- When $f_{xy} = f_{yx}$ ? Clairaut's thm: If $f$ is $C^2$-function $f_{xy} = f_{yx}$.
- $C^r$-function $\Rightarrow$ can change the order of partial derivatives up to order $r$.
- differentiability: $\exists$ linear approximation $L(\vec{x})$ s.t. $\varepsilon(\vec{x}) = f(\vec{x}) - L(\vec{x})$

$$\lim_{\vec{x} \to \vec{a}} \frac{\varepsilon(\vec{x})}{\|\vec{x} - \vec{a}\|} = 0.$$

<u>Example</u> $f(x,y) = \sqrt{|xy|}$ is not differentiable at $(0,0)$

(sol) partial derivatives?

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

$$= \lim_{h \to 0} \frac{0 - 0}{h} = 0.$$

Similarly, $\frac{\partial f}{\partial y}(0,0) = 0$.

$$\therefore L(x,y) = f(0,0) + \frac{\partial f}{\partial x}(0,0)(x-0) + \frac{\partial f}{\partial y}(0,0)(y-0)$$

$$= 0 + 0 + 0$$

$$= 0.$$

$L(x,y)$ is the zero function.

$$\varepsilon(x,y) = f(x,y) - L(x,y) = \sqrt{|xy|}$$

To be differentiable, $\lim\limits_{(x,y)\to(0,0)} \dfrac{\varepsilon(x,y)}{\|(x,y)-(0,0)\|} = 0$.

$$\lim_{(x,y)\to(0,0)} \frac{\varepsilon(x,y)}{\|(x,y)-(0,0)\|} = \lim_{(x,y)\to(0,0)} \frac{\sqrt{|xy|}}{\sqrt{x^2+y^2}}$$

$$\left(\begin{matrix} x = r\cos\theta \\ y = r\sin\theta \end{matrix}\right) = \lim_{r\to 0} \frac{\sqrt{|r^2\cos\theta\sin\theta|}}{r}$$

$$= \lim_{r\to 0} \sqrt{|\cos\theta\sin\theta|}$$

this depends on $\theta$.

i.e. This limit does not exist.

$\therefore$ $f$ is not differentiable at $(0,0)$.

**Rmk** In this example $f(x,y) = \sqrt{|xy|}$, $L(x,y) = 0$.

Restric to x-axis : $f(x,0) = 0 = L(x,0)$

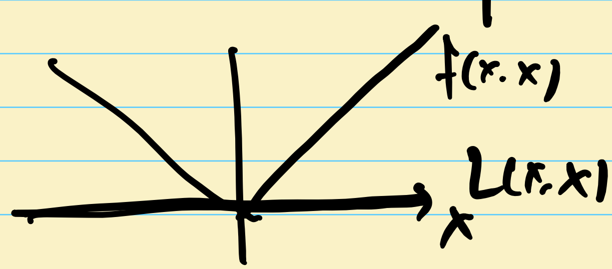
( $L(x,0)$ good approximation for $f(x,0)$ )

Restrict to y-axis : $f(0,y) = 0 = L(0,y)$

( $L(0,y)$ good appx for $f(0,y)$ )

Restrict to line $y=x$: $f(x,x) = |x|$

$L(x,x) = 0$.




( $L(x,x)$ is bad approximation for $f(x,x)$ )

differentiability implies :
information from coordinate directions ($\frac{\partial f}{\partial x_i}$)
can tell information on every direction

strong ——— ∘ ——— ∘ ———

**Thm** If $f(\vec{x})$ is differentiable at $\vec{a}$,
then $f(\vec{x})$ is continuous at $\vec{a}$.

(proof) Suppose $f(\vec{x})$ is differentiable at $\vec{a}$.
Then $\exists$ linear approximation

$$L(\vec{x}) = f(\vec{a}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{a})\,(x_i - a_i)$$

s.t. $\varepsilon(\vec{x}) = f(\vec{x}) - L(\vec{x})$ , $\displaystyle\lim_{\vec{x}\to\vec{a}} \frac{\varepsilon(\vec{x})}{\|\vec{x}-\vec{a}\|} = 0.$

