映射的微分

徐海峰整理

March 5, 2025

数学科学学院，扬州大学

此幻灯片的数学内容面向非数学专业的理工科本科生.

内容完全基于下面的教材:

梅加强编著《数学分析》, 高等教育出版社.

其他参考文献.

利用线性化的思想研究多元函数

一元函数可微的定义

回忆, 对于一元函数而言, 可微是指该函数可以被线性函数一阶逼近.

一元函数可微的定义

回忆, 对于一元函数而言, 可微是指该函数可以被线性函数一阶逼近.

定义

设 $f(x)$ 在 $(x_0-\delta, x_0+\delta)\in\mathbb{R}$ 上有定义, 若存在常数 A, 使得

$$f(x)=f(x_0)+A(x-x_0)+o(x-x_0),\quad (x\to x_0)$$

一元函数可微的定义

回忆, 对于一元函数而言, 可微是指该函数可以被线性函数一阶逼近.

定义

设 $f(x)$ 在 $(x_0-\delta, x_0+\delta) \in \mathbb{R}$ 上有定义, 若存在常数 A, 使得

$$f(x) = f(x_0) + A(x-x_0) + o(x-x_0), \quad (x \to x_0)$$

即

$$|f(x) - [f(x_0) + A(x-x_0)]| = o(|x-x_0|),$$

一元函数可微的定义

回忆, 对于一元函数而言, 可微是指该函数可以被线性函数一阶逼近.

定义

设 $f(x)$ 在 $(x_0-\delta, x_0+\delta)\in\mathbb{R}$ 上有定义, 若存在常数 A, 使得

$$f(x)=f(x_0)+A(x-x_0)+o(x-x_0),\quad (x\to x_0)$$

即

$$|f(x)-[f(x_0)+A(x-x_0)]|=o(|x-x_0|),$$

则称 f 在 x_0 处可微, 线性映射 $x\mapsto Ax$ 称为 f 在 x_0 处的微分, 记为 $\mathrm{d}f(x_0)$.

一元函数可微的定义

回忆, 对于一元函数而言, 可微是指该函数可以被线性函数一阶逼近.

定义

设 $f(x)$ 在 $(x_0-\delta, x_0+\delta)\in\mathbb{R}$ 上有定义, 若存在常数 A, 使得

$$f(x)=f(x_0)+A(x-x_0)+o(x-x_0), \quad (x\to x_0)$$

即

$$|f(x)-[f(x_0)+A(x-x_0)]|=o(|x-x_0|),$$

则称 f 在 x_0 处可微, 线性映射 $x\mapsto Ax$ 称为 f 在 x_0 处的微分, 记为 $\mathrm{d}f(x_0)$.

现在设 $D\subset\mathbb{R}^n$ 为开集, 向量值函数 $f:D\to\mathbb{R}^m$ 写成分量形式为

$$f(x_1,\dots,x_n)=\big(f_1(x_1,\dots,x_n),\dots,f_m(x_1,\dots,x_n)\big),$$

一元函数可微的定义

回忆, 对于一元函数而言, 可微是指该函数可以被线性函数一阶逼近.

定义

设 $f(x)$ 在 $(x_0-\delta, x_0+\delta)\in\mathbb{R}$ 上有定义, 若存在常数 A, 使得

$$f(x)=f(x_0)+A(x-x_0)+o(x-x_0), \quad (x\to x_0)$$

即

$$|f(x)-[f(x_0)+A(x-x_0)]|=o(|x-x_0|),$$

则称 f 在 x_0 处可微, 线性映射 $x\mapsto Ax$ 称为 f 在 x_0 处的微分, 记为 $\mathrm{d}f(x_0)$.

现在设 $D\subset\mathbb{R}^n$ 为开集, 向量值函数 $f:D\to\mathbb{R}^m$ 写成分量形式为

$$f(x_1,\dots,x_n)=\big(f_1(x_1,\dots,x_n),\dots,f_m(x_1,\dots,x_n)\big),$$

为方便起见, 在下面的内容中欧氏空间里的向量有时以列向量来表示.

