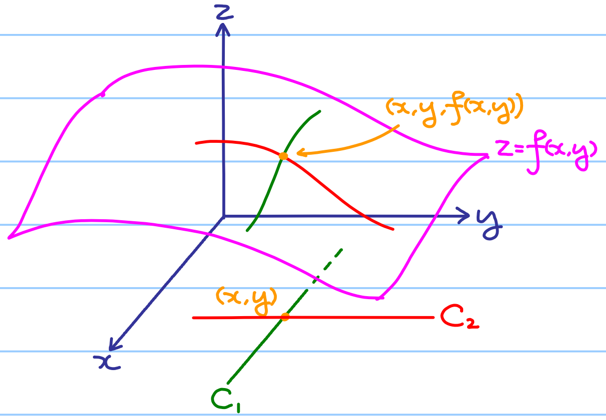
§ 8 Partial Derivatives

Partial Derivatives

Let $D \subseteq \mathbb{R}^2$ and let $f: D \to \mathbb{R}$ be a function. Suppose that $(x,y) \in D$.



Along C_1: y is fixed (regard y as a constant)

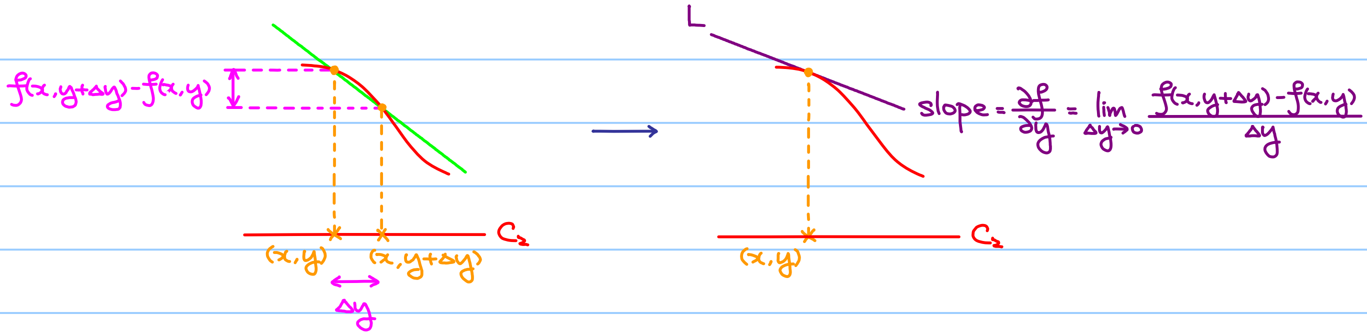
Then $f(x,y)$ becomes a function depending on x only.

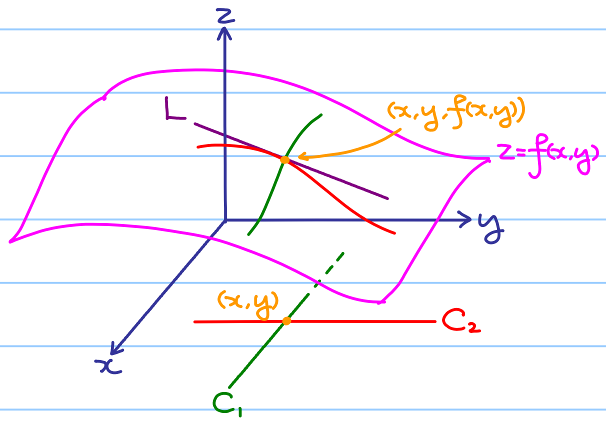
$$\frac{\partial f}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, y) - f(x,y)}{\Delta x} \quad \text{" } = \frac{df}{dx} \text{ regarding } y \text{ as constant"}$$

Along C_2: x is fixed (regard x as a constant)

Then $f(x,y)$ becomes a function depending on y only.

$$\frac{\partial f}{\partial y} = \lim_{\Delta y \to 0} \frac{f(x, y+\Delta y) - f(x,y)}{\Delta y} \quad \text{" } = \frac{df}{dy} \text{ regarding } x \text{ as constant"}$$





Definition 8.1

Let $D \subseteq \mathbb{R}^n$, $\vec{x} \in D$ and let $f: D \to \mathbb{R}$.