$$\Rightarrow \lim_{\vec{x}\to\vec{a}} \varepsilon(\vec{x}) = \underbrace{\lim_{\vec{x}\to\vec{a}} \frac{\varepsilon(\vec{x})}{\|\vec{x}-\vec{a}\|}}_{0} \cdot \underbrace{\lim_{\vec{x}\to\vec{a}} \|\vec{x}-\vec{a}\|}_{0}$$

$$= 0$$

$$f(\vec{x}) = L(\vec{x}) + \varepsilon(\vec{x})$$

$$\lim_{\vec{x}\to\vec{a}} f(\vec{x}) = \lim_{\vec{x}\to\vec{a}} L(\vec{x}) + \varepsilon(\vec{x})$$

$$= \lim_{\vec{x}\to\vec{a}} L(\vec{x}) + \lim_{\vec{x}\to\vec{a}} \varepsilon(\vec{x})$$

$$= \lim_{\vec{x}\to\vec{a}} \left(f(\vec{a}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_i - a_i)\right) + 0$$

$$= f(\vec{a}).$$

$\therefore \lim_{\vec{x}\to\vec{a}} f(\vec{x}) = f(\vec{a})$ hence $f$ is continuous at $\vec{a}$.

---

<u>Rmk</u> differentiability = linear approximation

What if $f$ is linear?

$$f(\vec{x}) = c + b_1x_1 + \cdots + b_nx_n$$

$$\frac{\partial f}{\partial x_i}(\vec{x}) = b_i \qquad \forall \vec{x} \in \mathbb{R}^n.$$

$$L(\vec{x}) = f(\vec{a}) + \sum \frac{\partial f}{\partial x_i}(\vec{a})(x_i - a_i)$$

$$= f(\vec{a}) + \sum b_i (x_i - a_i)$$

$$= c + b_1x_1 + \cdots + b_nx_n = f(\vec{x})$$

$$\varepsilon(\vec{x}) = f(\vec{x}) - L(\vec{x}) = 0.$$

The linearization of $f(\vec{x})$ at any point is $L(\vec{x}) = f(\vec{x})$ itself.

**Thm** If $f, g : \Omega (\subseteq \mathbb{R}^n) \to \mathbb{R}$ are differentiable at $\vec{a} \in \Omega$, then

① $f(\vec{x}) \pm g(\vec{x})$, $cf(\vec{x})$, $f(\vec{x})g(\vec{x})$ are differentiable at $\vec{a}$.

② If $g(\vec{a}) \neq 0$, then $f(\vec{x})/g(\vec{x})$ is differentiable at $\vec{a}$.

③ Let $h(x)$ be a one-variable function. Suppose $h$ is differentiable at $f(\vec{a})$. Then $h \circ f$ is differentiable at $\vec{a}$.

**Rmk** A formula for computing differential of composition of two functions is called a chain rule. ③ is special case of chain rule. We will see this later.

(proof of Thm) Similar to one-variable case.

By the Thm, we can produce many examples of differentiable functions.

We know $\left[\begin{array}{l} \text{constant function } f(\vec{x}) = c \\ \text{Coordinate function } f(\vec{x}) = x_i \end{array}\right.$

are differentiable.

By the theorem,

- polynomials eg $f(x,y,z) = x^3z + xyz + y + 1$
- rational functions eg $\dfrac{x^3y+z}{xz+y+z}$
- If $f(\vec{x})$ is differentiable, then $e^{f(\vec{x})}$, $\sin(f(\vec{x}))$, $\cos f(\vec{x})$ etc. are differentiable.

Also $\left.\begin{array}{lll} \ln(f(\vec{x})) & \text{for} & f(\vec{x}) > 0 \\ \sqrt{f(\vec{x})} & \text{"} & \text{"} \\ |f(\vec{x})| & \text{for} & f(\vec{x}) \neq 0 \\ \ln|f(\vec{x})| & \text{"} & \text{"} \end{array}\right\}$ are all differentiable.

---

Another way to check differentiability.

**Thm** Let $\Omega \subseteq \mathbb{R}^n$ open. If $f$ is a $C^1$-function on $\Omega$, then $f$ is differentiable on $\Omega$.

eg $f: \mathbb{R}^2 \to \mathbb{R}$ s.t. $f_x$, $f_y$ exist and are continuous on a small open ball $B_\varepsilon(0,0)$

$\Rightarrow f$ is differentiable on $B_\varepsilon(0,0)$

$\Rightarrow f$ is $\sim$ at $(0,0)$.