可微的定义

定义

设 f 如上, $x^0 = (x_1^0, \ldots, x_n^0) \in D$. 如果存在 $m \times n$ 阶的矩阵 $A = (a_{ij})_{m\times n}$, 使得对于 x^0 附近的点 x, 有

$$\|f(x) - [f(x^0) + A(x - x^0)]\| = o(\|x - x^0\|), \quad (x \to x^0)$$

则称 f **在** x^0 **处可微**,

定义

设 f 如上, $x^0=(x_1^0,\dots,x_n^0)\in D$. 如果存在 $m\times n$ 阶的矩阵 $A=(a_{ij})_{m\times n}$, 使得对于 x^0 附近的点 x, 有

$$\|f(x)-[f(x^0)+A(x-x^0)]\|=o(\|x-x^0\|),\quad (x\to x^0)$$

则称 f **在** x^0 **处可微**,线性映射

$$\begin{aligned} \mathrm{d}f(x^0):\ \mathbb{R}^n &\to \mathbb{R}^m \\ v &\mapsto Av \end{aligned}$$

称为 f **在** x^0 **处的微分**.

可微必可导

命题

如果 $f: D \to \mathbb{R}^m$ 在 x^0 处可微, 则其分量 f_i $(1 \leqslant i \leqslant m)$ 在 x^0 处存在方向导数, 并且

$$A = \left(\frac{\partial f_i}{\partial x_j}(x^0) \right)_{m \times n}.$$

命题

如果 $f: D \to \mathbb{R}^m$ 在 x^0 处可微, 则其分量 f_i $(1 \leqslant i \leqslant m)$ 在 x^0 处存在方向导数, 并且

$$A = \left(\frac{\partial f_i}{\partial x_j}(x^0) \right)_{m \times n}.$$

Proof.

从微分的定义可以看出, 如果 f 在 x^0 处可微, 则 f 在 x^0 处连续.

命题

如果 $f: D \to \mathbb{R}^m$ 在 x^0 处可微, 则其分量 f_i $(1 \leqslant i \leqslant m)$ 在 x^0 处存在方向导数, 并且

$$A = \left(\frac{\partial f_i}{\partial x_j}(x^0) \right)_{m \times n}.$$

Proof.

从微分的定义可以看出, 如果 f 在 x^0 处可微, 则 f 在 x^0 处连续.下面以 $m = 1$ 为例说明方向导数的存在性.

命题

如果 $f : D \to \mathbb{R}^m$ 在 x^0 处可微, 则其分量 f_i $(1 \leqslant i \leqslant m)$ 在 x^0 处存在方向导数, 并且

$$A = \left(\frac{\partial f_i}{\partial x_j}(x^0) \right)_{m \times n}.$$

Proof.

从微分的定义可以看出, 如果 f 在 x^0 处可微, 则 f 在 x^0 处连续.下面以 $m = 1$ 为例说明方向导数的存在性.

为此, 取单位向量 u, 由定义, 我们有

$$\begin{aligned} f(x^0 + tu) - f(x^0) &= A(x^0 + tu - x^0) + o(\|x^0 + tu - x^0\|) \\ &= tAu + o(|t|), \end{aligned}$$

命题

如果 $f: D \to \mathbb{R}^m$ 在 x^0 处可微, 则其分量 f_i $(1 \leqslant i \leqslant m)$ 在 x^0 处存在方向导数, 并且

$$A = \left(\frac{\partial f_i}{\partial x_j}(x^0) \right)_{m \times n}.$$

Proof.

从微分的定义可以看出, 如果 f 在 x^0 处可微, 则 f 在 x^0 处连续.下面以 $m = 1$ 为例说明方向导数的存在性.

为此, 取单位向量 u, 由定义, 我们有

$$\begin{aligned} f(x^0 + tu) - f(x^0) &= A(x^0 + tu - x^0) + o(\|x^0 + tu - x^0\|) \\ &= tAu + o(|t|), \end{aligned}$$

(注意这里 $m = 1$, 即 f 是多元函数, 故第一个等号成立.)

