9.5.3 多元函数的极值

定义 1 设 $p_0 \in D \subset \mathbb{R}^n$, $f : D \to \mathbb{R}$. 如果存在 $\delta > 0$, 使得

$$f(p) \geqslant f(p_0) \quad (f(p) \leqslant f(p_0)), \ (\forall\, p \in B(p_0, \delta) \cap D),$$

则称 p_0 是 f 的一个极小 (大) 值点. $f(p_0)$ 是一个极小 (大) 值. 若上面的不等式是严格的, 则 p_0 称为严格极小 (大) 值点.

定义 2 设 $D \subset \mathbb{R}^n$, $f : D \to \mathbb{R}$, $p_0 \in D^\circ$. 如果 f 在 p_0 的各一阶偏导数存在, 且 $Jf(p_0) = 0$, 则称 p_0 是 f 的一个驻点.

驻点不一定是极值点.

例 1 设 $f(x, y) = xy$. 则

$$\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0,$$

所以 $(0, 0)$ 是 f 的一个驻点. 但在 $(0, 0)$ 附近 x, y 同号时, f 为正, x, y 异号时, f 为负. 因此, $(0, 0)$ 不是极值点.

定理 1 设 $D \subset \mathbb{R}^n$, $f: D \to \mathbb{R}$, $p \in D^\circ$. 如果 f 在 p 取极值, 且 f 在 p 的各一阶偏导数存在, 则 p 是 f 的一个驻点.

证明 令 $p = (p_1, \cdots, p_n)$, 则一元函数 $\varphi(t) = f(t, p_2, \cdots, p_n)$ 在内点 p_1 取极值. 又因为 $\varphi'(p_1) = \frac{\partial f}{\partial x_1}(p_1, p_2, \cdots, p_n)$ 存在, 所以由一元函数的极值定理, 知 $\varphi'(p_1) = 0$, 即,

$$\frac{\partial f}{\partial x_1}(p) = 0.$$

同理,

$$\frac{\partial f}{\partial x_i}(p) = 0, \ i = 2, 3 \cdots, n.$$

所以 $Jf(p) = 0$, 即 p 是一个驻点. 证毕.

在一元函数的情形, 若 f 在驻点的邻域中有二阶导数, 且驻点处有正 (负) 的二阶导数, 则驻点是严格极小 (大) 值点. 由于多元函数的 Hessian 对应着一元函数的二阶导数, 因此, Hessian 的 “正” 或 “负” 在极值点的讨论中就很重要.

定义 3 设 $A=(a_{ij})$ 是一个 n 阶对称方阵, $x=(x_1,\cdots,x_n)\subset\mathbb{R}^n$. 称

$$Q(x)=xAx^T=\sum_{i,j=1}^{n}a_{ij}x_ix_j$$

为 $x_1,\cdots,x_n$ 的一个二次型. 方阵 A 称为二次型 $Q(x)$ 的系数方阵.

1° 如果对一切非零 x, 有 $Q(x)>0$ ($Q(x)<0$), 则称 $Q(x)$ 是正定的 (负定的). 也称 A 是正定的 (负定的).

2° 如果 $Q(x)$ 既有正值也有负值, 则称 $Q(x)$ 为不定的. 也称 A 是不定的

3° 如果对一切 x, 有 $Q(x)\geqslant 0$ ($Q(x)\leqslant 0$), 且对某些非零 x 有 $Q(x)=0$, 则称 $Q(x)$ 是半正定的 (半负定的). 也称 A 是半正定的 (半负定的).

例 2 $Q(x,y)=2x^2+2xy+y^2$ 是一个二元正定二次型. 系数方阵

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

是正定的.

定理 2 对称方阵 $A=(a_{ij})$ 是正定的 $\Longleftrightarrow$ 顺序主子式都大于零, 即

$$a_{11}>0,\quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}>0,\quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}>0,\ \cdots\ \det A>0.$$

$A=(a_{ij})$ 是负定的 $\Longleftrightarrow$ $-A$ 是正定的, 即

$$a_{11}<0,\quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}>0,\quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}<0,\ \cdots\ (-1)^n\det A>0.$$

例 3 设 $A=\begin{pmatrix} 3 & -2 \\ -2 & 2 \end{pmatrix}$, $B=\begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix}$, $C=\begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$. 则 A 是正定的, B 是负定的, C 是不定的.

定理 3 设 n 元函数 f 在 x_0 的一个邻域内有连续的二阶偏导数, 且 x_0 是 f 的一个驻点.

1° 如果 Hessian $Hf(x_0)$ 是正定的 (负定的), 那么 x_0 是 f 的一个严格极小值点 (极大值点).

2° 如果 Hessian $Hf(x_0)$ 是不定的, 那么 x_0 不是 f 的极值点.