If all $\frac{\partial f}{\partial x_i}$ can be easily checked that they are continuous, the theorem implies the differentiability of $f$.

eg $f(x,y,z) = xe^{x+y} - \log(x+z)$

Domain $f = \{(x,y,z) \in \mathbb{R}^3 \mid x+z > 0\}$

is open.

$$\left.\begin{aligned} \frac{\partial f}{\partial x} &= e^{x+y} + xe^{x+y} - \frac{1}{x+z} \\ \frac{\partial f}{\partial y} &= xe^{x+y} \\ \frac{\partial f}{\partial z} &= -\frac{1}{x+z} \end{aligned}\right\}$$

all are continuous on the domain of $f$.

$\therefore f$ is $C^1$

$\therefore f$ is differentiable.

(proof of Thm)

We will prove for 2-variable function $f(x,y)$.
Similar proof works for $n$-variables.
idea: MVT.

Suppose $(a,b) \in \Omega$.
Choose $\delta > 0$ s.t. $B_\delta(a,b) \subseteq \Omega$.
(Recall $\Omega$ is open).
For $(x,y) \in B_\delta(a,b)$,

$$f(x,y) - f(a,b)$$

$$= \underbrace{f(x,y) - f(x,b)} + \underbrace{f(x,b) - f(a,b)}$$

(MVT)

$$= f_y(x,k)(y-b) + f_x(h,b)(x-a)$$

for some $k$ between $b, y$
$h$ " $a, x$.

We need to show that

$$\lim_{\vec{x} \to (a,b)} \frac{\varepsilon(\vec{x})}{\|\vec{x} - \vec{a}\|} = 0.$$

where $L(x,y) = f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b)$

$\varepsilon(x,y) = f(x,y) - L(x,y)$

$$\left| \frac{\varepsilon(x,y)}{\|\vec{x}-\vec{a}\|} \right|$$

$$= \left| \frac{f(x,y) - f(a,b) - \underline{f_x(a,b)(x-a)} - f_y(a,b)(y-b)}{\sqrt{(x-a)^2+(y-b)^2}} \right|$$

$$= \left| \frac{f_y(x,k)(y-b) + f_x(h,b)(x-a) - f_x(a,b)(x-a) - f_y(a,b)(y-b)}{\sqrt{(x-a)^2+(y-b)^2}} \right|$$

$$= \left| \frac{(f_y(x,k) - f_y(a,b))(y-b) + (f_x(h,b) - f_x(a,b))(x-a)}{\sqrt{(x-a)^2+(y-b)^2}} \right|$$

$$\leq \left| \frac{(f_y(x,k) - f_y(a,b))(y-b)}{\sqrt{(x-a)^2+(y-b)^2}} \right| + \left| \frac{(f_x(h,b) - f_x(a,b))(x-a)}{\sqrt{(x-a)^2+(y-b)^2}} \right|$$

$\leq 1$ $\leq 1$

$$\leq |f_y(x,k) - f_y(a,b)| + |f_x(h,b) - f_x(a,b)|$$

Take $(x,y) \to (a,b)$, then

$(x,c)$, $(h,b) \to (a,b)$

By continuity of $f_x$, $f_y$, $\lim_{(x,y)\to(a,b)}$ RHS $\to 0$

By sandwich thm,

$$\lim_{(x,y)\to(a,b)} \frac{\varepsilon(x,y)}{\|(x,y)-(a,b)\|} = 0.$$

$\therefore f$ is differentiable at $(a,b)$ $\square$

---

Gradient & directional derivative

**Def** $\Omega \subseteq \mathbb{R}^n$ open. $\vec{a} \in \Omega$. $f: \Omega \to \mathbb{R}$.

The gradient vector of $f$ at $\vec{a}$ is

$$\nabla f(\vec{a}) = \left(\frac{\partial f}{\partial x_1}(\vec{a}), \cdots, \frac{\partial f}{\partial x_n}(\vec{a})\right)$$

(or $\vec{\nabla} f$)

