数学分析讲义（省身班）

段华贵

数学科学学院

2022年5月5日

第11.7节

极值理论

设$f(X)$是定义在区域D上的函数, X_0是D的内点. 如果存在X_0的邻域$B(X_0)\subseteq D$, 使得

$$f(X)\leqslant f(X_0),\quad \forall X\in B(X_0),$$

则称X_0是$f(X)$的**极大值点**, $f(X_0)$称为$f(X)$的**极大值**. 如果

$$f(X)\geqslant f(X_0),\quad \forall X\in B(X_0),$$

则称X_0是$f(X)$的**极小值点**, $f(X_0)$称为$f(X)$的**极小值**. 极大值与极小值统称极值，极大值点与极小值点统称极值点.

Theorem (必要条件)

设X_0是$f(X)$的极值点. 如果$\dfrac{\partial f}{\partial x_i}(X_0)$存在, 则

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

若$f(X)$在X_0的邻域内二阶偏导数连续, 则

(i) 若X_0是$f(X)$的极小值点，则$H_f(X_0)\geqslant 0.$

(ii) 若X_0是$f(X)$的极大值点，则$H_f(X_0)\leqslant 0.$

若X_0是$f(X)$的极小值点, 则

$$\begin{aligned}
& f(X_0+\Delta X)-f(X_0) \\
= & \langle \nabla f(X_0), \Delta X\rangle + \frac{1}{2}\Delta X \cdot H_f(X_0)\cdot \Delta X^T + o\left(|\Delta X|^2\right) \\
= & \frac{1}{2}\Delta X \cdot H_f(X_0)\cdot \Delta X^T + o\left(|\Delta X|^2\right) \geq 0,
\end{aligned}$$

若存在$Y_0 \in \mathbb{R}^n$, 使得$Y_0 \cdot H_f(X_0)\cdot Y_0^T = -\lambda$, 其中$\lambda > 0$, $|Y_0| = 1$, 则取$\Delta X = tY_0 \in \mathbb{R}^n$使得

$$\Delta X \cdot H_f(X_0)\cdot \Delta X^T = -\lambda t^2, \quad |\Delta X| = |t|.$$

当$|t|$充分小时,可知

$$f(X_0+\Delta X)-f(X_0) = \frac{-\lambda}{2}t^2 + o\left(t^2\right) < 0,$$

矛盾. 于是$H_f(X_0) \geqslant 0$.

设$f(X)$在X_0的一个邻域内所有一阶偏导数连续, 如果$\dfrac{\partial f}{\partial x_i}(X_0)=0, i=1,\cdots,n$, 即

$$\nabla f(X_0)=0,$$

则称X_0是f的**临界点**, $f(X_0)$为**临界值**.

极值点必为临界点, 但临界点未必是极值点.

Theorem (充分条件)

设$f(X)$在X_0的某邻域二阶偏导数连续，且X_0是$f(X)$的临界点，则

(i) $H_f(X_0) > 0 \Rightarrow X_0$是$f(X)$的极小值点.

(ii) $H_f(X_0) < 0 \Rightarrow X_0$是$f(X)$的极大值点.

(iii) $H_f(X_0)$为不定矩阵(即特征根有正有负) $\Rightarrow X_0$不是$f(X)$的极值点.

注记： 其余情形，即$H_f(X_0) \geqslant 0$但$H_f(X_0) > 0$不成立，或$H_f(X_0) \leqslant 0$但$H_f(X_0) < 0$不成立时，则需要进一步判定.

证明: 由于X_0是$f(X)$的临界点, 故

$$f(X_0+\Delta X)=f(X_0)+\frac{1}{2}\Delta X\cdot H_f(X_0)\cdot\Delta X^T+o\left(|\Delta X|^2\right).$$

(1) 如果$H_f(X_0)>0$, （断言）则存在$\lambda>0$使得

$$\Delta X\cdot H_f(X_0)\cdot\Delta X^T\geq\lambda|\Delta X|^2,\quad\forall\,\Delta X\in\mathbb{R}^n.$$

事实上, 可设$|\Delta X|\neq 0$. 记$S^{n-1}=\{X\in\mathbb{R}^n|\ |X|=1\}$,紧集, 函数

$$g(X)=X\cdot H_f(X_0)\cdot X^T,\ X\in S^{n-1}$$

连续, 存在最小值$\lambda>0$, 取$X=\frac{1}{|\Delta X|}\Delta X\in S^{n-1}$, 即得断言.

证明(续)

由于

$$\lim_{|\Delta X|\to 0} \frac{o\left(|\Delta X|^2\right)}{|\Delta X|^2} = 0,$$

从而存在$\eta > 0$, 使得

$$\left|o\left(|\Delta X|^2\right)\right| \leqslant \frac{1}{4}\lambda|\Delta X|^2, \quad \Delta X \in \mathbb{R}^n, \quad |\Delta X| < \eta.$$

从而当$\Delta X \in \mathbb{R}^n$且$0 < |\Delta X| < \eta$时

$$\begin{aligned} & f(X_0 + \Delta X) - f(X_0) \\ = \quad & \frac{1}{2}\Delta X \cdot H_f(X_0) \cdot \Delta X^T + o\left(|\Delta X|^2\right) \geq \frac{1}{4}\lambda|\Delta X|^2 > 0, \end{aligned}$$

即$f(X)$在X_0达到严格极小.

