《微积分A2》第七讲

教师　杨利军

清华大学数学科学系

2020年03月09日

两个曲面交线(曲线)的切线

设两个隐式曲面 $S_1 : F(x, y, z) = 0$ 和 $S_2 : G(x, y, z) = 0$ 的交线为曲线 C, 这里函数 F, G 假设在开区域 $D \subset \mathbb{R}^3$ 上是 C^1 的. 设 $P_0 = (x_0, y_0, z_0) \in C = S_1 \cap S_2$ 是曲线 C 上的一点. 以下说明曲线 C 在点 P_0 处的切线必位于这两个曲面在点 P_0 处的切平面上. 设曲线 C 有正则表示 $r(t) = (x(t), y(t), z(t))$, 并且 $(x(t_0), y(t_0), z(t_0)) = (x_0, y_0, z_0)$. 由于曲线 C 位于曲面 S_1 上, 故 $F(x(t), y(t), z(t)) \equiv 0$. 求导并令 $t = t_0$ 即得

$$F_x^0 x'(t_0) + F_y^0 y'(t_0) + F_z^0 z'(t_0) = 0.$$

这说明曲线 C 在点 P_0 处的切线

$$\frac{x-x_0}{x'(t_0)}=\frac{y-y_0}{y'(t_0)}=\frac{z-z_0}{z'(t_0)} \qquad (*)$$

位于曲面 $S_1: F(x,y,z)=0$ 在点 (x_0,y_0,z_0) 处的切平面

$$F_x^0(x-x_0)+F_y^0(y-y_0)+F_z^0(z-z_0)=0$$

上. 同理, 切线(*)也位于曲面 $S_2: G(x,y,z)=0$ 在点 (x_0,y_0,z_0) 处的切平面

$$G_x^0(x-x_0)+G_y^0(y-y_0)+G_z^0(z-z_0)=0$$

上.

因此曲线 C 在点 P_0 处的切线可表示为这两个切平面的交线，

即

$$\begin{cases} F_x^0(x-x_0)+F_y^0(y-y_0)+F_z^0(z-z_0)=0, \\ G_x^0(x-x_0)+G_y^0(y-y_0)+G_z^0(z-z_0)=0. \end{cases}$$

这里自然应假设 $\nabla F^0 \times \nabla G^0 \neq 0.$

例: 求曲线 C

$$\begin{cases} x^2 + y^2 + z^2 = 6, \\ z - x^2 - y^2 = 0, \end{cases}$$

在点 $P_0 = (1, 1, 2)$ 处的切线方程.

解: 记 $F(x, y, z) = x^2 + y^2 + z^2 - 6$, $G(x, y, z) = z - x^2 - y^2$.

简单计算得

$$\begin{cases} \nabla F = (2x, 2y, 2z), & \nabla F^0 = (2, 2, 4), \\ \nabla G = (-2x, -2y, 1), & \nabla G^0 = (-2, -2, 1). \end{cases}$$

例子续

于是所求的切线方程为

$$\begin{cases} 2(x-1)+2(y-1)+4(z-2)=0, \\ -2(x-1)-2(y-1)+(z-2)=0. \end{cases}$$

化简得

$$\begin{cases} x+y+2z=6, \\ 2x+2y-z=2. \end{cases}$$

注意

$$\nabla F^0 \times \nabla G^0 = (2,2,4)\times(-2,-2,1) = (10,-10,0) \neq 0.$$

例子完毕.

定理：设一元函数 $\phi(t)$ 在开区间 $J=(a,b)$ 上是 C^n 的，则对任意 $t_0, t \in J$,

$$\phi(t)=\phi(t_0)+\frac{\phi'(t_0)(t-t_0)}{1!}+\frac{\phi''(t_0)(t-t_0)^2}{2!}$$

$$+\cdots+\frac{\phi^{(n)}(t_0)(t-t_0)^n}{n!}+R_n(t),$$

其中 $R_n(t)$ 称为余项，有两种常用的表达式，Peano 余项 $R_n(t) = o(|t-t_0|^n)$ 和 Lagrange 余项(要求 ϕ 是 C^{n+1})

$$R_n(t)=\frac{\phi^{(n+1)}(\xi)}{(n+1)!}(t-t_0)^{n+1}, \quad \xi \text{ 位于 } t_0 \text{ 和 } t \text{ 之间}.$$

多元函数的二阶 Taylor 展式

定理: 设 $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ 为 n 元函数, C^2, 定义域 Ω 既开且凸, (凸的意思是, 连接 Ω 中任意两点的线段均包含在 Ω 中.) 对任意点 $x_0, x \in \Omega$, 记 $h = x - x_0$, 则

$$f(x) = f(x_0) + \nabla f(x_0)h + \frac{1}{2}h^T H_f(\xi)h,$$

$$\text{或 } f(x) = f(x_0) + \nabla f(x_0)h + \frac{1}{2}h^T H_f(x_0)h + o(\|h\|^2),$$

其中 $\xi = x_0 + \theta h$, $\theta \in (0, 1)$,

$$H_f(x) = \left[\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\right]_{n\times n}.$$

注记

<u>注一</u>: 上述两个等式分别称为 $f(x)$ 在点 x_0 处, 带 Lagrange 余项的二阶展式, 以及带 Peano 余项的二阶展式.