命题

如果 $f: D \to \mathbb{R}^m$ 在 x^0 处可微, 则其分量 f_i $(1 \leqslant i \leqslant m)$ 在 x^0 处存在方向导数, 并且

$$A = \left(\frac{\partial f_i}{\partial x_j}(x^0)\right)_{m \times n}.$$

Proof.

从微分的定义可以看出, 如果 f 在 x^0 处可微, 则 f 在 x^0 处连续.下面以 $m = 1$ 为例说明方向导数的存在性.

为此, 取单位向量 u, 由定义, 我们有

$$\begin{aligned} f(x^0 + tu) - f(x^0) &= A(x^0 + tu - x^0) + o(\|x^0 + tu - x^0\|) \\ &= tAu + o(|t|), \end{aligned}$$

(注意这里 $m = 1$, 即 f 是多元函数, 故第一个等号成立.) 这说明 $\dfrac{\partial f}{\partial u}(x^0) = Au$, 即方向导数存在. $\square$

Proof.

特别地

$$\frac{\partial f}{\partial x_i}(x^0) = Ae_i, \quad A = \left(\frac{\partial f}{\partial x_1}(x^0), \ldots, \frac{\partial f}{\partial x_n}(x^0)\right).$$

Proof.

特别地

$$\frac{\partial f}{\partial x_i}(x^0) = Ae_i, \quad A = \Big(\frac{\partial f}{\partial x_1}(x^0), \ldots, \frac{\partial f}{\partial x_n}(x^0)\Big).$$

这里 $e_i = (0, \ldots, 0, 1, 0, \ldots, 1)^T$, 第 i 个分量为 1, 其余为 0 的单位列向量. $\square$

例子

例

研究函数

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

在 $(0,0)$ 处的可导性和可微性.

例

研究函数

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

在 $(0,0)$ 处的可导性和可微性.

解. 因为 $|f(x,y)|\leqslant\frac{1}{2}|x|$, 故 $f(x,y)$ 在 $(0,0)$ 处连续.

例

研究函数

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

在 $(0,0)$ 处的可导性和可微性.

解. 因为 $|f(x,y)|\leqslant\frac{1}{2}|x|$, 故 $f(x,y)$ 在 $(0,0)$ 处连续.易见 $f'_x(0,0)=f'_y(0,0)=0$.

例

研究函数

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

在 $(0,0)$ 处的可导性和可微性.

解. 因为 $|f(x,y)|\leqslant\frac{1}{2}|x|$, 故 $f(x,y)$ 在 $(0,0)$ 处连续.易见 $f'_x(0,0)=f'_y(0,0)=0$.事实上

$$f'_x(0,0)=\lim_{t\to0}\frac{f(0+t,0)-f(0,0)}{t}=\lim_{t\to0}\frac{0-0}{t}=0,$$

例

研究函数

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

在 $(0,0)$ 处的可导性和可微性.

解. 因为 $|f(x,y)|\leqslant\frac{1}{2}|x|$, 故 $f(x,y)$ 在 $(0,0)$ 处连续.易见 $f_x'(0,0)=f_y'(0,0)=0$.事实上

$$f_x'(0,0)=\lim_{t\to0}\frac{f(0+t,0)-f(0,0)}{t}=\lim_{t\to0}\frac{0-0}{t}=0,$$

$$f_y'(0,0)=\lim_{t\to0}\frac{f(0,0+t)-f(0,0)}{t}=\lim_{t\to0}\frac{0-0}{t}=0.$$

∎

解. 如果 $u=(u_1,u_2)$ 为单位向量, 则

$$\frac{\partial f}{\partial u}(0,0)=\lim_{t\to 0}\frac{f(0+tu_1,0+tu_2)-f(0,0)}{t}=\lim_{t\to 0}\frac{f(tu_1,tu_2)}{t}$$
$$=\lim_{t\to 0}\frac{1}{t}\frac{t^3u_1^2u_2}{t^2u_1^2+t^2u_2^2}=u_1^2u_2.$$

解. 如果 $u=(u_1,u_2)$ 为单位向量, 则

$$\begin{aligned}\frac{\partial f}{\partial u}(0,0) &= \lim_{t\to 0}\frac{f(0+tu_1,0+tu_2)-f(0,0)}{t} = \lim_{t\to 0}\frac{f(tu_1,tu_2)}{t} \\ &= \lim_{t\to 0}\frac{1}{t}\frac{t^3u_1^2u_2}{t^2u_1^2+t^2u_2^2} = u_1^2u_2.\end{aligned}$$

这说明 f 在 $(0,0)$ 处沿任何方向的方向导数都存在.