设$f(x,y)$在(x_0,y_0)的某邻域内二阶偏导数连续, 且

$$\frac{\partial f}{\partial x}(x_0,y_0)=0,\qquad \frac{\partial f}{\partial y}(x_0,y_0)=0.$$

记$A=\dfrac{\partial^2 f}{\partial x^2}(x_0,y_0)$, $B=\dfrac{\partial^2 f}{\partial x\partial y}(x_0,y_0)$, $C=\dfrac{\partial^2 f}{\partial y^2}(x_0,y_0)$, 则

$$H_f(x_0,y_0)=\begin{pmatrix} A & B \\ B & C \end{pmatrix}.$$

$H_f(x_0,y_0)>0 \ \Leftrightarrow\ \det H_f(x_0,y_0)=AC-B^2>0$且$A>0$;

$H_f(x_0,y_0)<0 \ \Leftrightarrow\ \det H_f(x_0,y_0)=AC-B^2>0$且$A<0$.

$H_f(x_0,y_0)$不定$\Leftrightarrow \det H_f(x_0,y_0)=AC-B^2<0$.

$$\det H_f(x_0,y_0)>0,\ (x_0,y_0)\text{是极值点,}\ \begin{cases} A>0, (x_0,y_0)\text{是极小值点,} \\ A<0, (x_0,y_0)\text{是极大值点,} \end{cases}$$

$$\det H_f(x_0,y_0)<0, (x_0,y_0)\text{不是极值点,}$$

$$\det H_f(x_0,y_0)=0, (x_0,y_0)\text{可能是为极值点也可能不是极值点.}$$

Example

求函数$f(x,y)=x^2+xy+y^2-x-y$ 的临界点,极值点和极值.

解 首先

$$\begin{cases} f'_x = 2x+y-1=0, \\ \\ f'_y = x+2y-1=0\ . \end{cases}$$

解得临界点$P=\left(\frac{1}{3},\frac{1}{3}\right)$, 而

$$H_f(P)=\begin{pmatrix} f''_{xx} & f''_{xy} \\ f''_{yx} & f''_{yy} \end{pmatrix}=\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},$$

这里$AC-B^2=3$, $A=2$, 故$\left(\frac{1}{3},\frac{1}{3}\right)$为极小值点, 极小值为$-\frac{1}{3}$.

例题2

Example

设$f_i(X) \in C^1(\mathbb{R}^n), i = 1, 2, \cdots, n$, 且

$$\lim_{|X|\to+\infty} \sum_{i=1}^{n} |f_i(X)|^2 = +\infty.$$

求证: 对于任给的$Y_0 = (y_{01}, \cdots, y_{0n})$, 存在$X_0 \in \mathbb{R}^n$满足

$$\sum_{i=1}^{n} \frac{\partial f_i}{\partial x_k}(X_0)(f_i(X_0) - y_{0i}) = 0, \quad k = 1, \cdots, n.$$

记$f = (f_1, \cdots, f_n)$.

证明 考虑$g(X)=|f(X)-Y_0|^2$, $\forall X\in\mathbb{R}^n$, 则

$$\frac{\partial g}{\partial x_k}(X)=2\sum_{i=1}^{n}\frac{\partial f_i}{\partial x_k}(X)(f_i(X)-y_{0i}),\quad k=1,\cdots,n,$$

即$g(X)\in C^1(\mathbb{R}^n)$.

由条件知$\lim_{|X|\to+\infty}g(X)=+\infty$. 故$g(x)$在$\mathbb{R}^n$上有最小值, 记$\inf_{X\in\mathbb{R}^n}g(X)=c$. 存在$r>0$使得当$|X|\geqslant r$ 时，有

$$g(X)\geqslant c+1.$$

从而有

$$\inf_{X\in\mathbb{R}^n}g(X)=\inf_{|X|\leqslant r}g(X)=c.$$

因$\{X\in\mathbb{R}^n|\ |X|\leqslant r\}$为有界闭集, 存在$X_0$使得

$$g(X_0)=\inf_{|X|\leqslant r}g(X)=\inf_{X\in\mathbb{R}^n}g(X).$$

由极值的必要条件知结论成立.

Example

设$f(x,y)=(x^2+1)\left[(x+\mathrm{e}^y)^3-3(x+\mathrm{e}^y)+3\right]$，求证:
$f(x,y)$在$\mathbb{R}^2$有唯一临界点，且为极小值点，但是它没有最小值.

证明 解方程组

$$\begin{cases} f'_x(x,y)=0 \\ f'_y(x,y)=0 \end{cases}$$

解得$x=0$, $y=0$, 所以$f(x,y)$有唯一临界点$(0,0)$. 另外

$$H_f(0,0)=\begin{pmatrix} 8 & 6 \\ 6 & 6 \end{pmatrix}>0,$$

故$(0,0)$是极小值点, 极小值为$f(0,0)=1$. 但极小值不是最小值, 事实上，当$y=0$, $x\to-\infty$时，$f(x,y)\to-\infty$.