推论 1 设二元函数 $f(x, y)$ 在 (x_0, y_0) 的一个邻域内有连续的二阶偏导数, 且 (x_0, y_0) 是 f 的一个驻点. 记

$$A = \frac{\partial^2 f}{\partial x^2}(x_0, y_0),\ B = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0),\ C = \frac{\partial^2 f}{\partial y^2}(x_0, y_0).$$

1° 如果 $A > 0$, $AC - B^2 > 0$, 那么 (x_0, y_0) 是 f 的一个严格极小值点;

2° 如果 $A < 0$, $AC - B^2 > 0$, 那么 (x_0, y_0) 是 f 的一个严格极大值点;

3° 如果 $AC - B^2 < 0$, 那么 (x_0, y_0) 不是 f 的极值点.

定理 3 的证明 1° 因为 x_0 是 f 的一个驻点, 且 f 在 x_0 的一个邻域内是 C^2 的, 所以当 $\|h\|$ 小时, 存在 $\theta \in (0,1)$ 使得

$$\begin{aligned} f(x_0+h) &= f(x_0) + \frac{1}{2} hHf(x_0+\theta h)h^T \\ &= f(x_0) + \frac{1}{2} hHf(x_0)h^T + \frac{1}{2} h\Big(Hf(x_0+\theta h) - Hf(x_0)\Big)h^T \end{aligned}$$

因为 $Hf(x_0)$ 是正定的, 所以当 $h \neq 0$ 时, $hHf(x_0)h^T > 0$. 注意到

$$\varphi(h) = hHf(x_0)h^T$$

是 h 的连续函数. 记 $\mathbb{R}^n$ 中的单位球面为

$$S = \{h \in \mathbb{R}^n : \|h\| = 1\}$$

这是 $\mathbb{R}^n$ 中的有界闭集. 令 $m = \min\limits_{h\in S} \varphi(h)$, 则 $m > 0$.

对一般的 $h \neq 0$, 有 $\frac{h}{\|h\|} \in S$. 因此

$$hHf(x_0)h^T = \left(\frac{h}{\|h\|}\right) Hf(x_0) \left(\frac{h}{\|h\|}\right)^T \cdot \|h\|^2 \geqslant m\|h\|^2.$$

因为 $Hf(x)$ 在 x_0 的邻域内连续, 所以存在 $\delta > 0$, 使得当 $\|h\| < \delta$ 时, 有

$$\|Hf(x_0+\theta h) - Hf(x_0)\| < m.$$

因而

$$\|h\Big(Hf(x_0+\theta h) - Hf(x_0)\Big)h^T\| < m\|h\|^2, \quad (n \neq 0).$$

于是当 $\|h\| < \delta$ 且 $h \neq 0$ 时, 有

$$f(x_0+h) - f(x_0) \geqslant \frac{1}{2}m\|h\|^2 - \frac{1}{2}\|h\Big(Hf(x_0+\theta h) - Hf(x_0)\Big)h^T\| > 0.$$

所以 x_0 是 f 的一个极小值点.

2° 根据 Taylor 公式, 在 x_0 的邻域内有

$$f(x_0+h) - f(x_0) = \frac{1}{2}hHf(x_0)h^T + o(\|h\|^2).$$

因为 $Hf(x_0)$ 是不定方阵, 所以存在 $p, q \in \mathbb{R}^n$ 使得

$$pHf(x_0)p^T < 0 < qHf(x_0)q^T.$$

对充分小的 $\varepsilon > 0$, 分别取 $h = \varepsilon p$ 和 $h = \varepsilon q$ 得

$$\begin{aligned} f(x_0 + \varepsilon p) - f(x_0) &= \frac{\varepsilon^2}{2} pHf(x_0)p^T + o(\varepsilon^2) \\ &= \frac{\varepsilon^2}{2}\left(pHf(x_0)p^T + o(1)\right) < 0 \end{aligned}$$

$$\begin{aligned} f(x_0 + \varepsilon q) - f(x_0) &= \frac{\varepsilon^2}{2} qHf(x_0)q^T + o(\varepsilon^2) \\ &= \frac{\varepsilon^2}{2}\left(qHf(x_0)q^T + o(1)\right) > 0 \end{aligned}$$

所以 x_0 不是 f 的极值点.

例 4 讨论 $f(x,y)=x^3+y^3-3xy$ 的极值点情况.

解 因为

$$f'_x(x,y)=3x^2-3y,\quad f'_y(x,y)=3y^2-3x,$$

所以 f 有两个驻点 $P_0=(0,0)$ 和 $P_1=(1,1)$. 又因为

$$\begin{aligned}A&=f''_{xx}(x,y)=6x,\\ B&=f''_{xy}(x,y)=-3,\\ C&=f''_{yy}(x,y)=6y,\end{aligned}$$

所以在 P_0 有 $\Delta=AC-B^2<0$, 因此 P_0 不是极值点. 在 P_1 有 $A=6>0$, $\Delta=27>0$, 因此 P_1 是极小值点.

