1 微分与偏导

1.1 全微分

如果函数$z = f(x, y)$在点(x_0, y_0)可微分，那么该函数在点(x_0, y_0)的偏导数$f_x(x_0, y_0)$、$f_y(x_0, y_0)$必定存在，且$z = f(x, y)$在点(x_0, y_0)的全微分为：

$$\mathrm{d}z = f_x(x_0, y_0)\mathrm{d}x + f_y(x_0, y_0)\mathrm{d}y$$

$$\frac{\partial f}{\partial x}\text{、}\frac{\partial f}{\partial y}\text{ 在 }(x_0, y_0)\text{ 连续 }\Longrightarrow f(x, y)\text{ 在 }(x_0, y_0)\text{ 点可微分}$$

$$f(x, y)\text{ 在 }(x_0, y_0)\text{ 点可微分 }\Longrightarrow f(x, y)\text{ 在 }(x_0, y_0)\text{ 点连续}$$

1.2 高阶偏导数

如果函数$z = f(x, y)$的两个混合偏导数在区域D内连续，那么必有：

$$f_{xy}(x, y) = f_{yx}(x, y), \quad (x, y) \in D$$

例题：

函数$z = \sqrt{|xy|}$在点$(0, 0)$的偏导数，证明偏导数存在，但不连续。

根据偏导数：

$$\frac{\partial z}{\partial x} = \lim_{h \to 0} \frac{f(x + h, 0) - f(x, 0)}{h} = \lim_{h \to 0} \frac{0}{h} = 0$$

同理

$$\frac{\partial z}{\partial y} = 0$$

因此偏导数存在。

假设函数在$(0, 0)$处可微，则全微分为：

$$\mathrm{d}f = f_x(0, 0)\Delta x + f_y(0, 0)\Delta y = 0$$

$$\Delta f = f(0 + \Delta x, 0 + \Delta y) - f(0, 0) = \sqrt{|\Delta x \cdot \Delta y|}$$

由全微分可知，若

$$\Delta z - \mathrm{d}z = o(\rho)$$

那么，函数在此点可微，其中$\rho = \sqrt{(\Delta x)^2 + (\Delta y)^2}$

令$\Delta x = \Delta y$

$$\lim_{\Delta x\to 0}\frac{\Delta f - df}{\rho}=\frac{1}{\sqrt{2}}$$

即

$$\Delta z-\mathrm{d}z\neq o(\rho)$$

函数在$(0,0)$点不可微，根据微分的关系不可微的函数没有连续的偏导数

综上所述，偏导数存在但不连续

1.3 方向导数

1.3.1 方向导数的定义

假设该单位向量$\boldsymbol{u}$在xy面上的坐标为：

$$\boldsymbol{u}=\begin{pmatrix}u_1\\u_2\end{pmatrix}$$

根据刚才的分析，割线的极限是切线，那么割线的斜率的极限也就是切线的斜率，即切线的斜率为：

$$\lim_{t\to 0}\frac{f(x_0+tu_1,y_0+tu_2)-f(x_0,y_0)}{t},\quad t\in\mathbb{R}$$

对于二元函数$z=f(x,y)$，沿某单位向量：

$$\boldsymbol{u}=\begin{pmatrix}u_1\\u_2\end{pmatrix}$$

在(x_0,y_0)点的方向导数为：

$$\left.\frac{\partial f}{\partial u}\right|_{(x_0,y_0)}=\lim_{t\to 0}\frac{f(x_0+tu_1,y_0+tu_2)-f(x_0,y_0)}{t},\quad t\in\mathbb{R}$$

以上是国际上的定义。

同济教程的定义有所不同：

$$\left.\frac{\partial f}{\partial u}\right|_{(x_0,y_0)}=\lim_{t\to 0^+}\frac{f(x_0+tu_1,y_0+tu_2)-f(x_0,y_0)}{t},\quad t\in\mathbb{R}$$

即从$\boldsymbol{u}$的**正方向**逼近(x_0,y_0)

这种差异导致了有些情况下国际定义中方向导数不存在，而按照同济教材的定义则存在。

1.3.2 方向导数的性质

可微分则方向导数一定存在，因为如果**可微分意味着点附近的曲面可以用切平面近似。**也意味着所有（与面垂直的）平面和曲面的相交曲线都能够被线性近似，也就是说方向导数存在；并且这些曲线的切线都在切平面上：

