第 17 章　多元函数微分学

§1 可微性

【一】 可微性与全微分

【定义】 设函数 $z=f(x,y)$ 在点 $P_0\left(x_0,y_0\right)$ 的某领域 $U(P_0)$ 内有定义, 对于 $U(P_0)$ 中的点 $P(x,y)=(x_0+\Delta x,y_0+\Delta y)$, 若函数 f 在点 P_0 处的全增量 Δz 可表示为:

$$\Delta z=f(x_0+\Delta x,y_0+\Delta y)-f(x,y)=A\Delta x+B\Delta y+o(\rho),$$

其中 A,B 是仅与点 P_0 有关的常数, $\rho=\sqrt{\Delta x^2+\Delta y^2}=\left\|P-P_0\right\|$，则称函数 f 在点 P_0 可微, 并称

$$dz\,|_{P_0}=df(x_0,y_0)=A\Delta x+B\Delta y$$

为函数 f 在点 P_0 的**全微分**。

【等价定义】

$$\Delta z=A\Delta x+B\Delta y+\alpha\Delta x+\beta\Delta y,$$

这里 $\lim\limits_{(\Delta x,\Delta y)\to(0,0)}\alpha=\lim\limits_{(\Delta x,\Delta y)\to(0,0)}\beta=0.$

这是因为：若 $\Delta z=A\Delta x+B\Delta y+\alpha\Delta x+\beta\Delta y$，则

$$\frac{\left|\alpha\Delta x+\beta\Delta y\right|}{\sqrt{\Delta x^2+\Delta y^2}}\le|\alpha|\frac{|\Delta x|}{\sqrt{\Delta x^2+\Delta y^2}}+|\beta|\frac{|\Delta y|}{\sqrt{\Delta x^2+\Delta y^2}}\le|\alpha|+|\beta|\to0,\left(\Delta x,\Delta y\right)\to\left(0,0\right)$$

若 $\Delta z=A\Delta x+B\Delta y+o(\rho)$，则

$$o(\rho)=\frac{o(\rho)}{\rho}\cdot\frac{\Delta x}{\rho}\Delta x+\frac{o(\rho)}{\rho}\cdot\frac{\Delta y}{\rho}\Delta y$$

令 $\alpha=\dfrac{o(\rho)}{\rho}\cdot\dfrac{\Delta x}{\rho},\beta=\dfrac{o(\rho)}{\rho}\cdot\dfrac{\Delta y}{\rho}$，则

$$|\alpha|=\left|\frac{o(\rho)}{\rho}\right|\cdot\left|\frac{\Delta x}{\rho}\right|\le\left|\frac{o(\rho)}{\rho}\right|\to0\text{，同理}|\beta|\to0\text{，}\left(\Delta x,\Delta y\right)\to\left(0,0\right)$$

中国矿业大学理学院胡建华

显然，f 在点 P_0 可微 $\Rightarrow$ f 在点 P_0 连续。

【例 1】（P115）　考察函数 $f(x,y)=xy$ 在点 (x_0,y_0) 处的可微性.

解　　$\Delta f\left(x_0,y_0\right)=(x_0+\Delta x,y_0+\Delta y)-x_0y_0=y_0\Delta x+x_0\Delta y+\Delta x\Delta y.$

由于

$$\frac{\left|\Delta x\Delta y\right|}{\rho}=\rho\cdot\frac{\left|\Delta x\right|}{\rho}\cdot\frac{\left|\Delta y\right|}{\rho}\le\rho\to0\left(\rho\to0\right),$$

因此 $\Delta x\Delta y=o(\rho)$. 从而函数 f 在 x_0,y_0 可微，且

$$df=y_0\Delta x+x_0\Delta y.$$

【二】 偏导数

【定义】 设函数 $z=f(x,y),(x,y)\in D.$ 若 $(x_0,y_0)\in D$，且 $f\left(x,y_0\right)$ 在 x_0 的某一邻域内有定义，则当极限

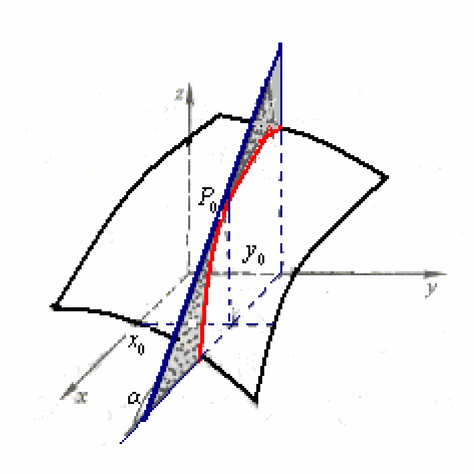
$$\lim_{\Delta x\to0}\frac{f(x_0+\Delta x,y_0)-f(x_0,y_0)}{\Delta x}$$

存在时，称这个极限为函数 f 在点 (x_0,y_0) 关于 x 的偏导数，记作

$$f_x\left(x_0,y_0\right)\quad 或\quad \left.\frac{\partial f}{\partial x}\right|_{(x_0,y_0)}.$$

类似定义 $f_y\left(x_0,y_0\right)$。

几何意义如图：



【例 2】（P117）求函数 $f(x,y)=x^3+2x^2y-y^3$ 在点 $(1,3)$ 关于 x 和关于 y 的偏导数.

解　先求 f 在点 $(1,3)$ 关于 x 的偏导数，为此，令 $y=3$，得到以 x 为自变量的函数 $f(x,3)=x^3+6x^2-27$，求它在 $x=1$ 的导数，即

$$f_x(1,3)=\left.\frac{df(x,3)}{dx}\right|_{x=1}=3x^2+12x\,|_{x=1}=15.$$

再求 f 在点 $(1,3)$ 关于 y 的偏导数，先令 $x=1$，得到以 y 为自变量的函数 $f(1,y)=1+2y-y^3$，求它在 $y=3$ 的导数，得

$$f_y(1,3)=\left.\frac{df(1,y)}{dy}\right|_{y=3}=2-3y^2\,|_{y=3}=-25.$$

通常为可分别先求出 f 关于 x 和 y 的**偏导函数**：

$$f_x(x,y)=3x^2+4xy, f_y(x,y)=2x^2-3y^2.$$

然后以 $(x,y)=(1,3)$ 代入，也能得到同样结果.

【例 3】（P118）求三元函数 $u=\sin(x+y^2-e^z)$ 的偏导数.

解　把 y 和 z 看作常数，得

$$\frac{\partial u}{\partial x}=\cos(x+y^2-e^z).$$

把 x,z 看作常数，得

$$\frac{\partial u}{\partial y}=2y\cos(x+y^2-e^z).$$

把 x,y 看作常数，得

$$\frac{\partial u}{\partial z}=-e^z\cos(x+y^2-e^z)$$

【例 4】（偏导数存在 $\not\Rightarrow$ 连续）

$$f(x,y)=\begin{cases}\dfrac{xy}{x^2+y^2}, & x^2+y^2\neq 0\\ 0, & x^2+y^2=0\end{cases}$$

由上一章，$\lim\limits_{(x,y)\to(0,0)}f(x,y)$ 不存在，所以 f 在点 $(0,0)$ 不连续，但

$$f_x\left(0,0\right)=\lim_{\Delta x\to 0}\frac{f\left(\Delta x,0\right)-f(0,0)}{\Delta x}=\lim_{\Delta x\to 0}\frac{0-0}{\Delta x}=0.$$

同理 $f_y\left(0,0\right)=0.$

【例 5】（连续$\nRightarrow$偏导数存在）

$$f(x,y)=\sqrt{x^2+y^2}$$

显然 f 在点 $(0,0)$ 连续, 但

$$f_x\left(0,0\right)=\lim_{\Delta x\to 0}\frac{f\left(\Delta x,0\right)-f(0,0)}{\Delta x}=\lim_{\Delta x\to 0}\frac{\left|\Delta x\right|}{\Delta x}\text{不存在。}$$

【三】 可微性条件

【定理 1】（可微的必要条件）　若二元函数 f 在点 (x_0,y_0) 处可微，则 f 在该点关于每个自变量的偏导数都存在，且

$$A=f_x(x_0,y_0),B=f_y(x_0,y_0)$$

证　$f_x(x_0,y_0)=\lim\limits_{\Delta x\to 0}\dfrac{f(x_0+\Delta x,y_0)-f(x_0,y_0)}{\Delta x}=\lim\limits_{\Delta x\to 0}\dfrac{A\Delta x+o(\Delta x)}{\Delta x}=A$