解. 如果 $u=(u_1,u_2)$ 为单位向量, 则

$$\frac{\partial f}{\partial u}(0,0)=\lim_{t\to 0}\frac{f(0+tu_1,0+tu_2)-f(0,0)}{t}=\lim_{t\to 0}\frac{f(tu_1,tu_2)}{t}$$
$$=\lim_{t\to 0}\frac{1}{t}\frac{t^3u_1^2u_2}{t^2u_1^2+t^2u_2^2}=u_1^2u_2.$$

这说明 f 在 $(0,0)$ 处沿任何方向的方向导数都存在.特别地,

$$\frac{\partial f}{\partial x}(0,0)=u_1^2u_2\Big|_{(1,0)}=0,\quad \frac{\partial f}{\partial y}(0,0)=u_1^2u_2\Big|_{(0,1)}=0,$$

∎

解. 使用反证法可以证明 f 在 $(0,0)$ 处不可微.

解. 使用反证法可以证明 f 在 $(0,0)$ 处不可微.若 f 在 $(0,0)$ 处可微, 则由定义,

$$f(tu_1, tu_2) - f(0,0) = A(tu) + o(\|tu\|)$$

这推出

$$\frac{t^2u_1^2tu_2}{t^2u_1^2 + t^2u_2^2} = tAu + o(|t|),$$

解. 使用反证法可以证明 f 在 $(0,0)$ 处不可微.若 f 在 $(0,0)$ 处可微, 则由定义,

$$f(tu_1, tu_2) - f(0,0) = A(tu) + o(\|tu\|)$$

这推出

$$\frac{t^2u_1^2tu_2}{t^2u_1^2 + t^2u_2^2} = tAu + o(|t|),$$

即

$$tu_1^2u_2 = tAu + o(|t|).$$

解. 使用反证法可以证明 f 在 $(0,0)$ 处不可微.若 f 在 $(0,0)$ 处可微, 则由定义,

$$f(tu_1, tu_2) - f(0,0) = A(tu) + o(\|tu\|)$$

这推出

$$\frac{t^2u_1^2tu_2}{t^2u_1^2 + t^2u_2^2} = tAu + o(|t|),$$

即

$$tu_1^2u_2 = tAu + o(|t|).$$

这里 Au 是矩阵乘法,

$$A = \left(\frac{\partial f}{\partial x_1}(0,0), \frac{\partial f}{\partial x_1}(0,0)\right) = (0,0),$$

而 $u = (u_1, u_2)^T$, 故得

$$tu_1^2u_2 = o(|t|),$$

矛盾. 因此 f 在 $(0,0)$ 处不可微. ∎

雅可比矩阵

定义

如果 f_i $(1 \leqslant i \leqslant m)$ 的偏导数均存在，则记

$$Jf = \left(\frac{\partial f_i}{\partial x_j}\right)_{m\times n} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \cdots & \cdots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}_{m\times n}$$

称为 f **的雅可比(Jacobian)矩阵**.

雅可比矩阵

定义

如果 f_i $(1 \leqslant i \leqslant m)$ 的偏导数均存在, 则记

$$Jf = \left(\frac{\partial f_i}{\partial x_j}\right)_{m\times n} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \cdots & \cdots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}_{m\times n}$$

称为 f **的雅可比(Jacobian)矩阵**.当 $m = 1$ 时, 又记 Jf 为 ∇f, 称为 f **的梯度**.

