第五章　多元微分

第一节、多元函数的偏导数

一、偏导数的定义

设二元函数 $z = f(x,y)$

1、在定点处的偏导数

(1) $\lim\limits_{\Delta x\to 0}\dfrac{f(x_0+\Delta x,y_0)-f(x_0,y_0)}{\Delta x}$ 叫做函数 $f(x,y)$ 在点 (x_0,y_0) 处对 x 的偏导数，记作 $f_x(x_0,y_0)$ ，

(2) $\lim\limits_{\Delta y\to 0}\dfrac{f(x_0,y_0+\Delta y)-f(x_0,y_0)}{\Delta y}$ 叫做函数 $f(x,y)$ 在点 (x_0,y_0) 处对 y 的偏导数，记作 $f_y(x_0,y_0)$

2、偏导数与一元函数的导数

设 $g(x)=f(x,y_0)$ ， $q(y)=f(x_0,y)$ ，则 $f_x(x_0,y_0)=g'(x_0)$ ， $f_y(x_0,y_0)=q'(y_0)$

3、一般的偏导数

函数 $f(x,y)$ 在点 (x,y) 处的偏导数 $f_x(x,y)$ ， $f_y(x,y)$ 也是 x,y 的二元函数

4、常用记号

$$f_x(x,y)=f_x=\frac{\partial f}{\partial x}=\frac{\partial}{\partial x}f(x,y)=\frac{\partial z}{\partial x}=D_x f，\quad f_y(x,y)=f_y=\frac{\partial f}{\partial y}=\frac{\partial}{\partial y}f(x,y)=\frac{\partial z}{\partial y}=D_y f$$

二、偏导数的几何意义

例、抛物面 $f(x,y)=x^2+y^2$ ，求(1) $f_x(x,y)$, $f_x(1,2)$　(2) $f_y(x,y)$, $f_y(1,2)$

解：(1) $f_x(x,y)=2x$, $f_x(1,2)=2$　　(2) $f_y(x,y)=2y$, $f_y(1,2)=4$

其中 $f_x(1,2)=2$ 表示抛物面 $z=x^2+y^2$ 被平面 $y=2$ 截得的曲线 $\begin{cases} z=x^2+4 \\ y=2 \end{cases}$ 在 x=1 处的切线的斜率,如图 1

其中 $f_y(1,2)=4$ 表示抛物面 $z=x^2+y^2$ 被平面 $x=1$ 截得的曲线 $\begin{cases} z=1+y^2 \\ x=1 \end{cases}$ 在 y=2 处的切线的斜率,如图 2

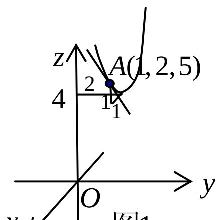


图1

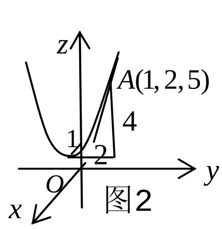


图2

三、高阶偏导数

1、求高阶偏导

例、求 $f(x,y)=x^3+x^2y^3-2y^2$ 的二阶导数

解： $f_x=3x^2+2xy^3, f_y=3x^2y^2-4y$， $f_{xx}=6x+2y^3, f_{yy}=6x^2y-4$

$f_{xy}=6xy^2, f_{yx}=6xy^2$

2、克莱罗定理

在 f 的定义域 D 内含有点 (a,b)，若 f_{xy}, f_{yx} 都在 D 内连续，则 $f_{xy}(a,b)=f_{yx}(a,b)$

四、切平面：设点 A (x_0,y_0,z_0) 在曲面 C： $z=f(x,y)$ 上，则

曲面 C 被平面 $y=y_0$ 截得的曲线 $\begin{cases} z=x^2+{y_0}^2 \\ y=y_0 \end{cases}$ 在点 A 处的切线是 $l_1:\begin{cases} z=z_0+f_x(x_0,y_0)(x-x_0) \\ y=y_0 \end{cases}$

曲面 C 被平面 $x=x_0$ 截得的曲线 $\begin{cases} z={x_0}^2+y^2 \\ x=x_0 \end{cases}$ 在点 A 处的切线是 $l_2:\begin{cases} z=z_0+f_y(x_0,y_0)(y-x_0) \\ x=x_0 \end{cases}$

切线 l_1 与 l_2 相交于 A 点，它们所确定的平面 $z=z_0+f_x(x_0,y_0)(x-x_0)+f_y(x_0,y_0)(y-x_0)(*)$ 就是曲面 C 的切平面方程。

证明：首先方程 $(*)$ 是平面方程，其次在 $(*)$ 中令 $y=y_0$ 就得切线 l_1，令 $x=x_0$ 就得切线 l_2，于是方程 $(*)$ 表示的平面过 l_1 与 l_2，因此它为切平面