<u>注二</u>: $\nabla F(x_0)h$ 可看作梯度向量 $\nabla F(x_0)$ 和向量 h 的内积, 或者 $\nabla F(x_0)$ 可看作 Jacobian 矩阵(行向量), h 看作列向量.

<u>注三</u>: 对称矩阵 $H_f(x)$ 称为函数 $f(x)$ 的 Hesse 矩阵.

考虑二元函数 $f(x, y)$ 情形. 记 $h = x - x_0$, $k = y - y_0$, 则 $f(x, y)$ 的 Lagrange 二阶展式如下

$$f(x, y) = f^0 + f_x^0 h + f_y^0 k + \frac{1}{2}[h, k]\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}_{(\xi,\eta)} \begin{bmatrix} h \\ k \end{bmatrix},$$

其中 $(\xi, \eta) = (x_0, y_0) + \theta(h, k)$, $\theta \in (0, 1)$, f^0, f_x^0 和 f_y^0 表示 f, f_x, f_y 在点 (x_0, y_0) 处取值.

Peano 展式, 情形 $n = 2$

二元函数 $f(x, y)$ 的 Peano 二阶展式如下

$$f(x, y) = f^0 + f_x^0 h + f_y^0 k + \frac{1}{2}[h, k]\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}^0 \begin{bmatrix} h \\ k \end{bmatrix} + o(\rho^2),$$

其中矩阵右上标 0 表示在点 (x_0, y_0) 处取值, $\rho^2 = h^2 + k^2$.

例子

例: 求二元函数 $f(x,y)=x/y$ 在点 $(1,1)$ 处的二阶 Taylor 展式, 带 Peano 余项 $o(\rho^2)$, $\rho^2=(x-1)^2+(y-1)^2$.

解法一: 求出一阶偏导数 f_x, f_y, 以及二阶偏导数 f_{xx}, f_{xy} 和 f_{yy}, 并计算出它们在点 $(1,1)$ 处的值, 就可得到所要求的 Peano 展式. 细节略.

解法二: 记 $h = x - 1$, $k = y - 1$, 则

$$f(x,y) = \frac{x}{y} = \frac{1+h}{1+k} = (1+h)(1-k+k^2-k^3+\cdots)$$

$$= (1+h)(1-k+k^2) + o(\rho^2) = 1 + h - k - hk + k^2 + o(\rho^2)$$

$$= 1 + (x-1) - (y-1) - (x-1)(y-1) + (y-1)^2 + o(\rho^2).$$

解法二的依据是如下二元函数的 Taylor 展式唯一性. 这个结论对一般 n 元高阶 Taylor 展式同样成立.

定理：设 $f(x,y)$ 在开圆盘 $\Omega:(x-x_0)^2+(y-y_0)^2<r^2$ 上是 C^2 的. 若 $f(x,y)$ 在 Ω 上可表为

$$f(x,y)=a_0+a_1h+a_2k+\frac{1}{2}[h,k]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} h \\ k \end{bmatrix}+o(\rho^2),$$

其中 $h=x-x_0$, $k=y-y_0$, $a_{12}=a_{21}$, $\rho^2=h^2+k^2$, 则上述展式就是 $f(x,y)$ 在点 (x_0,y_0) 处的二阶 Taylor展式，带 Peano 余项. 换言之，展式中的系数满足 $a_0=f^0$, $a_1=f_x^0$, $a_2=f_y^0$, $a_{11}=f_{xx}^0$, $a_{12}=f_{xy}^0$, $a_{21}=f_{yx}^0$, $a_{22}=f_{yy}^0$.

定理证明留作习题.

例: 近似计算
$$\frac{1.03^2}{\sqrt{0.98}\sqrt[3]{1.06}}.$$

解: 利用一阶全微分, 我们已经近似计算过上式, 见 Feb24讲义的补充习题. 以下利用带 Peano 余项的二阶 Taylor 展式, 作更精确的近似计算. 记 $f(x,y,z)=x^2y^{-1/2}z^{-1/3}$, $(x_0,y_0,z_0)=(1,1,1)$, $(h_1,h_2,h_3)=(0.03,-0.02,0.06)$. 于是我们要计算 $f(1.03,0.98,1.06)=f(x_0+h_1,y_0+h_2,z_0+h_3)$ 的近似值.