雅可比矩阵

定义

如果 f_i $(1 \leqslant i \leqslant m)$ 的偏导数均存在, 则记

$$Jf = \left(\frac{\partial f_i}{\partial x_j}\right)_{m\times n} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \cdots & \cdots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}_{m\times n}$$

称为 f **的雅可比(Jacobian)矩阵**.当 $m = 1$ 时, 又记 Jf 为 ∇f, 称为 f **的梯度**.Jf 在每一点的值构成一个映射 $Jf : D \to \mathbb{R}^{mn}$, 这里我们将 $m \times n$ 矩阵视为 $\mathbb{R}^{mn}$ 中的点.

可微的充分条件

定理

如果 Jf 在 D 上存在, 且它作为映射在 x^0 处连续, 则 f 在 x^0 处可微.

可微的充分条件

定理

如果 Jf 在 D 上存在, 且它作为映射在 x^0 处连续, 则 f 在 x^0 处可微.

Proof.

仍以 $m = 1$ 为例来证明.

可微的充分条件

定理

如果 Jf 在 D 上存在, 且它作为映射在 x^0 处连续, 则 f 在 x^0 处可微.

Proof.

仍以 $m = 1$ 为例来证明.由条件, f'_{x_i} 在 x^0 处连续, $i = 1, 2, \ldots, n$.

可微的充分条件

定理

如果 Jf 在 D 上存在, 且它作为映射在 x^0 处连续, 则 f 在 x^0 处可微.

Proof.

仍以 $m = 1$ 为例来证明.由条件, f'_{x_i} 在 x^0 处连续, $i = 1, 2, \ldots, n$.根据微分中值定理, 有

$$\begin{aligned}
f(x) - f(x^0) &= \sum_{i=1}^{n} \left[f(x_1^0, \ldots, x_{i-1}^0, x_i, x_{i+1}, \ldots, x_n) - f(x_1^0, \ldots, x_i^0, x_{i+1}, \ldots, x_n) \right] \\
&= \sum_{i=1}^{n} f'_{x_i}(x_1^0, \ldots, x_{i-1}^0,\ x_i^0 + \theta_i(x_i - x_i^0)\,, x_{i+1}, \ldots, x_n) \cdot (x_i - x_i^0) \\
&= \sum_{i=1}^{n} f'_{x_i}(x^0)(x_i - x_i^0) + \sum_{i=1}^{n} \alpha_i (x_i - x_i^0),
\end{aligned}$$

$\square$

Proof.

其中, 当 $x_i \to x_i^0$ 时,

$$\alpha_i = f'_{x_i}(x_1^0, \ldots, x_{i-1}^0, x_i^0 + \theta_i(x_i - x_i^0), x_{i+1}, \ldots, x_n) - f'_{x_i}(x_1^0, \ldots, x_n^0) \ \to \ 0,$$

Proof.

其中, 当 $x_i \to x_i^0$ 时,

$$\alpha_i = f'_{x_i}(x_1^0, \ldots, x_{i-1}^0, x_i^0 + \theta_i(x_i - x_i^0), x_{i+1}, \ldots, x_n) - f'_{x_i}(x_1^0, \ldots, x_n^0) \ \to \ 0,$$

从而

$$\begin{aligned} \left\| f(x) - \left[f(x^0) + \sum_{i=1}^n f'_{x_i}(x^0)(x_i - x_i^0) \right] \right\| &= \left\| \sum_{i=1}^n \alpha_i (x_i - x_i^0) \right\| \\ &\leqslant \left(\sum_{i=1}^n \alpha_i^2 \right)^{\frac{1}{2}} \|x - x^0\| = o(\|x - x^0\|), \end{aligned}$$

即 f 在 x^0 处可微. □

矩阵的范数

如果我们将 $m \times n$ 的矩阵视为 $\mathbb{R}^{mn}$ 中的点, 则矩阵也可定义自然的范数.