因为如果**可微分意味着(x_0,y_0)点附近的曲面可以用切平面近似**；也意味着所有（与$\boldsymbol{xy}$面垂直的）平面和曲面的相交曲线都能够被线性近似，也就是说方向导数存在；并且这些曲线的切线都在切平面上：

可得以下结论：

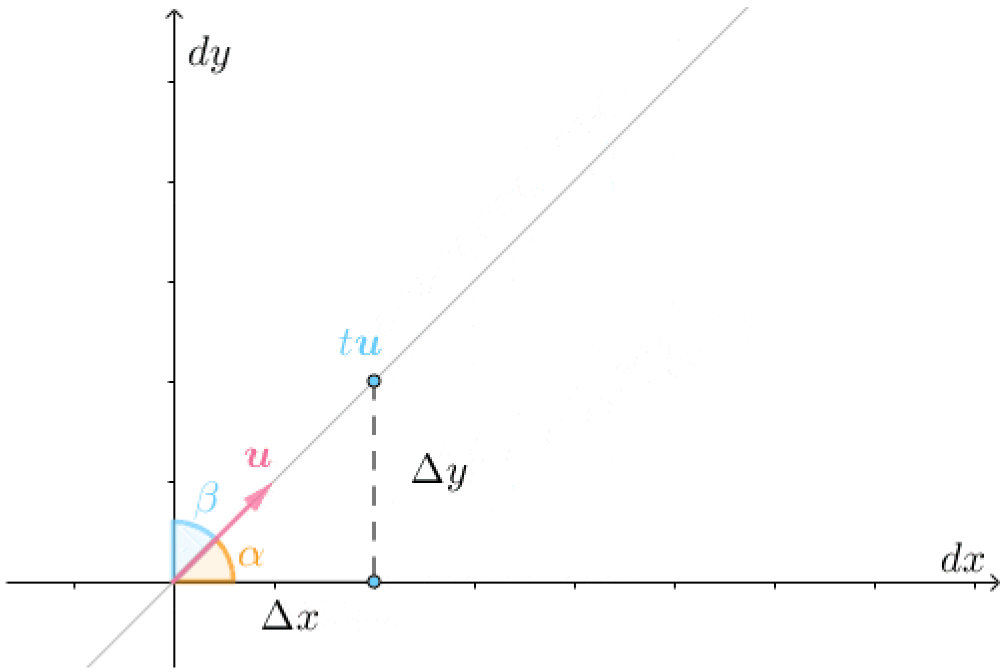
如果函数$z=f(x,y)$在(x_0,y_0)点可微分，那么函数在该点沿任意单位向量：

$$\boldsymbol{u}=\begin{pmatrix}\cos\alpha\\ \cos\beta\end{pmatrix}$$

的方向导数存在，且：

$$\left.\frac{\partial f}{\partial u}\right|_{(x_0,y_0)}=f_x(x_0,y_0)\cos\alpha+f_y(x_0,y_0)\cos\beta$$

**由这些方向导数和偏导数决定的切线都在同一个平面上，也就是在切平面上。*：



例题：

对于二元函数：

$$f(x,y)=x^2+y^2$$

在$(1,2)$点沿下向量的方向导数：

$$\boldsymbol{u}=\begin{pmatrix}1\\0\end{pmatrix}$$

的方向导数分别为2，那么沿着$-\boldsymbol{u}$，即反方向的方向导数为?

套用下公式来计算（国际通用的定义）：

$$\left.\frac{\partial f}{\partial u}\right|_{(x_0,y_0)}=\lim_{t\to0}\frac{f(1+t\cdot1,2)-f(1,2)}{t}=\lim_{t\to0}\frac{t^2+2t}{t}=2$$

$$\left.\frac{\partial f}{\partial(-u)}\right|_{(x_0,y_0)}=\lim_{t\to0}\frac{f(1+t\cdot(-1),2)-f(1,2)}{t}=\lim_{t\to0}\frac{t^2-2t}{t}=-2$$

