标 $s = n$;
4. A 合同于单位阵 E;
5. 存在可逆矩阵 C 使得 $A = C^TC$.

思路: $[1] \Rightarrow [2] \Rightarrow [3] \Rightarrow [4] \Rightarrow [5] \Rightarrow [1]$.

推论 10

正定矩阵的行列式大于零.

半正定矩阵性质

定理 11

对 n 阶实对称矩阵 A, $\text{rank}(A)=r\le n$, $f(X)=X^TAX$, 以下命题等价:

1. A 半正定;
2. A 的特征值非负;
3. $f(X)$ 的正惯性指标 $s=r\le n$;
4. A 合同于对角阵 $\begin{bmatrix} E_r & 0 \\ 0 & 0 \end{bmatrix}$;
5. 存在 n 阶矩阵 C 使得 $A=C^TC$.

Hurwitz 定理

定义 8 (顺序主子式)

设 $A=(a_{ij})_n$, A 的位于左上角的子式

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

称为 A 的 k 阶顺序主子式.

定理 12 (赫尔维茨定理)

n 阶实对称矩阵 A 为正定矩阵的充要条件是 A 的各顺序主子式大于零.

Hurwitz 定理

必要性证明.

若 A 正定, 令 $A_k = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix}$, 则对任意 $X_k = (x_1, \cdots, x_k)^T \neq 0$,

$$X_k^T A_k X_K = X^T A X > 0,$$

其中 $X = (x_1, \cdots, x_k, 0, \cdots, 0)^T$. 故 A_k 正定, 从而 $\det A_k > 0$. $\square$

Hurwitz 定理

我们用归纳法证明充分性.

充分性证明

对 $n=1$ 时, 显然. 设对 $n-1$ 阶方阵命题成立, 下面将证明对 n 阶方阵命题成立.

对 n 阶矩阵 A 分块 $A=\begin{bmatrix} A_{n-1} & \alpha \\ \alpha^T & a_{nn} \end{bmatrix}$, 其中 A_{n-1} 的各顺序主子式均大于 0, 由归纳假设, A_{n-1} 正定, 故可逆. 对 A 作合同变换, 令 $P=\begin{bmatrix} E_{n-1} & -A_{n-1}^{-1}\alpha \\ 0 & 1 \end{bmatrix}$, 则

$$P^TAP=\begin{bmatrix} A_{n-1} & 0 \\ 0 & a_{nn}-\alpha^TA_{n-1}^{-1}\alpha \end{bmatrix}.$$

注意到 $\det P=1$, 故 $\det A=(a_{nn}-\alpha^TA_{n-1}^{-1}\alpha)\det A_{n-1}$. A_{n-1} 正定, 故 $\det A_{n-1}>0$, 故 $a_{nn}-\alpha^TA_{n-1}^{-1}\alpha>0$.

Hurwitz 定理

充分性续.

对任意非零向量 $X=(x_1,\cdots,x_n)^T$, 记 $X_1=(x_1,\cdots,x_{n-1})^T$

$$
\begin{aligned}
&X^T\begin{bmatrix} A_{n-1} & 0 \\ 0 & a_{nn}-\alpha^T A_{n-1}^{-1}\alpha \end{bmatrix} X \\
=&X_1^T A_{n-1} X_1 + x_n^2(a_{nn}-\alpha^T A_{n-1}^{-1}\alpha) > 0,
\end{aligned}
$$

故 $\begin{bmatrix} A_{n-1} & 0 \\ 0 & a_{nn}-\alpha^T A_{n-1}^{-1}\alpha \end{bmatrix}$ 正定, 从而 A 正定. □

Hurwitz 定理

推论 13

A 为 n 阶实对称矩阵, 则 $f(X)=X^TAX$ 负定的充要条件是 A 的奇数阶顺序主子式小于零, 偶数阶顺序主子式大于零.

证明.