矩阵的范数

如果我们将 $m \times n$ 的矩阵视为 $\mathbb{R}^{mn}$ 中的点, 则矩阵也可定义自然的范数.即, 如果 $A = (a_{ij})_{m\times n}$, 则其范数定义为

$$\|A\| = \Big[\sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}^2\Big]^{\frac{1}{2}}.$$

如果我们将 $m \times n$ 的矩阵视为 $\mathbb{R}^{mn}$ 中的点, 则矩阵也可定义自然的范数.即, 如果 $A = (a_{ij})_{m\times n}$, 则其范数定义为

$$\|A\| = \Big[\sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}^2\Big]^{\frac{1}{2}}.$$

由 Schwarz 不等式可得

$$\|Av\| \leqslant \|A\| \cdot \|v\|, \quad \forall\ v \in \mathbb{R}^n.$$

定理

设 Δ 为 $\mathbb{R}^l$ 中开集, D 为 $\mathbb{R}^m$ 中开集, $g : \Delta \to D$ 及 $f : D \to \mathbb{R}^n$ 为映射. 如果 g 在 $u^0 \in \Delta$ 处可微, f 在 $x^0 = g(u^0)$ 处可微, 则复合映射 $h = f \circ g : \Delta \to \mathbb{R}^n$ 在 u^0 处可微, 且

$$Jh(u^0) = Jf(x^0) \cdot Jg(u^0).$$

复合求导的链式法则

定理

设 Δ 为 $\mathbb{R}^l$ 中开集, D 为 $\mathbb{R}^m$ 中开集, $g:\Delta\to D$ 及 $f:D\to\mathbb{R}^n$ 为映射. 如果 g 在 $u^0\in\Delta$ 处可微, f 在 $x^0=g(u^0)$ 处可微, 则复合映射 $h=f\circ g:\Delta\to\mathbb{R}^n$ 在 u^0 处可微, 且

$$Jh(u^0)=Jf(x^0)\cdot Jg(u^0).$$

Proof.

因为 g 在 u^0 可微, 故

$$g(u)-g(u^0)=Jg(u^0)\cdot(u-u^0)+R_g(u,u^0), \tag{1}$$

其中 $R_g(u,u^0)=o(\|u-u_0\|)$. □

续

Proof.

同理, 因为 f 在 $x^0 = g(u^0)$ 处可微, 故

$$f(x) - f(x^0) = Jf(x^0) \cdot (x - x^0) + R_f(x, x^0), \tag{2}$$

其中 $R_f(x, x^0) = o(\|x - x_0\|)$.

Proof.

同理, 因为 f 在 $x^0 = g(u^0)$ 处可微, 故

$$f(x) - f(x^0) = Jf(x^0) \cdot (x - x^0) + R_f(x, x^0), \tag{2}$$

其中 $R_f(x, x^0) = o(\|x - x_0\|)$.

由 (1) 式知, 当 $u \to u^0$ 时, $g(u) \to g(u^0) = x^0$.

Proof.

同理, 因为 f 在 $x^0 = g(u^0)$ 处可微, 故

$$f(x) - f(x^0) = Jf(x^0) \cdot (x - x^0) + R_f(x, x^0), \tag{2}$$

其中 $R_f(x, x^0) = o(\|x - x_0\|)$.

由 (1) 式知, 当 $u \to u^0$ 时, $g(u) \to g(u^0) = x^0$.以 $x = g(u)$ 代入 (2), 得

$$\begin{aligned} f \circ g(u) - f \circ g(u^0) &= Jf(x^0)(g(u) - g(u^0)) + R_f(g(u), g(u^0)) \\ &= Jf(x^0) \cdot \left[Jg(u^0) \cdot (u - u^0) + R_g(u, u^0) \right] + R_f(g(u), g(u^0)) \\ &= Jf(x^0) \cdot Jg(u^0) \cdot (u - u^0) + R_{f \circ g}(u, u^0), \qquad (3) \end{aligned}$$

其中

$$R_{f \circ g}(u, u^0) = Jf(x^0) \cdot R_g(u, u^0) + R_f(g(u), g(u^0)).$$

Proof.