例、求 $f(x,y)=2x^2+y^2$ 在点(1,1,3)的切平面与法线

解：(1) $f_x(x,y)=4x, f_y(x,y)=2y, f_x(1,1)=4, f_y(1,1)=2$，于是切平面为

$z=3+4(x-1)+2(y-1)$，即 $z=4x+2y-3$

(2)切平面方程可化为 $4(x-1)+2(y-1)-(z-3)=0$ 于是法向量为 $(4,2,-1)$

因此法线方程为 $\dfrac{x-1}{4}=\dfrac{y-1}{2}=\dfrac{z-3}{-1}$

五、近似与全微分

1、线性近似：平面 $L(x,y)=f(x_0,y_0)+f_x(x_0,y_0)(x-x_0)+f_y(x_0,y_0)(y-x_0)$ 是 $f(x,y)$ 在点 $(x_0,y_0,f(x_0,y_0))$ 处的切平面 $L(x,y)$ 叫做 $f(x,y)$ 在点 (x_0,y_0) 处的线性近似函数。

当 (x,y) 很接近 (x_0,y_0) 时， $f(x,y)\approx T(x,\mathrm{y})$，例如

$f(x_0+0.03,y_0+0.02)\approx L(x_0+0.03,y_0+0.02)=f(x_0,y_0)+0.03f_x(x_0,y_0)+0.02f_y(x_0,y_0)$

值得注意的是这种近似不如一元函数可靠

2、可微的定义：设 $z=f(x,y)$， $\Delta z=f(x_0+\Delta x,y_0+\Delta y)-f(x_0,y_0)$

若 $\Delta z=f_x(x_0,y_0)\Delta x+f_y(x_0,y_0)\Delta y+\varepsilon_1\Delta x+\varepsilon_2\Delta y$，这里 $(\Delta x,\Delta y)\to(0,0)\Rightarrow\varepsilon_1\to 0,\varepsilon_2\to 0$，则称 $z=f(x,y)$ 在 (x_0,y_0) 可微

由可微的定义知当 $z=f(x,y)$ 在点 (x_0,y_0) 可微，则线性近似是可靠的

{国内的定义：设 $z=f(x,y)$，若存在只与 (x_0,y_0) 相关的常数 A、B 使

$\Delta z=A\Delta x+B\Delta y+o(\sqrt{(\Delta x)^2+(\Delta y)^2})$ 则称 $z=f(x,y)$ 在 (x_0,y_0) 可微}

3、可微的判定：若 f_x, f_y 在 (x_0, y_0) 附近存在，并且在 (x_0, y_0) 点连续，则 f 在 (x_0, y_0) 点是可微的。

4、全微分：当 $z = f(x, y)$ 在点 (x, y) 可微时，$\Delta z = f_x(x, y)\Delta x + f_y(x, y)\Delta y + \varepsilon_1\Delta x + \varepsilon_2\Delta y$

这样，$f_x(x, y)\Delta x + f_y(x, y)\Delta y$ (*)就可为 Δz 的好的近似值，把(*)叫做 z 在点 (x, y) 处的全微分，记着 $dz = f_x(x, y)dx + f_y(x, y)dy$ ($dx = \Delta x, dy = \Delta y$)，当 $|\Delta x|,|\Delta y|$ 很小时 $\Delta z \approx dz$

例、若 $z = f(x, y) = x^2 + 3xy - z^2$，求(1) dz (2) $f(2.05, 2.96)$ 的近似值

解：(1) $dz = f_x(x, y)dx + f_y(x, y)dy = (2x + 3y)dx + (3x - 2y)dy$

(2) $f(2.05, 2.96) \approx f(2,3) + f_x(2,3)dx + f_y(2,3)dy = 13 + 13\times 0.05 + 0 = 13.65$

第二节　梯度向量

一、求导法则

1、链式法则 1：若 $z = f(x, y)$ 可微，$x = x(t), y = y(t)$ 可微，则

(1) z 是 t 的可微函数(2) $\dfrac{dz}{dt} = \dfrac{\partial f}{\partial x}\dfrac{dx}{dt} + \dfrac{\partial f}{\partial y}\dfrac{dy}{dt}$ {或$z_t' = f_x x_t' + f_y y_t'$}

证：因 $z = f(x, y)$ 可微，

于是 $\Delta z = f_x(x, y)\Delta x + f_y(x, y)\Delta y + \varepsilon_1\Delta x + \varepsilon_2\Delta y$，这里 $(\Delta x, \Delta y) \to 0 \Rightarrow \varepsilon_1 \to 0, \varepsilon_2 \to 0$

故 $\dfrac{\Delta z}{\Delta t} = f_x(x, y)\dfrac{\Delta x}{\Delta t} + f_y(x, y)\dfrac{\Delta y}{\Delta t} + \varepsilon_1\dfrac{\Delta x}{\Delta t} + \varepsilon_2\dfrac{\Delta y}{\Delta t}$

因 $x = x(t), y = y(t)$ 可微，故 $\text{当}\Delta t \to 0\text{时}, \dfrac{\Delta x}{\Delta t} \to x_t', \dfrac{\Delta y}{\Delta t} \to y_t', \Delta x \to 0, \Delta y \to 0$

所以 $\dfrac{dz}{dt} = \lim\limits_{\Delta t\to 0}\dfrac{\Delta z}{\Delta t} = f_x(x, y)x_t' + f_y(x, y)y_t'$