Example (最小二乘法)

设$(x_1, y_1), \cdots, (x_n, y_n)$是平面上的$n$个点，求直线$y = ax + b$使得$f(a, b) = \sum_{i=1}^n (ax_i + b - y_i)^2$最小.

$$\begin{vmatrix} x & y & 1 \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i y_i & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n y_i & n \end{vmatrix} = 0.$$

设D是$\mathbb{R}^n$中的开区域，$f(X), g_1(X), \cdots, g_k(X)$都定义在$D$内．考虑约束条件:

$$S:\begin{cases} g_1(X)=0, \\ \quad\cdots \\ g_k(X)=0. \end{cases} \tag{5}$$

设$X_0 \in D$满足(5)，即$g_i(X_0)=0, i=1,\cdots,k$．如果存在X_0的某邻域$B(X_0)\subseteq D$，使得对任何满足条件(5)且在$B(X_0)$内的X都有

$$f(X) \leqslant f(X_0),$$

则称X_0是目标函数$f(X)$在约束条件(5)下的**条件极大值点**，$f(X_0)$称为**条件极大值**.

$$\frac{\partial(g_1,\cdots,g_k)}{\partial(x_1,\cdots,x_n)} = \begin{pmatrix} \dfrac{\partial g_1}{\partial x_1} & \cdots & \dfrac{\partial g_1}{\partial x_n} \\ & \cdots & \\ \dfrac{\partial g_k}{\partial x_1} & \cdots & \dfrac{\partial g_k}{\partial x_n} \end{pmatrix}$$

假设$g_i(X) \in C^1(D), i = 1, \cdots, k,\ k < n$, 且

$$\operatorname{rank}\frac{\partial(g_1, \cdots, g_k)}{\partial(x_1, \cdots, x_n)} = k.$$

不妨设

$$\frac{D(g_1, \cdots, g_k)}{D(x_{n-k+1}, \cdots, x_n)}(X_0) \neq 0,$$

由隐函数定理得,在X_0某邻域内

$$\begin{cases} x_{n-k+1} & = & x_{n-k+1}(x_1, \cdots, x_{n-k}), \\ \cdots & & \cdots, \\ x_n & = & x_n(x_1, \cdots, x_{n-k}). \end{cases}$$

则 $\quad G(x_1, \cdots, x_{n-k})$

$$\equiv \quad f(x_1, \cdots, x_{n-k}, x_{n-k+1}(x_1, \cdots, x_{n-k}), \cdots, x_n(x_1, \cdots, x_{n-k})).$$

令$X=(X',X'')$, $X'=(x_1,\cdots,x_{n-k})$, $X''=(x_{n-k+1},\cdots,x_n)$, 则

$$\begin{aligned}\nabla G(X') &= \nabla_{X'}f(X)+\nabla_{X''}f(X)\frac{\partial(x_{n-k+1},\cdots,x_n)}{\partial(x_1,\cdots,x_{n-k})}(X')\\ &= \nabla_{X'}f(X)-\nabla_{X''}f(X)\left(\frac{\partial(g_1,\cdots,g_k)}{\partial(x_{n-k+1},\cdots,x_n)}(X)\right)^{-1}\\ &\quad\frac{\partial(g_1,\cdots,g_k)}{\partial(x_1,\cdots,x_{n-k})}(X).\end{aligned}$$

由极值的必要条件, $\nabla G(X_0')=0$, 从而

$$\begin{aligned}&\nabla_{X'}f(X_0)\\ =\ &\boxed{\nabla_{X''}f(X_0)\left(\frac{\partial(g_1,\cdots,g_k)}{\partial(x_{n-k+1},\cdots,x_n)}(X_0)\right)^{-1}}\\ &\frac{\partial(g_1,\cdots,g_k)}{\partial(x_1,\cdots,x_{n-k})}(X_0).\qquad \equiv \boxed{-(\lambda_1,\cdots,\lambda_k)}\end{aligned}$$

则

$$\begin{cases} \nabla_{X'} f(X_0) + (\lambda_1, \cdots, \lambda_k) \frac{\partial(g_1, \cdots, g_k)}{\partial(x_1, \cdots, x_{n-k})}(X_0) = 0, \\ \nabla_{X''} f(X_0) + (\lambda_1, \cdots, \lambda_k) \frac{\partial(g_1, \cdots, g_k)}{\partial(x_{n-k+1}, \cdots, x_n)}(X_0) = 0. \end{cases}$$

合并得

$$\nabla f(X_0) + (\lambda_1, \cdots, \lambda_k) \frac{\partial(g_1, \cdots, g_k)}{\partial(x_1, \cdots, x_n)}(X_0) = 0,$$

即

$$\nabla f(X_0) + \lambda_1 \nabla g_1(X_0) + \cdots + +\lambda_k \nabla g_k(X_0) = 0.$$

Theorem (条件极值的必要条件)

设f及$g_1,\cdots,g_k$均属于C^1, $\dfrac{\partial(g_1,\cdots,g_k)}{\partial(x_1,\cdots,x_n)}$ 的秩在D内处处为k. 如果X_0是$f(X)$在条件(5)下的极值点, 则存在实数$\lambda_1,\cdots,\lambda_k$使得

$$\nabla f(X_0)+\lambda_1\nabla g_1(X_0)+\cdots+\lambda_k\nabla g_k(X_0)=0, \tag{6}$$

即

$$\nabla f(X_0)\in \operatorname{span}\{\nabla g_1(X_0),\cdots,\nabla g_k(X_0)\}.$$

S在X_0的所有切向量连同零向量组成的集合称为S在X_0点的切空间，记为$T_{X_0}S$. 则

$$\mathrm{span}\{\nabla g_1(X_0), \cdots, \nabla g_k(X_0)\} \perp T_{X_0}S.$$

在上述定理的条件下，$\dim T_{X_0}S = n-k$. 称

$$N_{X_0}S = \mathrm{span}\{\nabla g_1(X_0), \cdots, \nabla g_k(X_0)\}$$

为S在点X_0的法向量空间.