我们有如下估计:

$$\begin{aligned}\|R_{f\circ g}(u,u^0)\| &\leqslant \|Jf(x^0)\cdot R_g(u,u^0)\| + \|R_f(g(u),g(u^0))\| \\ &\leqslant \|Jf(x^0)\|\cdot\|R_g(u,u^0)\| + o(\|g(u)-g(u^0)\|) \\ &= o(\|u-u^0\|) + o(O(\|u-u^0\|)) = o(\|u-u^0\|).\end{aligned}$$

Proof.

我们有如下估计:

$$\begin{aligned}\|R_{f\circ g}(u,u^0)\| &\leqslant \|Jf(x^0)\cdot R_g(u,u^0)\| + \|R_f(g(u),g(u^0))\| \\ &\leqslant \|Jf(x^0)\|\cdot\|R_g(u,u^0)\| + o(\|g(u)-g(u^0)\|) \\ &= o(\|u-u^0\|) + o(O(\|u-u^0\|)) = o(\|u-u^0\|).\end{aligned}$$

由 (3) 及微分的定义知 $f\circ g$ 在 u^0 处可微, 且 $J(f\circ g)(u^0) = Jf(x^0)\cdot Jg(u^0)$. $\square$

链式法则

如果将 f, g 分别表示成分量形式:

$$y_i = f_i(x_1, \ldots, x_m), \quad i = 1, 2, \ldots, n,$$

$$x_j = g_j(u_1, \ldots, u_l), \quad j = 1, 2, \ldots, m,$$

链式法则

如果将 f, g 分别表示成分量形式:

$$y_i = f_i(x_1, \dots, x_m), \quad i = 1, 2, \dots, n,$$
$$x_j = g_j(u_1, \dots, u_l), \quad j = 1, 2, \dots, m,$$

则 $J(f \circ g)(u^0) = Jf(x^0) \cdot Jg(u^0)$ 写成矩阵的形式为

$$\begin{pmatrix} \frac{\partial y_1}{\partial u_1}(u^0) & \cdots & \frac{\partial y_1}{\partial u_l}(u^0) \\ \vdots & & \vdots \\ \frac{\partial y_n}{\partial u_1}(u^0) & \cdots & \frac{\partial y_n}{\partial u_l}(u^0) \end{pmatrix}_{n\times l}$$
$$= \begin{pmatrix} \frac{\partial y_1}{\partial x_1}(u^0) & \cdots & \frac{\partial y_1}{\partial x_m}(u^0) \\ \vdots & & \vdots \\ \frac{\partial y_n}{\partial x_1}(u^0) & \cdots & \frac{\partial y_n}{\partial x_m}(u^0) \end{pmatrix}_{n\times m} \begin{pmatrix} \frac{\partial x_1}{\partial u_1}(u^0) & \cdots & \frac{\partial x_1}{\partial u_l}(u^0) \\ \vdots & & \vdots \\ \frac{\partial x_m}{\partial u_1}(u^0) & \cdots & \frac{\partial x_m}{\partial u_l}(u^0) \end{pmatrix}_{m\times l},$$

链式法则

即

$$\frac{\partial y_i}{\partial u_j}(u^0) = \sum_{s=1}^{m} \frac{\partial y_i}{\partial x_s}(g(u^0)) \cdot \frac{\partial x_s}{\partial u_j}(u^0).$$

链式法则

即

$$\frac{\partial y_i}{\partial u_j}(u^0) = \sum_{s=1}^{m} \frac{\partial y_i}{\partial x_s}(g(u^0)) \cdot \frac{\partial x_s}{\partial u_j}(u^0).$$

这就是所谓的**链式法则**.

例

设 $u = f(x, y)$ 可微, $x = r\cos\theta$, $y = r\sin\theta$, 证明

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

例

设 $u = f(x, y)$ 可微, $x = r\cos\theta$, $y = r\sin\theta$, 证明

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

Proof.

由链式法则,

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

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

□

续

Proof.

