DOI:10.13718/j.cnki.xsxb.2019.09.002

条件极值的充分条件与一类椭圆方程[①]

张 驰[1], 程 功[2], 钟海全[3], 黄华飞[2]

1. 成都长益西联软件有限公司 智能订货平台部,成都 610200;2. 广州华立科技职业学院 城建学院,广州 511325;
3. 西南石油大学 油气藏地质及开发工程国家重点实验室,成都 610500

摘要: 通过分析无条件极值与条件极值的充分条件具有不同判别矩阵的原因,推导出条件极值充分条件的判别方法,得出其自变量增量间的关系式,得到了多维多约束状态下条件极值充分条件的一种更精确的判别矩阵,并举正反例说明判别驻点时可能出现的情况. 有助于理解两种充分条件的关联及差别,提供了一种寻找更精确的条件极值充分条件的判别矩阵的方法.

关 键 词: 多元函数;条件极值;充分条件;Hesse 矩阵;椭圆方程

中图分类号: O172.1 **文献标志码:** A **文章编号:** 1000-5471(2019)09-0006-09

在很多数学分析的教材中,对于多元函数无条件极值的充分必要条件以及用拉格朗日乘数法解条件极值的必要条件都有详尽的阐述;但对于条件极值的充分条件,则没有详细描述. 本文根据文献[1]中的例子,提出一类椭圆方程,说明采用 Hesse 矩阵判别条件极值的驻点是错误的,接着阐述了为什么采用 Hesse 矩阵判别无条件极值的驻点是正确的,分析了两种充分条件的判别矩阵的关联与差别,并给出条件极值的充分条件的推导过程. 在此基础上,提出一种更精确的判别矩阵—Hesse 变形矩阵 $\boldsymbol{H}_{n-m}^{*}$,推导出其表达式,并举例说明采用 $\boldsymbol{H}_{n-m}^{*}$ 判别比采用 $\boldsymbol{H}_{n}$ 更可靠. 最后举出反例,说明本文给出的 $\boldsymbol{H}_{n-m}^{*}$ 仍不能将所有驻点的状态均正确判别,并阐明其原因及解决办法.

1 引例:一类椭圆方程

例 1 设在笛卡尔坐标系中有椭圆方程:

$$\begin{cases} ax + b \cdot (y^2 + z^2) = 0 & (1) \\ ax + b \cdot (y + z) + c = 0 & (2) \end{cases}$$

且 $a \neq 0$, $b \neq 0$, $b^2 + 2bc \geqslant 0$,则该椭圆上点到原点的距离必存在极大值和极小值,其中:式(1) 是以坐标原点为顶点、x 轴为转轴的旋转抛物面;式(2) 是以 (a, b, b) 为一法向,过 $\left(\frac{-c}{a}, 0, 0\right)$, $\left(0, \frac{-c}{b}, 0\right)$, $\left(0, 0, \frac{-c}{b}\right)$ 3 点的平面. 该椭圆则是两图形的交线,易知其上点到原点的距离必存在极大值和极小值. 也可用初等方法证明极值存在,下面给出证明过程.

证 设有二次方程 $Ax^2 + Bx + C = 0$,且 $B^2 - 4AC \geqslant 0$,所以有实根 $x_{1,2} = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}$,则

① 收稿日期:2017-01-14
作者简介:张 驰(1990-),男,硕士,主要从事最优化方法及时间序列预测算法研究.
通信作者:程 功,硕士,高级工程师.

$x_1+x_2=\frac{-B}{A}$，$x_1\cdot x_2=\frac{C}{A}$. 令 $A=1$，则 $x_1+x_2=-B$，$x_1\cdot x_2=C$. 由式(2)可得，$y+z=\frac{-c-ax}{b}$，将其代入式(1)可得 $y\cdot z=\frac{\left(\frac{ax+c}{b}\right)^2+\frac{ax}{b}}{2}$. 因为 y,z 为椭圆上的坐标，也是实数，所以 y,z 可看作 $Ax^2+Bx+C=0$ 的两实数根，其中 $-B=\frac{-c-ax}{b}$，$C=\frac{\left(\frac{ax+c}{b}\right)^2+\frac{ax}{b}}{2}$，则二次方程变为 $X^2+\frac{ax+c}{b}\cdot X+\frac{\left(\frac{ax+c}{b}\right)^2+\frac{ax}{b}}{2}=0$. 此时满足条件 $B^2-4AC=\left(\frac{ax+c}{b}\right)^2-2\left[\left(\frac{ax+c}{b}\right)^2+\frac{ax}{b}\right]\geqslant 0$，即 $a^2x^2+2a\cdot(b+c)\cdot x+c^2\leqslant 0$. 又因为 $a^2>0$，即二次方程开口向上，所以有

$$\frac{b+c}{-a}-\frac{\sqrt{b^2+2bc}}{|a|}\leqslant x\leqslant\frac{b+c}{-a}+\frac{\sqrt{b^2+2bc}}{|a|} \tag{3}$$

因为式(1)为一旋转抛物面，其上点到原点的距离 d 为 x 的严格单调函数，而式(3)表明 x 存在一个极大值和一个极小值，所以 d 必存在一个极大值和一个极小值，证毕.

特别地，当 $c=0$，不等式(3)变为 $\frac{b}{-a}-\frac{|b|}{|a|}\leqslant x\leqslant\frac{b}{-a}+\frac{|b|}{|a|}$；当 $\frac{a}{b}>0$ 时，$\frac{2b}{-a}\leqslant x\leqslant 0$；当 $\frac{a}{b}<0$ 时，$0\leqslant x\leqslant\frac{2b}{-a}$. 又当 $x=0$ 时，代入式(1)可得 $y=z=0$，此时椭圆上点(0，0，0)到原点的距离 $d_1=0$；当 $x=\frac{2b}{-a}$ 时，代入式(1),(2)可得 $y=z=1$，此时椭圆上点 $\left(\frac{2b}{-a},1,1\right)$ 到原点的距离 $d_2=\sqrt{\frac{4b^2}{a^2}+2}>0$. 所以可知点(0，0，0)为极小值点，距离的极小值 $d_{\min}=d_1=0$；点 $\left(\frac{2b}{-a},1,1\right)$ 为极大值点，距离的极大值 $d_{\max}=d_2=\sqrt{\frac{4b^2}{a^2}+2}$.