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

$f_x = 2x+2y$. $f_y = 2x$

$\nabla f = (2x+2y, 2x)$ $\quad \nabla f(1,2) = (6, 2)$

**Rmk** Using $\nabla f$, linearization of $f$ at $\vec{a}$ can be written as

$$L(\vec{x}) = f(\vec{a}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{a})(x_i - a_i)$$

$$= f(\vec{a}) + \underbrace{\nabla f(\vec{a})}_{\left(\frac{\partial f}{\partial x_1}(\vec{a}), \cdots, \frac{\partial f}{\partial x_n}(\vec{a})\right)} \cdot \underbrace{(\vec{x} - \vec{a})}_{(x_1 - a_1, \cdots, x_n - a_n)}$$

**Def** $\Omega \subseteq \mathbb{R}^n$, open. $\vec{a} \in \Omega$. $f : \Omega \to \mathbb{R}$.

Let $\vec{u} \in \mathbb{R}^n$ be an unit vector. (i.e. $\|\vec{u}\| = 1$)

The directional derivative of $f$ in the direction $\vec{u}$ at $\vec{a}$

$$D_{\vec{u}} f(\vec{a}) = \lim_{t \to 0} \frac{f(\vec{a} + t\vec{u}) - f(\vec{a})}{t}$$

= the rate of change of $f$ in the direction of $\vec{u}$ at the point $\vec{a}$.

**Rmk** $e_i = (0, \cdots, 1, \cdots 0) \in \mathbb{R}^n$

$\qquad\qquad\qquad\quad\uparrow$ $i$-th

Then $\partial f / \partial x_i (\vec{a}) = D_{e_i} f(\vec{a})$.

**Thm** Suppose $f$ is differentiable at $\vec{a}$. $\vec{u} \in \mathbb{R}^n$ unit vector

Then $D_{\vec{u}} f(\vec{a}) = \nabla f(\vec{a}) \cdot \vec{u}$.

**eg** Let $f(x,y) = \sin^{-1}\left(\frac{x}{y}\right)$.

Find the rate of change of $f$ at $(1, \sqrt{2})$ in the direction of $\vec{v} = (1, -1)$.

(Recall if $\vec{v} (\neq 0) \in \mathbb{R}^n$, the direction $\vec{v}$ is defined to be the unit vector $\frac{\vec{v}}{\|\vec{v}\|}$ )

(sol) $\frac{\vec{v}}{\|\vec{v}\|} = \frac{1}{\sqrt{2}}(1, -1) = \vec{u}$.

We want compute $D_{\vec{u}} f(1, \sqrt{2})$.

Recall that $(\sin^{-1} z)' = \frac{1}{\sqrt{1-z^2}}$. Hence

$$\frac{\partial f}{\partial x} = \frac{1}{y} \cdot \frac{1}{\sqrt{1-\left(\frac{x}{y}\right)^2}}, \quad \frac{\partial f}{\partial y} = -\frac{x}{y^2} \cdot \frac{1}{\sqrt{1-\left(\frac{x}{y}\right)^2}}$$

Note that $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ are both continuous at $(1, \sqrt{2})$

$\therefore f$ is $C^1$ at $(1, \sqrt{2})$.

$\therefore f$ is differentiable at $(1, \sqrt{2})$.

By Thm, $D_{\vec{u}} f(1, \sqrt{2})$

$$= \nabla f(1, \sqrt{2}) \cdot \vec{u}$$

$$= \left( \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{1-\left(\frac{1}{\sqrt{2}}\right)^2}}, -\frac{1}{2} \cdot \frac{1}{\sqrt{1-\left(\frac{1}{\sqrt{2}}\right)^2}} \right)$$

$$\cdot \left( \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right)$$

$$= \left( \frac{1}{\sqrt{2}} \cdot \sqrt{2}, -\frac{1}{2} \cdot \sqrt{2} \right) \cdot \left( \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right)$$

$$= \frac{1}{\sqrt{2}} + \frac{1}{2} \quad \text{①}$$

(proof of Theorem) : differentiable $\Rightarrow D_{\vec{u}} f(\vec{a}) = \nabla f(\vec{a}) \cdot \vec{u}$.