应用续一

先计算一阶近似，这是Feb24讲义的补充习题.

$$f(1.03, 0.98, 1.06) = f(P_0 + h) \approx f(P_0) + \nabla f(P_0) \cdot h$$

$$= f(1,1,1) + f_x^0 h_1 + f_y^0 h_2 + f_z^0 h_3.$$

简单计算得

$$\begin{aligned}
f(x,y,z) &= x^2 y^{-1/2} z^{-1/3}, & f(1,1,1) &= 1,\\
f_x(x,y,z) &= 2xy^{-1/2} z^{-1/3}, & f_x(1,1,1) &= 2,\\
f_y(x,y,z) &= -\tfrac{1}{2} x^2 y^{-3/2} z^{-1/3}, & f_y(1,1,1) &= -1/2,\\
f_z(x,y,z) &= -\tfrac{1}{3} x^2 y^{-1/2} z^{-4/3}, & f_z(1,1,1) &= -1/3.
\end{aligned}$$

于是所求一阶近似为

$$f(1.03, 0.98, 1.06) \approx 1 + 2 \times 0.03 + (-1/2) \times (-0.02)$$

$$+(-1/3) \times (0.06) = 1 + 0.06 + 0.01 - 0.02 = 1.05.$$

为计算二阶近似，我们需要计算 Hesse 矩阵 $H_f(P_0)$，其中

$$H_f(x, y, z) = \begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{bmatrix}.$$

简单计算得

$$\mathbf{H_f}(1,1,1) = \begin{bmatrix} 2 & -1 & -2/3 \\ -1 & 3/4 & 1/6 \\ -2/3 & 1/6 & 4/9 \end{bmatrix}.$$

于是二阶项为 $\frac{1}{2}\mathbf{h}^T\mathbf{H_f^0}\mathbf{h} =$

$$\frac{1}{2}[0.03, -0.02, 0.06] \begin{bmatrix} 2 & -1 & -2/3 \\ -1 & 3/4 & 1/6 \\ -2/3 & 1/6 & 4/9 \end{bmatrix} \begin{bmatrix} 0.03 \\ -0.02 \\ 0.06 \end{bmatrix}$$

$$= \frac{23}{20000}.$$

于是所求近似值为

$$\frac{1.03^2}{\sqrt{0.98}\sqrt[3]{1.06}} \approx 1.05 + \frac{23}{20000}.$$

解答完毕.

定理：设 $f:\Omega\subset \mathbb{R}^n\to\mathbb{R}$ 为 n 元函数，C^2 的，定义域 Ω 既开且凸，则对任意点 $x_0, x\in\Omega$，记 $h=x-x_0$，成立

$$f(x)=f(x_0)+\nabla f(x_0)h+\frac{1}{2}h^T H_f(\xi)h,$$

$$\text{或 } f(x)=f(x_0)+\nabla f(x_0)h+\frac{1}{2}h^T H_f(x_0)h+o(\|h\|^2),$$

其中 $\xi=x_0+\theta h$，$\theta\in(0,1)$，且

$$H_f(x)=\left[\frac{\partial^2 f(x)}{\partial x_i\partial x_j}\right]_{n\times n}.$$

证: 由假设 Ω 凸, 故对于任意两点 $x_0, x \in \Omega$, 连接它们的线段包含在 Ω 之中, 即 $[x_0, x] \triangleq \{x_0 + t(x - x_0), t \in [0, 1]\} \subset \Omega$. 令 $\phi(t) = f(x_0 + t(x - x_0))$, 则 $\phi(0) = f(x_0)$, $\phi(1) = f(x)$. 由于函数 f 是 C^2 的, 故函数 $\phi(t)$ 在开区间 $(-\delta, 1 + \delta)$ 上也是 C^2 的. 对一元函数 $\phi(t)$, 应用一元函数的 Taylor 公式应用即可得到展式 $\phi(1) = \phi(0) + \phi'(0)1 + \frac{1}{2}\phi''(\theta)1^2$, 其中 $\theta \in (0, 1)$. 以下来计算 $\phi'(0)$ 和 $\phi''(\theta)$,

记 $h = x - x_0$, 则由链规则得

$$\phi'(t) = [f(x_0 + th)]' = [f(x_{01} + th_1, \cdots, x_{0n} + th_n)]'$$

$$= \sum_{i=1}^{n} \frac{\partial f(\cdots)}{\partial x_i} h_i = \nabla f(\cdots) h,$$

这里 $x_0 = (x_{01}, \cdots, x_{0n})$, $h = (h_1, \cdots, h_n)$. 对上式再次求导得

$$\phi''(t) = \sum_{i=1}^{n} \left[\frac{\partial f(x_{01} + th_1, \cdots, x_{0n} + th_n)}{\partial x_i} \right]_t' h_i$$

$$= \sum_{i=1}^{n} \left[\sum_{j=1}^{n} \frac{\partial^2 f(\cdots)}{\partial x_i \partial x_j} h_j \right] h_i = h^T H_f(x_0 + th) h.$$

于是由 $\phi(1)=\phi(0)+\phi'(0)1+\frac{1}{2}\phi''(\theta)$ 得

$$f(x)=f(x_0)+\nabla f(x_0)h+\frac{1}{2}h^T H_f(\xi)h,\quad \xi=x_0+\theta h.\quad (*)$$

