

§ 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.

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

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

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

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

研究极值问题的根本方法是Taylor展开. 例如

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

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

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

1.极值的定义与必要性

Def. n 元函数 f 在 $x_0 (\in \mathbb{R}^n)$ 的某个邻域 U 中有定义,
若 $\forall x \in U, x \neq x_0$,都有

$$f(x)(>) \geq f(x_0),$$

则称 $f(x_0)$ 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 $\forall x \in U, x \neq x_0$,都有

$$f(x)(<) \leq f(x_0),$$

则称 $f(x_0)$ 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 $x_0 = (x_1^0, x_2^0, \dots, x_n^0) \in \mathbb{R}^n$ 的某个邻域中可微, $f(x_0)$ 极小(大), 则 x_0 为 f 的一个驻点, 即 $\text{grad}f(x_0) = 0$.

Proof. $f(x_0)$ 极小, 一元函数 $f(x_1, x_2^0, \dots, x_n^0)$ 在 x_1^0 取到极小值, 从而 $\frac{\partial f}{\partial x_1}(x_0) = 0$. 同理, $\frac{\partial f}{\partial x_k}(x_0) = 0, k = 2, 3, \dots, n$. 于是 $\text{grad}f(x_0) = 0$. \square

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

Remark: 定义域内部的极大(小)值不一定是最大(小)值; 反之, 若最大(小)值在定义域内部取到, 则一定是极大(小)值.

2. 矩阵的正定性

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

$$\mathbf{x} = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n.$$

$$\text{二次型: } \mathbf{x}^T P \mathbf{x} = \sum_{i,j=1}^n a_{ij} x_i x_j$$

Def. 设 $P \in M_{n \times n}$, $\textcolor{red}{P}^T = \textcolor{red}{P}$,

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

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

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

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 $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$, 则 $\lambda_1 \|\mathbf{x}\|^2 \leq \mathbf{x}^T A \mathbf{x} \leq \lambda_n \|\mathbf{x}\|^2, \forall \mathbf{x} \in \mathbb{R}^n$.

Proof. A 实对称阵, 则存在正交矩阵 $Q, s.t.$

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

令 $\mathbf{x} = Q\mathbf{y}$, 则 $\|\mathbf{x}\|^2 = \mathbf{x}^T \mathbf{x} = \mathbf{y}^T Q^T Q \mathbf{y} = \mathbf{y}^T \mathbf{y} = \|\mathbf{y}\|^2,$

$$\mathbf{x}^T A \mathbf{x} = \mathbf{y}^T Q^T A Q \mathbf{y} = \mathbf{y}^T \Lambda \mathbf{y} = \lambda_1 y_1^2 + \dots + \lambda_n y_n^2,$$

$$\lambda_1 \|\mathbf{y}\|^2 \leq \mathbf{x}^T A \mathbf{x} \leq \lambda_n \|\mathbf{y}\|^2,$$

故 $\lambda_1 \|\mathbf{x}\|^2 \leq \mathbf{x}^T A \mathbf{x} \leq \lambda_n \|\mathbf{x}\|^2, \forall \mathbf{x} \in \mathbb{R}^n. \square$

Thm. n 元函数 f 在 \mathbf{x}_0 的邻域中二阶连续可微,
 $\text{grad}f(\mathbf{x}_0) = 0$,

(1)若 $H_f(\mathbf{x}_0)$ 正定,则 $f(\mathbf{x}_0)$ 严格极小.

(2)若 $H_f(\mathbf{x}_0)$ 负定,则 $f(\mathbf{x}_0)$ 严格极大.

(3)若 $H_f(\mathbf{x}_0)$ 不定,则 $f(\mathbf{x}_0)$ 不是极值.

Proof:记 $\Delta\mathbf{x} = \mathbf{x} - \mathbf{x}_0$,因 $\text{grad}f(\mathbf{x}_0) = 0$,由Taylor公式,

$$\Delta f(\mathbf{x}_0) = f(\mathbf{x}_0 + \Delta\mathbf{x}) - f(\mathbf{x}_0)$$

$$= \frac{1}{2}(\Delta\mathbf{x})^T H_f(\mathbf{x}_0)\Delta\mathbf{x} + o(\|\Delta\mathbf{x}\|^2), \Delta\mathbf{x} \rightarrow 0 \text{时}.$$

$H_f(\mathbf{x}_0)$ 为实对称阵,故其所有特征值都是实的,
设为 $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$.

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

$$\Delta f(\mathbf{x}_0) \geq \frac{1}{2} \lambda_1 \|\Delta \mathbf{x}\|^2 + o(\|\Delta \mathbf{x}\|^2), \Delta \mathbf{x} \rightarrow 0 \text{ 时},$$

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

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

(3)若 $H_f(\mathbf{x}_0)$ 不定,则 $\lambda_1 < 0 < \lambda_n$.设 α, β 为对应于 λ_1, λ_n 的单位长度的特征向量,则

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

$$\text{令 } \Delta \mathbf{x} = t\alpha, \text{ 则 } \Delta f(\mathbf{x}_0) = \frac{1}{2} \lambda_1 t^2 + o(t^2), \quad t \rightarrow 0 \text{ 时,}$$

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

$$\text{令 } \Delta \mathbf{x} = t\beta, \text{ 则 } \Delta f(\mathbf{x}_0) = \frac{1}{2} \lambda_n t^2 + o(t^2), \quad t \rightarrow 0 \text{ 时,}$$

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

综上, \mathbf{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) $f(x_0)$ 极小, 则 $H_f(x_0)$ 的所有特征值均 ≥ 0 .

