§9. 多元函数的(无条件)极值

首先回顾一元函数的极值问题.设f充分光滑.

$$f(x_0)\text{极小} \Rightarrow f'(x_0)=0,$$

$$f(x_0)\text{极大} \Rightarrow f'(x_0)=0,$$

$$\left.\begin{array}{l} f'(x_0)=0 \\ f''(x_0)>0 \end{array}\right\} \Rightarrow f(x_0)\text{极小}, \quad \left.\begin{array}{l} f'(x_0)=0 \\ f''(x_0)<0 \end{array}\right\} \Rightarrow f(x_0)\text{极大}.$$

$$f(x_0)\text{极小} \Rightarrow f''(x_0)\geq 0, \quad f(x_0)\text{极大} \Rightarrow f''(x_0)\leq 0.$$

研究极值问题的根本方法是Taylor展开. 例如

$$\left.\begin{array}{l} f'(x_0)=f''(x_0)=f'''(x_0)=0 \\ f^{(4)}(x_0)>0 \end{array}\right\} \Rightarrow f(x_0)\text{极小},$$

$$\left.\begin{array}{l} f'(x_0)=f''(x_0)=f'''(x_0)=0 \\ f^{(4)}(x_0)<0 \end{array}\right\} \Rightarrow f(x_0)\text{极大},$$

$$\left.\begin{array}{l} f'(x_0)=f''(x_0)=0 \\ f'''(x_0)\neq 0 \end{array}\right\} \Rightarrow f(x_0)\text{不是极值}.$$

1.极值的定义与必要性

Def. n元函数f在$\mathrm{x}_0\left(\in\mathbb{R}^n\right)$的某个邻域$U$中有定义,

若$\forall \mathrm{x}\in U, \mathrm{x}\neq \mathrm{x}_0$, 都有

$$f(\mathrm{x})(>)\geq f(\mathrm{x}_0),$$

则称$f(\mathrm{x}_0)$为f的一个(严格)极小值, 称x_0为f的一个(严格)极小值点. 若$\forall \mathrm{x}\in U, \mathrm{x}\neq \mathrm{x}_0$, 都有

$$f(\mathrm{x})(<)\leq f(\mathrm{x}_0),$$

则称$f(\mathrm{x}_0)$为f的一个(严格)极大值, 称x_0为f的一个(严格)极大值点.

Thm. n元函数f在$\mathrm{x}_0=\left(x_1^0,x_2^0,\cdots,x_n^0\right)\in\mathbb{R}^n$的某个邻域中可微, $f(\mathrm{x}_0)$极小(大),则x_0为f的一个驻点,即$\mathrm{grad}f(\mathrm{x}_0)=0$.

Proof. $f(\mathrm{x}_0)$极小,一元函数$f\left(x_1,x_2^0,\cdots,x_n^0\right)$在$x_1^0$取到极小值,从而$\dfrac{\partial f}{\partial x_1}(\mathrm{x}_0)=0$.同理,$\dfrac{\partial f}{\partial x_k}(\mathrm{x}_0)=0,k=2,3,\cdots n$.于是$\mathrm{grad}f(\mathrm{x}_0)=0$.□

Remark: 对一般的函数f, $f(x_0)$极小, x_0不一定为驻点. 例如$f(x,y)=\sqrt{x^2+y^2}$, $f(0,0)$极小,但$(0,0)$不是f的驻点(偏导数不存在).

Remark: $\mathrm{grad} f(x_0)=0$, 但x_0不一定是f的极值点.例如, $f(x,y)=x^2-y^2$, $\mathrm{grad} f(0,0)=0$, 但$(0,0)$不是f的极值点.

Remark: 极大(小)值不一定是最大(小)值.反之,最大(小)值一定是极大(小)值.