从上述分析已知该类椭圆上点到原点的距离必存在极大极小值，且给出了 $c=0$ 时极大极小值的解析表达式. 现采用 Hesse 矩阵来判别 $c=0$ 时该类椭圆上点到原点的距离是否存在极值.

原问题等价为求目标函数 $f(x,y,z)=x^2+y^2+z^2$ 在约束条件式(1),(2)，当 $c=0$ 时的极值，即此时约束条件为：

$$\begin{cases} g_1=ax+b\cdot(y^2+z^2)=0 \\ g_2=ax+b\cdot(y+z)=0 \end{cases}$$

作拉格朗日函数 $L(x,y,z,\lambda_1,\lambda_2)=x^2+y^2+z^2+\lambda_1[ax+b\cdot(y^2+z^2)]+\lambda_2[ax+b\cdot(y+z)]$，由条件极值的必要性条件[2] $\frac{\partial L}{\partial x_i}=0$，$\frac{\partial L}{\partial \lambda_j}=g_j=0$ 可得：

$$\begin{cases} \frac{\partial L}{\partial x}=2x+a\lambda_1+a\lambda_2=0 & (4) \\ \frac{\partial L}{\partial y}=2y+2b\lambda_1 y+b\lambda_2=0 & (5) \\ \frac{\partial L}{\partial z}=2z+2b\lambda_1 z+b\lambda_2=0 & (6) \\ \frac{\partial L}{\partial \lambda_1}=g_1=ax+b\cdot(y^2+z^2)=0 & (7) \\ \frac{\partial L}{\partial \lambda_2}=g_2=ax+b\cdot(y+z)=0 & (8) \end{cases}$$

由式(5),(6)的对称性可得 $y=z$；将其代入式(8)可得 $x=\frac{2by}{-a}=\frac{2bz}{-a}$；将 x,y,z 的关系式代入式(7)可得

$y_{1,2}=z_{1,2}=0,1$，$x_{1,2}=0$，$\dfrac{2b}{-a}$；将 $x_{1,2}$，$y_{1,2}$ 代入式(4)，(5) 可得 $\lambda_{11,12}=0$，$-\dfrac{2}{b}-\dfrac{4b}{a^2}$，$\lambda_{21,22}=0$，$\dfrac{2}{b}+\dfrac{8b}{a^2}$. 所以 L 函数的两个驻点为：

$$\begin{cases} x_1=0 \\ y_1=0 \\ z_1=0 \\ \lambda_{11}=0 \\ \lambda_{12}=0 \end{cases} \qquad \begin{cases} x_2=\dfrac{2b}{-a} \\ y_2=1 \\ z_2=1 \\ \lambda_{21}=-\dfrac{2}{b}-\dfrac{4b}{a^2} \\ \lambda_{22}=\dfrac{2}{b}+\dfrac{8b}{a^2} \end{cases}$$

由之前的分析已知点(0，0，0)为极小值点，点 $\left(\dfrac{2b}{-a},1,1\right)$ 为极大值点，现用 Hesse 矩阵 $\boldsymbol{H}_n$ 来判别这两个驻点. 此时 $\boldsymbol{H}_3=\begin{bmatrix} \dfrac{\partial^2 L}{\partial x^2} & \dfrac{\partial^2 L}{\partial x\partial y} & \dfrac{\partial^2 L}{\partial x\partial z} \\ \dfrac{\partial^2 L}{\partial y\partial x} & \dfrac{\partial^2 L}{\partial y^2} & \dfrac{\partial^2 L}{\partial y\partial z} \\ \dfrac{\partial^2 L}{\partial z\partial x} & \dfrac{\partial^2 L}{\partial z\partial y} & \dfrac{\partial^2 L}{\partial z^2} \end{bmatrix}=\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2+2b\lambda_1 & 0 \\ 0 & 0 & 2+2b\lambda_1 \end{bmatrix}$，当 $\lambda_1=\lambda_{11}=0$ 时，在驻点(0，0，0)处，$2>0$，$2+2b\lambda_1>0$，$\boldsymbol{H}_3$ 正定，则该驻点为极小值点[3]，与证明相符；当 $\lambda_1=\lambda_{21}=-\dfrac{2}{b}-\dfrac{4b}{a^2}$ 时，$2>0$，$2+2b\lambda_1<0$，在驻点 $\left(\dfrac{2b}{-a},1,1\right)$ 处，$\boldsymbol{H}_3$ 的特征值有正有负，则该驻点为鞍点，与证明矛盾. 所以对于该类型的椭圆，当 $c=0$ 时，采用 Hesse 矩阵判别驻点状态是错误的. 由前述证明过程可类似推导出，当 $c\neq 0$ 时，采用 Hesse 矩阵判别驻点状态也是错误的.