所以沿反方向的方向导数为-2。

也就是说沿着$\boldsymbol{u}$和沿着$-\boldsymbol{u}$得到的方向导数是不一样的，虽然从几何上看，最终得到的切线是一样的，只是因为$\boldsymbol{u}$和$-\boldsymbol{u}$两者定义的正方向是反的，所以导致了求出来的导数值是相反数。

1.4 导数与雅可比矩阵

1.4.1 函数导数与矩阵

一般的多元函数：

$$y=f(x_1,x_2,\cdots,x_n),\quad (y,x_1,x_2,\cdots,x_n\in\mathbb{R})$$

如果可微的话，那么微分方程为：

$$\mathrm{d}y=f_{x_1}\mathrm{d}x_1+f_{x_2}\mathrm{d}x_2+\cdots+f_{x_n}\mathrm{d}x_n$$

依然可以改写为矩阵的形式：

$$\underbrace{\begin{pmatrix}\mathrm{d}y\end{pmatrix}}_{\boldsymbol{d_y}}=\underbrace{\begin{pmatrix}f_{x_1} & f_{x_2} & \cdots & f_{x_n}\end{pmatrix}}_{\boldsymbol{T}}\underbrace{\begin{pmatrix}\mathrm{d}x_1\\ \mathrm{d}x_2\\ \vdots\\ \mathrm{d}x_b\end{pmatrix}}_{\boldsymbol{d_x}}$$

或者写作：

$$\boldsymbol{d_y}=\boldsymbol{T}\boldsymbol{d_x}$$

其中矩阵$\boldsymbol{T}$就是该多元函数的导数。

1.4.2 方程组导数与矩阵

下面这样的方程组：

$$\begin{cases}z=f(x,y)\\ w=g(x,y)\end{cases}$$

其实完成了如下映射，所以它就是一个函数：

$$\begin{pmatrix}x\\ y\end{pmatrix}\xrightarrow{\text{方程组}}\begin{pmatrix}z\\ w\end{pmatrix}$$

将方程组代表的函数记为$\boldsymbol{F}$，那么方程组写作下面的形式就更容易看出是函数了：

$$\underbrace{\begin{pmatrix}z\\ w\end{pmatrix}}_{\boldsymbol{v}}=\boldsymbol{F}\underbrace{\begin{pmatrix}x\\ y\end{pmatrix}}_{\boldsymbol{u}}$$

或者再写简单点，这样可以看出该函数$\boldsymbol{F}$的自变量为$\boldsymbol{u}$，因变量为$\boldsymbol{v}$：

$$\boldsymbol{v}=\boldsymbol{F}\boldsymbol{u}$$

方程组既然是函数，那么函数就会有微分。对方程组的每个方程求微分，就可以得到该方程组所代表的函数的微分：

$$\begin{cases} \mathrm{d}z = \dfrac{\partial f}{\partial x}\mathrm{d}x + \dfrac{\partial f}{\partial y}\mathrm{d}y \\ \\ \mathrm{d}w = \dfrac{\partial g}{\partial x}\mathrm{d}x + \dfrac{\partial g}{\partial y}\mathrm{d}y \end{cases}$$

也可以写成矩阵的形式：

$$\underbrace{\begin{pmatrix} \mathrm{d}z \\ \mathrm{d}w \end{pmatrix}}_{\boldsymbol{d_v}} = \underbrace{\begin{pmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} \\ \dfrac{\partial g}{\partial x} & \dfrac{\partial g}{\partial y} \end{pmatrix}}_{T} \underbrace{\begin{pmatrix} \mathrm{d}x \\ \mathrm{d}y \end{pmatrix}}_{\boldsymbol{d_u}}$$

或者写作：

$$\boldsymbol{d_v} = T\boldsymbol{d_u}$$

其中矩阵T就是该方程组所代表的函数的导数。

1.4.3 雅可比矩阵

由于矩阵T的存在，所有的微分都具备了相同的形式：

一元函数	$\boldsymbol{d_y} = T\boldsymbol{d_x}$
二元函数	$\boldsymbol{d_z} = T\boldsymbol{d_{xy}}$
多元函数	$\boldsymbol{d_y} = T\boldsymbol{d_x}$
方程组	$\boldsymbol{d_v} = T\boldsymbol{d_u}$