记 $\Delta x=dx,\Delta y=dy$, 则

$$dz=f_x(x_0,y_0)dx+f_y(x_0,y_0)dy.$$

【例 6】（P118）（连续＋偏导数存在$\nRightarrow$可微）考察函数

$$f(x,y)=\begin{cases}\dfrac{xy}{\sqrt{x^2+y^2}}, & x^2+y^2\neq 0\\ 0, & x^2+y^2=0\end{cases}$$

令 $x=r\cos\theta,y=r\sin\theta$，则 $f(x,y)=\dfrac{1}{2}r\sin 2\theta\to 0(r\to 0)$，

$$\lim_{(x,y)\to(0,0)}f(x,y)=f(0,0)=0$$

说明 f 在原点连续。而且

$$f_x\left(0,0\right)=\lim_{\Delta x\to 0}\frac{f\left(\Delta x,0\right)-f(0,0)}{\Delta x}=\lim_{\Delta x\to 0}\frac{0-0}{\Delta x}=0.$$

同理可得 $f_y\left(0,0\right)=0$ 。

若函数 f 在原点可微，则

$$\Delta z - dz = f(0+\Delta x, 0+\Delta y) - f(0,0) - \left[f_x(0,0)\Delta x + f_y(0,0)\Delta y\right] = \frac{\Delta x + \Delta y}{\sqrt{\Delta x^2 + \Delta y^2}}$$

应是较 $\rho = \sqrt{\Delta x^2 + \Delta y^2}$ 高阶的无穷小量．为此，考察极限

$$\lim_{\rho\to 0}\frac{\Delta z - dz}{\rho} = \lim_{\rho\to 0}\frac{\Delta x\Delta y}{\Delta x^2 + \Delta y^2}$$

由上一章例题知，上述极限不存在，因而函数 f 在原点不可微。

【定理 2】（可微的充分条件）　若函数 $z = f(x,y)$ 的偏导数在点 (x_0, y_0) 的某邻域内存在，且 f_x 与 f_y 在点 (x_0, y_0) 处连续，则函数 f 在点 (x_0, y_0) 可微.

证　我们把全增量 Δz 写作

$$\Delta z = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0)$$

$$= \left[f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0 + \Delta y)\right] + \left[f(x_0, y_0 + \Delta y) f(x_0, y_0)\right].$$

拉格朗日中值定理，得

$$\Delta z = f_x(x_0 + \theta_1\Delta x, y_0 + \Delta y)\Delta x + f_y(x_0, y_0 + \theta_2\Delta y)\Delta y,\ 0 < \theta_1, \theta_2 < 1$$

由于 f_x 与 f_y 在点 (x_0, y_0) 连续，因此有

$$f_x(x_0 + \theta_1\Delta x, y_0 + \Delta y) = f_x(x_0, y_0) + \alpha,$$

$$f_y(x_0, y_0 + \theta_2\Delta y) = f_y(x_0, y_0) + \beta,$$

其中当 $(\Delta x, \Delta y) \to (0,0)$ 时，$\alpha \to 0, \beta \to 0$

因此

$$\Delta z = f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y + \alpha\Delta x + \beta\Delta y.$$

由可微的等价定义，函数 f 在点 (x_0, y_0) 可微。

若 $z = f(x,y)$ 在点 (x_0, y_0) 偏导数 f_x, f_y 连续，则称在点 (x_0, y_0) **连续可微**。

【注】偏导数连续并不是函数可微的必要条件，见下例。

【例 7】（P125 习题 7）

$$f(x,y)=\begin{cases}\left(x^2+y^2\right)\sin\dfrac{1}{\sqrt{x^2+y^2}}, & x^2+y^2\neq 0\\ 0, & x^2+y^2=0\end{cases}$$

在原点$(0,0)$处可微，但f_x与f_y却在$(0,0)$处不连续。

证　$$\Delta z=f\left(\Delta x,\Delta y\right)-f(0,0)=0\cdot\Delta x+0\cdot\Delta y+\left(\Delta x^2+\Delta y^2\right)\sin\frac{1}{\sqrt{\Delta x^2+\Delta y^2}}$$

$$\Delta z=f\left(\Delta x,\Delta y\right)-f(0,0)=0\cdot\Delta x+0\cdot\Delta y+o(\rho)$$

所以f在点$(0,0)$处可微，且

$$f_x(0,0)=f_y(0,0)=0$$

但f_x,f_y在点$(0,0)$处不连续。易求得

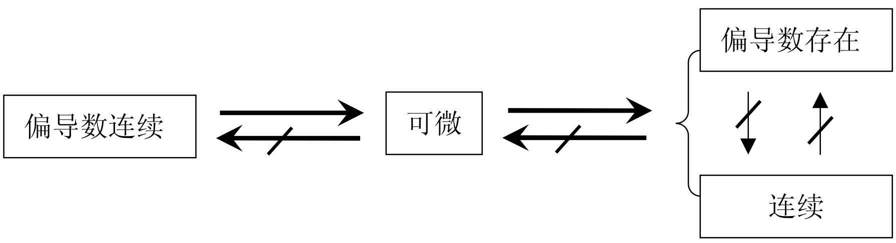
$$f_x\left(x,y\right)=\begin{cases}2x\sin\dfrac{1}{\sqrt{x^2+y^2}}-\dfrac{x}{\sqrt{x^2+y^2}}\cos\dfrac{1}{\sqrt{x^2+y^2}}, & x^2+y^2\neq 0\\ 0, & x^2+y^2=0\end{cases}$$

由于$\lim\limits_{(x,y)\to(0,0)}2x\sin\dfrac{1}{\sqrt{x^2+y^2}}=0$，$\lim\limits_{(x,y)\to(0,0)}\dfrac{x}{\sqrt{x^2+y^2}}\cos\dfrac{1}{\sqrt{x^2+y^2}}$不存在[这是因为

$\lim\limits_{\substack{x\to0^+\\ y=x}}\dfrac{x}{\sqrt{x^2+y^2}}\cos\dfrac{1}{\sqrt{x^2+y^2}}=\lim\limits_{x\to0^+}\dfrac{1}{\sqrt{2}}\cos\dfrac{1}{\sqrt{2}x}$不存在]

因此，$\lim\limits_{(x,y)\to(0,0)}f_x(x,y)$不存在，$f_x$在点$(0,0)$处不连续.

【总结】（连续，可偏导，可微之间的关系）



§2　复合函数微分法

【一】 复合函数的求导法则（链式法则）

【定理 1】 若函数 $x=\varphi(s,t), y=\psi(s,t)$ 在点 $(s,t)\in D$ 可微， $z=f(x,y)$ 在点 $(x,y)=(\varphi(s,t),\psi(s,t))$ 可微，则复合函数 $z=f(\varphi(s,t),\psi(s,t))$ 在点 (s,t) 可微，且它关于 s 与 t 的偏导数分别为

$$\left.\frac{\partial z}{\partial s}\right|_{(s,t)}=\left.\frac{\partial z}{\partial x}\right|_{(x,y)}\left.\frac{\partial x}{\partial s}\right|_{(s,t)}+\left.\frac{\partial z}{\partial y}\right|_{(x,y)}\left.\frac{\partial y}{\partial s}\right|_{(s,t)}$$

$$\left.\frac{\partial z}{\partial s}\right|_{(s,t)}=\left.\frac{\partial z}{\partial x}\right|_{(x,y)}\left.\frac{\partial x}{\partial t}\right|_{(s,t)}+\left.\frac{\partial z}{\partial y}\right|_{(x,y)}\left.\frac{\partial y}{\partial t}\right|_{(s,t)}$$

证　由假设 $x=\varphi(s,t)$， $y=\psi(s,t)$ 在点 (s,t) 可微，于是

$$\Delta x=\frac{\partial x}{\partial s}\Delta s+\frac{\partial x}{\partial t}\Delta t+\alpha_1\Delta s+\beta_1\Delta t$$
$$\Delta y=\frac{\partial y}{\partial s}\Delta s+\frac{\partial y}{\partial t}\Delta t+\alpha_2\Delta s+\beta_2\Delta t,$$

其中当 $\Delta s,\Delta t$ 趋于零时，$\alpha_1,\alpha_2,\beta_1,\beta_2$ 都趋向于零．又由 $z=f(x,y)$ 在点 (x,y) 可微，所以

$$\Delta z=\frac{\partial z}{\partial x}\Delta x+\frac{\partial z}{\partial y}\Delta y+\alpha\Delta x+\beta\Delta y$$

其中当 $\Delta x,\Delta y\to 0$ 时，$\alpha,\beta\to 0$.