即带 Lagrange 余项的展式成立. 为证带 Peano 余项的展式, 记

$$G(x)=[g_{ij}(x)]\triangleq H_f(x)-H_f(x_0),$$

则由展式(*)得

$$f(x)=f(x_0)+\nabla f(x_0)h+\frac{1}{2}h^T H_f(x_0)h+\frac{1}{2}h^T G(\xi)h.$$

因此只需证 $h^T G(\xi) h = o(\|h\|^2)$. 由于函数 $f(x)$ 是 C^2 的, 故 $g_{ij}(x) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x) - \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0)$ 连续, 从而 $g_{ij}(x) \to g_{ij}(x_0) = 0$, $x \to x_0$. 令 $g(x) = \max\{|g_{ij}(x)|\}$, 则不难证明函数 $g(x)$ 也连续. 故 $g(x) \to 0$, $x \to x_0$. 于是

$$|h^T G(x) h| = \left|\sum_{i,j=1}^{n} g_{ij}(x) h_i h_j\right| \le \sum_{i,j}^{n} |g_{ij}(x)||h_i||h_j|$$

$$\le g(x) \sum_{i,j=1}^{n} |h_i||h_j| \le \frac{1}{2} g(x) \sum_{i,j=1}^{n} (h_i^2 + h_j^2) = n g(x) \|h\|^2.$$

故 $\frac{1}{2} h^T G(\xi) h = o(\|h\|^2)$. 即带 Peano 余项的展式成立. 证毕.

定理：设二元函数 $f(x,y)$ 在平面开凸域 Ω 上是 C^3 的，则对任意点 $(x_0,y_0),(x,y)\in\Omega$，记 $h=x-x_0$，$k=y-y_0$，

$$f(x,y)=f^0+f^0_x h+f^0_y k+\frac{1}{2}\left[f^0_{xx}h^2+2f^0_{xy}hk+f^0_{yy}k^2\right]$$

$$+\frac{1}{3!}\left[f^0_{xxx}h^3+3f^0_{xxy}h^2k+3f^0_{xyy}hk^2+f^0_{yyy}k^3\right]+o(\rho^3),$$

其中 $\rho^2=h^2+k^2$，f^0，f^0_x 等表示函数 f，及其偏导数 f_x 等在点 (x_0,y_0) 处取值.

证明基本思想同二阶展式. 细节略.

例子

例: 求函数 x^y 在点 $(1,1)$ 处的三阶 Taylor 多项式, 并近似计算 $1.1^{1.02}$. (课本习题1.8 题2(1), page 82).

解: 记 $f(x,y)=x^y$, $(x_0,y_0)=(1,1)$. 于是

$$\begin{aligned}
f &= x^y, & f^0 &= 1,\\
f_x &= yx^{y-1}, & f_x^0 &= 1,\\
f_y &= x^y \ln x, & f_y^0 &= 0,\\
f_{xx} &= (y-1)yx^{y-2}, & f_{xx}^0 &= 0,\\
f_{xy} &= x^{y-1} + x^{y-1}\ln x, & f_{xy}^0 &= 1,\\
f_{yy} &= x^y(\ln x)^2, & f_{yy}^0 &= 0,
\end{aligned}$$

例子续一

$$
\begin{aligned}
f_{xxx} &= (y-2)(y-1)yx^{y-3}, & f^0_{xxx} &= 0,\\
f_{xxy} &= (2y-1)x^{y-2} + (y-1)yx^{y-2}\ln x, & f^0_{xxy} &= 1,\\
f_{xyy} &= 2(\ln x)x^{y-1} + (\ln x)^3 x^y, & f^0_{xyy} &= 0,\\
f_{yyy} &= (\ln x)^3 x^y, & f^0_{yyy} &= 0.
\end{aligned}
$$

由上述计算结果可知，函数 $f(x, y) = x^y$ 在点 $(1, 1)$ 处的三阶 Taylor 多项式为

$$
p_3(x, y) = f^0 + f^0_x h + f^0_y k + \frac{1}{2}\left[f^0_{xx}h^2 + 2f^0_{xy}hk + f^0_{yy}k^2\right]
$$

$$
+\frac{1}{3!}\left[f^0_{xxx}h^3 + 3f^0_{xxy}h^2k + 3f^0_{xyy}hk^2 + f^0_{yyy}k^3\right],
$$

例子续二

$$p_3(x,y) = 1 + h + \frac{1}{2}\Big(2hk\Big) + \frac{1}{3!}\Big(3h^2k\Big)$$

$$= 1 + (x-1) + (x-1)(y-1) + \frac{1}{2}(x-1)^2(y-1).$$

在点 $(x_0, y_0) = (1, 1)$ 附近，可以用 $p_3(x,y)$ 作为 $f(x,y)$ 的近似值．点 $(1.1, 1.02)$ 可看作 (x_0, y_0) 附近的点，即 $(1.1, 1.02) = (x_0 + h, y_0 + k)$，其中 $(h, k) = (0.1, 0.02)$．于是 $1.1^{1.02} \approx$

$$p_3(x_0 + h, y_0 + k) = 1 + 0.1 + 0.1 \times 0.02 + \frac{1}{2}(0.1)^2(0.02)$$

$$= 1 + 0.1 + 0.002 + 0.0001 = 1.1021.$$

例子续三

注: 也可用如下方法求得函数 x^y 在点 $(1,1)$ 处的三阶 Taylor 多项式:

$$
\begin{aligned}
x^y = (1+h)^{1+k} &= e^{(1+k)\ln(1+h)} \\
&= 1 + (1+k)\ln(1+h) + \frac{1}{2}(1+k)^2\ln^2(1+h) \\
&\quad + \frac{1}{3!}(1+k)^3\ln^3(1+h) + o(\rho^3) \\
&= 1 + (1+k)\left(h - \frac{h^2}{2} + \frac{h^3}{3}\right) \\
&\quad + \frac{1}{2}(1+2k+k^2)\left(h - \frac{h^2}{2} + \frac{h^3}{3}\right)^2 + \frac{h^3}{6} + o(\rho^3)
\end{aligned}
$$