2、链式法则 2：若 $z = f(x, y)$ 可微，$x = x(s,t), y = y(s,t)$ 可微，则

(1) z 是关于 s,t 的可微函数(2) $\dfrac{\partial z}{\partial s} = \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial s} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial s}, \dfrac{\partial z}{\partial t} = \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial t} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial t}$

证：因 $z = f(x, y)$ 可微，$x = x(s,t), y = y(s,t)$ 可微于是

$$\Delta z = f_x(x,y)\Delta x + f_y(x,y)\Delta y + o(\Delta x) + o(\Delta y) = f_x(x,y)[x_s(s,t)\Delta s + x_t(s,t)\Delta t + o(\Delta s) + o(\Delta t)]$$

$$+ f_y(x, y)[y_s(s,t)\Delta s + y_t(s,t)\Delta t + o(\Delta s) + o(\Delta t)] + o(\Delta x) + o(\Delta y)$$

$$= [f_x(x, y)x_s(s,t) + f_y(x, y)y_s(s,t)]\Delta s + [f_x(x, y)x_t(s,t) + f_y(x, y)y_t(s,t)]\Delta t + o(\Delta s) + o(\Delta t)$$

两边除以 Δs，令 $\Delta s \to 0$ 得 故 $\dfrac{\partial z}{\partial s} = \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial s} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial s}$

两边除以 Δt，令 $\Delta t \to 0$ 得 故 $\dfrac{\partial z}{\partial t} = \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial t} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial t}$

链式法 1、2 可推广到更多元的情况

如若 $w = f(x, y, z)$ 可微，$x = x(s,t), y = y(s,t), z = y(s,t)$

则 $\dfrac{\partial w}{\partial s} = \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial s} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial s} + \dfrac{\partial f}{\partial z}\dfrac{\partial z}{\partial s}$

3、隐函数的导数：由方程 $F(x, y) = 0$ 定义了一个函数 $y = y(x)$

我们可由 $F(x, y) = 0$ 求出 $\dfrac{dy}{dx}$，在方程 $F(x, y) = 0$ 都对 x 求导得

$$\frac{\partial F}{\partial x}\frac{dx}{dx} + \frac{\partial F}{\partial y}\frac{dy}{dx} = 0 \Rightarrow F_y\cdot\frac{dy}{dx} = -F_x \Rightarrow \frac{dy}{dx} = -\frac{F_x}{F_y}$$

例、当 $x^3+y^3=6xy$ 时，求 y'

解：$x^3+y^3-6xy=0$，设 $F(x,y)=x^3+y^3-6xy$ 则 $y'=-\frac{F_x}{F_y}=-\frac{3x^2-6y}{3y^2-6x}=-\frac{x^2-2y}{y^2-2x}$

二、方向导数

1、方向导数的认识

方向导数是偏导数的推广，若 $z=f(x,y)$，则 $\frac{\partial f}{\partial x}$ 与 $\frac{\partial f}{\partial y}$ 分别是 $z=f(x,y)$ 沿平面 xoy 内的单位向量(1,0)和(0,1)方向上的导数。

函数 $z=f(x,y)$ 沿平面 xoy 内的向量沿单位向量 $\vec{u}=(\frac{\sqrt{3}}{2},\frac{1}{2})$ 方向上的导数该如何规定呢？设平面 xoy 内点 $A(x_0,y_0)$，沿向量 $\vec{u}$ 的方向移动一个增量 h(长度添符号的实数)，则它在 $\vec{i}$ 与 $\vec{j}$ 方向的分量分别是 $\frac{\sqrt{3}}{2}h$和$\frac{1}{2}h$，于是 A 点就移到了 $B(x_0+\frac{\sqrt{3}}{2}h,y_0+\frac{1}{2}h)$，于是 z 产生的增量为 $\Delta z=f(x_0+\frac{\sqrt{3}}{2}h,y_0+\frac{1}{2}h)-f(x_0,y_0)$，我们把 $\lim_{h\to 0}\frac{f(x_0+\frac{\sqrt{3}}{2}h,y_0+\frac{1}{2}h)-f(x_0,y_0)}{h}$ 叫做函数 $f(x,y)$ 在点 (x_0,y_0) 处关于向量 $\vec{u}=(\frac{\sqrt{3}}{2},\frac{1}{2})$ 的方向导数，记作 $D_u f(x_0,y_0)$

如果 $z=f(x,y)$ 在点 (x_0,y_0) 可微，于是

$$\Delta z=f(x_0+\frac{\sqrt{3}}{2}h,y_0+\frac{1}{2}h)-f(x_0,y_0)=f_x(x,y)\frac{\sqrt{3}}{2}h+f_y(x,y)\frac{1}{2}h+o(h)$$

两边都除以 h，令 $h\to 0$，得 $D_u f(x_0,y_0)=\frac{\sqrt{3}}{2}f_x(x,y)+\frac{1}{2}f_y(x,y)$

2、定义：在极限存在的条件下，函数 $z=f(x,y)$ 在点 (x_0,y_0) 处关于单位向量 $\vec{u}=(a,b)$ 的方向导数是 $D_u f(x_0,y_0)=\lim_{h\to 0}\frac{f(x_0+ah,y_0+bh)-f(x_0,y_0)}{h}$