补充定义当 $\Delta x=0,\Delta y=0$ 时，$\alpha=\beta=0$，得

$$\Delta z=\left(\frac{\partial z}{\partial x}+\alpha\right)\left(\frac{\partial x}{\partial s}\Delta s+\frac{\partial x}{\partial t}\Delta t+\alpha_1\Delta s+\beta_1\Delta t\right)$$
$$+\left(\frac{\partial z}{\partial y}+\beta\right)\left(\frac{\partial y}{\partial s}\Delta s+\frac{\partial y}{\partial t}\Delta t+\alpha_2\Delta s+\beta_2\Delta t\right)$$

整理后

$$\Delta z=\left(\frac{\partial z}{\partial x}\frac{\partial x}{\partial s}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial s}\right)\Delta s+\left(\frac{\partial z}{\partial x}\frac{\partial x}{\partial t}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial t}\right)\Delta t+\overline{\alpha}\Delta s+\overline{\beta}\Delta t$$

其中

$$\overline{\alpha}=\frac{\partial z}{\partial x}\alpha_1+\frac{\partial z}{\partial y}\alpha_2+\frac{\partial x}{\partial s}\alpha+\frac{\partial y}{\partial s}\beta+\alpha\alpha_1+\beta\alpha_2$$

$$\overline{\beta}=\frac{\partial z}{\partial x}\beta_1+\frac{\partial z}{\partial y}\beta_2+\frac{\partial x}{\partial t}\alpha+\frac{\partial y}{\partial t}\beta+\alpha\beta_1+\beta\beta_2$$

由于 $\varphi(s,t),\psi(s,t)$ 在点 (s,t) 可微，因此它们在点 (s,t) 都连续，即当 $\Delta s,\Delta t\to 0$ 时，有 $\Delta x,\Delta y\to 0$. 从而也有 $\alpha\to 0,\beta\to 0$，以及 $\alpha_1,\alpha_2,\beta_1,\beta_2\to 0$。于是，当 $\Delta s,\Delta t\to 0$，有 $\overline{\alpha}\to 0,\overline{\beta}\to 0$。

【注】链式法则中，f 的可微性假设是不能省略的，否则可能导致错误。

$$f(x,y)=\begin{cases}\dfrac{x^2y}{x^2+y^2}, & x^2+y^2\neq 0\\ 0, & x^2+y^2=0\end{cases},\begin{cases}x=t\\ y=t\end{cases}$$

由 §1 习题 6 知 $f_x(0,0)=f_y(0,0)=0$，但 $f(x,y)$ 在 $(0,0)$ 不可微。复合函数

$$z=f(t,t)=\frac{t}{2}$$

因此

$$\left.\frac{dz}{dt}\right|_{t=0}=\frac{1}{2}$$

若用链式法则，将得出错误结果：

$$\left.\frac{dz}{dt}\right|_{t=0}=\left.\frac{\partial z}{\partial x}\right|_{(0,0)}\left.\frac{dx}{dt}\right|_{t=0}+\left.\frac{\partial z}{\partial y}\right|_{(0,0)}\left.\frac{dy}{dt}\right|_{t=0}=0$$

这个例子说明在使用复合函数求导公式时，必须注意外函数 f 可微这一重要条件.

【例 1】（P128）　设 $z=\ln(u^2+v)$，而 $u=e^{x+y^2},v=x^2+y$, 求 $\dfrac{\partial z}{\partial x},\dfrac{\partial z}{\partial y}$.

解　$$\frac{\partial z}{\partial x}=\frac{\partial z}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial z}{\partial v}\frac{\partial v}{\partial x}$$

$$=\frac{2u}{u^2+v}e^{x+y^2}+\frac{1}{u^2+v}2x=\frac{2}{u^2+v}(ue^{x+y^2}+x),$$

$$\frac{\partial z}{\partial y}=\frac{\partial z}{\partial u}\frac{\partial u}{\partial y}+\frac{\partial z}{\partial v}\frac{\partial v}{\partial y}$$

$$=\frac{2u}{u^2+v}2ye^{x+y^2}+\frac{1}{u^2+v}=\frac{1}{u^2+v}(4uye^{x+y^2}+1).$$

【例 2】（P129 改造）　设 $u=u(x,y)$ 可微，在极坐标变换 $x=r\cos\theta, y=r\sin\theta$ 下，求 $\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2$ 的表达式。

证　u 可以看作 r,θ 的复合函数 $u=u(r\cos\theta, r\sin\theta)$，因此

$$\frac{\partial u}{\partial r}=\frac{\partial u}{\partial x}\frac{\partial x}{\partial r}+\frac{\partial u}{\partial y}\frac{\partial y}{\partial r}=\frac{\partial u}{\partial x}\cos\theta+\frac{\partial u}{\partial y}\sin\theta$$

$$\frac{\partial u}{\partial \theta}=\frac{\partial u}{\partial x}\frac{\partial x}{\partial \theta}+\frac{\partial u}{\partial y}\frac{\partial y}{\partial \theta}=\frac{\partial u}{\partial x}(-r\sin\theta)+\frac{\partial u}{\partial y}(r\cos\theta)$$

解方程组得

$$\frac{\partial u}{\partial x}=\frac{1}{r}\left(r\cos\theta\frac{\partial u}{\partial r}-\sin\theta\frac{\partial u}{\partial \theta}\right)$$

$$\frac{\partial u}{\partial y}=\frac{1}{r}\left(r\sin\theta\frac{\partial u}{\partial r}+\cos\theta\frac{\partial u}{\partial \theta}\right)$$

代入整理得

$$\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2=\left(\frac{\partial u}{\partial r}\right)^2+\frac{1}{r^2}\left(\frac{\partial u}{\partial \theta}\right)^2$$

【例 3】　设 $z=uv+\sin t$, 其中 $u=e^t, v=\cos t$，求 $\frac{dz}{dt}$.

解　画树状图

$$\begin{tikzcd}[row sep=small] & u \arrow[r] & t \\ z \arrow[ur] \arrow[r] \arrow[dr] & v \arrow[r] & t \\ & t & \end{tikzcd}$$

$$\frac{dz}{dt}=\frac{\partial z}{\partial u}\frac{du}{dt}+\frac{\partial z}{\partial v}\frac{dv}{dt}+\frac{\partial z}{\partial t}$$

$$=ve^t+u(-\sin t)+\cos t=e^t(\cos t-\sin t)+\cos t.$$

【例 4】（P130）　用多元复合微分法计算下列一元函数的导数：

$$(1)y=x^x;\quad (2)y=\frac{(1+x^2)\ln x}{\sin x+\cos x}.$$

解　（1）令 $y=u^v, u=x, v=x$, 则有

$$\frac{dy}{dx}=y_u\frac{du}{dx}+y_v\frac{dv}{dx}=vu^{v-1}+u^v\ln v$$

$$=x\cdot x^{x-1}+x^x\ln x=x^x(1+\ln x).$$

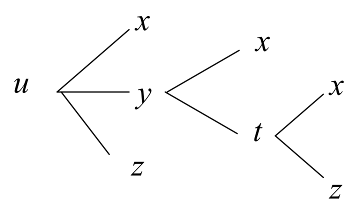
（2）令 $y=\dfrac{vw}{u},u=\sin x+\cos x,v=1+x^2,w=\ln x,$ 则有

$$\frac{dy}{dx}=\frac{\partial y}{\partial u}\frac{du}{dx}+\frac{\partial y}{\partial v}\frac{dv}{dx}+\frac{\partial y}{\partial w}\frac{dw}{dx}=\frac{-vw}{u^2}(\cos x-\sin x)+\frac{w}{u}(2x)+\frac{v}{u}\frac{1}{x}$$

$$=\frac{1}{(\sin x+\cos x)^2}\left[(\sin x+\cos x)(2x\ln x+\frac{1+x^2}{x})-(\cos x-\sin x)(1+x^2)\ln x\right]$$