$f(X)$ 负定当且仅当 $-f(X)=X^T(-A)X$ 正定, 当且仅当 $-A$ 的顺序主子式均大于零, 当且仅当 A 的奇数阶顺序主子式小于零, 偶数阶顺序主子式大于零. $\square$

Hurwitz 定理

例 13 (P174 例 6.3.2)

求 t 的值, 使得
$f(x_1, x_2, x_3) = tx_1^2 + 2x_2^2 + tx_3^3 + 2x_1x_2 + 4x_1x_3 + 2x_2x_3$ 正定.

解. 二次型的矩阵为 $A = \begin{bmatrix} t & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & t \end{bmatrix}$.

二次型正定, 当且仅当 A 的顺序主子式均 > 0, 即:

$$t > 0; \begin{vmatrix} t & 1 \\ 1 & 2 \end{vmatrix} = 2t - 1 > 0; \begin{vmatrix} t & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & t \end{vmatrix} = 2(t-2)(t+1) > 0.$$

这三个不等式的解的公共部分为 $t > 2$.
故当且仅当 $t > 2$ 时, f 为正定二次型

正定与特征值

例 14 (P175 例 6.3.3)

设 A 是 n 阶正定矩阵, E 为 n 阶单位矩阵, 证明 $|A+E|>1$.

证明. 已知 A 为 n 阶正定矩阵, 故其特征值 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 均为正. 从而 $A+E$ 的特征值为 $\lambda_1+1, \lambda_2+1, \cdots, \lambda_n+1$, 它们均 >1. 故

$$\det(A+E)=(\lambda_1+1)(\lambda_2+1)\cdots(\lambda_n+1)>1.$$

正定与特征值

课堂练习

已知 A 为实对称矩阵且 $6A^2-7A+2E=0$. 证明: A 正定.

提示: 考虑 A 的特征值.

证明. 设 λ 为 A 的特征值, 则 $6\lambda^2-7\lambda+2$ 为 $6A^2-7A+2E=0$ 的特征值, 故

$$6\lambda^2-7\lambda+2=0.$$

故 $\lambda=\frac{2}{3}$ 或 $\frac{1}{2}$.
A 的特征值均大于零, 故 A 正定.

正定二次型

课堂练习

已知 A 正定, 证明 A^T, A^{-1}, A^* 均正定.

证明. A 正定, 故对称, 从而 $A^T = A$ 正定.

A 正定, 故 $\det A > 0$, 从而 A 可逆. 由 $AA^{-1} = E$ 可得, $(A^{-1})^T A^T = E^T$, 即 $(A^{-1})^T A = E$. 故 $(A^{-1})^T = A^{-1}$, 即 A^{-1} 对称. 又因 A 正定, 故特征值 $\lambda_1, \cdots, \lambda_n$ 均为正, 故 A^{-1} 的特征值 $\lambda_1^{-1}, \cdots, \lambda_n^{-1}$ 均为正, 从而正定.

$A^* = (\det A)A^{-1}$ 故对称且特征值 $\frac{\det}{\lambda_1}, \cdots, \frac{\det A}{\lambda_n}$ 均为正, 从而正定.

正定二次型

课堂练习

已知 A, B 正定, 证明 $A+B$ 正定.

证明. A, B 正定, 故对称, 从而 $A+B$ 对称. 对任意非零向量 $X \in \mathbb{R}^n$,

$$X^T(A+B)X = X^TAX + X^TBX$$

由于 A, B 均正定, 故 $X^TAX > 0$, $X^TBX > 0$, 从而

$$X^T(A+B)X > 0,$$

故 $A+B$ 正定.

正定二次型

课堂练习

已知 A, B 分别为 m 和 n 阶正定矩阵, 证明 $C=\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ 正定.