3、如果 $z=f(x,y)$ 可微，则 f 关于单位向量 $\vec{u}=(a,b)$ 的方向导数 $D_u f(x,y)=af_x(x,y)+bf_y(x,y)$

例、求 $f(x,y)=x^3+y^3-6xy$ 关于方向角为 $\frac{\pi}{3}$ 的单位向量 $\vec{u}$ 的方向导数

解：$f_x(x,y)=3x^2-6y,\ f_y(x,y)=3y^2-6x$，$\vec{u}=(\cos\frac{\pi}{3},\sin\frac{\pi}{3})=(\frac{1}{2},\frac{\sqrt{3}}{2})$

于是 $D_u f(x,y)=\frac{1}{2}(3x^2-6y)+\frac{\sqrt{3}}{2}(3y^2-6x)=\frac{3}{2}x^2+\frac{3\sqrt{3}}{2}y^2-3\sqrt{3}x-3y$

四、梯度向量

1、梯度向量：$z=f(x,y)$ 关于向量 $\vec{u}=(a,b)$ 的方向导数

$D_u f(x,y)=af_x(x,y)+bf_y(x,y)$是向量$(f_x(x,y),\ f_y(x,y))$与$\vec{u}=(a,b)$的数量积向量$(f_x(x,y),\ f_y(x,y))$ 叫做 $z=f(x,y)$ 的梯度向量，记作 $\nabla f(x,y)=(f_x(x,y),\ f_y(x,y))$,或记为grad$f$

2、方向导数：$D_u f(x,y)=\nabla f(x,y)\cdot\vec{u}$

例、求 $f(x,y)=x^2y^3-4y$ 在点 $(2,-1)$ 对于向量 $\vec{v}=2i+5j$ 的方向导数

解：$\nabla f(x,y)=\left(f_x(x,y),\ f_y(x,y)\right)=(2xy^3,3x^2y^2-4)$，

故 $\nabla f(2,-1)=(-4,8)$，$\vec{u}=\dfrac{\vec{v}}{|\vec{v}|}=(\dfrac{2}{\sqrt{29}},\dfrac{5}{\sqrt{29}})$

于是 $D_u f(2,-1)=\nabla f(2,-1)\cdot\vec{u}=\dfrac{-8}{\sqrt{29}}+\dfrac{40}{\sqrt{29}}=\dfrac{32}{\sqrt{29}}$

3、梯度向量的意义

在方向导数 $D_u f(x,y)=\nabla f(x,y)\cdot\vec{u}$ 中，设 $\nabla f(x,y)$ 与 $\vec{u}$ 的夹角为 $\boldsymbol{\theta}$，则 $D_u f(x,y)=|\nabla f(x,y)||\vec{u}|\cos\boldsymbol{\theta}=|\nabla f(x,y)|\cos\boldsymbol{\theta}$，当 x,y 固定时，在 $\boldsymbol{\theta}=0$ 即，$\vec{u}$ 与 $\nabla f(x,y)$ 同向时 $D_u f(x,y)_{\max}=|\nabla f(x,y)|$ 这一结论说明了，函数 $z=f(x,y)$ 沿着 $\nabla f(x,y)$ 方向的导数最大。也就是说定义域中的点(x,y)沿着 $\nabla f(x,y)$ 方向变化时 z 的值增长的最快。

4、函数的定义域

一元函数 $y=f(x)$ 是定义在数轴 x 上的函数，二元函数 $z=f(x,y)$ 是定义在平面 xoy 上的函数，三元函数 $w=f(x,y,z)$ 是定义在三维空间 O−xyz 上的函数。

例 1、函数 $f(x)=x^2$ 是点定在数轴 x 上的函数，在点 3 处的函数值 $f(3)=9$

例 2、函数 $f(x,y)=x^2+y^2$ 是定义在平面 xoy 上的函数，在点(1,2)处的函数值 $f(1,2)=5$

例 3、函数 $w=x^2+y^2+z^2$ 定义在空间 O−xyz 上的函数，在点(1,2,1)处的函数值 $w=6$

5、等值线(等高线)

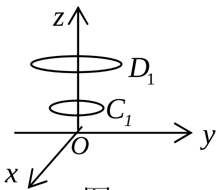


图3

函数 $z=x^2+y^2$ 是定义在平面 xoy 上的函数，其图象是抛物面。

当 z=1 时，得 $x^2+y^2=1$ 表示 z 坐标为 1 半径为 1 的圆 C_1

当 z=4 时，得 $x^2+y^2=4$ 表示 4 坐标为 2 半径为 1 的圆 D_1，如图 3

把圆 C_1 和 D_1 都投影到平面 xoy 上得圆 C 和 D，如图 4

当点(x,y)在圆 C 上时函数 z 的值都为 1，当点(x,y)在圆 D 上时函数 z 的值都为 4，我们把圆 C 就叫做函数值为 1 的等值线(等高线)，圆 D 就叫做函数值为 4 的等值线(等高线)