Lemma

在上述定理的条件下，对任何$X_0 \in S$有

$$T_{X_0}S = \operatorname{span}\{\nabla g_1(X_0), \cdots, \nabla g_k(X_0)\}^{\perp},$$

即线性方程组

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial g_k}{\partial x_1} & \cdots & \frac{\partial g_k}{\partial x_n} \end{pmatrix} \begin{pmatrix} \ell_1 \\ \vdots \\ \ell_n \end{pmatrix} = 0$$

的解空间为$(n-k)$维子空间.

证明 (i) 任取S上过X_0的光滑曲线

$$\Gamma : X = X(t), \quad t \in (a, b),$$

$X_0 = X(t_0)$, $t_0 \in (a, b)$. Γ在S上的充分必要条件为

$$g_i(X(t)) = 0, \quad \forall t \in (a, b), \quad i = 1, \cdots, k.$$

则

$$\langle \nabla g_i(X_0), X'(t_0) \rangle = 0, \quad i = 1, \cdots, k.$$

从而Γ在X_0的切向量$X'(t_0) \in \{\nabla g_1(X_0), \cdots, \nabla g_k(X_0)\}^{\perp}$. 由曲线$\Gamma$的任意性得

$$T_{X_0} S \subseteq \{\nabla g_1(X_0), \cdots, \nabla g_k(X_0)\}^{\perp}.$$

(ii) 反之, 任取非零向量

$$\vec{\ell} = (\ell_1, \cdots, \ell_n) \in \operatorname{span}\{\nabla g_1(X_0), \cdots, \nabla g_k(X_0)\}^{\perp},$$

$$\langle \nabla g_i(X_0), \vec{\ell} \rangle = 0, \quad i = 1, \cdots, k.$$

用矩阵表示等价于

$$\begin{pmatrix} \dfrac{\partial g_1}{\partial x_1} & \cdots & \dfrac{\partial g_1}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \dfrac{\partial g_k}{\partial x_1} & \cdots & \dfrac{\partial g_k}{\partial x_n} \end{pmatrix} \begin{pmatrix} \ell_1 \\ \vdots \\ \ell_n \end{pmatrix} = 0. \tag{7}$$

由于rank$\dfrac{\partial(g_1, \cdots, g_k)}{\partial(x_1, \cdots, x_n)}(X_0) = k$, 不妨设

$$\frac{D(g_1, \cdots, g_k)}{D(x_1, \cdots, x_k)}(X_0) \neq 0. \tag{8}$$

一方面，利用分块(7)可写为

$$\frac{\partial(g_1,\cdots,g_k)}{\partial(x_1,\cdots,x_k)}(X_0)\begin{pmatrix}\ell_1\\ \vdots\\ \ell_k\end{pmatrix}+\frac{\partial(g_1,\cdots,g_k)}{\partial(x_{k+1},\cdots,x_n)}(X_0)\begin{pmatrix}\ell_{k+1}\\ \vdots\\ \ell_n\end{pmatrix}=0,$$

$$\Rightarrow\ \begin{pmatrix}\ell_1\\ \vdots\\ \ell_k\end{pmatrix}=-\left[\frac{\partial(g_1,\cdots,g_k)}{\partial(x_1,\cdots,x_k)}(X_0)\right]^{-1}\frac{\partial(g_1,\cdots,g_k)}{\partial(x_{k+1},\cdots,x_n)}(X_0)\begin{pmatrix}\ell_{k+1}\\ \vdots\\ \ell_n\end{pmatrix}. \tag{9}$$

另一方面，由隐函数定理知，在X_0邻近S可表示为

$$\begin{cases} x_1=h_1(x_{k+1},\cdots,x_n),\\ \quad\cdots\\ x_k=h_k(x_{k+1},\cdots,x_n),\end{cases} \tag{10}$$

构造S上过X_0的光滑曲线$\Gamma(t)\in S$：

$$\Gamma:\begin{cases} x_1 = h_1(x_{0(k+1)}+\ell_{k+1}t,\cdots,x_{0n}+\ell_n t), \\ \quad\cdots \\ x_k = h_k(x_{0(k+1)}+\ell_{k+1}t,\cdots,x_{0n}+\ell_n t), \\ x_{k+1} = x_{0(k+1)}+\ell_{k+1}t, \\ \quad\cdots \\ x_n = x_{0n}+\ell_n t, \end{cases} \quad t\in(-\delta,\delta)$$

则$x_j'(t_0)=l_j,\ j=k+1,\cdots,n$, 且

$$\begin{pmatrix} x_1'(t_0) \\ \vdots \\ x_k'(t_0) \end{pmatrix} = \frac{\partial(h_1,\cdots,h_k)}{\partial(x_{k+1},\cdots,x_n)}(x_{0(k+1)},\cdots,x_{0n})\begin{pmatrix} \ell_{k+1} \\ \vdots \\ \ell_n \end{pmatrix}. \quad (11)$$

由$g_i(h_1(x_{k+1},\cdots,x_n),\cdots,h_k(x_{k+1},\cdots,x_n),x_{k+1},\cdots,x_n)=0.$