例子续四

$$= 1 + (1+k)\left(h - \frac{h^2}{2} + \frac{h^3}{3}\right)$$

$$+ \frac{1}{2}(1 + 2k + k^2)\left(h - \frac{h^2}{2} + \frac{h^3}{3}\right)^2 + \frac{h^3}{6} + o(\rho^3)$$

$$= 1 + h - \frac{h^2}{2} + \frac{h^3}{3} + hk - \frac{h^2 k}{2} + \frac{1}{2}(h^2 - h^3 + 2h^2 k) + \frac{h^3}{6} + o(\rho^3)$$

$$= 1 + h + hk + \frac{1}{2}h^2 k + o(\rho^3).$$

故所求三阶 Taylor 多项式为 $p_3(x, y) = 1 + h + hk + \frac{1}{2}h^2 k$

$$= 1 + (x-1) + (x-1)(y-1) + \frac{1}{2}(x-1)^2(y-1).$$

例子完毕.

多元函数的局部极大值和局部极小值

Definition

定义: 设 $f:\Omega\subset\mathbb{R}^n\to\mathbb{R}$, Ω 开, $x_0\in\Omega$. 若存在 $\delta>0$, 使得

(i) $f(x)\le f(x_0)$, $\forall x\in B_\delta(x_0)$, 则称 f 在 x_0 处取(局部)极大值, 点 x_0 称为(局部)极大值点;

(ii) $f(x)\ge f(x_0)$, $\forall x\in B_\delta(x_0)$, 则称 f 在 x_0 处取(局部)极小值, 点 x_0 称为(局部)极小值点;

(iii) 如果将(i)和(ii)中的非严格不等式改为相应的严格不等式, 在 x_0 的某个去心邻域 $B^\circ(x_0)$ 成立, 则称函数 f 在点 x_0 处取得(局部)严格极大值和严格极小值.

Definition

定义：设 $f:\Omega\subset\mathbb{R}^n\to\mathbb{R}$，$\Omega$ 开或闭，若存在 $x_0\in\Omega$，使得

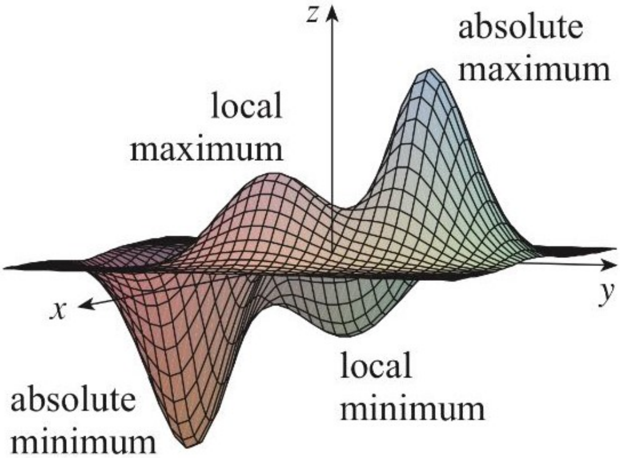
(i) $f(x)\le f(x_0)$，$\forall x\in\Omega$，则称 f 在 x_0 处取绝对最大值，点 x_0 称为绝对最大值点；

(ii) $f(x)\ge f(x_0)$，$\forall x\in\Omega$，则称 f 在 x_0 处取绝对最小值，点 x_0 称为绝对最小值点；

(iii) 如果将 (i)和 (ii)中的非严格不等式改为相应的严格不等式，点 x_0 除外，则称函数 f 在点 x_0 处，取得绝对严格最大值和绝对严格最小值.

绝对最大(小)值(点)也常称作整体或全局(globally)最大(小)值(点).

局部和绝对极值之图示



z
absolute maximum
local maximum
x
y
local minimum
absolute minimum

例子

Example

例: (i) 函数 $f(x,y)=x^2+y^2$ 在原点 $(0,0)$ 处，在全平面上取得绝对严格极(最)小值;

(ii) 函数 $g(x,y)=x^2-y^2$ 在原点 $(0,0)$ 处无极值;

(iii) 函数 $h(x,y)=-x^4-y^4$ 在原点 $(0,0)$ 处，在全平面上取得绝对严格最大值.

Theorem

<u>定理</u>: 设 n 元函数 $f(x)$ 在点 x_0 处有(局部)极值(极大值或极小值), 且在点 x_0 处可微, 则 $\nabla f(x_0) = 0$.

Proof.