【例 5】（P130）设 $u=f(x,y,z),y=\varphi(x,t),t=\psi(x,z)$ 都有一阶连续偏导数，求 $\dfrac{\partial u}{\partial x},\dfrac{\partial u}{\partial z}$。

解　$u=f\left(x,\varphi[x,\psi(x,z)],z\right)$



$$\frac{\partial u}{\partial x}=\frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}\frac{\partial \varphi}{\partial x}+\frac{\partial f}{\partial y}\frac{\partial \varphi}{\partial t}\frac{\partial \psi}{\partial x}$$

$$\frac{\partial u}{\partial z}=\frac{\partial f}{\partial z}+\frac{\partial f}{\partial y}\frac{\partial \varphi}{\partial t}\frac{\partial \psi}{\partial z}$$

【二】 复合函数的全微分

若以 x 和 y 为自变量的函数 $z=f(x,y)$ 可微，则其全微分为

$$dz=\frac{\partial z}{\partial x}dx+\frac{\partial z}{\partial y}dy.$$

如果 x,y 作为中间变量又是自变量 s,t 的可微函数

$$x=\varphi(s,t),y=\psi(s,t),$$

则复合函数 $z=f(\varphi(s,t),\psi(s,t))$ 是可微的，其全微分为

$$dz=\frac{\partial z}{\partial s}ds+\frac{\partial z}{\partial t}dt=\left(\frac{\partial z}{\partial x}\frac{\partial x}{\partial s}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial s}\right)ds+(\frac{\partial z}{\partial x}\frac{\partial x}{\partial t}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial t})dt$$

$$=\frac{\partial z}{\partial x}(\frac{\partial x}{\partial t}ds+\frac{\partial x}{\partial t}dt)+\frac{\partial z}{\partial y}(\frac{\partial y}{\partial s}ds+\frac{\partial y}{\partial t}dt)=\frac{\partial z}{\partial x}dx+\frac{\partial z}{\partial y}dy$$

这就是关于多元函数的**一阶（全）微分形式不变性**.

【例 6】（P131）　设 $z=e^{xy}\sin(x+y)$，利用微分形式不变性求 dz，并由此导出 $\frac{\partial z}{\partial x}$ 与 $\frac{\partial z}{\partial y}$.

解　令 $z=e^{u}\sin v, u=xy, v=x+y$. 由于

$$dz=z_u du+z_v dv=e^{u}\sin v du+e^{u}\cos v dv,$$

$$du=ydx+xdy, dv=dx+dy,$$

因此

$$dz=e^{u}\sin v(ydx+xdy)+e^{u}\cos v(dx+dy)$$

$$=e^{xy}[y\sin(x+y)+\cos(x+y)]dx+e^{xy}[x\sin(x+y)+\cos(x+y)]dy,$$

并由此得到

$$z_x=e^{xy}[y\sin(x+y)+\cos(x+y)], z_y=e^{xy}[x\sin(x+y)+\cos(x+y)].$$

§3 方向导数与梯度

【一】 方向导数

偏导数是函数沿坐标轴方向上的变化率：

$$f_x(x_0,y_0)=\lim_{\Delta x\to 0}\frac{f(x_0+\Delta x,y_0)-f(x_0,y_0)}{\Delta x}.$$

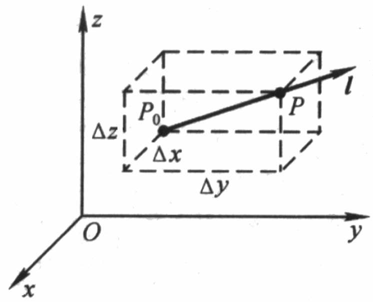
现讨论沿任一给定方向上的变化率。

【定义】 设三元函数 f 在点 $P_0(x_0,y_0,z_0)$ 的某邻域 $U(P_0)\subset R^3$ 内有定义，$\vec{l}$ 为从点 P_0 出发的射线，$P(x,y,z)$ 为 $\vec{l}$ 上且含于 $U(P_0)$ 内的任一点，以 ρ 表示 P 与 P_0 两点间的距离。若极限

$$\lim_{\rho\to 0^+}\frac{f(P)-f(P_0)}{\rho}$$

存在，则称此极限为函数 f 在点 P_0 沿方向 $\vec{l}$ 的**方向导数**，记作

$$\left.\frac{\partial f}{\partial \vec{l}}\right|_{P_0},\ f_{\vec{l}}(P_0) \text{ 或 } f_{\vec{l}}(x_0,y_0,z_0).$$



设 $\|\vec{l}\|=1$，则方向导数又可写为：

$$\left.\frac{\partial f}{\partial \vec{l}}\right|_{P_0}=\lim_{t\to 0^+}\frac{f(P_0+t\vec{l})-f(P_0)}{t}$$

容易看到，f 在点 P_0 沿 x 轴正向的方向导数恰为

$$\left.\frac{\partial f}{\partial \vec{l}}\right|_{P_0}=\left.\frac{\partial f}{\partial x}\right|_{P_0}.$$

沿 x 轴负向的方向导数为

$$\left.\frac{\partial f}{\partial \vec{l}}\right|_{P_0}=-\left.\frac{\partial f}{\partial x}\right|_{P_0}.$$

【定义】　设 $f(x,y,z)$ 在点 $P_0(x_0,y_0,z_0)$ 的存在对所有自变量的偏导数，则称向量

$$\operatorname{grad} f(P_0)=\left(f_x(P_0),f_y(P_0),f_z(P_0)\right)$$

为 f 在点 P_0 的**梯度**。

【定理】　若函数 f 在点 $P_0(x_0,y_0,z_0)$ 可微，则 f 在点 P_0 处沿任一方向 $\vec{l}$ 的方向导数都存在，且

$$f_{\vec{l}}(P_0)=f_x(P_0)\cos\alpha+f_y(P_0)\cos\beta+f_z(P_0)\cos\gamma=\operatorname{grad} f(P_0)\cdot\vec{l}_0$$

这里 $\vec{l}_0=\dfrac{\vec{l}}{\|\vec{l}\|}=(\cos\alpha,\cos\beta,\cos\gamma)$。

证　设 $P(x,y,z)$ 为 l 上任一点，于是

$$x-x_0=\Delta x=\rho\cos\alpha,\ y-y_0=\Delta y=\rho\cos\beta,\ z-z_0=\Delta z=\rho\cos\gamma.$$

由假设 f 在点 P_0 可微，则有

$$f\left(P\right)-f\left(P_0\right)=f_x(P_0)\Delta x+f_y(P_0)\Delta y+f_z(P_0)\Delta z+o\left(\rho\right)$$

于是

$$\frac{f(P)-f(P_0)}{\rho}=f_x(P_0)\frac{\Delta x}{\rho}+f_y(P_0)\frac{\Delta y}{\rho}+f_z(P_0)\frac{\Delta z}{\rho}+\frac{o(\rho)}{\rho}$$

$$=f_x(P_0)\cos\alpha+f_y(P_0)\cos\beta+f_z(P_0)\cos\gamma+\frac{o(\rho)}{\rho}.$$

$$f_{\vec{l}}\left(P_0\right)=\lim_{\rho\to0^+}\frac{f(P)-f(P_0)}{\rho}=f_x(P_0)\cos\alpha+f_y(P_0)\cos\beta+f_z(P_0)\cos\gamma.$$

【例 1】　设 $f(x,y,z)=x+y^2+z^3$，求 f 在点 $P_0\ (1,1,1)$ 沿方向 $\vec{l}=(2,-2,1)$ 的方向导数.