得

$$\frac{\partial(g_1,\cdots,g_k)}{\partial(x_1,\cdots,x_k)}(X_0)\frac{\partial(h_1,\cdots,h_k)}{\partial(x_{k+1},\cdots,x_n)}+\frac{\partial(g_1,\cdots,g_k)}{\partial(x_{k+1},\cdots,x_n)}(X_0)=0,$$

即

$$\frac{\partial(h_1,\cdots,h_k)}{\partial(x_{k+1},\cdots,x_n)}=-\left[\frac{\partial(g_1,\cdots,g_k)}{\partial(x_1,\cdots,x_k)}(X_0)\right]^{-1}\frac{\partial(g_1,\cdots,g_k)}{\partial(x_{k+1},\cdots,x_n)}(X_0).$$

从而比较(9)与(11)得

$$x_i'(t_0)=\ell_i,\quad i=1,\cdots,k,$$

由$\vec{\ell}$的任意性可知

$$\{\nabla g_1(X_0),\cdots,\nabla g_k(X_0)\}^{\perp}\subseteq T_{X_0}S.$$

在S上任取通过X_0的光滑曲线

$$\Gamma: \ X = X(t), \quad t \in (-\delta, \delta), \quad X_0 = X(0).$$

由于X_0是$f(X)$的条件极值点, 所以$t=0$是$f(X(t))$在$(-\delta,\delta)$内的极值点,从而

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(X_0) x_i'(0) = 0,$$

即$\langle \nabla f(X_0), X'(0) \rangle = 0$. 由$\Gamma$的任意性可得

$$\nabla f(X_0) \perp T_{X_0} S.$$

则由引理可知$\nabla f(X_0) \in \mathrm{span}\{\nabla g_1(X_0), \cdots, \nabla g_k(X_0)\}$, 即存在$\lambda_1, \cdots, \lambda_k$使得(6)成立.

条件临界点：对于$X=(x_1,\cdots,x_n)\in D$, 如果存在$\lambda_1,\cdots,\lambda_k$使得(5)和(6)成立, 则称$X$是函数$f$在条件(5)下的临界点. 条件极值点必是条件临界点.

$$\begin{cases} \dfrac{\partial f}{\partial x_i}(x_1,\cdots,x_n)+\displaystyle\sum_{l=1}^{k}\lambda_l\dfrac{\partial g_l}{\partial x_i}(x_1,\cdots,x_n)=0, i=1,\cdots,n, \\ g_j(x_1,\cdots,x_n)=0,\ j=1,\cdots,k. \end{cases}$$

这里$\lambda_1,\cdots,\lambda_k$称为**拉格朗日乘子**.

令$L(X)=f(X)+\sum_{i=1}^{k}\lambda_i g_i(X)$(**拉格朗日函数**), 则条件(6) 即为

$$\nabla L(X_0)=0.$$

这种求条件极值的方法称为**拉格朗日乘子法**.

Example

求函数$f(x,y)=2x^2+12xy+y^2$在有界闭区域

$$\overline{D}=\{(x,y)\in\mathbb{R}^2|\ x^2+4y^2\leqslant 25\}$$

上的最大值与最小值.

解　首先求$f(x,y)$在$D=\{(x,y)\in\mathbb{R}^2|\ x^2+4y^2<25\}$内的临界值. 令

$$\begin{cases} \dfrac{\partial f}{\partial x}=4x+12y=0, \\ \dfrac{\partial f}{\partial y}=12x+2y=0, \end{cases}$$

解之得$x=y=0$, 所以$f(x,y)$在D内唯一的临界值为0.

其次求$f(x,y)$在D的边界$x^2+4y^2=25$上的临界值. 用拉格朗日乘子法(令$L(x,y,\lambda)=2x^2+12xy+y^2+\lambda(x^2+4y^2-25)$):

$$\begin{cases} 4x+12y+2\lambda x=0, \\ 12x+2y+8\lambda y=0, \\ x^2+4y^2-25=0, \end{cases}$$

由于$(x,y)\neq(0,0)$,故

$$\begin{vmatrix} 4+2\lambda & 12 \\ 12 & 2+8\lambda \end{vmatrix} = 16\lambda^2+36\lambda-136=(4\lambda-8)(4\lambda+17)=0$$

解得$\lambda=-\dfrac{17}{4}$或2.

最后可得, $f(x,y)$在$\overline{D}$上的最大值是$\dfrac{425}{4}$, 最小值是-50.

Example

设$A=(a_{ij})$为n阶实对称方阵, 记其特征值为$\lambda_1<\cdots<\lambda_k$.
设l_i为λ_i的重数, 则$l_1+\cdots+l_k=n$. 定义$\mathbb{R}^n$上函数

$$h(X)=XAX^T=\sum_{i=1}^{n}\sum_{j=1}^{n}x_ia_{ij}x_j,\quad \forall X\in\mathbb{R}^n,$$

则

$$\lambda_1=\min_{|X|=1}XAX^T=\min_{X\neq 0}\frac{XAX^T}{|X|^2},$$

$$\lambda_k=\max_{|X|=1}XAX^T=\max_{X\neq 0}\frac{XAX^T}{|X|^2}.$$

证明 由连续性可知$h(X)$在$|X|=1$下的最小值存在.