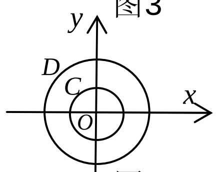


图4

6、梯度向量的几何意义

函数 $z=x^2+y^2$ 对平面 xoy 上的每个点 P(x,y)都可得到唯一的 z 值设为 m，于是 P(x,y) 在平面 xoy 上有一条等值线 $x^2+y^2=m$ 如图 5，此函数的梯度向量 $\nabla z=(2x,2y)$，于是 P(x,y)的梯度向量 $\nabla z=\overrightarrow{OP_1}=2\overrightarrow{OP}$ 由第 1 条的结论知当 P(x,y)沿 $\overrightarrow{OP_1}$ 的方向运动时，z 的值增长的最快。

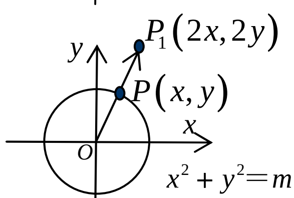


图5

在等值线 $x^2+y^2=m$ 两边对 x 取导数得 $2x+2y\dfrac{dy}{dx}=0,\dfrac{dy}{dx}=-\dfrac{x}{y}$

等值线在 P(x,y)处的斜率为 $k_{切}=-\dfrac{x}{y}$，$\overrightarrow{OP_1}$ 的斜率 $k_{梯度}=\dfrac{2y}{2x}=\dfrac{y}{x}$，于是 $k_{梯度}k_{切}=-1$

由此可见梯度向量与曲线垂直的。下面证明这个结论普遍成立。

设 $z=f(x,y)$ 在点 P(x,y)处的等高线为 $f(x,y)-m=0$，设 $F(x,y)=f(x,y)-m$ 则 $k_{切}=\dfrac{dy}{dx}=-\dfrac{F_x(x,y)}{F_y(x,y)}=-\dfrac{f_x(x,y)}{f_y(x,y)}$，而梯度向量 $\nabla f(x,y)=\left(f_x(x,y),\ f_y(x,y)\right)$ 的斜率为 $k_{梯度}=\dfrac{f_y(x,y)}{f_x(x,y)}$，因此 $k_{梯度}k_{切}=-1$

7、切平面与法线

三元方程$F(x,y,z)=m$确定的曲面 S 可以看成函数$w=F(x,y,z)$的$w=m$一个等值面。我们在曲面 S 上任意取一条曲线 C 设它的向方程是$r(t)=(x(t),y(t),z(t))$

因为点 M$(x(t),y(t),z(t))$在曲面 S 上，所以$F(x(t),y(t),z(t))=m$方程两边对 t 求导得

$\frac{\partial F}{\partial x}\frac{dx}{dt}+\frac{\partial F}{\partial y}\frac{dy}{dt}+\frac{\partial F}{\partial z}\frac{dz}{dt}=0$，即$(\frac{\partial F}{\partial x},\frac{\partial F}{\partial y},\frac{\partial F}{\partial z})\bullet(\frac{dx}{dt},\frac{dy}{dt},\frac{dz}{dt})=0$

于是$\nabla F(x,y,z)\bullet r'(t)=0$，由此函数$w=F(x,y,z)$在等值面 S 上任意一点$P(x,y,z)$的梯度$\nabla F(x,y,z)$与曲面 S 上任意一条过 P 点的曲线 C 都垂直。因此曲面 S 在$P(x,y,z)$处的切平面的法向量是$\nabla F(x,y,z)$，这一结论也可从二元函数类比出来。

例 1、求椭球$\frac{x^2}{4}+y^2+\frac{z^2}{9}=3$过$(-2,1,-3)$的切平面与法线方程

解：$F(x,y,z)=\frac{x^2}{4}+y^2+\frac{z^2}{4}$，取 m=3 时的等值面就是椭球面

因为$\nabla F(x,y,z)=(F_x,F_y,F_Z)=(\frac{x}{2},2y,\frac{2z}{9})$ 所以$\nabla F(-2,1,-3)=(-1,2,-\frac{2}{3})$

于是切平面方程是$-1(x+2)+2(y-1)-\frac{2}{3}(z+3)=0$即$3x-6y+2z+18=0$

法线方程是$\frac{x+2}{-1}=\frac{y-1}{2}=\frac{z+3}{-\frac{2}{3}}$

例 2、求曲面$z=f(x,y)$在点 A(x_0,y_0,z_0)处的切平面方程

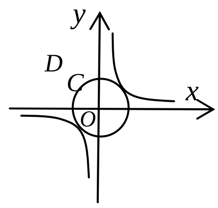
解：$F(x,y,z)=z-f(x,y)$，取 m=0 的等值面就是曲面$z=f(x,y)$

$\nabla F(x,y,z)=(F_x,F_y,F_Z)=(-f_x,-f_y,1)$，于是$\nabla F(x_0,y_0,z_0)=(-f_x(x_0,y_0),-f_y(x_0,y_0),1)$