2. 矩阵的正定性

$$P=(a_{ij})_{n\times n}\in M_{n\times n},(P^T=P),$$

$$\mathrm{x}=\left(x_1,x_2,\cdots,x_n\right)^T\in\mathbb{R}^n.$$

$$二次型:\mathrm{x}^T P\mathrm{x}=\sum_{i,j=1}^{n}a_{ij}x_ix_j$$

Def. 设$P\in M_{n\times n}, P^T=P$,

称P正定(负定),若$\forall \mathrm{x}\in\mathbb{R}^n, \mathrm{x}\neq 0, \mathrm{x}^T P\mathrm{x}>(<)0$.

称P半正定(半负定),若$\forall \mathrm{x}\in\mathbb{R}^n, \mathrm{x}\neq 0, \mathrm{x}^T P\mathrm{x}\geq(\leq)0$.

称P不定,若$\exists \mathrm{x},\mathrm{y}\in\mathbb{R}^n, \mathrm{x}^T P\mathrm{x}>0, \mathrm{y}^T P\mathrm{y}<0$.

Thm. $P \in M_{n\times n}, P^T = P$,则

P正定 $\Leftrightarrow P$的每个主子式 > 0

$\Leftrightarrow P$的每个顺序主子式 > 0

$\Leftrightarrow P$的所有特征值都 > 0

P半正定 $\Leftrightarrow P$的每个主子式 ≥ 0

$\Leftrightarrow P$的所有特征值都 ≥ 0.

3.极值的充分条件

Lemma1.设n阶实对称阵A的所有特征值为$\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$,则 $\lambda_1 \|\mathrm{x}\|^2 \le \mathrm{x}^{\mathrm{T}}\mathrm{A}\mathrm{x} \le \lambda_n \|\mathrm{x}\|^2$, $\forall \mathrm{x} \in \mathbb{R}^n$.

Proof.A实对称阵, 则存在正交矩阵Q,$s.t.$

$$\mathrm{Q}^{\mathrm{T}}\mathrm{A}\mathrm{Q} = \Lambda = diag(\lambda_1, \lambda_2, \cdots, \lambda_n).$$

令$\mathrm{x} = \mathrm{Q}\mathrm{y}$,则

$$\|\mathrm{x}\|^2 = \mathrm{x}^{\mathrm{T}}\mathrm{x} = \mathrm{y}^{\mathrm{T}}\mathrm{Q}^{\mathrm{T}}\mathrm{Q}\mathrm{y} = \mathrm{y}^{\mathrm{T}}\mathrm{y} = \|\mathrm{y}\|^2,$$

$$\mathrm{x}^{\mathrm{T}}\mathrm{A}\mathrm{x} = \mathrm{y}^{\mathrm{T}}\mathrm{Q}^{\mathrm{T}}\mathrm{A}\mathrm{Q}\mathrm{y} = \mathrm{y}^{\mathrm{T}}\Lambda\mathrm{y} = \lambda_1 \mathrm{y}_1^2 + \cdots + \lambda_n \mathrm{y}_n^2,$$

$$\lambda_1 \|\mathrm{y}\|^2 \le \mathrm{x}^{\mathrm{T}}\mathrm{A}\mathrm{x} \le \lambda_n \|\mathrm{y}\|^2,$$

故$\lambda_1 \|\mathrm{x}\|^2 \le \mathrm{x}^{\mathrm{T}}\mathrm{A}\mathrm{x} \le \lambda_n \|\mathrm{x}\|^2$, $\forall \mathrm{x} \in \mathbb{R}^n$.□

Thm. n元函数f在x_0的邻域中二阶连续可微,

$\mathrm{grad}f(x_0)=0$,

(1)若$H_f(x_0)$正定,则$f(x_0)$严格极小.

(2)若$H_f(x_0)$负定,则$f(x_0)$严格极大.

(3)若$H_f(x_0)$不定,则$f(x_0)$不是极值.

Proof:记$\Delta x=x-x_0$,因$\mathrm{grad}f(x_0)=0$,由Taylor公式,

$$\Delta f(x_0)=f(x_0+\Delta x)-f(x_0)$$

$$=\frac{1}{2}(\Delta x)^{T}H_f(x_0)\Delta x+o(\|\Delta x\|^2),\ \Delta x\to 0\text{时}.$$

$H_f(\mathrm{x}_0)$为实对称阵,故其所有特征值都是实的,

设为$\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$.