设$X_0=(x_{01},\cdots,x_{0n})\in\mathbb{R}^n$, $|X_0|=1$, 使得

$$h(X_0)=\min_{|X|=1} h(X). \tag{12}$$

条件$|X|=1$即为$\sum\limits_{i=1}^{n} x_i^2=1$, 由必要条件可知存在实数$\lambda$使得

$$\frac{\partial h}{\partial x_i}(X_0)-2\lambda x_{0i}=0,\quad i=1,\cdots,n,$$

注意到$\nabla h(X)=XA+XA^T=X(A+A^T)=2XA$, 即得

$$2AX_0^T-2\lambda X_0^T=0.$$

从而λ是A的特征值, 且

$$h(X_0)=X_0AX_0^T=\lambda\langle X_0,X_0\rangle=\lambda|X_0|^2=\lambda.$$

对于A的任一特征值λ_i, 由于存在$X_i \in \mathbb{R}^n$, $|X_i| = 1$, 使得$AX_i^T = \lambda_i X_i^T$, 则$\lambda_i = h(X_i)$. 由(12)得$\lambda_i \geqslant \lambda$, 所以$\lambda = \lambda_1$.

同理可证, A的最大特征值λ_k是$h(X)$在$|X| = 1$下的最大值.

注记: 令$E_{\lambda_1,\cdots,\lambda_i} = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_i}$, 则

$$\lambda_{i+1} = \min_{X \in E^{\perp}_{\lambda_1,\cdots,\lambda_i}, |X|=1} XAX^T = \min_{X \in E^{\perp}_{\lambda_1,\cdots,\lambda_i} \setminus \{0\}} \frac{XAX^T}{|X|^2}.$$

Example

设$\alpha_1 > 0, x_i > 0, i = 1, \cdots, n$. 证明

$$x_1^{\alpha_1} \cdots x_n^{\alpha_n} \leq \left(\frac{\alpha_1 x_1 + \cdots + \alpha_n x_n}{\alpha_1 + \cdots + \alpha_n} \right)^{\alpha_1 + \cdots + \alpha_n}$$

等号成立当且仅当$x_1 = \cdots = x_n$.

提示：考虑函数$f(x_1, \cdots, x_n) = \sum_{i=1}^n \alpha_i \ln x_i$在条件$\sum_{i=1}^n \alpha_i x_i = c$下的条件极值.

Example

设$\alpha_1>0, x_i>0, i=1,\cdots,n$. 证明

$$x_1^{\alpha_1}\cdots x_n^{\alpha_n}\le\left(\frac{\alpha_1x_1+\cdots+\alpha_nx_n}{\alpha_1+\cdots+\alpha_n}\right)^{\alpha_1+\cdots+\alpha_n}$$

等号成立当且仅当$x_1=\cdots=x_n$.

提示： 考虑函数$f(x_1,\cdots,x_n)=\sum_{i=1}^n\alpha_i\ln x_i$在条件$\sum_{i=1}^n\alpha_ix_i=c$下的条件极值.

Theorem

设$D\subseteq\mathbb{R}^n$为开区域, $f,g_1,\cdots,g_k\in C^2(D)$,且$\operatorname{rank}\frac{\partial(g_1,\cdots,g_k)}{\partial(x_1,\cdots,x_n)}=k$.
设$X_0\in D$是$f(X)$在条件(5)下的极值点,
记$L(X)=f(X)+\lambda_1 g_1(X)+\cdots+\lambda_k g_k(X)$, 则
(i) 若X_0是$f(X)$在条件(5)下的极小值点，则

$$VH_L(X_0)V^T\geqslant 0,\quad \forall V\in T_{X_0}S.$$

(ii) 若X_0是$f(X)$在条件(5)下的极大值点，则

$$VH_L(X_0)V^T\leqslant 0,\quad \forall V\in T_{X_0}S.$$

证明 任取$V\in T_{X_0}S$, 在S上任取通过X_0的光滑曲线

$$\Gamma: X=X(t),\quad t\in(-\delta,\delta),$$

$X_0=X(0)$, $X'(0)=V\in T_{X_0}S$. 记$\Delta X=X(t)-X(0)$, 由条件(5)和泰勒公式，对于$i=1,\cdots,k$有

$$\begin{aligned}0 &= g_i(X(t))-g_i(X_0)=g_i(X_0+\Delta X)-g_i(X_0)\\ &= \langle\nabla g_i(X_0),\Delta X\rangle+\frac{1}{2}\Delta X H_{g_i}(X_0)\Delta X^T+o(|\Delta X|^2).\end{aligned}$$

$$f(X(t))-f(X_0)=\langle\nabla f(X_0),\Delta X\rangle+\frac{1}{2}\Delta X H_f(X_0)\Delta X^T+o(|\Delta X|^2),$$

上面两式相加得到

$$f(X(t))-f(X_0)=\langle\nabla L(X_0),\Delta X\rangle+\frac{1}{2}\Delta X H_L(X_0)\Delta X^T+o(|\Delta X|^2),$$

由$\nabla L(X_0)=\nabla f(X_0)+\lambda_1\nabla g_1(X_0)+\cdots+\lambda_k\nabla g_k(X_0)=0$得

$$f(X(t))-f(X_0)=\frac{1}{2}\Delta X H_L(X_0)\Delta X^T+o(|\Delta X|^2).$$

由于$\Delta X=X(t)-X(0)=X'(0)t+o(|t|)$, 所以

$$f(X(t))-f(X_0)=\frac{1}{2}X'(0)H_L(X_0)X'(0)^Tt^2+o(|t|^2). \tag{13}$$

若X_0是$f(X)$在条件(5)下的极小值点, 则存在正数$\eta>0$ 使得

$$f(X(t))-f(X_0)\geqslant 0, \quad \forall t\in(-\eta,\eta).$$

由(13)可得

$$VH_L(X_0)V^T\geqslant 0,$$

即(i)成立.