在点 A(x_0,y_0,z_0)处的切平面方程为$f_x(x_0,y_0)(x-x_0)+f_y(x_0,y_0)(y-x_0)-(z-z_0)=0$

第三节、极值与最值

一、拉格日数乘法



1、引例；在双曲线$xy=6$上求一点使它到原点的距离最小

（注本题用初数很容易，下面通过此题讲拉格日算子）

目标函数是$\sqrt{x^2+y^2}$可改为$f(x,y)=x^2+y^2$，使以原点为圆心的圆

与其几何意义就双曲线相切。圆是$f(x,y)$的等高线，设$g(x,y)=xy$，则双曲线是$g(x,y)$

的等高线，当$f(x,y)$与$g(x,y)$的等高线相切时$f(x,y)$最小。

此时梯度向量共线即$\nabla f(x,y)//\nabla g(x,y)$

由于$\nabla f(x,y)=(f_x(x,y),\ f_y(x,y))=(2x,2y)$，$\nabla g(x,y)=(g_x(x,y),\ g_y(x,y))=(y,x)$

于是可设$(2x,2y)=\lambda(y,x)$这有两个方程，再添上$xy=6$得方程组

$\begin{cases}2x=\lambda y\\2y=\lambda x\\xy=6\end{cases}$解得$x=y=\sqrt{6}$或$x=y=-\sqrt{6}$ 于是最小值为$\sqrt{f(x,y)}=\sqrt{x^2+y^2}=\sqrt{12}=2\sqrt{3}$

这种方法叫做拉格日数乘法，λ 叫做拉格日算子

2、拉格日数乘法：**要求在约束条件** $g(x,y)=m(m\text{为常数})$ **下，函数** $z=f(x,y)$ **的极值**

应使唤 $f(x,y)$ 与 $g(x,y)$ 的等高线相切，即 $\nabla f(x,y)//\nabla g(x,y)$，设 $\nabla f(x,y)=\lambda\nabla g(x,y)$

可由方程组 $\begin{cases} f_x(x,y)=\lambda g_x(x,y) \\ f_y(x,y)=\lambda g_y(x,y) \\ g(x,y)=m \end{cases}$ 求出极值点进而求出 $z=f(x,y)$ **的极值**

例、求旋转抛物面 $z=x^2+y^2$ 与平面 $x+y-2z=2$ 之间的最短距离.

解：设 $P(x,y,z)$ 为抛物面 $z=x^2+y^2$ 上任一点,则
P 到平面 $x+y-2z-2=0$ 的距离为 d,

$$d=\frac{1}{\sqrt{6}}\left|x+y-2z-2\right|$$

设 $f(x,y,z)=(x+y-2z-2)^2$，$g(x)=x^2+y^2-z$

因为 $\nabla f(x,y,z)=(2(x+y-2z-2),2(x+y-2z-2),-4(x+y-2z-2))$

$\nabla g(x)=(2x,2y,-1)$

于是

$$\begin{cases} 2(x+y-2z-2)=2\lambda x & (1) \\ 2(x+y-2z-2)=2\lambda y & (2) \\ -4(x+y-2z-2)=-\lambda z & (3) \\ z=x^2+y^2 & (4) \end{cases}$$

解得 $x=\frac{1}{4},y=\frac{1}{4},z=\frac{1}{8}$ 即得唯一驻点 $(\frac{1}{4},\frac{1}{4},\frac{1}{8})$,

根据题意距离的最小值一定存在，且有唯一驻点，故必在 $(\frac{1}{4},\frac{1}{4},\frac{1}{8})$ 处取得最小值.

$$d_{\min}=\frac{1}{\sqrt{6}}\left|\frac{1}{4}+\frac{1}{4}-\frac{1}{4}-2\right|=\frac{7}{4\sqrt{6}}.$$

二、求多元函数的最值

1、一元函数

(1)几何理解

我们知道可微一元函数 $y=f(x)$ 的极值要找到平行于 x 轴的切线，于是可从 $f'(x)=0$ 求临界点 $x=x_0$，再看切线有没有被函数 $y=f(x)$ 的图象穿越定出极大值点，极小值点，或不是极值点。

(2)代数理解

从 $f'(x)=0$ 求临界点 $x=x_0$ 后，由 $y=f(x)$ 的二次近似得

$$f(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{1}{2!}f''(x_0)(x-x_0)^2+o(x-x_0)^2$$

$$=f(x_0)+\frac{1}{2!}f''(x_0)(x-x_0)^2+o(x-x_0)^2$$

由于$o(x-x_0)^2$可忽略，于是

当$f''(x_0)>0$时$f(x_0)$是局部最小，即$f(x)_{极小}=f(x_0)$

当$f''(x_0)<0$时$f(x_0)$是局部最大，即$f(x)_{极大}=f(x_0)$

当$f''(x_0)=0$时就无极极或要有更高次近似判定

2、二元函数

(1)几何理解

类比上面可微二元函数 $z=f(x,y)$ 的极值要找到平行于平面 xoy 的切面，于是可从 $f_x(x,y)=0, f_y(x,y)=0$ 求驻点 $(x,y)=(x_0,y_0)$，再看切面有没有被函数 $z=f(x,y)$ 的图象穿越定出极大值点，极小值点，或鞍点。