(1)若$H_f(\mathrm{x}_0)$正定,则$0 < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$,

$$\Delta f(\mathrm{x}_0) \ge \frac{1}{2}\lambda_1 \|\Delta \mathrm{x}\|^2 + o(\|\Delta \mathrm{x}\|^2), \Delta \mathrm{x} \to 0\text{时},$$

因此,$\exists \delta > 0$,当$\|\Delta \mathrm{x}\| < \delta$时,$\Delta f(\mathrm{x}_0) > 0$,即$\mathrm{x}_0$为$f$的严格极小值点.

(2)若$H_f(\mathrm{x}_0)$负定,则$\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n < 0$. 同上可证$\mathrm{x}_0$为$f$的严格极大值点.

(3)若$H_f(\mathrm{x}_0)$不定,则$\lambda_1 < 0 < \lambda_n$. 设$\alpha,\beta$为对应于$\lambda_1,\lambda_n$的单位长度的特征向量,则

$$\alpha^{\mathrm{T}} H_f(\mathrm{x}_0)\alpha = \lambda_1\|\alpha\|^2 = \lambda_1, \quad \beta^{\mathrm{T}} H_f(\mathrm{x}_0)\beta = \lambda_n\|\beta\|^2 = \lambda_n.$$

令$\Delta\mathrm{x} = t\alpha$,则 $\Delta f(\mathrm{x}_0) = \dfrac{1}{2}\lambda_1 t^2 + o(t^2),\ t \to 0$时,

故在x_0的任意小邻域中,总$\exists\mathrm{x}$, $s.t.\ f(\mathrm{x}) < f(\mathrm{x}_0)$.

令$\Delta\mathrm{x} = t\beta$,则 $\Delta f(\mathrm{x}_0) = \dfrac{1}{2}\lambda_n t^2 + o(t^2),\ t \to 0$时,

故在x_0的任意小邻域中,总$\exists\mathrm{x}$, $s.t.\ f(\mathrm{x}) > f(\mathrm{x}_0)$.

综上, x_0不是f的极值点.□

Thm. n元函数f在x_0的邻域中二阶连续可微.

(1) $f(\mathrm{x}_0)$极小, 则$H_f(\mathrm{x}_0)$的所有特征值均≥ 0.

(2) $f(\mathrm{x}_0)$极大, 则$H_f(\mathrm{x}_0)$的所有特征值均≤ 0.

Proof: (1) $f(\mathrm{x}_0)$极小,则$\mathrm{grad}f(\mathrm{x}_0)=0$. 当$\Delta\mathrm{x}\to 0$时,

$$f(\mathrm{x}_0+\Delta\mathrm{x})-f(\mathrm{x}_0)=\frac{1}{2}(\Delta\mathrm{x})^{\mathrm{T}}H_f(\mathrm{x}_0)\Delta\mathrm{x}+o(\|\Delta\mathrm{x}\|^2).$$

若$H=H_f(\mathrm{x}_0)$有特征值$\lambda<0$, 设$H\alpha=\lambda\alpha,\|\alpha\|=1$.则

$$f(\mathrm{x}_0+t\alpha)-f(\mathrm{x}_0)=\frac{1}{2}\lambda t^2+o(t^2),\quad (t\to 0\text{时}).$$

$|t|$充分小时, $f(\mathrm{x}_0+t\alpha)-f(\mathrm{x}_0)<0$,与$f(\mathrm{x}_0)$极小矛盾.

同理可证(2).□

Remark 判断多元函数的驻点是否为极值点, 关键在于研究函数在这一点的*Hasse*矩阵的正定性.

Thm. 设$f(x, y)$在$M_0(x_0, y_0)$的邻域中二阶连续可微, $\mathrm{grad} f(x_0, y_0) = 0$, 记

$$A = \frac{\partial^2 f(M_0)}{\partial x^2}, B = \frac{\partial^2 f(M_0)}{\partial x \partial y}, C = \frac{\partial^2 f(M_0)}{\partial y^2},$$

则1)若$A > 0, AC - B^2 > 0$, 则$f(x_0, y_0)$严格极小.