Let $L(\vec{x})$ be the linearization of $f(\vec{x})$ at $\vec{a}$.

$$f(\vec{x}) = L(\vec{x}) + \varepsilon(\vec{x})$$

$$= f(\vec{a}) + \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a}) + \varepsilon(\vec{x}).$$

If we let $\vec{x} = \vec{a} + t\vec{u}$

$$f(\vec{a}+t\vec{u}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot (t\vec{u}) + \varepsilon(\vec{a}+t\vec{u})$$

$$D_{\vec{u}} f(\vec{a})$$

$$= \lim_{t\to 0} \frac{f(\vec{a}+t\vec{u}) - f(\vec{a})}{t}$$

$$= \lim_{t\to 0} \frac{\nabla f(\vec{a}) \cdot (t\vec{u}) + \varepsilon(\vec{a}+t\vec{u})}{t}$$

$$= \nabla f(\vec{a}) \cdot \vec{u} + \lim_{t\to 0} \frac{\varepsilon(\vec{a}+t\vec{u})}{t}$$

By differentiability of $f$ at $\vec{a}$,

$\displaystyle \lim_{\vec{x}\to\vec{a}} \frac{|\varepsilon(\vec{x})|}{\|\vec{x}-\vec{a}\|} = 0$. In particular,

$$\lim_{t\to 0} \frac{|\varepsilon(\vec{a}+t\vec{u})|}{\|\vec{a}+t\vec{u}-\vec{a}\|} = 0$$

$$= \lim_{t\to 0} \frac{|\varepsilon(\vec{a}+t\vec{u})|}{\|t\vec{u}\|} \quad (\|\vec{u}\| = 1)$$

i.e. $\displaystyle \lim_{t\to 0} \frac{|\varepsilon(\vec{a}+t\vec{u})|}{|t|} = 0.$

$$\therefore D_{\vec{u}} f(\vec{a}) = \nabla f(\vec{a}) \cdot \vec{u} + 0$$
$$= \nabla f(\vec{a}) \cdot \vec{u}. \qquad \square$$

---

## Geometric meaning of $\nabla f$.

If $f$ is differentiable at $\vec{a}$, and $\|\vec{u}\| = 1$.

$$D_{\vec{u}} f(\vec{a}) = \nabla f(\vec{a}) \cdot \vec{u}$$

By Cauchy - Schwarz Inequality,

$$|\nabla f(\vec{a}) \cdot \vec{u}| \leq \|\nabla f(\vec{a})\| \, \|\vec{u}\|$$
$$= \|\nabla f(\vec{a})\|$$

If $\nabla f(\vec{a}) \neq 0$,

$$-\|\nabla f(\vec{a})\| \leq \nabla f(\vec{a}) \cdot \vec{u} \leq \|\nabla f(\vec{a})\|$$

equality holds $\Leftrightarrow \nabla f(\vec{a}) = k\vec{u}$ for some $k<0$

equality holds $\Leftrightarrow \nabla f(\vec{a}) = k\vec{u}$ for some $k>0$

At $\vec{a}$, $f$ increases most rapidly in the direction of $\nabla f(\vec{a})$.

$f$ decreases most rapidly in the

direction of $-\nabla f(\vec{a})$.

at a rate of change $\|\nabla f(\vec{a})\|$.

Prop (properties of gradient)

$f, g : \Omega \to \mathbb{R}^n$, differentiable.

① $\nabla(f \pm g) = \nabla f \pm \nabla g$

$\nabla(cf) = c\nabla f$.

② $\nabla(fg) = g\nabla f + f\nabla g$

③ $\nabla(f/g) = \dfrac{g\nabla f - f\nabla g}{g^2}$ if $g \neq 0$.

(proof) follows easily from the properties of partial differentiations.

Rmk In the definition of $D_{\vec{u}}f$, we assumed $\vec{u}$ is a unit vector.

However, we can drop this condition.

i.e. consider any vector $\vec{v}$ of any length. we can define

$$D_{\vec{v}}f(\vec{a}) = \lim_{t \to 0} \frac{f(\vec{a}+t\vec{v}) - f(\vec{a})}{t}$$

and $D_{\vec{v}} f(\vec{a}) = Df(\vec{a}) \cdot \vec{v}$.

Note that $D_{c\vec{v}} f(\vec{a}) = c \cdot D_{\vec{v}} f(\vec{a})$

$$D_{\vec{v}} f(\vec{a}) = \begin{cases} \|\vec{v}\| \, D_{\vec{u}} f(\vec{a}) & \vec{v} \neq 0, \ \vec{u} = \dfrac{\vec{v}}{\|\vec{v}\|} \\ 0 & \text{if } \vec{v} = \vec{0}. \end{cases}$$