证明. A, B 正定, 故对称, 从而 $C=\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ 对称. 任意非零的 $Z\in\mathbb{R}^{m+n}$, 将其写成 $Z=\begin{bmatrix} X \\ Y \end{bmatrix}$, 其中 $X\in\mathbb{R}^m$, $Y=\mathbb{R}^n$, 则 X, Y 不同时为零.

$$Z^TCZ=X^TAX+Y^TBY>0.$$

故 C 正定.

正定二次型

小结: 常用判断二次型正定的办法:

1. 定义,
2. 对称 + 特征值大于零,
3. Hurwitz 定理.

第六章 二次型

多元函数的极值

回顾: 对光滑的一元函数, 极值点为临界点: 即 $f'(x) = 0$ 的点. 且对临界点 x_0,

- 若 $f''(x_0) > 0$, 则 x_0 为极小值点;
- 若 $f''(x_0) < 0$, 则 x_0 为极大值点;
- 若 $f''(x_0) = 0$, 我们无法据此判断 x_0 是否为极值点.

问题

对光滑的多元函数, 我们怎么找极值和极值点?

多元函数的极值

记 $X=(x_1,x_2,\cdots,x_n)^T\in\mathbb{R}^n$, 设 n 元函数 $y=f(X)$ 在点 $X_0=(x_1^0,x_2^0,\cdots,x_n^0)^T$ 的邻域有二阶连续偏导数

$$f_{ij}(X)=\frac{\partial^2 f}{\partial x_i\partial x_j},\quad (i,j=1,2,\cdots,n).$$

$y=f(X)$ 的梯度: $\quad \operatorname{grad}f(X)=[\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2},\cdots,\frac{\partial f}{\partial x_n}]^T,$

多元函数的极值

$y = f(X)$ 的二阶偏导数构成的矩阵 (海森（Hessian）矩阵)

$$H(X) = \begin{bmatrix} f_{11}(X) & f_{12}(X) & \cdots & f_{1n}(X) \\ f_{21}(X) & f_{22}(X) & \cdots & f_{2n}(X) \\ \vdots & \vdots & & \vdots \\ f_{n1}(X) & f_{n2}(X) & \cdots & f_{nn}(X) \end{bmatrix}$$

注意到, $f_{ij}(X) = f_{ji}(X)$, 故 Hessian 矩阵 $H(X)$ 为对称矩阵.

多元函数的极值

类似于一元函数，对光滑的多元函数，其极值点 X_0 为临界点，即若 X_0 为极值点，则 $\mathrm{grad} f(X_0) = 0$.

证明. 我们仅证明 X_0 为极小值点的情况.
对 $i = 1, \cdots, n$, 均有

$$\frac{\partial f}{\partial x_i}(X_0) = \lim_{h \to 0+} \frac{f(X_0 + he_i) - f(X_0)}{h} \geq 0$$

$$\frac{\partial f}{\partial x_i}(X_0) = \lim_{h \to 0+} \frac{f(X_0 - he_i) - f(X_0)}{-h} \leq 0$$

故

$$\frac{\partial f}{\partial x_i}(X_0) = 0.$$

故要找光滑的多元函数的极值点，应该首先找临界点，再判断临界点是否为极值点.

多元函数的极值

类似于一元函数，对光滑的多元函数，要判断临界点 X_0 是否为极值点，我们需要对 X_0 的小邻域中的任意点 X，判断

$$f(X) - f(X_0)$$

的正负.

我们构造辅助函数: $\varphi(t) = f(X_0 + t\Delta X) - f(X_0)$, 其中 $\Delta X = X - X_0 = (h_1, h_2, \cdots, h_n)^T$, 则

$$f(X) - f(X_0) = \varphi(1) - \varphi(0)$$

$\varphi'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(X_0 + t\Delta X)h_i = (\mathrm{grad} f(X_0 + t\Delta X), \Delta X)$,

$\varphi'(0) = 0.$

$\varphi''(t) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(X_0 + t\Delta X)h_j h_i$

$= (\Delta X)^T H(X_0 + t\Delta X)\Delta X.$

多元函数的极值

运用二阶 (拉格朗日余项) 的泰勒公式, 存在 $\theta \in [0,1]$, 使得

$$\varphi(1)-\varphi(0)=\frac{1}{2}\varphi''(\theta)=\frac{1}{2}(\Delta X)^T H(X_0+\theta\Delta X)\Delta X$$

如果 $H(X_0+\theta\Delta X)$ 的正定/ 负定/ 不定, 则可以判断 $f(X)-f(X_0)$ 的正负.