2)若$A < 0, AC - B^2 > 0$, 则$f(x_0, y_0)$严格极大.

3)若$AC - B^2 < 0$, 则$f(x_0, y_0)$不是f的极值.

Remark: 当$AC-B^2=0$时, $f(x_0,y_0)$可能不是f的极值, 也可能是f的极大值或极小值. 例如:

$f(x,y)$	$H_f(0,0)$	$(0,0)$
x^2+y^3	$\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$	不是f的极值点
$x^2+x^2y^2$	$\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$	是f的极小值点.
$-x^2-x^2y^2$	$\begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix}$	是f的极大值点.

Remark: $f \in C^2(D)$,(x_0, y_0)为D的内点, 则

$$f(x_0, y_0)\text{极小} \Rightarrow \begin{cases} f''_{xx}(x_0, y_0) \ge 0 \\ f''_{yy}(x_0, y_0) \ge 0 \end{cases}$$

$$f(x_0, y_0)\text{极大} \Rightarrow \begin{cases} f''_{xx}(x_0, y_0) \le 0 \\ f''_{yy}(x_0, y_0) \le 0 \end{cases}$$

(Hint: 考虑一元函数$f(x, y_0)$和$f(x_0, y)$的极值.)

Remark:求函数f的极值, 先求出f的所有驻点, 再逐个判断他们是否为极值点.

4.例题

例: $x^2+y^2+z^2-2x+2y-4z-10=0$确定隐函数 $z=z(x,y)$.求$z(x,y)$的极值.

分析:Step1. 求$z=z(x,y)$的驻点..

Step2.求驻点处的Hasse矩阵, 判断是否为极值点.

解:视方程

$$x^2+y^2+z^2-2x+2y-4z-10=0$$

中$z=z(x,y)$,分别对x和y求偏导,得

$$2x+2zz'_x-2-4z'_x=0 \qquad (1)$$

$$2y+2zz'_y+2-4z'_y=0 \qquad (2)$$

于是

$$z'_x=\frac{x-1}{2-z},\ z'_y=\frac{y+1}{2-z}.$$

驻点为$(x,y)=(1,-1)$,对应$z=-2$,或$z=6$.

(1)式分别对x,y求偏导,得

$$2+2z_x'^2+2zz_{xx}''-4z_{xx}''=0,$$

$$2z_x'z_y'+2zz_{xy}''-4z_{xy}''=0,$$

(2)对y求偏导,得

$$2+2z_y'^2+2zz_{yy}''-4z_{yy}''=0.$$

当$(x,y,z)=(1,-1,-2)$时,

$$z_{xx}''=1/4,z_{xy}''=0,z_{yy}''=1/4.$$

$H=\begin{pmatrix}1/4 & 0\\ 0 & 1/4\end{pmatrix}$正定,故$z=-2$为极小值.

当$(x,y,z)=(1,-1,6)$时,

$$z''_{xx}=-1/4,\ z''_{xy}=0,\ z''_{yy}=-1/4.$$

$H=\begin{pmatrix}-1/4 & 0\\ 0 & -1/4\end{pmatrix}$负定,故$z=6$为极大值. □

例. 求$f=(x^2+y^2)e^{-(x^2+y^2)}$的极值.