<u>证明</u>: 只证二维情形. 设 $f(x, y)$ 在点 (x_0, y_0) 处有极值, 则一元函数 $f(x, y_0)$ 在 $x = x_0$ 处有极值. 根据一元函数的极值理论可知 $f_x(x_0, y_0) = 0$. 同理可证 $f_y(x_0, y_0) = 0$. 故 $\nabla f(x_0, y_0) = 0$. 证毕. □

临界点 (驻点)

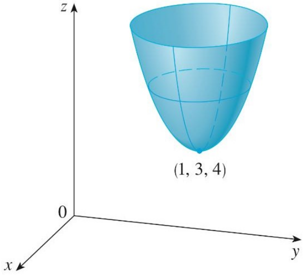
Definition

定义: 设 $f:\Omega\subset \mathbb{R}^n\to\mathbb{R}$ 可微, Ω 开, 使得 $\nabla f(x_0)=0$ 的点 $x_0\in\Omega$ 均称为函数 $f(x)$ 的临界点(critical points) 或驻点 (stationary points)

Example

例: 原点 $(0,0)$ 是三个函数 x^2+y^2, $-x^4-y^4$ 和 x^2-y^2 共同的临界点. 前两个函数在原点处有极值, 但第三个函数处无极值. 由此可见临界点不一定是极值点.

例: 设 $f(x,y)=x^2+y^2-2x-6y+14$, 则 $f_x=2x-2$, $f_y=2y-6$. 令 $f_x=f_y=0$ 可解得唯一的临界点 $(1,3)$. 另一方面用配方法可知 $f(x,y)=(x-1)^2+(y-3)^2+4$. 由此可知函数 $f(x,y)$ 在这个临界点处有全局最小值.



Theorem

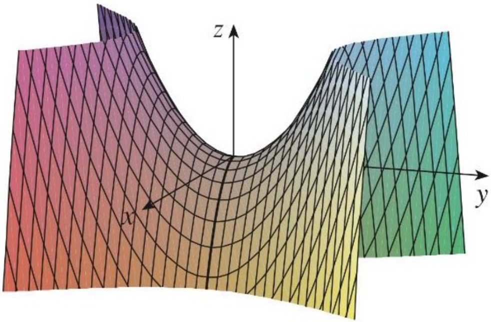
定理: 设 $f:\Omega\subset \mathbb{R}^n\to\mathbb{R}$ 是 C^2 函数, Ω 开. 设 $x_0\in\Omega$ 是 $f(x)$ 的临界点, 即 $\nabla f(x_0)=0$. 记 $H(x)$ 为函数 $f(x)$ 的 Hesse 矩阵, 则以下结论成立.

(i) 若 $H(x_0)$ 正定, 则 $f(x)$ 在点 x_0 处有严格极小值;

(ii) 若 $H(x_0)$ 负定, 则 $f(x)$ 在点 x_0 处有严格极大值;

(iii) 若 $H(x_0)$ 不定, 则 $f(x)$ 在点 x_0 处无极值.

对于二元函数而言，情形 (iii) 时的临界点称为鞍点(saddle points)．一个典型的例子是 $z=y^2-x^2$．易见原点是函数的鞍点．如图．



例子

例1: 函数 x^2+y^2 有临界点 $(0,0)$, $H^0=\text{diag}(2,2)$ 正定. 函数在原点处有全局最小值;

例2: 函数 $-x^2-y^2$ 有临界点 $(0,0)$, $H^0=\text{diag}(-2,-2)$ 负定. 函数在原点处有极大值. 实际上还是全局最大值;

例3: 函数 x^2-y^2 有临界点 $(0,0)$, $H^0=\text{diag}(2,-2)$ 不定. 函数在原点处无极值; 此时临界点是鞍点.

例4: 函数 x^2+y^3 有临界点 $(0,0)$, $H^0=\text{diag}(2,0)$ 半正定. 函数在原点处无极值;

例5: 函数 x^2+y^4 有临界点 $(0,0)$, $H^0=\text{diag}(2,0)$ 半正定. 函数在原点处有极小值. 实际上还是全局最小值.

注: 极值充分性定理考虑了 Hesse 矩阵 $H(x_0)$ 的三种情况: 正定, 负定和不定. 还有一种情况就是半正定或半负定. 此时函数 $f(x)$ 在点 x_0 处的有极值情况不确定. 上述例 4和例 5说明, 此时函数有极值和无极值的两种可能性都存在.

例一：求函数 $f(x,y)=x^4+y^4-4xy+1$ 的极大值和极小值，以及鞍点。

解：计算得 $f_x=4x^3-4y$，$f_y=4y^3-4x$。解方程组 $f_x=f_y=0$，得三个解 $(0,0)$，$(1,1)$，$(-1,-1)$。进一步计算得