Theorem

在必要条件定理的条件下,

(i) 若$H_L(X_0)$在$T_{X_0}S$上正定, 则X_0是$f(X)$满足(5)的极小值点.

(ii) 若$H_L(X_0)$在$T_{X_0}S$上负定, 则X_0是$f(X)$满足(5)的极大值点.

(iii) 若存在$V_1, V_2 \in T_{X_0}S$使得

$$V_1 H_L(X_0) V_1^T > 0, \quad V_2 H_L(X_0) V_2^T < 0,$$

则X_0不是$f(X)$在条件(5)下的极值点.

(iv) 其余情况不能判定X_0是否为$f(X)$在条件(5)下的极值点.

易得$f(X_0+Y)-f(X_0)=\frac{1}{2}YH_L(X_0)Y^T+o(|Y|^2)$.
不妨设

$$\frac{D(g_1,\cdots,g_k)}{D(x_1,\cdots,x_k)}(X_0)\neq 0.$$

记$Y=(Y_1,Y_2)=(y_1,\cdots,y_k,y_{k+1},\cdots,y_n)$, 对于方程组

$$\begin{cases} g_1(X_0+Y)=0 \\ \cdots \\ g_k(X_0+Y)=0 \end{cases}$$

由隐函数定理，存在$\delta,\eta>0$及向量值函数Φ使得

$$Y_1^T=\Phi(Y_2)=\begin{pmatrix} \phi_1(Y_2) \\ \vdots \\ \phi_k(Y_2) \end{pmatrix},\ \Phi(O)=0,\ |Y_2|<\delta,\ |Y_1|<\eta$$

和

$$\begin{aligned} J_\Phi(O) &= \frac{\partial(\phi_1,\cdots,\phi_k)}{\partial(y_{k+1},\cdots,y_n)}(O) \\ &= -\left(\frac{\partial(g_1,\cdots,g_k)}{\partial(x_1,\cdots,x_k)}(X_0)\right)^{-1}\frac{\partial(g_1,\cdots,g_k)}{\partial(x_{k+1},\cdots,x_n)}(X_0). \end{aligned}$$

由$\Phi(Y_2)=\Phi(O)+J_\Phi(O)Y_2^T+o(|Y_2|)$得到

$$Y_1^T=-\left(\frac{\partial(g_1,\cdots,g_k)}{\partial(x_1,\cdots,x_k)}(X_0)\right)^{-1}\frac{\partial(g_1,\cdots,g_k)}{\partial(x_{k+1},\cdots,x_n)}(X_0)Y_2^T+o(|Y_2|). \quad (14)$$

令$Z=(Z_1,Z_2)$,其中

$$\begin{cases} Z_1^T=-\left(\dfrac{\partial(g_1,\cdots,g_k)}{\partial(x_1,\cdots,x_k)}(X_0)\right)^{-1}\dfrac{\partial(g_1,\cdots,g_k)}{\partial(x_{k+1},\cdots,x_n)}(X_0)Y_2^T, \\ Z_2^T=Y_2^T. \end{cases}$$

由(9)及其后的计算，易知

$$Z \in \{\nabla g_1(X_0), \cdots, \nabla g_k(X_0)\}^{\perp} = T_{X_0}S,$$

且$|Y - Z| = o(|Y_2|)$.

由(14)知$|Y|$与$|Y_2|$同阶无穷小(因$|Y_1|$比$|Y_2|$高阶或同阶), 同理$|Y|$与$|Z|$同阶无穷小, 即存在常数$M > 0$使得

$$M^{-1}|Z| \leqslant |Y| \leqslant M|Z|. \tag{15}$$

从而由$f(X_0 + Y) - f(X_0) = \dfrac{1}{2}YH_L(X_0)Y^T + o(|Y|^2)$得

$$f(X_0 + Y) - f(X_0) = \frac{1}{2}ZH_L(X_0)Z^T + o(|Z|^2). \tag{16}$$

若$H_L(X_0)$在$T_{X_0}S$上正定,则存在正数$a>0$使得

$$VH_F(X_0)V^T \geq a|V|^2, \quad \forall V \ \in T_{X_0}S.$$

由(15)和(16)可知存在$\epsilon>0$使得

$$f(X_0+Y)-f(X_0) \geq \frac{a}{4}|Z|^2, \qquad |Y|<\epsilon.$$

于是当$X_0+Y\in S,\ 0<|Y|<\epsilon$时

$$f(X_0+Y)-f(X_0)>0,$$

即X_0是$f(X)$在条件(5)下的严格极小值点.

第一步：利用$T_{X_0}S=\text{span}\{\nabla g_1(X_0),\cdots,\nabla g_k(X_0)\}^{\perp}$求出切空间,即$T_{X_0}S$为方程组$\frac{\partial(g_1,\cdots,g_k)}{\partial(x_1,\cdots,x_n)}(X_0)(y_1,\cdots,y_n)^T=0$ 的解空间.

第二步：利用充分条件和$H_L(X_0)|_{T_{X_0}S}$的符号判定是否为极值.

Example

考察函数$f(x,y,z)=x^2+y^2-z^2$在条件

$$S: g_1(x,y,z)=z-1=0 \tag{17}$$

下的极值.

