Lecture 15

## §1 Chain rule

### 1. Special Case 1 : $f(x(t), y(t))$

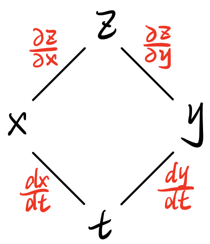
令 $z=f(x,y)$ 在 $(x_0,y_0)$ 处 differentiable，且 $x=x(t)$ 与 $y=y(t)$ 在 $t_0$ 处均 differentiable，

$(x_0=x(t_0), \ y_0=y(t_0))$

1° 链式法则:

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

2° branch diagram :




3° 证明

- Fix $t=t_0$，令 $t$ 有一个改变量 $\Delta t$，由此导致 $\Delta x, \Delta y, \Delta z$
- 因为 $f$ 在 $(x_0,y_0)$ 处可微，

  $$\Delta z=(f_x(x_0,y_0)+\varepsilon_1)\Delta x+(f_y(x_0,y_0)+\varepsilon_2)\Delta y \quad ①$$

  当 $(\Delta x,\Delta y)\to(0,0)$ 时，$\varepsilon_1\to0$，$\varepsilon_2\to0$.
- 因为 $x, y$ 在 $t_0$ 处可导，

  $$\Delta x=(x'(t_0)+\varepsilon_x)\Delta t$$

  $$\Delta y=(y'(t_0)+\varepsilon_y)\Delta t$$

  当 $t\to0$ 时，$\varepsilon_x\to0$，$\varepsilon_y\to0$
- 由此，① 式变为

  $$\Delta z=(f_x(x_0,y_0)+\varepsilon_1)\cdot(x'(t_0)+\varepsilon_x)\cdot\Delta t+(f_y(x_0,y_0)+\varepsilon_2)\cdot(y'(t_0)+\varepsilon_y)\Delta t$$

  因此，

  $$\frac{\Delta z}{\Delta t}=(f_x(x_0,y_0)+\varepsilon_1)\cdot(x'(t_0)+\varepsilon_x)+(f_y(x_0,y_0)+\varepsilon_2)\cdot(y'(t_0)+\varepsilon_y) \quad ②$$
- 因为随着 $\Delta t\to0$，$(\Delta x,\Delta y)\to(0,0)$，

  因此 $\varepsilon_1\to0$，$\varepsilon_2\to0$
- 对 ② 式两侧同时取 $\lim\limits_{\Delta t\to0}$ 得：

  $$\left.\frac{dz}{dt}\right|_{t=t_0}=f_x(x_0,y_0)\cdot x'(t_0)+f_y(x_0,y_0)\cdot y'(t_0)$$

例: Suppose that $z=x^2y+3xy^4$, $x=\sin(2t)$ and $y=\cos t$. Find $\frac{dz}{dt}$ when $t=0$.

$\frac{\partial z}{\partial x}=2yx+3y^4$ ， $\frac{\partial z}{\partial y}=x^2+12xy^3$ ， $\frac{dx}{dt}=2\cos(2t)$ ， $\frac{dy}{dt}=-\sin t$

$\left.\frac{dz}{dt}\right|_{t=0}=3\times2+0\times0=6$

2. General Case 1: $f(x_1(t), x_2(t), \cdots, x_n(t))$

令 $w = f(x_1, x_2, \cdots, x_n)$, $x_1 = x_1(t)$, $\cdots$ $x_n = x_n(t)$ 均 differentiable. 则 $w$ 对 $t$ 可微，且

$$\frac{dw}{dt} = \frac{\partial w}{\partial x_1}\cdot\frac{dx_1}{dt} + \cdots + \frac{\partial w}{\partial x_n}\cdot\frac{dx_n}{dt}$$

其中 $x_i$ 为 $f$ 的第 $i$ 个变量

3. Special Case 2: $z = f(x,y)$, $x = x(s,t)$, $y = y(s,t)$

令 $z = f(x,y)$, $x = x(s,t)$, $y = y(s,t)$ 均 differentiable, 则 $z$ 对 $(s,t)$ 均 differentiable, 且

$$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x}\cdot\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\cdot\frac{\partial y}{\partial s}$$

$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\cdot\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\cdot\frac{\partial y}{\partial t}$$

例: Suppose that $z = e^x \sin y$, $x = st^2$ and $y = s^2 t$. Find $\frac{\partial z}{\partial s}$