2　Hesse 矩阵 $\boldsymbol{H}_n$ 与 Hesse 变形矩阵 $\boldsymbol{H}^*_{n-m}$ 的不同

对于无条件极值，设目标函数为 $f(x)$，$x\in\mathbb{R}^n$，$n\in\mathbb{N}^+$，且设本文中 $f(x)$ 对每个 x_i 具有二阶连续偏导数. 由无条件极值的必要性条件[4]，解 n 个方程 $\dfrac{\partial f}{\partial x_i}=0$ 可得驻点，在任一驻点 $\boldsymbol{X}^0$ 处对 $f(x)$ 作二阶泰勒展开可得：

$$\begin{aligned} \Delta f &= \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\mathrm{d}x_i^0+\frac{1}{2!}\left(\sum_{i=1}^{n}\frac{\partial^2 f}{\partial x_i^2}(\mathrm{d}x_i^0)^2+2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_j}\mathrm{d}x_i^0\mathrm{d}x_j^0\right)+o\left(\sum_{i=1}^{n}(\mathrm{d}x_i^0)^2\right)= \\ &\frac{1}{2}\begin{bmatrix}\mathrm{d}x_1^0 & \mathrm{d}x_2^0 & \cdots & \mathrm{d}x_n^0\end{bmatrix}\begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1\partial x_n} \\ \dfrac{\partial^2 f}{\partial x_2\partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n\partial x_1} & \dfrac{\partial^2 f}{\partial x_n\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{bmatrix}\begin{bmatrix}\mathrm{d}x_1^0 \\ \mathrm{d}x_2^0 \\ \vdots \\ \mathrm{d}x_n^0\end{bmatrix}+o\left(\sum_{i=1}^{n}(\mathrm{d}x_i^0)^2\right)= \\ &\frac{1}{2}\mathrm{d}\boldsymbol{X}^0\boldsymbol{H}_n(\mathrm{d}\boldsymbol{X}^0)^{\mathrm{T}}+o\left(\sum_{i=1}^{n}(\mathrm{d}x_i^0)^2\right) \end{aligned} \tag{9}$$

当 $\mathrm{d}x_i^0$ 的取值很小时，$o\left(\sum_{i=1}^{n}(\mathrm{d}x_i^0)^2\right)$ 是二次型 $\mathrm{d}\boldsymbol{X}^0\boldsymbol{H}_n(\mathrm{d}\boldsymbol{X}^0)^T$ 的高阶无穷小，则 Δf 与 $\mathrm{d}\boldsymbol{X}^0\boldsymbol{H}_n(\mathrm{d}\boldsymbol{X}^0)^{\mathrm{T}}$ 同号. 若此时 $\boldsymbol{H}_n$ 正定，则 $\boldsymbol{H}_n$ 的特征值均大于0，所有 $\mathrm{d}\boldsymbol{X}^0\boldsymbol{H}_n(\mathrm{d}\boldsymbol{X}^0)^{\mathrm{T}}>0$，$\Delta f(x)>0$，$f(\boldsymbol{X}^0)$ 取极小值，$\boldsymbol{X}^0$ 为极

小值点；若此时 $\boldsymbol{H}_n$ 负定，则 $\boldsymbol{H}_n$ 特征值均小于 0，所有 $\mathrm{d}\boldsymbol{X}^0\boldsymbol{H}_n(\mathrm{d}\boldsymbol{X}^0)^{\mathrm{T}}<0$，$\Delta f(x)<0$，$f(\boldsymbol{X}^0)$ 取极大值，$\boldsymbol{X}^0$ 为极大值点；若此时 $\boldsymbol{H}_n$ 的特征值或正或负，则 $\mathrm{d}\boldsymbol{X}^0\boldsymbol{H}_n(\mathrm{d}\boldsymbol{X}^0)^{\mathrm{T}}$ 也或正或负，$\Delta f(x)$ 也或正或负，$\boldsymbol{X}^0$ 为鞍点．除此之外的情况需另行讨论驻点 $\boldsymbol{X}^0$ 的状态．

下面采用降维思想[5]，将等式约束条件极值问题用拉格朗日乘数法来求解．只是这时自变量及其增量均受等式约束限制，因而判别极值的矩阵 $\boldsymbol{H}^*_{n-m}$ 与 $\boldsymbol{H}_n$ 有所不同．现设 $m\in\mathbb{N}^+$，$m<n$，等式约束有 m 个：

$$g_j(x)=0,\ j=1,\cdots,m \tag{10}$$

作拉格朗日函数：

$$L=f+\sum_{j=1}^{m}\lambda_j g_j \tag{11}$$

在驻点及其某一邻域内满足条件雅克比矩阵满秩(式(12))，并且满足式(13)

$$\mathrm{Rank}(J_{(x_1,x_2,\cdots,x_n)})=\mathrm{Rank}\left(\frac{\partial(g_1,g_2,\cdots,g_m)}{\partial(x_1,x_2,\cdots,x_n)}\right)=m \tag{12}$$

$$\frac{\partial g_j}{\partial x_i}\neq 0,\ i=1,2,\cdots,n \tag{13}$$

由条件极值的必要性条件，解 $n+m$ 个方程：$\dfrac{\partial L}{\partial x_i}=0$，$\dfrac{\partial L}{\partial \lambda_j}=g_j=0$，可得驻点．可知任一驻点 $\boldsymbol{X}^0=(x_1^0,\cdots,x_n^0;\lambda_1^0,\cdots,\lambda_m^0)$ 显然满足 m 个等式约束(10)，即 $g_j(x)=0$；且该驻点处的微小改变量$(x_1^0+\mathrm{d}x_1^0,\cdots,x_n^0+\mathrm{d}x_n^0)$ 也必须满足这 m 个等式约束(10)，即 $g_j(x_1^0+\mathrm{d}x_1^0,\cdots,x_n^0+\mathrm{d}x_n^0)=\Delta g_j=0$，则有 $\Delta L=L(x_1^0+\mathrm{d}x_1^0,\cdots,x_n^0+\mathrm{d}x_n^0;\lambda_1^0+\mathrm{d}\lambda_1^0,\cdots,\lambda_m^0+\mathrm{d}\lambda_m^0)-L(x_1^0,\cdots,x_n^0;\lambda_1^0,\cdots,\lambda_m^0)=\Delta f+\sum\limits_{j=1}^{m}\lambda_j\Delta g_j=\Delta f$．