解　易见 f 在点 P_0 可微.

$$\operatorname{grad} f(P_0)=\left(f_x(P_0),f_y(P_0),f_z(P_0)\right)=\left(1,2,3\right)$$

$$\vec{l}_0=\frac{\vec{l}}{\|\vec{l}\|}=\left(\frac{2}{3},-\frac{2}{3},\frac{1}{3}\right)$$

$$f_{\vec{l}}\left(P_0\right)=1\cdot\frac{2}{3}+2\cdot(-\frac{2}{3})+3\cdot\frac{1}{3}=\frac{1}{3}$$

【例 2】 $f(x,y)=\begin{cases}\dfrac{xy}{\sqrt{x^2+y^2}}, & x^2+y^2\neq 0\\ 0, & x^2+y^2=0\end{cases}$,

求 f 在点 $(0,0)$ 沿任意方向 $\vec{l}_0=(\cos\alpha,\cos\beta)$ 的方向导数。

解　由第 1 节例 6 知，$f_x\left(0,0\right)=f_y\left(0,0\right)=0$，如果按上面公式计算，得

$$f_{\vec{l}_0}\left(0,0\right)=0$$

但这是错误的。因为

$$f_{\vec{l}_0}\left(0,0\right)=\lim_{\rho\to 0^+}\frac{f(\Delta x,\Delta y)-f(0,0)}{\rho}=\lim_{\rho\to 0^+}\frac{\Delta x}{\rho}\cdot\frac{\Delta y}{\rho}=\cos\alpha\cos\beta$$

原因在于，f 在点 $(0,0)$ 不可微，不满足定理中的条件。

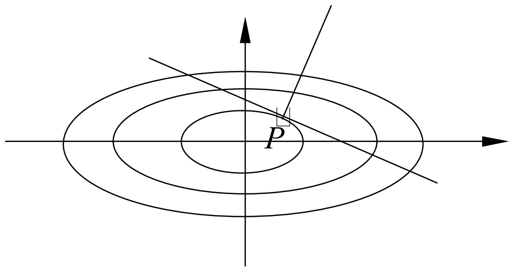
【二】 梯度

设 $\|\vec{l}\|=1$，则

$$f_{\vec{l}}(P)=\operatorname{grad}f(P)\cdot\vec{l}=\|\operatorname{grad}f(P)\|\cos(\operatorname{grad}f(P),\vec{l})$$

因此，当 $\vec{l}$ 取方向 $\operatorname{grad}f(P)$ 时，函数值（局部）增加最快。当 $\vec{l}$ 取方向 $-\operatorname{grad}f(P)$ 时，函数值（局部）下降最快。

设 $z=f(x,y)$ 看作空间曲面，令 $z=f(x,y)=c$，则得到 xy 平面的一条曲线，c 变化则得一簇曲线，称为 $z=f(x,y)$ 的等高线（或等值线）。



如图，从 P 点出发，移动相同的距离，函数值 z 增加的值不相等。$f(x,y)=c$ 两边微分 $f_x dx+f_y dy=0$，则 $\frac{dy}{dx}=-\frac{f_y}{f_x}$，从而 $\vec{t}=\pm(f_y,-f_x)$ 为曲线 $f(x,y)=c$ 的切线方向，而 $\mathrm{grad}\, f=(f_x,f_y)$ 正是法线方向。

§4　泰勒公式与极值问题

【一】 高阶偏导数

【例 1】（P137）求 $z=e^{x+2y}$ 的所有二阶偏导数和 $\dfrac{\partial^3 z}{\partial y\partial x^2}$

解
$$\frac{\partial u}{\partial x}=e^{x+2y}，\ \frac{\partial u}{\partial y}=2e^{x+2y}$$

$$\frac{\partial^2 u}{\partial x^2}=\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right)=e^{x+2y}，\ \frac{\partial^2 u}{\partial y^2}=\frac{\partial}{\partial y}\left(\frac{\partial u}{\partial y}\right)=4e^{x+2y}$$

$$\frac{\partial^2 u}{\partial x\partial y}=\frac{\partial}{\partial y}\left(\frac{\partial u}{\partial x}\right)=2e^{x+2y}，\ \frac{\partial^2 u}{\partial y\partial x}=\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial y}\right)=2e^{x+2y}$$

$$\frac{\partial^3 z}{\partial y\partial x^2}=\frac{\partial}{\partial x}\left(\frac{\partial^2 z}{\partial y\partial x}\right)=2e^{x+2y}$$

【例 2】（P138）

$$f(x,y)=\begin{cases} xy\dfrac{x^2-y^2}{x^2+y^2}, & x^2+y^2\neq 0\\ 0, & x^2+y^2=0\end{cases}$$

它的一阶偏导数为

$$f_x(x,y)=\begin{cases} \dfrac{y\left(x^4+4x^2y^2-y^4\right)}{\left(x^2+y^2\right)^2}, & x^2+y^2\neq 0\\ 0, & x^2+y^2=0\end{cases}$$

$$f_y(x,y)=\begin{cases} \dfrac{x\left(x^4-4x^2y^2-y^4\right)}{\left(x^2+y^2\right)^2}, & x^2+y^2\neq 0\\ 0, & x^2+y^2=0\end{cases}$$

进而

$$f_{xy}(0,0)=\lim_{\Delta y\to 0}\frac{f_x(0,\Delta y)-f_x(0,0)}{\Delta y}=\lim_{\Delta y\to 0}\frac{-\Delta y}{\Delta y}=-1$$

$$f_{yx}(0,0)=\lim_{\Delta x\to 0}\frac{f_y(\Delta x,0)-f_y(0,0)}{\Delta x}=\lim_{\Delta x\to 0}\frac{\Delta x}{\Delta x}=1$$

【定理 1】若 $f_{xy}(x,y)$ 和 $f_{yx}(x,y)$ 都在点 (x_0,y_0) 连续，则

$$f_{xy}(x_0,y_0)=f_{yx}(x_0,y_0)$$

（证明略）

这个定理的结论对 n 元函数的混合偏导数也成立。如三元函数 $u=f(x,y,z)$，若下述六个三阶混合偏导数

$$f_{xyz}(x,y,z),f_{yzx}(x,y,z),f_{zxy}(x,y,z),$$

$$f_{xzy}(x,y,z),f_{yxz}(x,y,z),f_{zyx}(x,y,z)$$

在某一点都连续，则在这一点六个混合偏导数都相等；同样，若二元函数 $z=f(x,y)$ 在点 (x,y) 存在直到 n 阶的连续混合偏导数，则在这一点 $m(\le n)$ 阶混合偏导数都与顺序无关.

今后除特别指出外，都假设相应阶数的混合偏导数连续，从而混合偏导数与求导顺序无关.

【复合函数的高阶偏导数】

设 z 是通过中间变量 x,y 而成为 s,t 的函数，即

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

其中 $x=\varphi(s,t),y=\psi(s,t)$，若函数 f,φ,ψ 都具有连续的二阶偏导数，则作为复合函数的 z 对 s,t 同样存在二阶连续偏导数。具体计算如下：

$$\frac{\partial z}{\partial s}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial s}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial s},\quad \frac{\partial z}{\partial t}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial t}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial t}.$$

显然 $\frac{\partial z}{\partial s}$ 与 $\frac{\partial z}{\partial t}$ 仍是 s,t 的复合函数，其中 $\frac{\partial z}{\partial x},\frac{\partial z}{\partial y}$ 是 x,y 的函数，$\frac{\partial x}{\partial s},\frac{\partial x}{\partial t},\frac{\partial y}{\partial s},\frac{\partial y}{\partial t}$ 是 s,t 的函数。继续求 z 关于 s,t 的二阶偏导数

$$\frac{\partial^2 z}{\partial s^2}=\frac{\partial}{\partial s}\left(\frac{\partial z}{\partial x}\right)\frac{\partial x}{\partial s}+\frac{\partial z}{\partial x}\cdot\frac{\partial}{\partial s}\left(\frac{\partial x}{\partial s}\right)+\frac{\partial}{\partial s}\left(\frac{\partial z}{\partial y}\right)\frac{\partial y}{\partial s}+\frac{\partial z}{\partial y}\cdot\frac{\partial}{\partial s}\left(\frac{\partial y}{\partial s}\right)$$

$$=\left(\frac{\partial^2 z}{\partial x^2}\frac{\partial x}{\partial s}+\frac{\partial^2 z}{\partial x\partial y}\frac{\partial y}{\partial s}\right)\frac{\partial x}{\partial s}+\frac{\partial z}{\partial x}\cdot\frac{\partial^2 x}{\partial s^2}+\left(\frac{\partial^2 z}{\partial y\partial x}\frac{\partial x}{\partial s}+\frac{\partial^2 z}{\partial y^2}\frac{\partial y}{\partial s}\right)\frac{\partial y}{\partial s}+\frac{\partial z}{\partial y}\frac{\partial^2 y}{\partial s^2}$$

$$=\frac{\partial^2 z}{\partial x^2}\left(\frac{\partial x}{\partial s}\right)^2+2\frac{\partial^2 z}{\partial x\partial y}\frac{\partial x}{\partial s}\cdot\frac{\partial y}{\partial s}+\frac{\partial^2 z}{\partial y^2}\left(\frac{\partial y}{\partial s}\right)^2+\frac{\partial z}{\partial x}\frac{\partial^2 x}{\partial s^2}+\frac{\partial z}{\partial y}\frac{\partial^2 y}{\partial s^2}.$$

同理可求$\frac{\partial^2 z}{\partial t^2},\frac{\partial^2 z}{\partial s\partial t}$.