$\frac{\partial z}{\partial x} = \sin y \cdot e^x$, $\frac{\partial z}{\partial y} = e^x \cos y$, $\frac{\partial x}{\partial s} = t^2$, $\frac{\partial y}{\partial s} = 2t\cdot s$

$$\frac{\partial z}{\partial s} = \sin y \cdot e^x \cdot t^2 + e^x \cdot \cos y \cdot 2ts$$

$$= \sin(s^2 t)\cdot e^{st^2}\cdot t^2 + e^{st^2}\cdot\cos(s^2 t)\cdot 2ts$$

4. General Case 2

**Theorem (Chain Rule)**

*Suppose that $u = f(x_1, x_2, \ldots, x_n)$ is a differentiable real-valued function with $n$ variables, and suppose that for each $i$, $x_i = g_i(t_1, t_2, \ldots, t_m)$ is a differentiable real-valued function with $m$ variables. Then $u$ is differentiable with respect to $(t_1, t_2, \ldots, t_m)$, and*

$$\frac{\partial u}{\partial t_j} = \frac{\partial u}{\partial x_1}\frac{\partial x_1}{\partial t_j} + \frac{\partial u}{\partial x_2}\frac{\partial x_2}{\partial t_j} + \cdots + \frac{\partial u}{\partial x_n}\frac{\partial x_n}{\partial t_j}.$$

§2 Implicit differentiation

1. 二元函数的隐函数求导

14.4.8

**THEOREM ~~8~~—A Formula for Implicit Differentiation** Suppose that $F(x, y)$ is differentiable and that the equation $F(x, y) = 0$ defines $y$ as a differentiable function of $x$. Then at any point where $F_y \neq 0$,

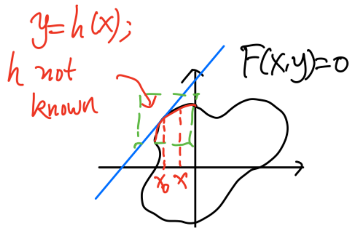
$$\frac{dy}{dx} = -\frac{F_x}{F_y}. \tag{1}$$

证明:

- 在 $(x_0, y_0)$ 附近，$y$ 可以视作 $x$ 的隐函数 $h(x)$, given by $F(x,y)=0$.
  则在 $(x_0,y_0)$ 附近，曲线上的任一点均满足 $F(x, h(x)) = 0$

- 令 $z = F(x,y) = F(x, h(x))$，则 $\frac{dz}{dx}\Big|_{(x,y)=(x_0,y_0)} = 0$
- 根据 chain rule，若 $F(x,y)$ 可微，则

$$\frac{dz}{dx} = F_x \cdot \frac{dx}{dx} + F_y \cdot \frac{dy}{dx} = F_x + F_y \frac{dy}{dx} = 0$$




由此：$\frac{dy}{dx} = -\frac{F_x}{F_y}$ (if $F_y \neq 0$)

例: Find the slope of the tangent line to the curve $x^3+y^3=6xy$ at the point $(3,3)$

$F(x,y) = x^3+y^3-6xy$

$F_x = 3x^2-6y$，$F_y = 3y^2-6x$

$\frac{dy}{dx}\Big|_{(x,y)=(3,3)} = -\frac{F_x}{F_y} = -1$

## 2、三元函数的隐函数求导

$$\boxed{\frac{\partial z}{\partial x} = -\frac{F_x}{F_z} \text{ and } \frac{\partial z}{\partial y} = -\frac{F_y}{F_z}, \text{ whenever } F_z \neq 0.}$$

例: Find $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ if $x^3+y^3+z^3+6xy=1$

$F_x = 3x^2+6y$

$F_y = 3y^2+6x$

$F_z = 3z^2$

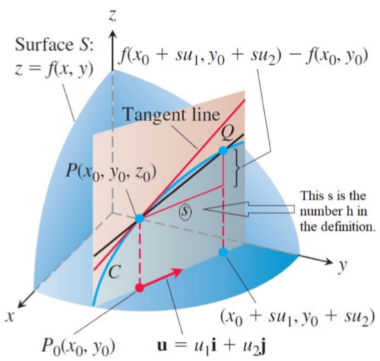
$\frac{\partial z}{\partial x} = -\frac{F_x}{F_z} = -\frac{x^2+2y}{z^2}$

$\frac{\partial z}{\partial y} = -\frac{F_y}{F_z} = -\frac{y^2+2x}{z^2}$

# §3 Directional Derivative and Gradient Vectors

## 1. Directional derivative (方向导数) 的定义

> **Definition**
> Let $(x_0, y_0)$ be an interior point of the domain of a two-variable function $f$, and let $\vec{u}$ be a unit vector in $\mathbb{R}^2$. The (directional) derivative of $f$ at $(x_0, y_0)$ in the direction of $\vec{u}$, denoted by $D_{\vec{u}}f(x_0, y_0)$, is defined by
>
> $$D_{\vec{u}}f(x_0,y_0) := \lim_{h\to 0}\frac{f(x_0+hu_1, y_0+hu_2) - f(x_0,y_0)}{h}.$$