现对 L 函数(11) 在驻点 $\boldsymbol{X}^0$ 处作关于 x_i^0 的二阶泰勒展开，此时需将 L 函数中的 λ_j^0 看作常量，可得：

$$\Delta f=\Delta L=\sum_{i=1}^{n}\frac{\partial L}{\partial x_i}\mathrm{d}x_i^0+\frac{1}{2!}\left(\sum_{i=1}^{n}\frac{\partial^2 L}{\partial x_i^2}(\mathrm{d}x_i^0)^2+2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\frac{\partial^2 L}{\partial x_i\partial x_j}\mathrm{d}x_i^0\mathrm{d}x_j^0\right)+o\left(\sum_{i=1}^{n}(\mathrm{d}x_i^0)^2\right)$$

因为在 $\boldsymbol{X}^0$ 处有必要性条件$\dfrac{\partial L}{\partial x_i}=0$，可得：

$$\Delta f=\frac{1}{2}\begin{bmatrix}\mathrm{d}x_1^0 & \mathrm{d}x_2^0 & \cdots & \mathrm{d}x_n^0\end{bmatrix}\begin{bmatrix}\dfrac{\partial^2 L}{\partial x_1^2} & \dfrac{\partial^2 L}{\partial x_1\partial x_2} & \cdots & \dfrac{\partial^2 L}{\partial x_1\partial x_n}\\ \dfrac{\partial^2 L}{\partial x_2\partial x_1} & \dfrac{\partial^2 L}{\partial x_2^2} & \cdots & \dfrac{\partial^2 L}{\partial x_2\partial x_n}\\ \vdots & \vdots & \ddots & \vdots\\ \dfrac{\partial^2 L}{\partial x_n\partial x_1} & \dfrac{\partial^2 L}{\partial x_n\partial x_2} & \cdots & \dfrac{\partial^2 L}{\partial x_n^2}\end{bmatrix}\begin{bmatrix}\mathrm{d}x_1^0\\ \mathrm{d}x_2^0\\ \vdots\\ \mathrm{d}x_n^0\end{bmatrix}+o\left(\sum_{i=1}^{n}(\mathrm{d}x_i^0)^2\right) \tag{14}$$

对于条件极值和无条件极值充分条件的差别，其关键在于式(9) 和式(14) 中$\begin{bmatrix}\mathrm{d}x_1^0 & \mathrm{d}x_2^0 & \cdots & \mathrm{d}x_n^0\end{bmatrix}$是不同的：式(9) 中 n 个 $\mathrm{d}x_i^0$ 均为独立变量；式(14) 的 n 个 $\mathrm{d}x_i^0$ 中，独立变量只有 $n-m$ 个，而另外 m 个非独立变量为这 $n-m$ 个变量的隐函数[1,6]．所以必须找到 m 个非独立变量与 $n-m$ 个独立变量间更精确的关系式才能更精确地判别极值点．

因此约束(10) 在满足式(12)，(13) 的驻点 $\boldsymbol{X}^0$ 处有泰勒级数：

$$\sum_{i=1}^{n}\left[\frac{\partial g_j}{\partial x_i}\mathrm{d}x_i^0+o\left(\sqrt{\sum_{i=1}^{n}(\mathrm{d}x_i^0)^2}\right)\right]=0,\ j=1,\cdots,m \tag{15}$$

当 $\mathrm{d}x_i^0$ 足够小，可令 $o\left(\sqrt{\sum\limits_{i=1}^{n}(\mathrm{d}x_i^0)^2}\right)=0$，于是解这 m 个线性方程可得：$\begin{bmatrix}\mathrm{d}x_1^0\\ \mathrm{d}x_2^0\\ \vdots\\ \mathrm{d}x_m^0\end{bmatrix}=$

$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n-m} \\ a_{21} & a_{22} & \cdots & a_{2n-m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn-m} \end{bmatrix} \begin{bmatrix} \mathrm{d}x_{m+1}^0 \\ \mathrm{d}x_{m+2}^0 \\ \vdots \\ \mathrm{d}x_{m+(n-m)}^0 \end{bmatrix}$，此时是将 $\mathrm{d}x_{m+1}^0$，…，$\mathrm{d}x_n^0$ 作为独立变量，其中 $\lim\limits_{\mathrm{d}x,\ \mathrm{d}x_i^0 \to 0} \dfrac{\mathrm{d}x}{\mathrm{d}x_i^0} = c \neq 0$，$i = 1,\cdots,n$. 记系数矩阵 $\boldsymbol{A}_{m\times(n-m)}$ 为：

$$\boldsymbol{A}_{m\times(n-m)} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n-m} \\ a_{21} & a_{22} & \cdots & a_{2n-m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn-m} \end{bmatrix}_{m\times(n-m)} \tag{16}$$

记单位阵 $\boldsymbol{I}_{n-m}$ 为：

$$\boldsymbol{I}_{n-m} = \begin{bmatrix} 1 & \cdots & 0 & 0 \\ \vdots & 1 & \cdots & 0 \\ 0 & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}_{n-m}$$

记 Hesse 变形矩阵 $\boldsymbol{H}_{n-m}^*$ 为：

$$\boldsymbol{H}_{n-m}^* = [\boldsymbol{A}^{\mathrm{T}} \quad \boldsymbol{I}_{n-m}]\boldsymbol{H}_n \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{I}_{n-m} \end{bmatrix} \tag{17}$$