$$H_f(x,y)=\begin{bmatrix} 12x^2 & -4 \\ -4 & 12y^2 \end{bmatrix}.$$

例一，续

显然

$$H_f(0,0)=\begin{bmatrix}0 & -4\\ -4 & 0\end{bmatrix}$$

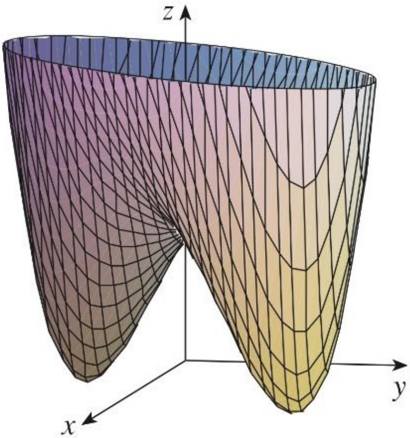
有特征值 ± 4. 因此原点为鞍点. 而

$$H_f(\pm 1,\pm 1)=\begin{bmatrix}12 & -4\\ -4 & 12\end{bmatrix}$$

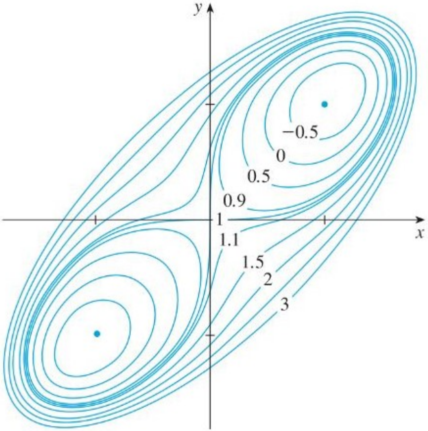
有两个正特征值 $8, 16$. 故两个 Hesse 矩阵 $H_f(\pm 1,\pm 1)$ 均正定.

因此这个两个点均为局部极小值点.

函数 $f(x,y)=x^4+y^4-4xy+1$ 的图像如下



函数 $f(x, y) = x^4 + y^4 - 4xy + 1$ 的水平线如图所示.



例二：求函数 $u = u(x, y, z) = x^3 + y^2 + z^2 + 6xy + 2z$ 的极值和极值点.

解：Step 1. 求临界点. 令 $u_x = u_y = u_z = 0$，即

$$\begin{cases} u_x = 3x^2 + 6y = 0, \\ u_y = 2y + 6x = 0, \\ u_z = 2z + 2 = 0. \end{cases}$$

解这个方程组得到两组解 $P_1 = (6, -18, -1)$，$P_2 = (0, 0, -1)$.

Step 2. 求 Hesse 矩阵. 计算得函数的 Hesse 矩阵为

$$H(x,y,z)=\begin{bmatrix} 6x & 6 & 0 \\ 6 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

在点 $P_1=(6,-18,-1)$ 处，

$$H(P_1)=\begin{bmatrix} 36 & 6 & 0 \\ 6 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

不难验证它是正定的. 因为它的三个顺序主子式均大于零. 因此临界点 P_1 是极小值点，且极小值为 -106.

在点 $P_2 = (0, 0, -1)$ 处，

$$H(P_2) = \begin{bmatrix} 0 & 6 & 0 \\ 6 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

计算得矩阵的特征值为 2, $1 \pm \sqrt{37}$. 两正一负. 这说明 Hesse 矩阵是不定的. 因此临界点 P_2 不是极值点. 例子完毕.

证明: 注意到 x_0 是驻点, 即 $\nabla f(x_0)=0$, 故函数 $f(x)$ 在点 x_0 处, 分别带 Lagrange 余项和 Peano 余项的二阶 Taylor 展开为

$$f(x)=f(x_0)+\frac{1}{2}h^T H(\xi)h, \quad (*)$$

$$f(x)=f(x_0)+\frac{1}{2}h^T H(x_0)h+o(\|h\|^2), \quad (**)$$

这里 $h=x-x_0$, $\xi=x_0+\theta h$, $\theta\in(0,1)$.

情形一: $H(x_0)$ 正定. 由 Hesse 矩阵的连续性可知, 当 $\|x-x_0\| = \|h\|$ 充分小时, $H(\xi)$ 也正定. 故由展式 (*) 知 $f(x)>f(x_0)$, $\forall x\in B_\delta^\circ(x_0)$. 此即函数 f 在点 x_0 处有严格极小值.

情形二: $H(x_0)$ 负定. 证明思想同情形一. 细节略.

情形三: $H(x_0)$ 不定. 此时必存在向量 $p, q \in \mathbb{R}^n$, 使得

$$p^T H(x_0) p < 0 < q^T H(x_0) q.$$

取 $h = \varepsilon p$, 即 $x = x_0 + \varepsilon p$, $\varepsilon > 0$ 充分小, 则由带 Peano 余项的展式 (**) 得

$$f(x) - f(x_0) = \frac{1}{2}\varepsilon^2 p^T H(x_0) p + o(\varepsilon^2)$$

$$= \frac{\varepsilon^2}{2}\left[p^T H(x_0) p + \frac{o(\varepsilon^2)}{\varepsilon^2}\right] < 0.$$

这说明在点 x_0 处的任意邻域内，存在点 $x = x_0 + \varepsilon p$，使得 $f(x) < f(x_0)$. 同理取 $h = \varepsilon q$, $\varepsilon > 0$ 充分小，我们有

$$f(x) - f(x_0) = \frac{\varepsilon^2}{2}\left[q^T H(x_0) q + \frac{o(\varepsilon^2)}{\varepsilon^2}\right] > 0.$$

这说明在点 x_0 处的任意邻域内，存在点 $x = x_0 + \varepsilon q$，使得 $f(x) > f(x_0)$. 这说明 x_0 不是函数 $f(x)$ 的极值点. 证毕. $\square$

记号: 设 $a, b \in \mathbb{R}^n$, 记

$$[a, b] = \{a + t(b - a), t \in [0, 1]\},$$

$$(a, b) = \{a + t(b - a), t \in (0, 1)\},$$

并分别称它们为连接 a, b 两点的闭线段和开线段. 注意依定义, 我们有 $[a, b] = [b, a]$, $(a, b) = (b, a)$.