解(按照注记1) 设$L(x,y,z)=f(x,y,z)+\lambda g_1(x,y,z)$，由

$$\begin{cases}\dfrac{\partial f}{\partial x}+\lambda\dfrac{\partial g_1}{\partial x}=2x=0,\\ \dfrac{\partial f}{\partial y}+\lambda\dfrac{\partial g_1}{\partial y}=2y=0,\\ \dfrac{\partial f}{\partial z}+\lambda\dfrac{\partial g_1}{\partial z}=-2z+\lambda=0,\\ z-1=0\end{cases}$$

解得$x=0,y=0,z=1,\lambda=2$，即$X_0=(0,0,1)$是唯一的临界点.

记$L(x,y,z)=f(x,y,z)+2g_1(x,y,z)=x^2+y^2-z^2+2z-2$,则

$$H_L(X_0)=\begin{pmatrix} L''_{xx} & L''_{xy} & L''_{xz} \\ L''_{yx} & L''_{yy} & L''_{yz} \\ L''_{zx} & L''_{zy} & L''_{zz} \end{pmatrix}=\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

显然

$$T_{X_0}S=\{(x,y,z)\in\mathbb{R}^3 |\ (x,y)\in\mathbb{R}^2, z=0\}.$$

$H_L(X_0)$在$T_{X_0}S$上正定, 但$H_L(X_0)$在$\mathbb{R}^3$上不定.

注记2：$(y_1,\cdots,y_k)^T=A(y_{k+1},\cdots,y_n)^T$,其中

$$A=-\left(\frac{\partial(g_1,\cdots,g_k)}{\partial(x_1,\cdots,x_k)}(X_0)\right)^{-1}\frac{\partial(g_1,\cdots,g_k)}{\partial(x_{k+1},\cdots,x_n)}(X_0),$$

则$Y=(Y_2A^T,Y_2)=Y_2(A^T,I_{n-k})=Y_2B^T$, 即得

$$\begin{aligned}T_{X_0}S &= \left\{Y=(y_1,\cdots,y_n)\middle|\,\frac{\partial(g_1,\cdots,g_k)}{\partial(x_1,\cdots,x_n)}Y^T=0\right\}\\ &= \{Y=(Y_2A^T,Y_2)\equiv Y_2B^T|\ Y_2\in\mathbb{R}^{n-k}\},\end{aligned}$$

这里$B=\begin{pmatrix}A\\ I_{n-k}\end{pmatrix}$. 对于$Z=Y_2B^T\in T_{X_0}S$,

$$ZH_L(X_0)Z^T=Y_2B^TH_L(X_0)BY_2^T.$$

因此$H_L(X_0)|_{T_{X_0}S}>0$等价于矩阵$B^TH_L(X_0)B>0$.

Example

求函数$f(x,y,z,t)=x+y+z+t$在限制条件$xyzt=1$下的极值.

解(按照注记2求解) 令$L=x+y+z+t+\lambda(xyzt-1)$, 解

$$\begin{cases} \nabla L=0 \\ xyzt=1 \end{cases}$$

可得$\lambda=\pm1$, $X_0=\mp(1,1,1,1,)$. 当$X_0=(1,1,1,1)$, $\lambda=-1$时,

$$H_L(X_0)=-\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

令$g(x,y,z,t)=xyzt-1$, $g'_x(X_0)=1$, $(g'_y,g'_z,g'_t)(X_0)=(1,1,1)$, 从而$A=-g'_x(X_0)^{-1}(g'_y,g'_z,g'_t)(X_0)=(-1,-1,-1)$, 故

$$B=\begin{pmatrix}-1&-1&-1\\1&0&0\\0&1&0\\0&0&1\end{pmatrix}, B^TH_L(X_0)B=\begin{pmatrix}2&1&1\\1&2&1\\1&1&2\end{pmatrix}.$$

显然,其一阶、二阶、三阶顺序主子式都大于零，从而其正定. 故$f(1,1,1,1)=4$为$f(x,y,z,t)$在条件$xyzt=1$下的极小值.

同理可得$f(-1,-1,-1,-1)=-4$为$f(x,y,z,t)$在条件$xyzt=1$下的极大值.

注记3：

令$Y=\Delta X=(\Delta x_1,\cdots,\Delta x_n)=(\mathrm{d}x_1,\cdots,\mathrm{d}x_n)=\mathrm{d}X$，从而

$$\begin{aligned} T_{X_0}S &= \{\Delta X\in\mathbb{R}^n|\ J_G(X_0)\Delta X^T=0\} \\ &= \{\mathrm{d}X\in\mathbb{R}^n|\ J_G(X_0)\mathrm{d}X^T=\mathrm{d}G(X_0)=0\}. \end{aligned}$$

因此$H_L(X_0)|_{T_{X_0}S}>0$ 等价于

$$\mathrm{d}^2L=\sum_{i,j}\frac{\partial^2 L}{\partial x_i\partial x_j}(X_0)\mathrm{d}x_i\mathrm{d}x_j>0,$$

$$\forall\ (\mathrm{d}x_1,\cdots,\mathrm{d}x_n)$$

$$\in\left\{(\mathrm{d}x_1,\cdots,\mathrm{d}x_n)\Big|(\mathrm{d}x_1,\cdots,\mathrm{d}x_n)\neq 0,\mathrm{d}G(X_0)=0\right\}.$$