则此时可记

$$\Delta f = \Delta L = \frac{1}{2}[\mathrm{d}x_{m+1}^0 + o(\mathrm{d}x) \quad \mathrm{d}x_{m+2}^0 + o(\mathrm{d}x) \quad \cdots \quad \mathrm{d}x_{m+(n-m)}^0 + o(\mathrm{d}x)]\boldsymbol{H}_{n-m}^* \begin{bmatrix} \mathrm{d}x_{m+1}^0 + o(\mathrm{d}x) \\ \mathrm{d}x_{m+2}^0 + o(\mathrm{d}x) \\ \vdots \\ \mathrm{d}x_{m+(n-m)}^0 + o(\mathrm{d}x) \end{bmatrix} +$$

$$o\Big(\sum_{i=1}^{n}(\mathrm{d}x_i^0)^2\Big) = \frac{1}{2}\boldsymbol{Y}\boldsymbol{H}_{n-m}^* Y^{\mathrm{T}} + o\Big(\sum_{i=1}^{n}(\mathrm{d}x_i^0)^2\Big)$$

当 $\mathrm{d}x_i^0$ 足够小时，$o\Big(\sum\limits_{i=1}^{n}(\mathrm{d}x_i^0)^2\Big)$是二次型 $\boldsymbol{Y}\boldsymbol{H}_{n-m}^*\boldsymbol{Y}^{\mathrm{T}}$ 的高阶无穷小，Δf 与 $\boldsymbol{Y}\boldsymbol{H}_{n-m}^*\boldsymbol{Y}^{\mathrm{T}}$ 同号. 此时若 $\boldsymbol{H}_{n-m}^*$ 正定，则 $\boldsymbol{H}_{n-m}^*$ 的特征值均大于 0，所有 $\boldsymbol{Y}\boldsymbol{H}_{n-m}^*\boldsymbol{Y}^{\mathrm{T}} > 0$，$\Delta f(x) > 0$，$f(\boldsymbol{X}^0)$ 取极小值，驻点 $\boldsymbol{X}^0$ 为极小值点；若 $\boldsymbol{H}_{n-m}^*$ 负定，$\boldsymbol{H}_{n-m}^*$ 的特征值均小于 0，所有 $\boldsymbol{Y}\boldsymbol{H}_{n-m}^*\boldsymbol{Y}^{\mathrm{T}} < 0$，$\Delta f(x) < 0$，$f(\boldsymbol{X}^0)$ 取极大值，驻点 $\boldsymbol{X}^0$ 为极大值点；若 $\boldsymbol{H}_{n-m}^*$ 的特征值或正或负，则 $\boldsymbol{Y}\boldsymbol{H}_{n-m}^*\boldsymbol{Y}^{\mathrm{T}}$ 或正或负，$\Delta f(x)$ 或正或负，驻点 $\boldsymbol{X}^0$ 为鞍点. 除此之外的情况需另行讨论驻点 $\boldsymbol{X}^0$ 的状态.

下面推导系数矩阵 $\boldsymbol{A}_{m\times(n-m)}$ 中元素 a_{jk} 的表达式以更精确地判别极值.

当 $n = 3$，$m = 2$ 时，有目标函数 $f(x_1, x_2, x_3)$ 及约束条件(18)，

$$g_1(x_1, x_2, x_3) = 0,\ g_2(x_1, x_2, x_3) = 0 \tag{18}$$

设此时雅克比行列式为 $J_{(x_1,\ x_2)} = \dfrac{D(g_1, g_2)}{D(x_1, x_2)}$，即以 x_3 作为独立变量，此时 $i = 1,2,3$，$j = 1,2$，则对等式约束(18) 在驻点处作一阶泰勒展开可得：

$$\begin{bmatrix} \mathrm{d}x_1^0 \\ \mathrm{d}x_2^0 \end{bmatrix} = \frac{1}{J} \begin{bmatrix} \dfrac{D(g_2, g_1)}{D(x_3, x_2)} \\ \dfrac{D(g_2, g_1)}{D(x_3, x_1)} \end{bmatrix} \mathrm{d}x_3^0 + \begin{bmatrix} o(\mathrm{d}x) \\ o(\mathrm{d}x) \end{bmatrix}$$

当 $n = 4$，$m = 2$ 时，有目标函数 $f(x_1, x_2, x_3, x_4)$ 及约束条件(19)，

$$g_1(x_1, x_2, x_3, x_4) = 0,\ g_2(x_1, x_2, x_3, x_4) = 0 \tag{19}$$

此时设雅克比行列式为 $J_{(x_1,\ x_2)} = \dfrac{D(g_1, g_2)}{D(x_1, x_2)}$，即以 x_3，x_4 作为独立变量，此时 $i = 1,2,3,4$，$j = 1,2$，则对

等式约束(19) 在驻点处作一阶泰勒展开可得：

$$\begin{bmatrix} \mathrm{d}x_1^0 \\ \mathrm{d}x_2^0 \end{bmatrix} = \frac{1}{J}\begin{bmatrix} \dfrac{D(g_2,\ g_1)}{D(x_3,\ x_2)} \\ \dfrac{D(g_2,\ g_1)}{D(x_1,\ x_3)} \end{bmatrix}\mathrm{d}x_3^0 + \frac{1}{J}\begin{bmatrix} \dfrac{D(g_2,\ g_1)}{D(x_4,\ x_2)} \\ \dfrac{D(g_2,\ g_1)}{D(x_1,\ x_4)} \end{bmatrix}\mathrm{d}x_4^0 + \begin{bmatrix} o(\mathrm{d}x) \\ o(\mathrm{d}x) \end{bmatrix}$$

继续求解多维多约束的情形，并由数学归纳法可得，$\forall\, n,\ m \in \mathbb{N}^+$，$m < n$，$k = 1, 2, \cdots, n - m$，当目标函数为 $f(x)$，且 $f(x)$ 在定义域内对每个 x_i 具有二阶连续偏导数，有等式约束(10)，L 函数(11) 于驻点满足条件(12)，(13) 时，有式(10) 的一阶泰勒级数如下：

1）当 $j = 1$，即所求系数 a_{1k} 为 $\boldsymbol{A}_{m\times(n-m)}$ 第一行的元素时，

$$a_{1k} = \frac{(-1)^1}{J}\frac{D(g_m,\ g_{m-1},\ \cdots g_1)}{D(x_{m+k},\ x_m,\ \cdots x_2)} \tag{20}$$