解: Step1, 求驻点.由

$$\begin{cases} f'_x=2x(1-x^2-y^2)e^{-(x^2+y^2)}=0\\ f'_y=2y(1-x^2-y^2)e^{-(x^2+y^2)}=0\end{cases}$$

得驻点$(0,0)$或$x^2+y^2=1.$

Step2. 求Hasse矩阵, 极值判断

$$f''_{xx}=[2(1-3x^2-y^2)-4x^2(1-x^2-y^2)]e^{-(x^2+y^2)}$$

$$f''_{yy}=[2(1-x^2-3y^2)-4y^2(1-x^2-y^2)]e^{-(x^2+y^2)}$$

$$f''_{xy}=-4xy(2-x^2-y^2)e^{-(x^2+y^2)}$$

• 当$(x,y)=(0,0)$时, $f''_{xx}=2$, $f''_{xy}=0$, $f''_{yy}=2.$

$$H=\begin{pmatrix}2 & 0\\ 0 & 2\end{pmatrix}\text{正定, } f(0,0)\text{极小.}$$

• 当$x^2+y^2=1$时,

$$f''_{xx}=-4x^2e^{-1},\ f''_{xy}=-4xye^{-1},\ f''_{yy}=-4y^2e^{-1}.$$

$\det\begin{pmatrix} f''_{xx} & f''_{xy} \\ f''_{xy} & f''_{yy} \end{pmatrix}=0$, 不能直接判断$f(x,y)$是否为极值.

令$t=x^2+y^2$, 则

$$f(x,y)=(x^2+y^2)e^{-(x^2+y^2)}=te^{-t}\triangleq g(t)$$

由$g'(t)=(1-t)e^{-t}=0$得驻点$t=1$. $g''(t)=(t-2)e^{-t}$,

$g''(1)=-e^{-1}<0$. $g(t)$在$t=1$时有极大值$g(1)=e^{-1}$.

从而$f(x,y)$当$x^2+y^2=1$时有极大值e^{-1}.□

例: 求$z=x^4+y^4-2x^2+4xy-2y^2$的极值.

解: $z'_x=4x^3-4x+4y,\ \ z'_y=4y^3+4x-4y.$

得驻点$(\sqrt{2},-\sqrt{2}),(-\sqrt{2},\sqrt{2}),(0,0).$

$$z''_{xx}=12x^2-4,\quad z''_{xy}=4,\quad z''_{yy}=12y^2-4.$$

(1)在$(\sqrt{2},-\sqrt{2})$,

$$A=C=20, B=4,\ AC-B^2>0,$$

取得极小值.

(2)同理$z(x,y)$在$(-\sqrt{2},\sqrt{2})$取得极小值.

(3)在$(0,0)$,

$$A=C=-4, B=4, \quad AC-B^2=0,$$

判别法失效. 注意到

$$z(x,x)=2x^4>0, \text{当} x\neq 0 \text{时}.$$

$$z(x,0)=x^4-2x^2$$

$$=x^2(x^2-2)<0, \text{当} 0<x^2<2 \text{时}.$$

故$(0,0)$不是极值点.□

例. $\lim\limits_{\substack{(x,y)\to(0,0)\\ x\neq\sin y}}\dfrac{f(x,y)-f(0,0)}{(x-\sin y)^2}=\mathrm{A}>0,\ f$连续.

$(0,0)$是否为f的极值点?

解:$\exists\delta>0, s.t.$

$$\frac{f(x,y)-f(0,0)}{(x-\sin y)^2}>\frac{\mathrm{A}}{2},\quad \forall x^2+y^2\le\delta,\ x\neq\sin y.$$

由f的连续性, $\forall x^2+y^2\le\delta$, 有

$$f(x,y)-f(0,0)\ge \mathrm{A}(x-\sin y)^2/2\ge 0.$$

故$(0,0)$为f的极小值点. □

Remark: 考虑 $f(x,y)=f(0,0)+(x-\sin y)^2$.

例:f连续, $\lim\limits_{(x,y)\to(0,0)} \dfrac{f(x,y)-xy}{(x^2+y^2)^2}=1$. $f(0,0)$是否极值?