【例 3】 设$z=f\left(x,\frac{x}{y}\right)$,，求$\frac{\partial^2 z}{\partial x^2}$，$\frac{\partial^2 z}{\partial x\partial y}$.

解　$z=f(u,v),u=x,v=\frac{x}{y}$.

$$\frac{\partial z}{\partial x}=\frac{\partial f}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial f}{\partial v}\frac{\partial v}{\partial x}=\frac{\partial f}{\partial u}+\frac{1}{y}\frac{\partial f}{\partial u}.$$

所以

$$\frac{\partial^2 \mathrm{z}}{\partial x^2}=\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial u}+\frac{1}{y}\frac{\partial f}{\partial v}\right)$$

$$=\frac{\partial^2 f}{\partial u^2}\frac{\partial u}{\partial x}+\frac{\partial^2 f}{\partial u\partial v}\frac{\partial v}{\partial x}+\frac{1}{y}\left(\frac{\partial^2 f}{\partial v\partial u}\frac{\partial u}{\partial x}+\frac{\partial^2 f}{\partial v^2}\frac{\partial v}{\partial x}\right)=\frac{\partial^2 f}{\partial u^2}+\frac{2}{y}\frac{\partial^2 f}{\partial u\partial v}+\frac{1}{y^2}\frac{\partial^2 f}{\partial v^2}$$

$$\frac{\partial^2 z}{\partial x\partial y}=\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial u}+\frac{1}{y}\frac{\partial f}{\partial v}\right)=\frac{\partial^2 f}{\partial u^2}\frac{\partial u}{\partial y}+\frac{\partial^2 f}{\partial u\partial v}\frac{\partial v}{\partial y}-\frac{1}{y^2}\frac{\partial f}{\partial v}+\frac{1}{y}\left(\frac{\partial^2 f}{\partial v\partial u}\frac{\partial u}{\partial y}+\frac{\partial^2 f}{\partial v^2}\frac{\partial v}{\partial y}\right)$$

$$=-\frac{x}{y^2}\frac{\partial^2 f}{\partial u\partial v}-\frac{x}{y^3}\frac{\partial^2 f}{\partial v^2}-\frac{1}{y^2}\frac{\partial f}{\partial v}.$$

【例 4】验证$u=\frac{1}{r}(r=\sqrt{x^2+y^2+z^2})$满足方程$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}=0$。

$$\frac{\partial u}{\partial x}=\frac{du}{dr}\cdot\frac{\partial r}{\partial x}=-\frac{1}{r^2}\cdot\frac{1}{2}\frac{2x}{\sqrt{x^2+y^2+z^2}}=-\frac{1}{r^2}\cdot\frac{x}{r}=-\frac{x}{r^3}$$

$$\frac{\partial^2 u}{\partial x^2}=-\frac{1}{r^3}-x\left(-\frac{3}{r^4}\cdot\frac{x}{r}\right)=-\frac{1}{r^3}+\frac{3x^2}{r^5}$$

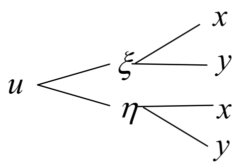
由对称性，

$$\frac{\partial^2 u}{\partial y^2}=-\frac{1}{r^3}+\frac{3y^2}{r^5},\quad\frac{\partial^2 u}{\partial z^2}=-\frac{1}{r^3}+\frac{3z^2}{r^5}$$

因此

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}=-\frac{3}{r^3}+\frac{3(x^2+y^2+z^2)}{r^5}=-\frac{3}{r^3}+\frac{3r^2}{r^5}=0$$

【例 5】设$u=f(x,y)$，在新坐标系$\begin{cases}\xi=x-y\\ \eta=x+y\end{cases}$下，求$\dfrac{\partial^2 u}{\partial x^2}-\dfrac{\partial^2 u}{\partial y^2}$的表达式。



$$\frac{\partial u}{\partial x}=\frac{\partial u}{\partial \xi}\frac{\partial \xi}{\partial x}+\frac{\partial u}{\partial \eta}\frac{\partial \eta}{\partial x}=\frac{\partial u}{\partial \xi}+\frac{\partial u}{\partial \eta}$$

$$\frac{\partial^2 u}{\partial x^2}=\frac{\partial^2 u}{\partial \xi^2}+\frac{\partial^2 u}{\partial \xi\partial \eta}+\frac{\partial^2 u}{\partial \eta\partial \xi}+\frac{\partial^2 u}{\partial \eta^2}=\frac{\partial^2 u}{\partial \xi^2}+2\frac{\partial^2 u}{\partial \xi\partial \eta}+\frac{\partial^2 u}{\partial \eta^2}$$

同理，

$$\frac{\partial^2 u}{\partial y^2}=\frac{\partial^2 u}{\partial \xi^2}-2\frac{\partial^2 u}{\partial \xi\partial \eta}+\frac{\partial^2 u}{\partial \eta^2}$$

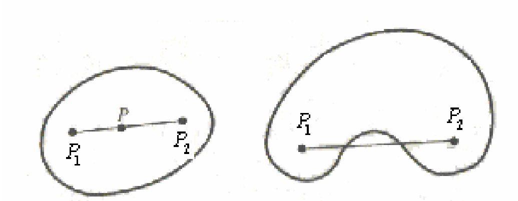
因此

$$\frac{\partial^2 u}{\partial x^2}-\frac{\partial^2 u}{\partial y^2}=4\frac{\partial^2 u}{\partial \xi\partial \eta}$$

【二】 中值定理和泰勒公式

凸区域：若区域D上任意两点的连线都含于D，则称D为凸区域．即任意两点$P_1\ (x_1,y_1),P_2(x_2,y_2)\in D$和一切$\lambda(0\le\lambda\le 1)$，恒有

$$P(x_1+\lambda(x_2-x_1),y_1+\lambda(y_2-y_1))\in D$$



凸区域　　非凸区域

【定理 2】（中值定理）　设二元函数f在凸开域$D\subset R^2$上连续，在D的所有内点都可微，则对D内任意两点$P_0(x_0,y_0),P_1(x_0+h,y_0+k)\in \operatorname{int}D$，在$P_0,P_1$的连线上存在某点

$P(x_0+\theta h, y_0+\theta k)(0<\theta<1)$，使得

$$f(P_1)-f(P_0)=f_x(P)h+f_y(P)k$$

证　令

$$\Phi(t)=f(x_0+th, y_0+tk)$$

它是定义在$[0,1]$上的一元函数，由定理中的条件知$\Phi(t)$在$[0,1]$上连续，在$(0,1)$内可微．于是根据一元函数中值定理，存在$\theta(0<\theta<1)$使得

$$\Phi(1)-\Phi(0)=\Phi'(\theta)$$

由复合函数的求导法则

$$\Phi'(\theta)=f_x(x_0+\theta h, y_0+\theta k)h+f_y(x_0+\theta h, y_0+\theta k)k$$

得证。

【定理 3】（泰勒定理 1）　若函数f在点$P_0(x_0,y_0)$的某邻域$U(P_0)$内有直到$n+1$阶的连续偏导数，则对$U(P_0)$内任一点(x_0+h, y_0+k)，存在$\theta\in(0,1)$，使得

$$\begin{aligned}f(x_0+h,y_0+k)=&f(x_0,y_0)+(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y})f(x_0,y_0)\\&+\frac{1}{2!}(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y})^2 f(x_0,y_0)+\cdots+\frac{1}{n!}(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y})^n f(x_0,y_0)\\&+\frac{1}{(n+1)!}(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y})^{n+1} f(x_0+\theta h,y_0+\theta k).\end{aligned}$$