2）当 $j = m$，即所求系数 a_{mk} 为 $\boldsymbol{A}_{m\times(n-m)}$ 最后一行的元素时，

$$a_{mk} = \frac{(-1)^m}{J}\frac{D(g_m,\ g_{m-1},\ \cdots g_1)}{D(x_1,\ x_{m-1},\ x_{m-2},\ \cdots x_2,\ x_{m+k})} \tag{21}$$

3）当 $j = 2, 3, \cdots, m - 1$，即所求系数 a_{jk} 为 $\boldsymbol{A}_{m\times(n-m)}$ 中间行的元素时，

$$a_{jk} = \frac{(-1)^j}{J}\frac{D(g_m,\ g_{m-1},\ \cdots g_1)}{D(x_{j+1},\ x_{j-1},\ x_{j-2},\ \cdots x_2,\ x_{m+k},\ x_m,\ x_{m-1},\ \cdots,\ x_{j+3},\ x_{j+2},\ x_1)}$$

现举一例来说明对于等式约束的条件极值采用 $\boldsymbol{H}^*_{n-m}$ 判别比采用 $\boldsymbol{H}_n$ 判别更精确.

例 2　对于式(1)，(2)，取 $a = 1$，$b = c = -1$，此时满足条件 $a \neq 0$，$b \neq 0$，$b^2 + 2bc \geqslant 0$，椭圆方程(1)，(2) 变为：

$$\begin{cases} x - y^2 - z^2 = 0 & (22) \\ x - y - z - 1 = 0 & (23) \end{cases}$$

求该椭圆上的点到坐标原点距离的极大值及极小值.

由前述证明可知该距离必存在一个极大值和一个极小值，现分别采用 $\boldsymbol{H}_n$ 和 $\boldsymbol{H}^*_{n-m}$ 判别极值.

解　原题等价于求目标函数 $f = x^2 + y^2 + z^2$ 在等式约束 $g_1 = x - y^2 - z^2 = 0$，$g_2 = x - y - z - 1 = 0$ 下的极大值和极小值. 作 L 函数 $L = f + \lambda_1 g_1 + \lambda_2 g_2$，于是有：

$$\begin{cases} \dfrac{\partial L}{\partial x} = 2x + \lambda_1 + \lambda_2 = 0 \\ \dfrac{\partial L}{\partial y} = 2y - 2\lambda_1 y - \lambda_2 = 0 \\ \dfrac{\partial L}{\partial z} = 2z - 2\lambda_1 z - \lambda_2 = 0 \\ \dfrac{\partial L}{\partial \lambda_1} = g_1 = x - y^2 - z^2 = 0 \\ \dfrac{\partial L}{\partial \lambda_2} = g_2 = x - y - z - 1 = 0 \end{cases}$$

采用与上文相同的求解步骤，得 L 函数的两个驻点：

$$\begin{cases} x_1 = 2 + \sqrt{3} \\ y_1 = \dfrac{1 + \sqrt{3}}{2} \\ z_1 = \dfrac{1 + \sqrt{3}}{2} \\ \lambda_{11} = 3 + \dfrac{5}{3}\sqrt{3} \\ \lambda_{12} = -7 - \dfrac{11}{3}\sqrt{3} \end{cases} \qquad \begin{cases} x_2 = 2 - \sqrt{3} \\ y_2 = \dfrac{1 - \sqrt{3}}{2} \\ z_2 = \dfrac{1 - \sqrt{3}}{2} \\ \lambda_{21} = 3 - \dfrac{5}{3}\sqrt{3} \\ \lambda_{22} = -7 + \dfrac{11}{3}\sqrt{3} \end{cases}$$

此时若采用 $\boldsymbol{H}_n$ 判别，则有 $\boldsymbol{H}_3=\begin{bmatrix}2 & 0 & 0\\ 0 & 2-2\lambda_1 & 0\\ 0 & 0 & 2-2\lambda_1\end{bmatrix}$，当 $\lambda_1=\lambda_{11}=3+\frac{5}{3}\sqrt{3}$ 时，$2-2\lambda_1<0$，$2>0$，则 $\boldsymbol{H}_3$ 不定，(x_1, y_1, z_1) 为 f 的鞍点；当 $\lambda_1=\lambda_{21}=3-\frac{5}{3}\sqrt{3}$ 时，$2-2\lambda_1>0$，$2>0$，则 $\boldsymbol{H}_3$ 正定，(x_2, y_2, z_2) 为 f 的极小值点. 但此结果与证明相悖，所以改用 $\boldsymbol{H}_{n-m}^{*}$ 判别. 因为此时 $n=3$，$m=2$，点 (x_1, y_1, z_1)，(x_2, y_2, z_2) 满足条件：$J_{(x, y)}=\frac{D(g_1, g_2)}{D(x, y)}=2y-1\neq 0$，$\frac{\partial g_1}{\partial x}=1\neq 0$，$\frac{\partial g_1}{\partial y}=-2y\neq 0$，$\frac{\partial g_2}{\partial x}=1\neq 0$，$\frac{\partial g_2}{\partial y}=-1\neq 0$，$\mathrm{Rank}(J_{(x, y)})=\mathrm{Rank}\begin{bmatrix}1 & -2y\\ 1 & -1\end{bmatrix}=2=m$，所以可采用 $\boldsymbol{H}_{n-m}^{*}$ 判别极值. 由式(17)可得，$\boldsymbol{H}_{3-2}^{*}=\begin{bmatrix}\boldsymbol{A}_{2\times(3-2)}^{T} & \boldsymbol{I}_{3-2}\end{bmatrix}\boldsymbol{H}_3\begin{bmatrix}\boldsymbol{A}_{2\times(3-2)}\\ \boldsymbol{I}_{3-2}\end{bmatrix}$，由式(16)可得 $\boldsymbol{A}_{2\times(3-2)}=\begin{bmatrix}a_{11}\\ a_{21}\end{bmatrix}$；当 $j=1$ 时，由式(20)可得，$a_{11}=\frac{(-1)^1}{J_{(x, y)}}\cdot\frac{D(g_2, g_1)}{D(x_3, x_2)}=\frac{2y-2z}{2y-1}=0$，当 $j=2=m$ 时，由式(21)可得，$a_{21}=\frac{(-1)^2}{J_{(x, y)}}\cdot\frac{D(g_2, g_1)}{D(x_1, x_3)}=\frac{-2z+1}{2y-1}=-1$；所以此时 $\boldsymbol{A}_{2\times(3-2)}=\begin{bmatrix}0\\ -1\end{bmatrix}$，$\boldsymbol{H}_{3-2}^{*}=\begin{bmatrix}0 & -1 & 1\end{bmatrix}\boldsymbol{H}_3\begin{bmatrix}0\\ -1\\ 1\end{bmatrix}=4(1-\lambda_1)$. 当 $\lambda_1=\lambda_{11}=3+\frac{5}{3}\sqrt{3}$ 时，$4(1-\lambda_1)<0$，则 $\boldsymbol{H}_{3-2}^{*}$ 负定，此时 $\Delta f=\mathrm{d}z\cdot 4(1-\lambda_1)\cdot\mathrm{d}z+o(\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2)<0$，$(x_1, y_1, z_1)$ 为 f 的极大值点，椭圆上的点到原点距离的极大值 $\mathrm{d}_{\max}=\sqrt{f_{\max}}=4.63$；当 $\lambda_1=\lambda_{21}=3-\frac{5}{3}\sqrt{3}$ 时，$4(1-\lambda_1)>0$，则 $\boldsymbol{H}_{3-2}^{*}$ 正定，此时 $\Delta f>0$，(x_2, y_2, z_2) 为 f 的极小值点，距离的极小值 $\mathrm{d}_{\min}=\sqrt{f_{\min}}=0.78$. 所以存在一个极大值和一个极小值，与前述证明相符.