例 5 求函数 $f(x,y)=\sin x\sin y\sin(x+y)$ 在闭区域 $D:\ x\geqslant 0, y\geqslant 0, x+y\leqslant\pi$ 上的最大值和最小值.

解 显然 f 在 D 上非负, 在 D 的边界上为零, 在 D 内部为正. 因为 f 连续, 因此在 D 上有最大值, 最大值点 M 必在 D 内部. 因为 f 可微, 所以 M 是 f 的驻点. 因为

$$f'_x=\sin y\sin(2x+y),\ f'_y=\sin x\sin(x+2y),$$

而 $\sin x$ 和 $\sin y$ 在 D 内部为正, $x+2y$ 和 $2x+y$ 在 $(0,2\pi)$ 中, 因此在内部驻点有

$$2x+y=x+2y=\pi,$$

即, $x=y=\frac{\pi}{3}$. 这是内部唯一的驻点. 于是 $M=(\frac{\pi}{3},\frac{\pi}{3})$. 所以 f 在 D 上的最大值为 $f(M)=\frac{3\sqrt{3}}{8}$. 最小值为零.

9.5.4 条件极值

条件极值是多元函数极值问题中的一种, 比如要求空间中一张曲面 $\varphi(x,y,z)=0$ 上到原点的距离最近的点, 就是要使得函数 $f(x,y,z)=\sqrt{x^2+y^2+z^2}$ 取得最小, 然而, 这里的量 x,y,z 不再是完全自由的, 而必须是在曲面上, 即它们受到曲面方程 $\varphi(x,y,z)=0$ 的约束. 再比如, 如果一个函数 $f(x,y)$ 的极值点在其定义域的边界上, 这时前面所讨论的求极值的方法就不适用了. 如果函数的定义域是 Oxy 平面上一条曲线 $\varphi(x,y)=0$ 所围成的区域, 那么函数在边界上取到极值的问题就化为求函数 $f(x,y)$ 的极值, 其中点 (x,y) 满足 $\varphi(x,y)=0$.

一般情况下, 设函数 $z=f(x,y)$, 在区域 $D\subset\mathbb{R}^2$ 上有连续的微商, 求函数 $f(x,y)$ 在限制条件

$$\varphi(x,y)=0$$

下的极值问题称为条件极值问题.

如果方程 $\varphi(x,y)=0$ 存在隐函数, 例如有隐函数 $y=y(x)$, 则它的微商是 $y'(x)=-\frac{\varphi_x'(x,y)}{\varphi_y'(x,y)}$. 代入 $z=f(x,y)$ 得

$$z=z(x)=f(x,y(x))$$

因此条件极值就变成了一元函数 $z=z(x)$ 的极值问题. 设 $x=a$ 是它的一个驻点, 记 $b=y(a)$, 则

$$z'(a)=f_x'(a,b))+f_y'(a,b)y'(a)=f_x'(a,b)-f_y'(a,b)\frac{\varphi_x'(a,b)}{\varphi_y'(a,b)}=0$$

记

$$\lambda=-\frac{f_x'(a,b)}{\varphi_x'(a,b)}=-\frac{f_y'(a,b)}{\varphi_y'(a,b)}$$

就有

$$(f_x'+\lambda\varphi_x')|_{(a,b)}=0,$$
$$(f_y'+\lambda\varphi_y')|_{(a,b)}=0.$$

也就是说存在数 λ 使得 (a,b) 是函数 $f(x,y)+\lambda\varphi(x,y)$ 的驻点.

因此在求函数 $z = f(x, y)$ 在约束条件 $\varphi(x, y) = 0$ 之下的条件极值时, 可以先引进辅助函数

$$F(x, y) = f(x, y) + \lambda\varphi(x, y).$$

此时, 条件极值点应满足驻点方程

$$\begin{cases} F_x'(x, y) = f_x'(x, y) + \lambda\varphi_x'(x, y) = 0, \\ F_y'(x, y) = f_y'(x, y) + \lambda\varphi_y'(x, y) = 0, \\ \varphi(x, y) = 0. \end{cases}$$

因此, 欲求条件极值, 可以首先从上述方程中解出驻点, 然后再根据题意, 去判别哪些驻点是极值点. 这种方法称为 Lagrange **乘数法**.

对于有更多变元的函数的条件极值问题, 也可以同样用 Lagrange 乘数法来讨论.

例 6 求在约束条件 $(x-1)^2+y^2-1=0$ 下 $z=xy$ 的极值.