则

$$\begin{aligned}\left(\frac{\partial u}{\partial r}\right)^2+\frac{1}{r^2}\left(\frac{\partial u}{\partial \theta}\right)^2&=\left(\frac{\partial u}{\partial x}\cos\theta+\frac{\partial u}{\partial y}\sin\theta\right)^2+\frac{1}{r^2}\left(-r\frac{\partial u}{\partial x}\sin\theta+r\frac{\partial u}{\partial y}\cos\theta\right)^2\\&=\left(\frac{\partial u}{\partial x}\cos\theta+\frac{\partial u}{\partial y}\sin\theta\right)^2+\left(-\frac{\partial u}{\partial x}\sin\theta+\frac{\partial u}{\partial y}\cos\theta\right)^2\\&=\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2.\end{aligned}$$

续

Proof.

则

$$\begin{aligned}\left(\frac{\partial u}{\partial r}\right)^2+\frac{1}{r^2}\left(\frac{\partial u}{\partial \theta}\right)^2 &= \left(\frac{\partial u}{\partial x}\cos\theta+\frac{\partial u}{\partial y}\sin\theta\right)^2+\frac{1}{r^2}\left(-r\frac{\partial u}{\partial x}\sin\theta+r\frac{\partial u}{\partial y}\cos\theta\right)^2\\ &= \left(\frac{\partial u}{\partial x}\cos\theta+\frac{\partial u}{\partial y}\sin\theta\right)^2+\left(-\frac{\partial u}{\partial x}\sin\theta+\frac{\partial u}{\partial y}\cos\theta\right)^2\\ &= \left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2.\end{aligned}$$

这说明 Δu 在极坐标系下的形式为

$$\Delta u = u_{rr}+\frac{1}{r^2}u_{\theta\theta}.$$

□

全微分

全微分(形式微分)

定义

设 $D \subset \mathbb{R}^m$, $f: D \to \mathbb{R}$ 为可微的 m 元函数, 由定义, f 在 x 处的微分 $\mathrm{d}f(x)$ 是一个线性映射:

$$\begin{aligned} \mathrm{d}f(x): \mathbb{R}^m &\to \mathbb{R} \\ \mathbf{v} &\mapsto \sum_{i=1}^{m} \frac{\partial f}{\partial x_i} \cdot v_i. \end{aligned}$$

这里 $\mathbf{v} = (v_1, \ldots, v_m) \in \mathbb{R}^m$. 称映射 $x \mapsto \mathrm{d}f(x)$ 为 f 的**全微分或形式微分**, 记为 $\mathrm{d}f$.

全微分之间的运算

由于

$$\mathrm{d}(\lambda f+\mu g)(x)=\lambda \mathrm{d}f(x)+\mu \mathrm{d}g(x), \quad \forall\ \lambda,\mu\in\mathbb{R},$$

因此，全微分之间也可以定义加法和数乘运算，在这个意义下，有

$$\mathrm{d}f=\sum_{i=1}^{m}\frac{\partial f}{\partial x_i}\mathrm{d}x_i \quad (*)$$

$$\mathrm{d}(\lambda f+\mu g)=\lambda \mathrm{d}f+\mu \mathrm{d}g,$$

$$\mathrm{d}(f\cdot g)=f\mathrm{d}g+g\mathrm{d}f,$$

$$\mathrm{d}\Big(\frac{f}{g}\Big)=\frac{g\mathrm{d}f-f\mathrm{d}g}{g^2} \quad (g\neq 0).$$

如果将 $(*)$ 写成矩阵形式:

$$\mathrm{d}f = Jf \cdot (\mathrm{d}x_1, \ldots, \mathrm{d}x_m)^T,$$

则复合映射的链式法则可写为

$$\begin{aligned}\mathrm{d}(f \circ g) &= J(f \circ g) \cdot (\mathrm{d}u_1, \ldots, \mathrm{d}u_\ell)^T \\ &= Jf(x) \cdot Jg(u)(\mathrm{d}u_1, \ldots, \mathrm{d}u_\ell)^T \qquad (x = g(u)) \\ &= Jf(g(u)) \cdot (\mathrm{d}g_1, \ldots, \mathrm{d}g_m)^T.\end{aligned}$$

这个等式称为**全微分的形式不变性**.