只是每一行的T有所不同，需要如下来定义：

假如$f_1, f_2, \cdots, f_n$都是$x_1, x_2, \cdots, x_n$的函数，并且相对于各个自变量的偏微分都存在，那么定义T为：

$$T = \frac{\partial(f_1, f_2, \cdots, f_n)}{\partial(x_1, x_2, \cdots, x_m)} = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \cdots & \dfrac{\partial f_1}{\partial x_m} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \cdots & \dfrac{\partial f_2}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f_n}{\partial x_1} & \dfrac{\partial f_n}{\partial x_2} & \cdots & \dfrac{\partial f_n}{\partial x_m} \end{pmatrix}$$

该矩阵T称为雅可比矩阵。因为雅可比矩阵的英文名为Jacobian Matrix，所以上述矩阵又常写作：

$$J = \frac{\partial(f_1, f_2, \cdots, f_n)}{\partial(x_1, x_2, \cdots, x_m)}$$

套用上面的定义，对于一元函数$y = f(x)$，那么雅可比矩阵为：

$$T=\frac{\partial(f)}{\partial(x)}=\frac{\mathrm{d}f}{\mathrm{d}x}$$

对于多元函数$y=f(x_1,x_2,\cdots,x_n)$，其雅可比矩阵为：

$$T=\frac{\partial(f)}{\partial(x_1,x_2,\cdots,x_n)}=\begin{pmatrix}\frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_n}\end{pmatrix}$$

而方程组（所代表的函数）：

$$\begin{cases}z=f(x,y)\\ \\ w=g(x,y)\end{cases}$$

其雅可比矩阵为：

$$T=\frac{\partial(f,g)}{\partial(x,y)}=\begin{pmatrix}\frac{\partial f}{\partial x} & \frac{\partial f}{\partial y}\\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y}\end{pmatrix}$$

这些雅可比矩阵都是对应的函数的导数。

弄明白导数是什么之后，就可以知道，微分和导数其实是一个概念，有了导数自然可以求出微分，而微分存在必然有导数，即：

$$\text{可微}\iff\text{可导}$$

之前总结的微分、偏导数、方向导数的关系可以再加上可导：

$$\text{偏导数连续}\implies\begin{matrix}\text{可微}\\ \text{可导}\end{matrix}\implies\begin{cases}\text{连续}\\ \text{偏导数都存在}\\ \text{方向导数都存在}\\ \text{全导数都存在（逼近路径可导的话）}\end{cases}$$

1.4.4 链式法则

1.4.4.1 函数的链式法则

假设$y=f(x)$，其中$x=g(t)$，那么可以根据一元函数的链式法则求出y关于t的导数为：

$$\frac{\mathrm{d}y}{\mathrm{d}t}=\frac{\mathrm{d}y}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t}$$

定义了雅可比矩阵之后，多元函数的链式法则也有了相同的形式。比如$w=f(x,y,z)$是可微分的，而x、y和z是t的可导函数，那么：

$$\frac{\mathrm{d}w}{\mathrm{d}t}=\frac{\partial(w)}{\partial(x,y,z)}\frac{\partial(x,y,z)}{\partial(t)}$$

查看详细

这样的形式比较好记忆，看上去似乎分子、分母可以消去：

$$\frac{\mathrm{d}w}{\mathrm{d}t} = \frac{\partial(w)}{\cancel{\partial(x,y,z)}} \frac{\cancel{\partial(x,y,z)}}{\partial(t)}$$

1.4.4.2 方程组的链式法则

更复杂的方程组的链式法则也有同样的形式。比如函数$z = f(x, y)$，其中x、y是中间变量，有：

$$\begin{cases} x = x(\theta, \phi) \\ y = y(\theta, \phi) \end{cases}$$

那么有：

$$\frac{\partial(z)}{\partial(\theta, \phi)} = \frac{\partial(z)}{\partial(x, y)} \frac{\partial(x, y)}{\partial(\theta, \phi)}$$

1.4.5 海森矩阵

一元导数$y = f(x)$的二阶导数就是连续两次使用$\dfrac{\mathrm{d}}{\mathrm{d}x}$：

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)$$

类似的，二元函数$z = f(x, y)$的二阶导数就是连续两次计算雅可比矩阵：

$$\frac{\partial^2 z}{\partial(x, y)^2} = \frac{\partial}{\partial(x, y)}\left(\frac{\partial z}{\partial(x, y)}\right) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} \tag{1}$$