其中

$$(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y})^m f(x_0,y_0)=\sum_{i=0}^{m}C_m^i\frac{\partial^m}{\partial x^i\partial y^{m-i}}f(x_0,y_0)h^i k^{m-i}.$$

证　作函数

$$\Phi(t)=f(x_0+th, y_0+tk).$$

$\Phi(t)$在$[0,1]$上满足一元函数泰勒定理条件, 于是有

$$\Phi(1)=\Phi(0)+\frac{\Phi'(0)}{1!}+\frac{\Phi''(0)}{2!}+\cdots+\frac{\Phi^{(n)}(0)}{n!}+\frac{\Phi^{(n+1)}(\theta)}{(n+1)!}\quad(0<\theta<1).$$

应用复合函数求导法则,可求得$\Phi(t)$的各阶导数：

$$\Phi^{(m)}(t)=(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y})^m f(x_0+th,y_0+tk).\quad (m=1,2,\cdots,n+1).$$

于是

$$\Phi^{(m)}(0)=(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y})^m f(x_0,y_0)\qquad (m=1,2,\cdots,n).$$

及

$$\Phi^{(n+1)}(\theta)=(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y})^{n+1} f(x_0+\theta h,y_0+\theta k).$$

代入便得证。

易见，中值公式正是泰勒公式在$n=0$时的特殊情形.

【定理 4】（泰勒定理 2）　若函数f在点$P_0(x_0,y_0)$的某邻域$U(P_0)$内有直到n阶的连续偏导数，则对$U(P_0)$内任一点(x_0+h,y_0+k)，有

$$f(x_0+h,y_0+k)=f(x_0,y_0)+\sum_{P=0}^{n}\frac{1}{p!}(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y})^P f(x_0,y_0)+o(\rho^n).$$

其中$\rho=\sqrt{h^2+k^2}$。

证 由泰勒定理 1，

$$f(x_0+h,y_0+k)=f(x_0,y_0)+\sum_{P=0}^{n-1}\frac{1}{p!}(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y})^P f(x_0,y_0)$$

$$+\frac{1}{n!}(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y})^n f(x_0+\theta h,y_0+\theta k).$$

因为f具有n阶的连续偏导数,所以（下面$\alpha_1+\alpha_2=n$）

$$h^{\alpha_1}k^{\alpha_2}\frac{\partial^n}{\partial x^{\alpha_1}\partial y^{\alpha_2}}f(x_0+\theta h,y_0+\theta k)=h^{\alpha_1}k^{\alpha_2}\left[\frac{\partial^n}{\partial x^{\alpha_1}\partial y^{\alpha_2}}f(x_0,y_0)+o(1)\right]$$

而$\dfrac{\left|h^{\alpha_1}k^{\alpha_2}\right|}{\rho^n}\le 1$，因此$h^{\alpha_1}k^{\alpha_2}o(1)=o(\rho^n)$，于是

$$\frac{1}{n!}(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y})^n f(x_0+\theta h,y_0+\theta k)=\frac{1}{n!}(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y})^n f(x_0,y_0)+o(\rho^n)$$

【注】n 元函数的 Taylor 公式（2 阶）：设 $\boldsymbol{x},\boldsymbol{x}_0\in\mathbf{R}^n$，$\Delta\boldsymbol{x}=\boldsymbol{x}-\boldsymbol{x}_0$，则

$$f(\boldsymbol{x}+\Delta\boldsymbol{x})=f(\boldsymbol{x}_0)+f'(\boldsymbol{x}_0)\Delta\boldsymbol{x}+\Delta\boldsymbol{x}^{\mathrm{T}}f''(\boldsymbol{x}_0)\Delta\boldsymbol{x}+o(\|\Delta\boldsymbol{x}\|^2)$$

这里

$$f'(\boldsymbol{x}_0)=\left[\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2},\cdots,\frac{\partial f}{\partial x_n}\right]_{\boldsymbol{x}_0}$$

$$f''(\boldsymbol{x}_0)=\boldsymbol{H}_f(\boldsymbol{x}_0)=\begin{bmatrix}\dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1\partial x_n}\\ \dfrac{\partial^2 f}{\partial x_2\partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2\partial x_n}\\ \vdots & \vdots & & \vdots\\ \dfrac{\partial^2 f}{\partial x_n\partial x_1} & \dfrac{\partial^2 f}{\partial x_n\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2}\end{bmatrix}_{\boldsymbol{x}_0}$$

$\boldsymbol{H}_f(\boldsymbol{x}_0)$ 称 Hesse 矩阵。

【三】 极值问题

多元函数极值的定义与一元函数完全类似。

若 f 在点 (x_0,y_0) 取得极值，则当固定 $y=y_0$ 时，一元函数 $f(x,y_0)$ 必定 $x=x_0$ 在取相同的极值上．同理，一元函数 $f(x_0,y)$ 在 $y=y_0$ 也取相同的极值．于是得到二元函数取极值的必要条件如下：

【定理 5】（极值必要条件）　若函数 f 在点 $P_0(x_0,y_0)$ 存在偏导数，且在 P_0 取得极值，则有

$$f_x(x_0,y_0)=f_y(x_0,y_0)=0.$$

若函数 f 在点 P_0 满足 $f_x(P_0)=f_y(P_0)=0$，则称点 P_0 为 f 的**稳定点**.

【注 1】稳定点并不都是极值点，如 $f(x,y)=xy$，原点为其稳定点，但它在原点不取极值.

【注 2】函数在偏导数不存在的点上也有可能取得极值。如 $f(x,y)=\sqrt{x^2+y^2}$ 在原点没有偏导数，但 $f(0,0)=0$ 是 f 的极小值.

【引理 1】设 A 是实对称矩阵，则

$$\lambda_1 x^T x \le x^T A x \le \lambda_n x^T x$$

其中 λ_1, λ_n 是 A 的最小与最大特征值。

【引理 2】二阶对称矩阵 $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$，则

（1） A 正定 $\Leftrightarrow a > 0, ac - b^2 > 0$

（2） A 负定 $\Leftrightarrow a < 0, ac - b^2 > 0$

（3） A 不定 $\Leftrightarrow ac - b^2 < 0$

【定理 6】（极值充分条件）　设二元函数 f 在点 $P_0(x_0, y_0)$ 的某邻域 $U(P_0)$ 内具有二阶连续导数，且 P_0 是 f 的稳定点。则

（1）当 $H_f(P_0)$ 是正定矩阵时，f 在 P_0 取得严格极小值；

（2）当 $H_f(P_0)$ 是负定矩阵时，f 在 P_0 取得严格极大值；

（3）当 $H_f(P_0)$ 是不定矩阵时，f 在 P_0 不取极值.

证　由 f 在 P_0 的二阶泰勒公式，并注意到条件 $f_x(P_0) = f_y(P_0) = 0,$，有

$$f(x,y) - f(x_0,y_0) = \frac{1}{2}(\Delta x, \Delta y) H_f(P_0)(\Delta x, \Delta y)^T + o\left(\Delta x^2 + \Delta y^2\right).$$

由于 $H_f(P_0)$ 正定，由引理 1

$$(\Delta x, \Delta y) H_f(P_0)(\Delta x, \Delta y)^T \ge \lambda_1 (\Delta x, \Delta y)(\Delta x, \Delta y)^T = \lambda_1(\Delta x^2 + \Delta y^2)$$

这里 $\lambda_1 > 0$。

$$f(x,y) - f(x_0,y_0) \ge \frac{1}{2}\lambda_1\left(\Delta x^2 + \Delta y^2\right) + o\left(\Delta x^2 + \Delta y^2\right) = \left(\Delta x^2 + \Delta y^2\right)\left(\frac{1}{2}\lambda_1 + o(1)\right)$$

因此，当 $\Delta x, \Delta y$ 充分小时（$(\Delta x, \Delta y) \ne (0,0)$），有

$$f(x,y) - f(x_0,y_0) > 0$$

即 f 在点 (x_0, y_0) 取得严格极小值.

同理可证 $H_f(P_0)$ 为负定矩阵时，f 在 P_0 取得严格极大值.

最后证明，当 $H_f(P_0)$ 不定时，f 在 P_0 不取极值.