而当 $H(X_0)$ 正定/ 负定/ 不定, 且 ΔX 充分小时, $H(X_0+\theta\Delta X)$ 也正定/ 负定/ 不定.

多元函数的极值

定理 14

设 $y = f(X)$ 在点 X_0 的邻域内有定义且有连续的二阶偏导数，且 $\mathrm{grad} f(X_0) = 0$，则

1. 当 $H(X_0)$ 正定时，$y = f(X)$ 在 X_0 取极小值;
2. 当 $H(X_0)$ 负定时，$y = f(X)$ 在 X_0 取极大值;
3. 当 $H(X_0)$ 不定时，X_0 不是 f 的极值点;
4. 当 $H(X_0)$ 半正定或半负定时，不能据此确定 X_0 是否为极值点.

多元函数的极值

证明.

存在 $\theta \in [0,1]$ 使得,

$$f(X) - f(X_0) = \frac{1}{2}\varphi''(\theta) = \frac{1}{2}(\Delta X)^T H(X_0 + \theta \Delta X)\Delta X$$

f 有连续的二阶偏导数, 故当 X 在 X_0 的充分小邻域内时,

1. 当 $H(X_0)$ 正定时,$H(X_0 + \theta\Delta X)$ 也正定,
   $f(X) - f(X_0) > 0, y = f(X)$ 在 X_0 取极小值.
2. 当 $H(X_0)$ 负定时,$H(X_0 + \theta\Delta X)$ 也负定,
   $f(X) - f(X_0) < 0, y = f(X)$ 在 X_0 取极大值.
3. 当 $H(X_0)$ 不定时,$H(X_0 + \theta\Delta X)$ 也不定, 存在 X 使得 $f(X) - f(X_0) > 0$, 也存在 X 使得 $f(X) - f(X_0) < 0, X_0$ 不是 f 的极值点.

□

多元函数的极值

当 $H(X_0)$ 半正定或半负定时, 不能确定 X_0 是否为极值点.
如:

a) $f(x_1,x_2)=x_1^2+x_2^4$, 在点 $X_0=(0,0)$, $\mathrm{grad}f(X_0)=0$,
$H(X_0)=\begin{bmatrix}2&0\\0&0\end{bmatrix}$ 半正定, X_0 为极小值点;

b) $f(x_1,x_2)=x_1^2-x_2^4$, 在点 $X_0=(0,0)$, $\mathrm{grad}f(X_0)=0$,
$H(X_0)=\begin{bmatrix}2&0\\0&0\end{bmatrix}$ 半正定, X_0 不为极值点;

c) $f(x_1,x_2)=-x_1^4-x_2^4$, 在点 $X_0=(0,0)$, $\mathrm{grad}f(X_0)=0$,
$H(X_0)=\begin{bmatrix}0&0\\0&0\end{bmatrix}$ 半正/半负定, X_0 为极大值点.

多元函数的极值

例 15 (P176 例 6.4.1)

求函数 $f(x,y,z)=2x^4-x+y^2+z^2-yz+y+z$ 的极值.