注: 区域 $\Omega \subset \mathbb{R}^n$ 为凸, 当且仅当对 $\forall a, b \in \Omega$, $[a, b] \subset \Omega$.

Theorem

定理: 设 $\mathbf{f}:\Omega\subset\mathrm{I\!R}^n\to\mathrm{I\!R}$ 为 C^1 函数, Ω 开. 设 $\mathbf{a},\mathbf{b}\in\Omega$ 且 $[\mathbf{a},\mathbf{b}]\subset\Omega$, 则存在 $\xi\in(\mathbf{a},\mathbf{b})$, 使得

$$\mathbf{f}(\mathbf{b})-\mathbf{f}(\mathbf{a})=\nabla\mathbf{f}(\xi)(\mathbf{b}-\mathbf{a}),$$

这里 $\xi=\mathbf{a}+\theta(\mathbf{b}-\mathbf{a})$, $\theta\in(0,1)$.

Proof.

证明: 记 $\phi(t) = f(a + t(b - a))$, $t \in [0, 1]$, 则 $\phi(0) = f(a)$, $\phi(1) = f(b)$, 且 $\phi(t)$ 在闭区间 $[0, 1]$ 上连续, 开区间 $(0, 1)$ 上可导. 由一元函数的中值定理知 $\phi(1) - \phi(0) = \phi'(\theta)$, 其中 $\theta \in (0, 1)$. 此即

$$f(b) - f(a) = \nabla f(\xi)(b - a),$$

这里 $\xi = a + \theta(b - a)$, $\theta \in (0, 1)$. 证毕. □

Mar04讲义选作题: 设函数 $f(x,y,z)$ 在全空间 $\mathbb{R}^3$ 上连续可微, 它的三个一阶偏导函数处处相等(恒同), 即

$$f_x(x,y,z)=f_y(x,y,z)=f_z(x,y,z),\quad \forall (x,y,z)\in\mathbb{R}^3$$

若 $f(x,0,0)>0$, $\forall x\in\mathbb{R}$. 证明 $f(x,y,z)>0$, $\forall (x,y,z)\in\mathbb{R}^3$.

证明: 任意固定一点 $P=(x,y,z)$, 记 $P_0=(x+y+z,0,0)$, 根据多元函数的中值定理得 $f(P)-f(P_0)=\nabla f(\xi)(P-P_0)$. 注意 $P-P_0=(x,y,z)-(x+y+z,0,0)=(-y-z,y,z)$. 由假设 $f_x=f_y=f_z$ 得 $f(P)-f(P_0)=f_x(\xi)\ (-y-z+y+z)=0$. 即 $f(x,y,z)=f(x+y+z,0,0)>0$. 命题得证.

中值定理不能直接推广到向量值函数情形

对于一般 C^1 的向量值函数 $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$, $m > 1$, 上述形式的中值定理不再成立. 例如设 $f : \mathbb{R} \to \mathbb{R}^2$, $t \mapsto (t^2, t^3)^T$, 则

$$f(1) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad f(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad f'(t) = \begin{bmatrix} 2t \\ 3t^2 \end{bmatrix},$$

但 $f(1) - f(0) \neq f'(t)(1-0)$, $\forall t \in (0,1)$, 即

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 2t \\ 3t^2 \end{bmatrix}, \quad \forall t \in (0,1).$$

一. 习题1.8 (pp.81-82): 1, 2.

二. 习题1.9: (page 93-94): 1(1)(3)(5), 2, 4(1), 5, 6.

三. 补充习题. 证明二元函数带 Peano 余项的 Taylor 展式的唯一性定理, 即第12页中的定理.

四. 选作题: 设函数 $f(x, y)$ 在全平面 $\mathbb{R}^2$ 上连续可微, 有且仅有一个极小值点 (x_0, y_0), 无其他驻点. 问这个极小值是否为全局极小值 (最小值), 即 $f(x_0, y_0) \leq f(x, y)$, $\forall (x, y) \in \mathbb{R}^2$. 若是, 请给予证明; 若否, 请提供一个反例.

注一：本选作题是课本第93 页习题1.9 题3 的一个修改版．原题第三行中的"驻点"二字似应改为"极值点"较好．

注二：对于一维情形，相应的结论成立．即如下结论成立：设函数 $f(x)$ 在全实轴 $\mathbb{R}$ 上连续可微，有且仅有一个驻点 x_0，并且函数在这个驻点处有极小值，则这个极小值是全局极小值(最小值)，即 $f(x_0) \leq f(x), \forall x \in \mathbb{R}$．可用反证法证明(利用 Rolle 定理)．请有兴趣的同学尝试一下证明．