(2)代数理解

求出驻点 (x_0,y_0) 后，由 $z=f(x,y)$ 的二次近似得

$$f(x,y)=f(x_0,y_0)+f_x(x_0,y_0)h+f_y(x_0,y_0)k$$

$$+\frac{1}{2}f_{xx}(x_0,y_0)h^2+f_{xy}(x_0,y_0)hk+\frac{1}{2}f_{yy}(x_0,y_0)k^2+o(h^2+k^2)$$

$$=f(x_0,y_0)+\frac{1}{2}f_{xx}(x_0,y_0)h^2+f_{xy}(x_0,y_0)hk+\frac{1}{2}f_{yy}(x_0,y_0)k^2+o(h^2+k^2)$$

这里$h=x-x_0, k=y-y_0$，由于$o(h^2+k^2)$可忽略，于是

只要看h，k的二次函数$\frac{1}{2}f_{xx}(x_0,y_0)h^2+f_{xy}(x_0,y_0)hk+\frac{1}{2}f_{yy}(x_0,y_0)k^2$

的符号来确定极值的情况

(3)二元二次齐次式的分类

$$f(x,y)=ax^2+bxy+cy^2=a(x^2+\frac{b}{a}xy+\frac{c}{a}y^2)=a[(x+\frac{b}{2a}y)^2+\frac{4ac-b^2}{4a^2}y^2]$$

1° 当中括号内是实数的平方和，这时 $4ac-b^2>0$

2° 当中括号内是实数的平方差，这时 $4ac-b^2<0$

3° 当中括号内是退化为 $(x+\frac{b}{2a}y)^2$，这时 $4ac-b^2=0$

(4)回到(2)去

可向微函数数 $f(x,y)$ 由 $f_x(x,y)=0, f_y(x,y)=0$ 求临界点 $(x,y)=(x_0,y_0)$

设 $f_{xx}(x,y)=A, f_{xy}(x,y)=B, f_{yy}(x,y)=C$

(1)当 $4ac-b^2=AC-B^2>0$ 时，

1° 若 $A>0$,则 (x_0,y_0) 是极小值点，2° 若 $A<0$,则 (x_0,y_0) 是极大值点

(2)当 $4ac-b^2=AC-B^2<0$ 时，(x_0,y_0) 是鞍点

(3)当 $AC-B^2=0$ 时，不能判定

例 1　求函数 $f(x,y)=x^4+y^4-x^2-2xy-y^2$ 的极值.

解　$f_x(x,y)=4x^3-2x-2y=0$，$f_y(x,y)=4y^3-2x-2y=0$，

得驻点（1，1)，(-1，-1)，（0，0）。

判断：求二阶偏导 $f_{xx}(x,y)=12x^2-2$，$f_{xy}(x,y)=-2$，$f_{yy}(x,y)=12y^2-2$，

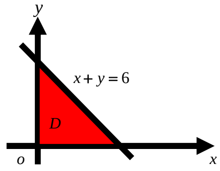
在点（1，1）处，$A=f_{xx}(1,1)=10$，$B=f_{xy}(1,1)=-2$，$C=f_{yy}(1,1)=10$.

因 $B^2—AC<0$，且 $A>0$，　故 $f(1,1)= -2$ 为极小值.　类似可得 $f(-1,-1)= -2$ 为极小值.

在点（0，0）处，$A=B=C= -2$，$B^2-AC=0$，此时应用极值定义判断 $f(0,0)=0$ 是否为极值.

对足够小的正数ε，有　$f(\varepsilon,0)=\varepsilon^2(\varepsilon^2-1)<0$，$f(\varepsilon,-\varepsilon)=2\varepsilon^4>0$

这说明在点（0，0）的任一邻域内，既有函数值大于 $f(0,0)$ 的点，又有函数值小于 $f(0,0)$ 的点，故 $f(0,0)$ 非极值.



例 2、求二元函数 $z=f(x,y)=x^2y(4-x-y)$ 在直线 $x+y=6$，x 轴和 y 轴所围成的闭区域 D 上的最大值与最小值.

解：先求函数在 D 内的驻点，

解方程组 $\begin{cases} f_x'(x,y)=2xy(4-x-y)-x^2y=0 \\ f_y'(x,y)=x^2(4-x-y)-x^2y=0 \end{cases}$

得区域 D 内唯一驻点 $(2,1)$，且 $f(2,1)=4$，

再求 $f(x,y)$ 在 D 边界上的最值，在边界 $x=0$ 和 $y=0$ 上 $f(x,y)=0$，

在边界 $x+y=6$ 上，即 $y=6-x$ 于是 $f(x,y)=x^2(6-x)(-2)$，

由　$f_x'=4x(x-6)+2x^2=0$，得 $x_1=0, x_2=4$（舍去 x_1）$\Rightarrow y=6-x|_{x=4}=2,\ f(4,2)=-64$,

比较后可知 $f(2,1)=4$ 为最大值，$f(4,2)=-64$ 为最小值.

例 3、最小二乘法

(1) 测量的预测与最小二乘法：

对一个零件长度 x 的 n 次测量假设测量值为 $x_1,x_2,\cdots x_n$，如何取 $\hat{x}$ 做为 x 的估计的值更合理呢？直观的想法是 $\hat{x}$ 的值应该最接近这些测量数据，数学描述就是：　$\hat{x}$ 的值应该使所有的误差平方和 $\sum_{i=1}^{n}(\hat{x}-x_i)^2$ 达到最小.

$$f(\hat{x})=\sum_{i=1}^{n}(\hat{x}-x_i)^2=n\hat{x}^2-2\hat{x}\sum_{i=1}^{n}x_i+\sum_{i=1}^{n}x_i^2$$

当 $\hat{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$ 时，$f(\hat{x})$ 达到最小. 即用测量数据的平均值作为零件估计值最好. 这种估计方法就叫最小二乘法. 最小二乘法的优点是：有效利用了全部测量数据，使误差平方和达到最小.