解. $\mathrm{grad}f(x,y,z)=(8x^3-1,2y-z+1,2z-y+1)^T$,
故临界点为 $X_0=(\frac{1}{2},-1,-1)^T$.

$$H(X_0)=\begin{bmatrix}24x^2 & 0 & 0\\ 0 & 2 & -1\\ 0 & -1 & 2\end{bmatrix}\Big|_{X_0}=\begin{bmatrix}6 & 0 & 0\\ 0 & 2 & -1\\ 0 & -1 & 2\end{bmatrix}.$$

$H(X_0)$ 正定, 故 X_0 为极小值点, $f(X)$ 在 X_0 取极小值

$$f(X_0)=-\frac{11}{8}.$$

最小二乘法

在大规模计算与工程计算中，常涉及用一个给定的模型，如线性函数，多项式或三角多项式等，去拟合一大批测量数据。如由于测量误差或实验误差的存在，使得得到的往往是一个不相容超定方程组，这时可以考虑用最小二乘原理求超定方程组的最佳逼近解. 如高斯在计算谷神星轨道时，就使用了最小二乘法.

超定方程组的最小二乘解

设 $A=(a_{ij})_{m\times n}$, $b\in\mathbb{R}^m$ 是已知数据, $X\in\mathbb{R}^n$ 是未知向量(待求的模型参数), $m>>n$, 按照模型应该有

$$AX=b,$$

但由于误差的存在, 该超定方程组一般是无解的, 即 $R(A)+1=R(A,b)$.
目标: 找合适的参数 X, 使得模型和数据之间的误差较小. 如我们可以考虑选择适当的参数 X, 使得 AX 和 b 之间的距离 $\|AX-b\|$ 最小. 最小化 $\|AX-b\|$, 等价于最小化 $\|AX-b\|^2$.

定义 9

使得 $\|AX-b\|^2$ 最小的点 X_0, 称为超定方程组 $AX=b$ 的最小二乘解.

超定方程组的最小二乘解

考虑 $f(X)=\|AX-b\|^2=(AX-b,AX-b)$
$f(X)=(AX-b)^T(AX-b)=(X^TA^T-b^T)(AX-b)$
$=X^TA^TAX-X^TA^Tb-b^TAX+b^Tb$
$=X^T(A^TA)X-2X^TA^Tb+b^Tb.$

$$\mathrm{grad} f(X)=2A^TAX-2A^Tb,$$

故 $f(X)$ 的极值点应满足 $\mathrm{grad} f(X)=0$, 即满足方程组

$$A^TAX=A^Tb.$$

超定方程组的最小二乘解

定理 15

对任意 $A_{m\times n}$, $b\in\mathbb{R}^n$, 方程组 $A^TAX=A^Tb$ 有解.

分析: 我们只需要证明 $r(A^TA)=r(A^TA,A^Tb)$.
我们注意到: $r(A^TA)\le r(A^TA,A^Tb)=r(A^T[A,b])\le r(A^T)$,
而 $r(A^T)=r(A)=r(A^TA)$.

证明.

我们首先证明 $r(A)=r(A^TA)$.
事实上, $A^TAX=0\Rightarrow X^TA^TAX=0\Rightarrow AX=0$,
而 $AX=0\Rightarrow A^TAX=0$,
故 $A^TAX=0$ 与 $AX=0$ 同解. 从而 $r(A)=r(A^TA)$.
故
$r(A^TA)\le r(A^TA,A^Tb)=r(A^T(A,b))\le r(A^T)=r(A^TA)$,
故 $r(A^TA)=r(A^TA,A^Tb)$, 方程组 $A^TAX=A^Tb$ 有解. $\square$

超定方程组的最小二乘解

定理 16

设 X_0 是方程组 $A^TAX = A^Tb$ 的解，则 X_0 是超定方程组 $AX = b$ 的最小二乘解.

我们需要证明，对函数 $f(X) = \|AX - b\|^2$，任意 $X \in \mathbb{R}^n$，$f(X) \geq f(X_0)$.

证明.