二阶导数又称为海森矩阵（Hessian Matrix），所以常用H来表示这个矩阵：

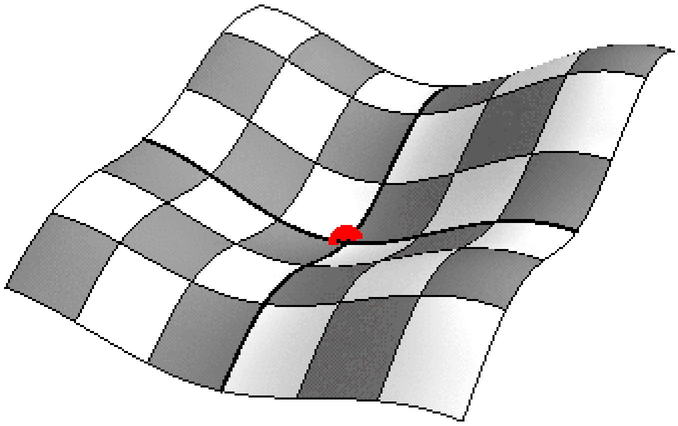
$$H = \frac{\partial^2 z}{\partial(x, y)^2} = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$$

1.4.6 海森矩阵几何意义

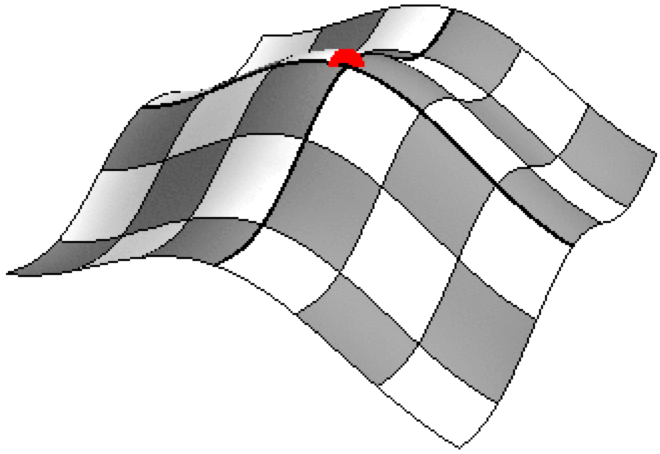
二元函数$z = f(x, y)$的二阶导数，也就是海森矩阵：

$$H = \frac{\partial^2 z}{\partial(x, y)^2} = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$$

同样决定了某点（比如下图中的红点）附近图像的凹凸性，当海赛矩阵的行列式$|H| > 0$时，红点附近的图像凹凸性是一致的：



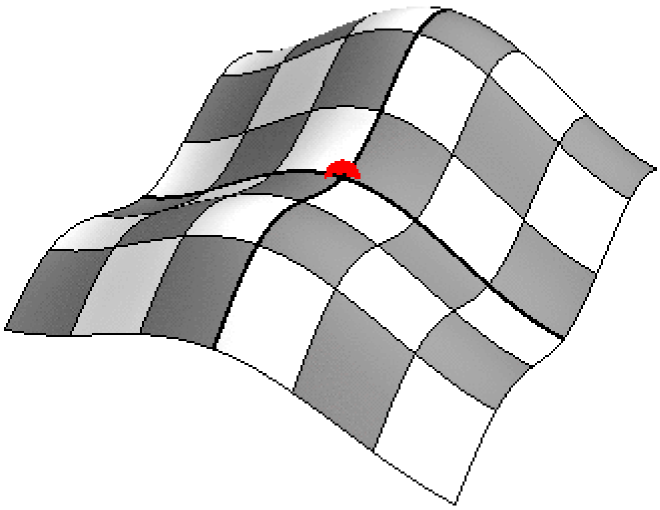
凹的，$f_{xx} > 0$且$|H| > 0$



凸的，$f_{xx} < 0$且$|H| > 0$

查看详细

当$|H| < 0$时，附近的凹凸性不一致（此时图像有点像马鞍，所以该红点也称为鞍点）：



图中的两根黑色曲线，一根朝下弯曲，一根朝上弯曲