注: 1° $\vec{u} = u_1\vec{i} + u_2\vec{j}$ 可以视作方向余弦

2° 数值上偏导可以视作方向导数的 special case：

$f_x = D_{\vec{u}_1}f$，$\vec{u}_1 = \langle 1, 0\rangle$

$f_y = D_{\vec{u}_2}f$，$\vec{u}_2 = \langle 0, 1\rangle$

3° Vector notation (generalize to n-variable functions)

$$D_{\vec{u}}f(\vec{x}_0) = \lim_{h\to 0}\frac{f(\vec{x}_0+h\vec{u}) - f(\vec{x}_0)}{h}$$

## 2、方向导数的性质与梯度

**Theorem**

*If f is a differentiable two-variable function, then*

$$D_{\vec{u}}f(x_0, y_0) = f_x(x_0, y_0)u_1 + f_y(x_0, y_0)u_2.$$

*In other words,*

$$D_{\vec{u}}f(x_0, y_0) = \nabla f(x_0, y_0) \cdot \vec{u},$$

*where* $\nabla f := \langle f_x, f_y \rangle$.

注：1° 向量 $\nabla f$ 被称为 gradient（梯度）或 gradient vector of $f$

2° 对于 n 元函数 $f(x_1, \cdots, x_n)$，$\nabla f := \langle f_{x_1}, \cdots, f_{x_n} \rangle$.

若 f 在一点 $\vec{x_0} \in \mathbb{R}^n$ 处可微，则有 $D_{\vec{u}}f(\vec{x_0}) = \nabla f(\vec{x_0}) \cdot \vec{u}$

证明：

- 定义 $g(t) := f(x_0 + tu_1, y_0 + tu_2)$. 则 $g(0) = f(x_0, y_0)$，且

$$D_{\vec{u}}f(x_0, y_0) = \lim_{h \to 0} \frac{g(h) - g(0)}{h} = g'(0)$$

- 根据链式法则，

$$g'(0) = f_x(x(0), y(0)) \cdot x'(0) + f_y(x(0), y(0)) \cdot y'(0)$$
$$= f_x(x_0, y_0) \cdot u_1 + f_y(x_0, y_0) \cdot u_2$$
$$= \nabla f(x_0, y_0) \cdot \vec{u}$$

where $\nabla f := \langle f_x, f_y \rangle$ is the gradient vector.

例：Find the directional derivative of $f(x,y) := xe^y + \cos(xy)$ at the point $(2,0)$ in the direction of $\vec{v} = \langle 3, -4 \rangle$.

- $\vec{u} = \langle \frac{3}{5}, -\frac{4}{5} \rangle$
- $f_x = e^y - y\sin(xy)$

  $f_y = xe^y - x\sin(xy)$

  $\nabla f(2,0) = \langle 1, 2 \rangle$
- $D_{\vec{u}}f(2,0) = \frac{3}{5} - \frac{8}{5} = -1$

## 3、方向导数与可微

existence of all directional derivatives $\not\longrightarrow$ differentiability

e.g. Consider $f(x,y) := \begin{cases} xy^2/(x^2+y^4) & , \text{ if } (x,y) \neq (0,0) \\ 0 & , \text{ if } (x,y) = (0,0) \end{cases}$

- Let $\vec{u}$ be any unit vector $\langle u_1, u_2 \rangle$

$$D_{\vec{u}}f(0,0) = \lim_{h\to 0} \frac{f(hu_1,hu_2)-f(0,0)}{h}$$

$$= \lim_{h\to 0} \frac{h^3u_1u_2^2/(h^2u_1^2+h^4u_2^4)}{h}$$

$$= \lim_{h\to 0} \frac{u_1u_2^2}{u_1^2+h^2u_2^4}$$

$$= \begin{cases} u_2^2/u_1, & \text{if } u_1\neq 0 \\ 0, & \text{if } u_1=0 \end{cases}$$

- 因此，f在(0,0)处的所有方向导数均存在。
  但f在(0,0)处不连续，因此不可微

## 4. 梯度向量的几何意义

1° 意义1：指向 fasest increase 的方向

> **Theorem**
> *Let f be a differentiable function (with multiple variables), and let $\vec{x_0}$ represent an interior point of the domain of f. Then, among all directional derivatives $D_{\vec{u}}f(\vec{x_0})$:*
> - *The maximum value of $D_{\vec{u}}f(\vec{x_0})$ is $|\nabla f(\vec{x_0})|$, and it occurs when $\vec{u}$ is the direction of $\nabla f(\vec{x_0})$.*
> - *The minimum value of $D_{\vec{u}}f(\vec{x_0})$ is $-|\nabla f(\vec{x_0})|$, and it occurs when $\vec{u}$ is the direction of $-\nabla f(\vec{x_0})$.*
>
> *If $\nabla f(\vec{x_0}) \neq \vec{0}$ and $\vec{u}$ is orthogonal to $\nabla f(\vec{x_0})$, then $D_{\vec{u}}f(\vec{x_0}) = 0$.*

注：上述定理表明，在任意一点，一个可微函数沿着 $\nabla f$ 方向增长最快，沿着 $-\nabla f$ 方向增长最慢，增长率为 $|\nabla f|$（梯度的模）

例 **Exercise**
Suppose that the temperature at a point $(x,y,z)$ in space is given by $T(x,y,z) = 80/(1+x^2+2y^2+3z^2)$, where temperature is measured in degrees Celsius and distance is measured in meters. In which direction does the temperature increase fastest at the point $(1,1,-2)$? What is the maximum rate of increase?