(2) 回归直线的方程

设有线性相关系系的变量 x,y 的一组测量值是

x	x_1	x_2	…	x_n
y	y_1	y_2	…	y_n

设回归直线的方程 $\hat{y}=bx+a$，则可由最小二乘法确定 b，a 的值

误差平方和 $M(b,a)=\sum_{k=1}^{n}(y_k-bx_k-a)^2$

$$\begin{cases} \dfrac{\partial M}{\partial b}=-2\sum_{k=1}^{n}(y_k-bx_k-a)x_k=0 \\ \dfrac{\partial M}{\partial a}=-2\sum_{k=1}^{n}(y_k-bx_k-a)=0 \end{cases} \qquad \begin{cases} (\sum_{k=1}^{n}x_k^2)b+n\bar{x}a=\sum_{k=1}^{n}x_ky_k(1) \\ n\bar{x}b+na=n\bar{y}\Rightarrow\bar{x}b+a=\bar{y}(2) \end{cases}$$

解得$b=\dfrac{\sum_{i=1}^{n}x_iy_i-n\overline{x}\overline{y}}{\sum_{i=1}^{n}x_i^2-n\overline{x}^2}$求出 b 后 a 的值由(2)确定 ，由(2)得$\overline{y}=b\overline{x}+a$，因此回归直线过样本中心$(\overline{x},\overline{y})$

4、附注泰勒公式

(1)一元函数的泰勒公式：

$$f(x)=f(x_0)+\frac{f'(x_0)}{1!}(x-x_0)+\frac{f^{(2)}(x_0)}{2!}(x-x_0)^2+\cdots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n+\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}$$

证明：$f(x)=f(x_0)+\int_{x_0}^{x}f'(t)dt=f(x_0)+\int_{x_0}^{x}f'(t)d(t-x)$

$$=f(x_0)+f'(t)(t-x)|_{x_0}^{x}-\int_{x_0}^{x}(t-x)f''(t)d$$

$$=f(x_0)+f'(x_0)(x-x_0)-\frac{1}{2!}\int_{x_0}^{x}f''(t)d(t-x)^2$$

$$=f(x_0)+f'(x_0)(x-x_0)-\frac{1}{2!}f''(t)(t-x)^2|_{x_0}^{x}+\frac{1}{2!}\int_{x_0}^{x}f'''(t)(t-x)^2dt$$

$$=f(x_0)+f'(x_0)(x-x_0)+\frac{1}{2!}f''(x_0)(x-x_0)^2+\frac{1}{2!}\int_{x_0}^{x}f'''(t)(t-x)^2dt$$

由积分中值定理得

$$\int_{x_0}^{x}f'''(t)(t-x)^2dt=f'''(\xi)\int_{x_0}^{x}(t-x)^2dt=\frac{1}{3}f'''(\xi)(x-x_0)^3(\xi在x_0与x之间)$$

于是$f(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{1}{2!}f''(x_0)(x-x_0)^2+\frac{1}{3!}f'''(\xi)(x-x_0)^3$

用数学归纳法可证出原命题成立

(2)二元函数的泰勒公式：

$$f(x_0+h,y_0+k)=f(x_0,y_0)+\frac{1}{1!}(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y})f(x_0,y_0)+\frac{1}{2!}(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y})^2f(x_0,y_0)+$$

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

证：设$\phi(t)=f(x_0+ht,y_0+kt)$

$$\phi(1)=\phi(0)+\phi'(0)+\frac{1}{2!}\phi''(0)+\frac{1}{3!}\phi'''(\theta)(0\le\theta\le1)$$

因为$\phi(1)=f(x_0+h,y_0+k)$

$$\phi'(t)=hf_x(x_0+ht,y_0+kt)+kf_y(x_0+ht,y_0+kt)$$

$$\phi'(0)=hf_x(x_0,y_0)+kf_y(x_0,y_0)$$

$$\phi''(t)=h^2f_{xx}(x_0+ht,y_0+kt)+khf_{xy}(x_0+ht,y_0+kt)$$

$$+hkf_{yx}(x_0+ht,y_0+kt)+k^2f_{yy}(x_0+ht,y_0+kt)$$

$$\phi''(0)=h^2f_{xx}(x_0,y_0)+2khf_{xy}(x_0,y_0)+k^2f_{yy}(x_0,y_0)$$

于是

$$f(x_0+h,y_0+k)=f(x_0,y_0)+\frac{1}{1!}[hf_x(x_0,y_0)+kf_y(x_0,y_0)]$$

$$+\frac{1}{2!}[h^2f_{xx}(x_0,y_0)+2khf_{xy}(x_0,y_0)+k^2f_{yy}(x_0,y_0)]+o(h^2+k^2)$$