任意 $X \in \mathbb{R}^n$，$f(X) = (AX - b, AX - b)$
$= (AX_0 - b + AX - AX_0, AX_0 - b + AX - AX_0)$
$= (AX_0 - b, AX_0 - b) + 2(AX_0 - b, AX - AX_0) + (AX - AX_0, AX - AX_0)$
$\geq f(X_0) + 2(AX_0 - b, AX - AX_0)$
$= f(X_0) + 2(X - X_0)^TA^T(AX_0 - b) = f(X_0)$. □

超定方程组的最小二乘解

定义 10

我们称恰定方程组

$$A^TAX = A^Tb$$

为超定方程组 $AX = b$ 的正则方程组.

小结

要求超定方程组 $AX = b$ 的最小二乘解, 只需求正则方程组

$$A^TAX = A^Tb$$

的解. 特别的, 如果 A 列满秩, 则 A^TA 可逆, 最小二乘解为

$$X = (A^TA)^{-1}A^Tb.$$

最小二乘问题的几何观点

我们要求超定方程 $AX=b$ 的最小二乘解, 其中 A 为 $m\times n$ 矩阵且 $m>>n$, 即求最小化 $\|AX-b\|$ 的 X.
将 A 按照列分块 $A=[\alpha_1,\alpha_2,\cdots,\alpha_n]$, 则
$AX=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n$ 为 $\alpha_1,\cdots,\alpha_n$ 的线性组合, 即 AX 表示 $\mathrm{col}A\subset\mathbb{R}^m$ 中的向量. $\|AX-b\|$ 即 $\mathrm{col}A$ 中的向量 AX 与空间 $\mathbb{R}^m$ 中的向量 b 之间的距离. 要使得距离 $\|AX-b\|$ 最小, 这等价于 $b-AX$ 正交 (垂直) 于 $\mathrm{col}A$. 此即

$$A^T(b-AX)=0,\quad \text{即正则方程组} A^TAX=A^Tb$$

特别的, 当 A 列满秩时, $\alpha_1,\cdots,\alpha_n$ 线性无关, 构成 $\mathrm{col}A$ 的一组基. 此时, b 在 $\mathrm{col}A$ 中的投影在基 $\alpha_1,\cdots,\alpha_n$ 下有唯一的坐标 X, 即为 $A^TAX=A^Tb$ 的唯一解 $X=(A^TA)^{-1}A^Tb$.
当 A 不列满秩时, $\alpha_1,\cdots,\alpha_n$ 线性相关, b 在 $\mathrm{col}A$ 中的投影可写成 $\alpha_1,\cdots,\alpha_n$ 的线性组合, 但写法不唯一; 此时 $A^TAX=A^Tb$ 有无穷多解.

最小二乘问题的几何观点

我们也可以求出 b 在空间 $\mathrm{col}A$ 中的投影, 并利用投影直接求最小二乘解: 即求使得 AX 为 b 在空间 $\mathrm{col}A$ 中的投影的 X.

首先, 求 $\mathrm{col}A$ 的一组标准正交基 $\varepsilon_1, \cdots, \varepsilon_r$.

方法: Schmidt 正交规范化 A 的列向量 $\alpha_1, \cdots, \alpha_n$. 当 $\alpha_1, \cdots, \alpha_n$ 线性无关时, 直接得标准正交基; 当 $\alpha_1, \cdots, \alpha_n$ 线性相关时, 得到 $\mathrm{col}A$ 一组标准正交基及若干个零向量.

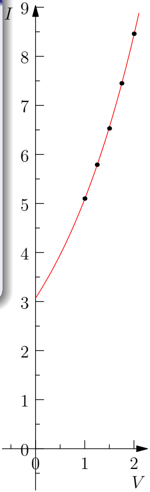
最小二乘法

例 16 (P 178 例 6.4.2)

再某电路实验中, 测得电压 V 与电流 I 的一组数据如下:

V_i	1.00	1.25	1.50	1.75	2.00
I_i	5.10	5.79	6.53	7.45	8.46

试用最小二乘法, 找出电流与电压的函数关系.