解 作辅助函数

$$F(x,y)=xy+\lambda\big((x-1)^2+y^2-1\big).$$

于是得到驻点方程

$$\begin{cases} y+2\lambda(x-1)=0, \\ x+2\lambda y=0, \\ (x-1)^2+y^2-1=0. \end{cases}$$

解得驻点 $M_1(0,0)$, $M_2\left(\frac{3}{2},\ \frac{\sqrt{3}}{2}\right)$, $M_3\left(\frac{3}{2},\ -\frac{\sqrt{3}}{2}\right)$.

显然 M_1 不是 $z=xy$ 的极值点. 又因为连续函数 $z=xy$ 在有界闭集 $D=\{(x,y)\,|\,(x-1)^2+y^2-1=0\}$ 上必能取到最大值和最小值, 而且最大值点和最小值点都是驻点. 所以 $z_2=\frac{3\sqrt{3}}{4}$ 是最大值, 也是极大值, 而 $z_3=-\frac{3\sqrt{3}}{4}$ 是最小值, 也是极小值.

例 7 求由原点到曲面 $(x-y)^2 - z^2 = 1$ 的最短距离.

解 考虑原点到曲面上的点 (x, y, z) 的距离平方 $d^2 = x^2 + y^2 + z^2$, 则问题就化成求函数

$$u = x^2 + y^2 + z^2$$

在限制条件 $(x-y)^2 - z^2 = 1$下的最小值. 按乘数法作辅助函数

$$F(x, y, z) = x^2 + y^2 + z^2 + \lambda[(x-y)^2 - z^2 - 1],$$

并求得驻点方程组

$$\begin{cases} \dfrac{\partial F}{\partial x} = 2x + 2\lambda(x-y) = 0, \\ \dfrac{\partial F}{\partial y} = 2y - 2\lambda(x-y) = 0, \\ \dfrac{\partial F}{\partial z} = 2z - 2\lambda z = 0, \\ (x-y)^2 - z^2 - 1 = 0. \end{cases}$$

由第三方程得 $z(1-\lambda) = 0$. 但当 $\lambda = 1$ 时容易看出, 这组方程不相容, 所以

只能有 $z = 0$. 从而解得

$$x = \pm\frac{1}{2}, \quad y = -x = \mp\frac{1}{2}.$$

于是点 $\left(\frac{1}{2}, -\frac{1}{2}, 0\right)$ 与 $\left(-\frac{1}{2}, \frac{1}{2}, 0\right)$ 便可能是极值点. 由问题本身的意义知最小值一定是存在, 而函数在这两点上取相同的值 $\frac{1}{2}$. 因此这两点都是函数的最小值点, 并得出所求的最短距离 $d = \frac{\sqrt{2}}{2}$.

例 8 试将正数 a 分成 n 个正数的和, 使这 n 个正数的乘积最大.

解 设 a 分成的 n 个正数为 $x_1, x_2, \cdots, x_n$, 则问题就成为在限制条件 $x_1 + x_2 + \cdots + x_n = a$ 下求函数 $u = x_1 x_2 \cdots x_n$ 的最大值. 作辅助函数

$$F(x_1, x_2, \cdots, x_n) = x_1 x_2 \cdots x_n + \lambda(x_1 + x_2 + \cdots + x_n - a),$$

得驻点方程组

$$\begin{aligned}
&\frac{\partial F}{\partial x_1} = x_2 x_3 \cdots x_n + \lambda = 0,\\
&\frac{\partial F}{\partial x_2} = x_1 x_3 \cdots x_n + \lambda = 0,\\
&\cdots\cdots,\\
&\frac{\partial F}{\partial x_n} = x_1 x_2 \cdots x_{n-1} + \lambda = 0.\\
&x_1 + x_2 + \cdots + x_n = a
\end{aligned}$$

比较这些等式可知

$$x_1 = x_2 = \cdots = x_n.$$

代入限制条件求得

$$x_1 = x_2 = \cdots = x_n = \frac{a}{n}.$$

从题意知道最大值一定存在, 因此这个唯一可能的极值点就是使函数取最大值的点. 从而推得, 若将正数 a 分成 n 个相等的正数, 则这 n 个正数的乘积最大, 其最大值为 $\left(\frac{a}{n}\right)^n$.

从上述结果还可以得到一个重要的不等式. 由于

$$x_1x_2\cdots x_n \leqslant \left(\frac{a}{n}\right)^n = \left(\frac{x_1+x_2+\cdots x_n}{n}\right)^n,$$

所以

$$\sqrt[n]{x_1x_2\cdots x_n} \leqslant \frac{x_1+x_2+\cdots x_n}{n},$$

即 n 个正数的几何平均值不大于它们的算术平均值.