解: $\lim\limits_{(x,y)\to(0,0)} \big(f(x,y)-xy\big)=0,\ f(0,0)=0.$

存在$\varepsilon>0$, 当$x^2+y^2<\varepsilon$时,

$$\frac{3}{2}(x^2+y^2)^2 > f(x,y)-xy > \frac{1}{2}(x^2+y^2)^2.$$

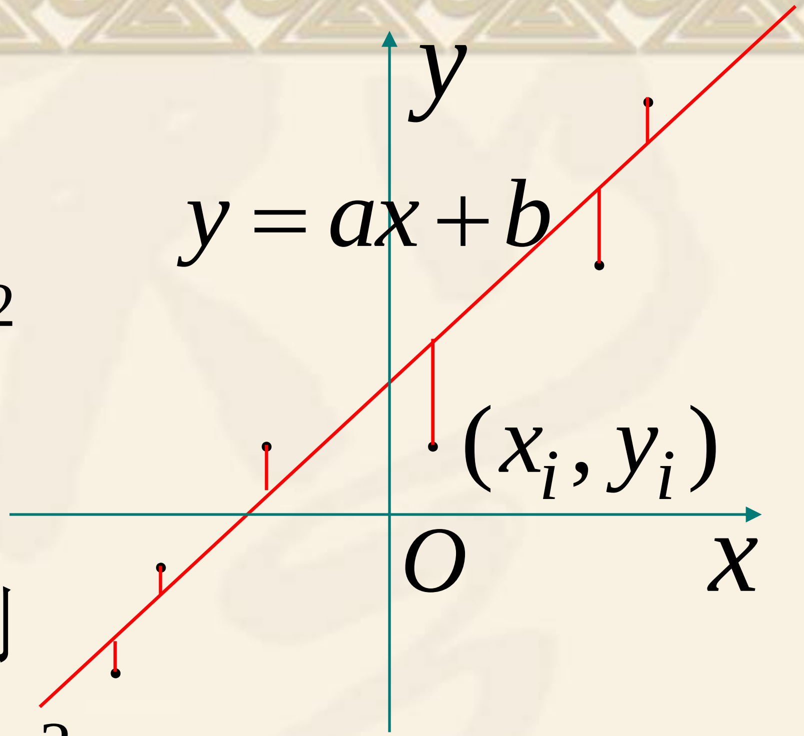
于是对充分大的n, $f\left(\dfrac{1}{n},\dfrac{1}{n}\right) > \dfrac{1}{n^2}+\dfrac{2}{n^4}>0,$

$$f\left(\frac{1}{n},-\frac{1}{n}\right) < -\frac{1}{n^2}+\frac{6}{n^4} = -\frac{1}{n^2}(1-\frac{6}{n^2})<0.$$

故$f(0,0)$不是极值. □

例.（最小二乘法）

分析: 使误差的平方和最小.



解: $f(a,b) = \sum_{i=1}^{n} [(ax_i + b) - y_i]^2$

Step1.证明$f(a,b)$有最小值.

记$A = \sum_{i=1}^{n} x_i^2, B = \sum_{i=1}^{n} x_i$,则

$$f(a,b) = Aa^2 + 2Bab + nb^2 + Da + Eb + G$$

且 $\lim\limits_{a^2+b^2 \to +\infty} f(a,b) = +\infty$.

故$\exists R > 0$,当$a^2 + b^2 > R^2$时,$f(a,b) > f(0,0)$.从而f在$a^2 + b^2 \le R^2$上的最小值就是全局最小值.

Step2.求$f(a,b)$的最小值点.由

$$\begin{cases} f_a' = -2\sum_{i=1}^{n}(y_i - ax_i - b)x_i = 0 \\ f_b' = -2\sum_{i=1}^{n}(y_i - ax_i - b) = 0 \end{cases}$$

得f的唯一驻点

$$a = \frac{n\sum_{i=1}^{n}x_i y_i - B\sum_{i=1}^{n}y_i}{nA - B^2}, b = \frac{A\sum_{i=1}^{n}y_i - B\sum_{i=1}^{n}x_i y_i}{nA - B^2}.$$

$Hasse$矩阵$\mathrm{H}_f = \begin{pmatrix} 2A & 2B \\ 2B & 2n \end{pmatrix}$正定, 驻点为极小值点.

而f的最小值点必为极小值点,因此f唯一的极小值点就是f的最小值点.□

作业:习题1.9 No.1,2