(2) $f(x_0)$ 极大, 则 $H_f(x_0)$ 的所有特征值均 ≤ 0 .

Proof: (1) $f(x_0)$ 极小, 则 $\text{grad}f(x_0) = 0$. 当 $\Delta x \rightarrow 0$ 时,

$$f(x_0 + \Delta x) - f(x_0) = \frac{1}{2}(\Delta x)^T H_f(x_0) \Delta x + o(\|\Delta x\|^2).$$

若 $H = H_f(x_0)$ 有特征值 $\lambda < 0$, 设 $H\alpha = \lambda\alpha, \|\alpha\| = 1$. 则

$$f(x_0 + t\alpha) - f(x_0) = \frac{1}{2}\lambda t^2 + o(t^2), \quad (t \rightarrow 0 \text{ 时}).$$

$|t|$ 充分小时, $f(x_0 + t\alpha) - f(x_0) < 0$, 与 $f(x_0)$ 极小矛盾.

同理可证(2). \square

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上的 *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, $\text{grad}f(x_0, y_0) = 0$, 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
$x^2 + x^2 y^2$	$\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$	是 f 的极小值点.
$-x^2 - x^2 y^2$	$\begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix}$	是 f 的极大值点.

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

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

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

(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Hesse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

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

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

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

$$2 + 2z'_x{}^2 + 2zz''_{xx} - 4z''_{xx} = 0,$$

$$2z'_xz'_y + 2zz''_{xy} - 4z''_{xy} = 0,$$

(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,

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

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

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Hesse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,

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

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

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} . \square

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),

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

判别法失效. 注意到

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

$$\begin{aligned} z(x, 0) &= x^4 - 2x^2 \\ &= x^2(x^2 - 2) < 0, \text{当} 0 < x^2 < 2 \text{时.} \end{aligned}$$

故(0,0)不是极值点.□

例. $\lim_{\substack{(x,y) \rightarrow (0,0) \\ x \neq \sin y}} \frac{f(x,y) - f(0,0)}{(x - \sin y)^2} = A > 0, f \text{ 连续.}$

$(0,0)$ 是否为 f 的极值点?

解: $\exists \delta > 0, s.t.$

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

由 f 的连续性, $\forall x^2 + y^2 \leq \delta$, 有

$$f(x,y) - f(0,0) \geq A(x - \sin y)^2 / 2 \geq 0.$$

故 $(0,0)$ 为 f 的极小值点. \square

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

例: f 连续, $\lim_{(x,y) \rightarrow (0,0)} \frac{f(x,y) - xy}{(x^2 + y^2)^2} = 1$. $f(0,0)$ 是否极值?

解: $\lim_{(x,y) \rightarrow (0,0)} (f(x,y) - xy) = 0, f(0,0) = 0$.

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 , $f\left(\frac{1}{n}, \frac{1}{n}\right) > \frac{1}{n^2} + \frac{2}{n^4} > 0$,

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

故 $f(0,0)$ 不是极值. \square

例. (最小二乘法)

分析: 使误差的平方和最小.

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

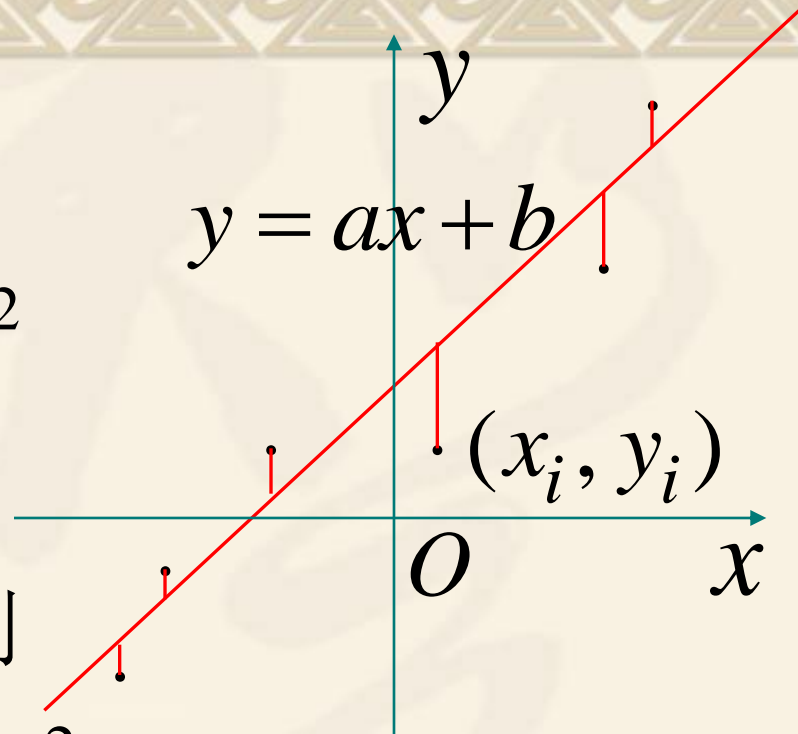
Step1. 证明 $f(a, b)$ 有最小值.

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 $\lim_{a^2 + b^2 \rightarrow +\infty} f(a, b) = +\infty$.

故 $\exists R > 0$, 当 $a^2 + b^2 > R^2$ 时, $f(a, b) > f(0, 0)$. 从而 f 在 $a^2 + b^2 \leq R^2$ 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

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

而 f 的最小值点必为极小值点, 从而是驻点, 因此 f 唯一的驻点就是 f 的最小值点. \square



作业：习题1.9 No. 1, 2