分析: 首先将数据画出, 猜测其关系, 再拟合. 猜测其为指数关系: $I = ae^{bV}$. 即:

$$\ln I = \ln a + bV.$$

令 $y = \ln I$, $\alpha = \ln a$, 则 $y = \alpha + bV$. 我们需要求出参数 α, b.

最小二乘法

解. 设 $I = ae^{bV}$, 则

$$\ln I = \ln a + bV,$$

令 $y = \ln I$, $\alpha = \ln a$, 我们有关系

$$y = \alpha + bV.$$

将数据带入, 得超定方程组:

$$\begin{cases} 1.629 = \alpha + 1.00b, \\ 1.756 = \alpha + 1.25b, \\ 1.876 = \alpha + 1.50b, \\ 2.008 = \alpha + 1.75b, \\ 2.135 = \alpha + 2.00b, \end{cases}, \quad \text{即} \begin{bmatrix} 1 & 1.00 \\ 1 & 1.25 \\ 1 & 1.50 \\ 1 & 1.75 \\ 1 & 2.00 \end{bmatrix} \begin{bmatrix} \alpha \\ b \end{bmatrix} = \begin{bmatrix} 1.629 \\ 1.756 \\ 1.876 \\ 2.008 \\ 2.135 \end{bmatrix}$$

最小二乘法

解续. 对应的正则方程组为

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1.00 & 1.25 & 1.50 & 1.75 & 2.00 \end{bmatrix} \begin{bmatrix} 1 & 1.00 \\ 1 & 1.25 \\ 1 & 1.50 \\ 1 & 1.75 \\ 1 & 2.00 \end{bmatrix} \begin{bmatrix} \alpha \\ b \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1.00 & 1.25 & 1.50 & 1.75 & 2.00 \end{bmatrix} \begin{bmatrix} 1.629 \\ 1.756 \\ 1.876 \\ 2.008 \\ 2.135 \end{bmatrix},$$

即

$$\begin{bmatrix} 5 & 7.5 \\ 7.5 & 11.875 \end{bmatrix} \begin{bmatrix} \alpha \\ b \end{bmatrix} = \begin{bmatrix} 9.404 \\ 14.422 \end{bmatrix}$$

最小二乘法

解续. 解得

$$\alpha = 1.1224, \quad b = 0.5056.$$

所以, $a = e^{\alpha} = 7.0722$,

$$I = 3.0722e^{0.5056\,V}.$$

三维空间中常见的二次曲面

1 椭球面 (Ellipsoid)/球面 (Sphere):

$$\frac{x_1^2}{a^2}+\frac{x_2^2}{b^2}+\frac{x_3^2}{c^2}=1, \quad (a, b, c>0)$$

特别的, 当 $a=b=c=R$ 时, 是半径为 R 的球面.

2 单叶双曲面 (Hyperboloid of one sheet)

$$\frac{x_1^2}{a^2}+\frac{x_2^2}{b^2}-\frac{x_3^2}{c^2}=1, \quad (a, b, c>0)$$

3 双叶双曲面 (Hyperboloid of two sheets)

$$\frac{x_1^2}{a^2}-\frac{x_2^2}{b^2}-\frac{x_3^2}{c^2}=1, \quad (a, b, c>0)$$

三维空间中常见的二次曲面

4 二次锥面 (cone of the second order)

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} - \frac{x_3^2}{c^2} = 0, \quad (a, b, c > 0)$$

5 椭圆抛物面 (Elliptic paraboloid)

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = x_3, \quad (a, b > 0)$$

6 双曲抛物面 (马鞍面)(Hyperbolic paraboloid)

$$\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} = x_3, \quad (a, b > 0)$$

三维空间中常见的二次曲面

7 柱面:

7.1 椭圆柱面 (Elliptic cylinder)/圆柱面 (Cylindrical surface)

$$\frac{x_1^2}{a^2}+\frac{x_2^2}{b^2}=1,\quad (a,b>0),$$

特别的，当 $a=b$ 时，为圆柱面.