3 Hesse 变形矩阵 $\boldsymbol{H}_{n-m}^{*}$ 的不足

但采用 $\boldsymbol{H}_{n-m}^{*}$ 并不能将所有等式约束条件极值的驻点状态准确判别，现再举一例.

例 3 设有椭圆方程：

$$\begin{cases}\frac{x+3}{2}-y^2-z^2=0\\ x-y-z=0\end{cases}$$

求该椭圆上点到原点距离的极大极小值.

解 该问题同样可用前文所述初等方法证明极大极小值的存在. 现继续采用 $\boldsymbol{H}_{n-m}^{*}$ 判别极值. 此时该问题等价为求目标函数 $f=x^2+y^2+z^2$ 在等式约束 $g_1=\frac{x+3}{2}-y^2-z^2=0$，$g_2=x-y-z=0$ 下的极大值和极小值. 作 L 函数 $L=f+\lambda_1 g_1+\lambda_2 g_2$，有：

$$\begin{cases}\frac{\partial L}{\partial x}=2x+\frac{\lambda_1}{2}+\lambda_2=0\\ \frac{\partial L}{\partial y}=2y-2\lambda_1 y-\lambda_2=0\\ \frac{\partial L}{\partial z}=2z-2\lambda_1 z-\lambda_2=0\\ \frac{\partial L}{\partial \lambda_1}=\frac{x+3}{2}-y^2-z^2=0\\ \frac{\partial L}{\partial \lambda_2}=x-y-z=0\end{cases}$$

解得 L 函数的两个驻点为：

$$\begin{cases} x_1 = \dfrac{1+\sqrt{13}}{2} \\ y_1 = \dfrac{1+\sqrt{13}}{4} \\ z_1 = \dfrac{1+\sqrt{13}}{4} \\ \lambda_{11} = 3+\dfrac{3}{13}\sqrt{13} \\ \lambda_{12} = -\dfrac{65}{26}-\dfrac{29}{26}\sqrt{13} \end{cases} \qquad \begin{cases} x_2 = \dfrac{1-\sqrt{13}}{2} \\ y_2 = \dfrac{1-\sqrt{13}}{4} \\ z_2 = \dfrac{1-\sqrt{13}}{4} \\ \lambda_{21} = 3-\dfrac{3}{13}\sqrt{13} \\ \lambda_{22} = -\dfrac{65}{26}+\dfrac{29}{26}\sqrt{13} \end{cases}$$