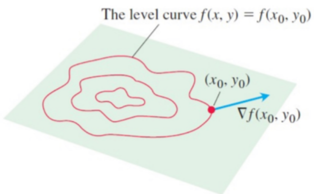
Sol:
- $$T_x(x,y,z) = -\frac{80\cdot 2x}{(1+x^2+2y^2+3z^2)^2}$$
  $$T_y(x,y,z) = -\frac{80\cdot 4y}{(1+x^2+2y^2+3z^2)^2}$$
  $$T_z(x,y,z) = -\frac{80\cdot 6z}{(1+x^2+2y^2+3z^2)^2}$$
  $$\nabla T(1,1,-2) = -\frac{5}{8}\langle 1,2,-6\rangle$$
- increases fastest in the direction $\vec{u} = \nabla T(1,1,2)/|\nabla T(1,1,2)| = -\frac{1}{\sqrt{41}}\langle 1,2,6\rangle$
- $\text{Max } D_{\vec{u}}T(1,1,2) = |\nabla T(1,1,2)| = \frac{5}{8}\sqrt{41}$ (°C/m)

2° 意义2：与 level curves 的联系

平面上一曲线(等位线)上 $(x_0,y_0)$ 处切线为：$f_x(x_0,y_0)\cdot(x-x_0)+f_y(x_0,y_0)\cdot(y-y_0)=0$

**Theorem**

*At any point $(x_0, y_0)$ in the domain of a differentiable function $f$, if $C$ is the level curve through $(x_0, y_0)$, then the gradient of $f$ at $(x_0, y_0)$ is perpendicular to the tangent vector to $C$ at $(x_0, y_0)$.*




This geometric fact allows us to describe the tangent line to the level curve of $f$ at a point $(x_0, y_0)$ using the equation

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) = 0.$$

证明：

- 令 $f(x,y)$ 在 $(x_0,y_0)$ 处可微.

  令一条曲线 $C$ 在定义域平面内经过 $(x_0,y_0)$，其参数方程为 $\vec{r}(t)=\langle x(t),y(t)\rangle$

  考虑沿曲线 $C$ 的 "height"：$z=f(x,y)$

  $$z=f(x(t),y(t)) = f(\vec{r}(t))$$

- By chain rule,

  $$\left.\frac{dz}{dt}\right|_{t=t_0} = f_x(x(t_0),y(t_0))\,x'(t_0) + f_y(x(t_0),y(t_0))\,y'(t_0),$$

  i.e. $(f\circ\vec{r})'(t_0) = \nabla f(\vec{r}(t_0))\cdot\vec{r}\,'(t_0)$

- 若令 $f(x,y)=K$，则对于 $C$ 上任一点 $(x_0,y_0)$，有 $\frac{dz}{dt}=0$，即

  $$\nabla f(x_0,y_0)\cdot\vec{r}\,'|_{(x_0,y_0)}=0$$

例：Find an equation for the tangent line to the ellipse $\frac{x^2}{8}+\frac{y^2}{2}=1$ at point $(-2,1)$

- $f(x,y)=\frac{x^2}{8}+\frac{y^2}{2}$

  $f_x=\frac{x}{4}$，$f_y=y$

- $f_x(-2,1)=-\frac{1}{2}$，$f_y(-2,1)=1$

- Tangent is $-\frac{1}{2}(x+2)+1\cdot(y-1)=0 \Rightarrow y=\frac{1}{2}x+2$

5. 梯度的性质

**Algebra Rules for Gradients**

| | | |
|---|---|---|
| **1.** *Sum Rule:* | $\nabla(f+g)=\nabla f+\nabla g$ | |
| **2.** *Difference Rule:* | $\nabla(f-g)=\nabla f-\nabla g$ | |
| **3.** *Constant Multiple Rule:* | $\nabla(kf)=k\nabla f$ (any number $k$) | |
| **4.** *Product Rule:* | $\nabla(fg)=f\nabla g+g\nabla f$ | Scalar multipliers on left of gradients |
| **5.** *Quotient Rule:* | $\nabla\left(\frac{f}{g}\right)=\frac{g\nabla f-f\nabla g}{g^2}$ | |