7.2 双曲柱面 (Hyperbolic cylinder)

$$\frac{x_1^2}{a^2}-\frac{x_2^2}{b^2}=1,\quad (a,b>0),$$

7.3 抛物柱面 (Parabolic cylinder)

$$x_1^2=2px_2,\quad (p>0).$$

三维空间中的二次曲面

我们知道, 三维空间中的二次曲面, 对应着一个三元二次方程:

$$a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3 \\ + b_1x_1 + b_2x_2 + b_3x_3 + c = 0$$

记 $A = (a_{ij})_{3\times 3}$, $b = (b_1, b_2, b_3)^T$, $X = (x_1, x_2, x_3)^T$, 则上述方程可以写成

$$X^TAX + b^TX + c = 0.$$

目标

将方程化成较简单的形式, 以判断其图像的形状.

三维空间中的二次曲面

方程:

$$X^TAX + b^TX + c = 0.$$

A 对称, 可以考虑对角化: 即有正交阵 P, 使得 $P^TAP = \Lambda$.
令 $X = PY$, 则方程化为:

$$Y^T\Lambda Y + b^TPY + c = 0.$$

即有形式:

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \lambda_3 y_3^2 + d_1 y_1 + d_2 y_2 + d_3 y_3 + c = 0.$$

对 $\lambda_i \neq 0$, 且 $d_i \neq 0$ 的项, 通过配方变为 $\lambda_i(y_i + \frac{d_i}{2\lambda_i})^2$, 将上式化简, 再通过平移变换, 可化为几种常见曲面的形式.
注: 正交变换, 平移, 均不改变曲面形状.

三维空间中的二次曲面

例 17 (P181 例 6.4.3)

判断 $x_1^2+2x_2^2+2x_3^2-4x_2x_3+2x_1-\sqrt{2}x_2-\sqrt{2}x_3+6=0$ 的形状.

解. 令 $A=\begin{bmatrix}1&0&0\\0&2&-2\\0&-2&2\end{bmatrix}$, $b=\begin{bmatrix}2\\-\sqrt{2}\\-\sqrt{2}\end{bmatrix}$, $X=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$, 则原方程可以写为

$$X^TAX+b^TX+6=0.$$

将 A 正交相似对角化:
A 的特征值为 $\lambda_1=0,\ \lambda_2=1,\ \lambda_3=-4$.
属于 $\lambda_1=0$ 的特征向量单位化后为: $\xi_1=(0,\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2})^T$,
属于 $\lambda_2=1$ 的特征向量单位化后为: $\xi_2=(1,0,0)^T$,
属于 $\lambda_3=-4$ 的特征向量单位化后为: $\xi_3=(0,\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2})^T$,

三维空间中的二次曲面

解续. 令 $P=[\xi_1,\xi_2,\xi_3]=\begin{bmatrix} 0 & 1 & 0 \\ \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} \end{bmatrix}$, 则

$P^TAP=\Lambda=\begin{bmatrix} 0 & & \\ & 1 & \\ & & 4 \end{bmatrix}$

作正交变换 $X=PY$, 其中 $Y=[y_1,y_2,y_3]^T$, 则

$$Y^T\Lambda Y+b^TPY+6=0,$$

即

$$y_2^2+4y_3^2-2y_1+2y_2+6=0,$$

$$(y_2+1)^2+4y_3^2=2(y_1-\frac{5}{2}).$$

三维空间中的二次曲面

解续. 作平移变换 $\begin{cases} z_1 = y_1 - \frac{5}{2}, \\ z_2 = y_2 + 1, \\ z_3 = y_3, \end{cases}$ 则

$$\frac{z_2^2}{2} + \frac{z_3^2}{1/2} = z_1,$$

这是一个椭圆抛物面.