此时有 $J_{(x,y)} = 2y-\dfrac{1}{2} \neq 0$，$\dfrac{\partial g_1}{\partial x} = \dfrac{1}{2} \neq 0$，$\dfrac{\partial g_1}{\partial y} = -2y \neq 0$，$\dfrac{\partial g_2}{\partial x} = 1 \neq 0$，$\dfrac{\partial g_2}{\partial y} = -1 \neq 0$，$\mathrm{Rank}(J_{(x,y)}) = \mathrm{Rank}\begin{bmatrix} \dfrac{1}{2} & -2y \\ 1 & -1 \end{bmatrix} = 2 = m$，所以可采用 $\boldsymbol{H}_{n-m}^{*}$ 判别极值. 当 $j=1$ 时，由式(20) 可得，$a_{11} = \dfrac{(-1)^1}{J_{(x,y)}} \cdot \dfrac{D(g_2, g_1)}{D(x_3, x_2)} = \dfrac{2y-2z}{-2y+\dfrac{1}{2}} = 0$，当 $j=2=m$ 时，由式(21) 可得，$a_{21} = \dfrac{(-1)^2}{J_{(x,y)}} \cdot \dfrac{D(g_2, g_1)}{D(x_1, x_3)} = \dfrac{-2z+\dfrac{1}{2}}{2y-\dfrac{1}{2}} = -1$；所以此时 $\boldsymbol{A}_{2\times(3-2)} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$，$\boldsymbol{H}_{3-2}^{*} = [0 \quad -1 \quad 1]\boldsymbol{H}_3 \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = 4(1-\lambda_1)$. 当 $\lambda_1 = \lambda_{11} = 3+\dfrac{3}{13}\sqrt{13}$ 时，$4(1-\lambda_1) < 0$，则 $\boldsymbol{H}_{3-2}^{*}$ 负定，此时 $\Delta f < 0$，(x_1, y_1, z_1) 为 f 的极大值点；当 $\lambda_1 = \lambda_{21} = 3-\dfrac{3}{13}\sqrt{13}$ 时，$4(1-\lambda_1) < 0$，则 $\boldsymbol{H}_{3-2}^{*}$ 负定，此时 $\Delta f < 0$，(x_2, y_2, z_2) 也为 f 的极大值点，与实际情况矛盾，此时使用 $\boldsymbol{H}_{n-m}^{*}$ 判定失效. 出现这种情况的原因在于式(15) 为寻找自变量增量间的解析表达式，对约束条件在驻点处作的一阶泰勒展开，省略了二阶及其以上的高阶无穷小量，因而不是自变量增量间最精确的关系式[7-8]. 随着目标函数维数增加，等式约束数目及复杂程度增大，即使只作一阶泰勒展开，求解式(15) 中非独立变量关于独立变量的解析表达式已相当复杂，如果将泰勒级数展开到二阶或者更高阶，将更难甚至无法求解出表达式. 尽管如此，采用 $\boldsymbol{H}_{n-m}^{*}$ 判别等式约束的条件极值确实比采用 $\boldsymbol{H}_n$ 更精确，因为用一阶泰勒展开求解了自变量增量间的关系式.

4　结　　论

采用 Hesse 矩阵 $\boldsymbol{H}_n$ 判别无条件极值的充分条件是准确的，而判别存在等式约束的条件极值的充分条件不够精确；采用本文提出的 Hesse 变形矩阵 $\boldsymbol{H}_{n-m}^{*}$ 判别后者更精确，因其可判别出一部分被 $\boldsymbol{H}_n$ 错误判定的驻点. 产生这种差别的原因是无条件极值问题的自变量微小增量 $\mathrm{d}x_i$ 间没有约束关系，即所有 $\mathrm{d}x_i$ 均是独立变量；而条件极值问题的自变量微小增量 $\mathrm{d}x_i$ 与 m 个等式约束有关，n 个 $\mathrm{d}x_i$ 中只有 $n-m$ 个是独立变量，另外 m 个非独立变量是这 $n-m$ 个变量的隐函数.

由于对等式约束在驻点处作一阶泰勒展开得到的关系式(15) 是在舍弃了二阶及其以上的高阶无穷小量时得到的，即本文得出的非独立变量与独立变量间的关系式并不是最准确的，所以 $\boldsymbol{H}_{n-m}^{*}$ 仍不能将所有驻点的状态均准确判定. 若将式(15) 式作更高阶的泰勒展开，则可得到更精确的 $\boldsymbol{H}_{n-m}^{*}$，从而正确判别出更多驻点的状态；但其解析表达式过于复杂不便推导及表述，寻找系数矩阵 $\boldsymbol{A}_{m\times(n-m)}$ 的数值解是更有实际意义的.

参考文献:

[1] 邱伯驺,李景功,潘　杰. 关于用拉格朗日乘数法求条件极值的充分条件[J]. 大学数学,1993(S2):26-31.

[2] JAMES S. Calculus[M]. 7th Ed. 北京:高等教育出版社,2014:946-962.

[3] STEVEN J L. Linear Algebra with Application[M]. 9th Ed. 北京:机械工业出版社,2015:351-370.

[4] 周民强,方企勤. 数学分析(第三册)[M]. 北京:科学出版社,2014:88-95.

[5] 张　芳,徐文雄. 关于约束极值问题降维法的探讨[J]. 大学数学,2008(6):130-133.

[6] 宁荣健. 也谈条件极值问题的充分条件[J]. 高等数学研究,2005(2):40-43.

[7] 张　驰,胡　博,秦　琴,等. 多元函数条件极值的一种较精确的充分条件[J]. 大理大学学报,2017,2(12):5-11.

[8] RICHARD C,FRITZ J. 微积分和数学分析引论(第二卷,第一分册)[M]. 北京:科学出版社,2001.

On Sufficient Condition of Conditional Extreme Value and a Kind of Ellipse Equations

ZHANG Chi[1], CHENG Gong[2],
ZHONG Hai-quan[3], HUANG Hua-fei[2]

1. Department of Artificial Intelligence Ordering Platform, Chengdu Changyi Western Union Software Co. Ltd, Chengdu 610200, China;

2. Institute of Urban Construction, Guangzhou Huali Science and Technology Vocational College, Guangzhou 511325, China;

3. State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Southwest Petroleum University, Chengdu 610500, China

Abstract: There is rarely discussion on sufficient condition of conditional extreme value of multi-variable function constraint to equalities in most mathematical analysis teaching materials, or with rather complex discussion instead, but conditional extreme value always follows unconditional extreme value tightly, which makes difficult to understand the relevance and difference between two kinds of discriminant matrices of sufficient conditions for students, or in mistake that both of discriminant matrices are in identity. To help students gasp the relevance and difference between both of them profoundly, understand sufficient condition of both problems of extreme value quicker and more comprehensive, sufficient conditions of unconditional and conditional extreme value have been discussed in detail, and different situations may encounter in distinguishing with various kinds of stationary points been explained by examples of validity and invalidity, this can help students understand the problem of extreme value quicker, more profoundly and comprehensively.

Key words: multi-variable function; conditional extreme value; sufficient condition; Hessian; ellipse equation

责任编辑　张　枸