这是因为倘若 f 取极值（例如取极大值），则沿任何过 P_0 的直线 $x = x_0 + t\Delta x$, $y = y_0 + t\Delta y$, $f(x,y) = f(x_0 + t\Delta x, y_0 + t\Delta y) = \varphi(t)$ 在 $t = 0$ 亦取极大值. 故 $\varphi''(0) \le 0$.（否则，如 $\varphi''(0) > 0$，则 φ 在 $t = 0$ 将取极小值)，而

$$\varphi'(t) = f_x\Delta x + f_y\Delta y, \varphi''(t) = f_{xx}\Delta x^2 + 2f_{xy}\Delta x\Delta y + f_{yy}\Delta y^2,$$

$$\varphi''(0) = (\Delta x, \Delta y) H_f(P_0) (\Delta x, \Delta y)^T.$$

这表明 $H_f(P_0)$ 必须是负半定的。同理，倘若 f 取极小值，则将导致 $H_f(P_0)$ 必须是正半定的。也就是说，当 f 在 P_0 取极值时，$H_f(P_0)$ 必须时正半定或负半定矩阵，但这与假设相矛盾.

对于二元函数 f，P_0 是 f 的稳定点，由引理 2

(ⅰ) 当 $f_{xx}(P_0) > 0, (f_{xx}f_{yy} - f_{xy}^2)(P_0) > 0$ 时，f 在点 P_0 取得严格极小值;

(ⅱ) 当 $f_{xx}(P_0) < 0, (f_{xx}f_{yy} - f_{xy}^2)(P_0) > 0$ 时，f 在点 P_0 取得严格极大值;

(iii) 当 $(f_{xx}f_{yy} - f_{xy}^2)(P_0) < 0$ 时，f 在点 P_0 不能取得极值;

(iv) 当 $(f_{xx}f_{yy} - f_{xy}^2)(P_0) = 0$ 时，不能肯定 f 在点 P_0 是否取得极值.

【例 6】　求 $f(x,y) = e^{x-y}(x^2 - 2y^2)$ 的极值.

解　由方程组

$$\begin{cases} f_x = e^{x-y}(x^2 - 2y^2 + 2x) = 0 \\ f_y = -e^{x-y}(x^2 - 2y^2 + 4y) = 0 \end{cases}$$

得 f 的稳定点 $(0,0), (-4,-2)$。

由于 $H_f(0,0) = \begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix}$ 是不定矩阵，故 f 在点 $(0,0)$ 不取极值。

由于 $H_f(-4,-2) = e^{-2}\begin{bmatrix} -6 & 8 \\ 8 & -12 \end{bmatrix}$ 是负正矩阵，故 f 在点 $(-4,-2)$ 取严格极大值。

又因 f 处处存在偏导数，故 $(-4,-2)$ 为 f 的唯一极大值点，也是最大值点.

$$f_{\max} = f(-4,-2) = 8e^2$$

【例 7】(P147）讨论 $f(x,y) = x^2 + xy$ 是否存在极值.

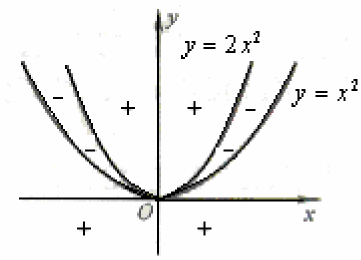
解　由方程组 $f_x = 2x + y = 0, f_y = x = 0$ 解得稳定点为原点 $O(0,0)$.

因 $f_{xx}f_{yy} - f_{xy}^2 = -1 < 0,$ ，故原点不是 f 的极值点。又因 f 处处可微，所以 f 没有极值点。

【例 8】(P148)　讨论 $f(x,y) = (y - x^2)(y - 2x^2)$ 在原点是否取得极值.

解　容易验证原点是 f 的稳定点，且在原点 $f_{xx}f_{yy} - f_{xy}^2 = 0,$ 故由上面定理无法判定 f 在原点是否取到极值.

$f(0,0) = 0$ ，但由于当 $x^2 < y < 2x^2$ 时 $f(x,y) < 0,$ 而当 $y > 2x^2$ 或 $y < x^2$ 时，$f(x,y) > 0,$ 所以函数 f 不可能在原点取得极值.



【例 9】(线性最小二乘问题)

给定一组测量数据 $(t_i, y_i)(i = 1,2,\cdots,m)$ （t_i 互不同）和 n 个函数 $\varphi_1(t),\cdots,\varphi_n(t)$ （一般地 $m \gg n$ ），记函数

$$\varphi(t) = x_1\varphi_1(t) + x_2\varphi_2(t) + \cdots + x_n\varphi_n(t)$$

求 $x = (x_1, x_2, \cdots, x_n)^T \in R^n$ 使得在点 $t_i (i = 1,2,\cdots,m)$ 上测量值 y_i 与函数值 $\varphi(t_i)$ 误差的平方和达到最小。即

$$\min r(x) = \sum_{i=1}^{m}\left[y_i - \varphi(t_i)\right]^2 = \sum_{i=1}^{m}\left[y_i - \sum_{k=1}^{n} x_k\varphi_k(t_i)\right]^2$$

记 $m \times n$ 的矩阵

$$A=\begin{bmatrix}\varphi_1(t_1) & \varphi_2(t_1) & \cdots & \varphi_n(t_1)\\ \varphi_1(t_2) & \varphi_2(t_2) & \cdots & \varphi_n(t_2)\\ \vdots & \vdots & & \vdots\\ \varphi_1(t_m) & \varphi_2(t_m) & \cdots & \varphi_n(t_m)\end{bmatrix}$$

并假设 A 的列向量线性无关。再记向量 $b=(y_1,y_2,\cdots,y_m)^T\in R^m$ 。则

$$r(x)=\|b-Ax\|^2=x^T(A^TA)x-2x^T(A^Tb)+b^Tb$$

令 $r'(x)=0$ ，即

$$2\left[(A^TA)x-(A^Tb)\right]=0$$

解得唯一的稳定点是（注意， A^TA 是正定矩阵）

$$x_0=(A^TA)^{-1}(A^Tb)$$

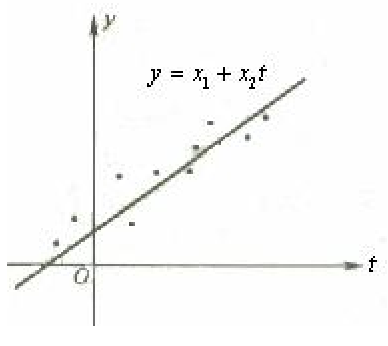
又

$$r''(x)=2(A^TA)$$

是正定矩阵。因此 $r(x)$ 在点 x_0 取得唯一的最小值。

特别地（P149），通过观测或实验得到一列点 $(t_i,y_i), i=1,2,\cdots,m$ 。它们大体上在一条直线上，取 $\varphi_1(t)=1,\varphi_2(t)=t$

$$y=x_1\varphi_1(t)+x_2\varphi_2(t)=x_1+x_2t$$



记

$$A=\begin{bmatrix}1 & t_1\\ 1 & t_2\\ \vdots & \vdots\\ 1 & t_m\end{bmatrix},\quad b=\begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_m\end{bmatrix}$$

注意 t_i 互不等，因此 A 的列向量线性无关。

$$A^T A=\begin{bmatrix} m & \sum_{i=1}^{m} t_i \\ \sum_{i=1}^{m} t_i & \sum_{i=1}^{m} t_i^2 \end{bmatrix},\left(A^T A\right)^{-1}=\frac{1}{m\sum_{i=1}^{m} t_i^2-\left(\sum_{i=1}^{m} t_i\right)^2}\begin{bmatrix} \sum_{i=1}^{m} t_i^2 & -\sum_{i=1}^{m} t_i \\ -\sum_{i=1}^{m} t_i & m \end{bmatrix},A^T b=\begin{bmatrix} \sum_{i=1}^{m} y_i \\ \sum_{i=1}^{m} t_i y_i \end{bmatrix}$$

由 $x=(A^T A)^{-1}(A^T b)$ 求得

$$x_1=\frac{\left(\sum t_i^{\,2}\right)\left(\sum y_i\right)-\left(\sum t_i y_i\right)\left(\sum t_i\right)}{m\sum t_i^2-(\sum t_i)^2},\ x_2=\frac{m\sum t_i y_i-(\sum t_i)(\sum y_i)}{m\sum t_i^2-(\sum t_i)^2}$$