If $\lim\limits_{h\to 0} \dfrac{f(\vec{x}+h\vec{e}_i)-f(\vec{x})}{h} = \lim\limits_{h\to 0} \dfrac{f(x_1,\cdots,x_{i-1},x_i+h,x_{i+1},\cdots,x_n)-f(x_1,\cdots,x_{i-1},x_i,x_{i+1},\cdots,x_n)}{h}$ exists,

then the limit is said to be the partial derivative of f with respect to x_i and it is denoted by $\dfrac{\partial f}{\partial x_i}$ (or f_{x_i}, $\partial_i f$, $\nabla_i f$, $D_i f$).

Example 8.1

Let $f(x,y) = x^2+2xy+y^3$. Find $\dfrac{\partial f}{\partial x}$ at $(1,2)$.

$$\lim_{h\to 0} \frac{f(1+h,2)-f(1,2)}{h} = \lim_{h\to 0} \frac{[(1+h)^2+2(1+h)(2)+(2)^3]-[(1)^2+2(1)(2)+(2)^3]}{h}$$

$$= \lim_{h\to 0} \frac{6h+h^2}{h}$$

$$= \lim_{h\to 0} 6+h$$

$$= 6$$

However, if we treat y as a constant, then $f(x,y)$ is a polynomial of x which is a differentiable function. Therefore, we differentiate $f(x,y)$ with respect to x (keeping y as a constant), we have

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}(x^2+2xy+y^3) = 2x+2y+0 = 2x+2y$$

$$\therefore \frac{\partial f}{\partial x}(1,2) = 2(1+2) = 6 \qquad \text{(same as before)}$$

Exercise : Find $\dfrac{\partial f}{\partial y}$ at $(1,2)$ by (i) using limit definition.

(ii) direct computation. Ans : $\dfrac{\partial f}{\partial y}(1,2) = 14$

Example 8.2

Let $f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{f(0+h,0)-f(0,0)}{h} = \lim_{h\to 0} \frac{\frac{h\cdot 0}{h^2+0^2}-0}{h} = 0$$

$$\frac{\partial f}{\partial y}(0,0) = \lim_{h\to 0} \frac{f(0,0+h)-f(0,0)}{h} = \lim_{h\to 0} \frac{\frac{0\cdot h}{0^2+h^2}-0}{h} = 0$$

Proposition 8.1 (Algebraic Rules)

1) $\frac{\partial}{\partial x_i}(f+g) = \frac{\partial f}{\partial x_i} + \frac{\partial g}{\partial x_i}$

2) $\frac{\partial}{\partial x_i}(f-g) = \frac{\partial f}{\partial x_i} - \frac{\partial g}{\partial x_i}$

3) $\frac{\partial}{\partial x_i}(fg) = \frac{\partial f}{\partial x_i} \cdot g + f \frac{\partial g}{\partial x_i}$

4) $\frac{\partial}{\partial x_i}\left(\frac{f}{g}\right) = \dfrac{\frac{\partial f}{\partial x_i} \cdot g - f \cdot \frac{\partial g}{\partial x_i}}{g^2}$

Results from single variable calculus

Example 8.3

Let $f(x,y,z) = xe^y \sin(2x+yz)$.

Find all first partial derivatives of f.

$\frac{\partial f}{\partial x} = e^y \sin(2x+yz) + 2xe^y \cos(2x+yz)$

$\frac{\partial f}{\partial y} = xe^y \sin(2x+yz) + xze^y \cos(2x+yz)$

$\frac{\partial f}{\partial z} = xye^y \cos(2x+yz)$.

There is a generalization of partial derivatives:

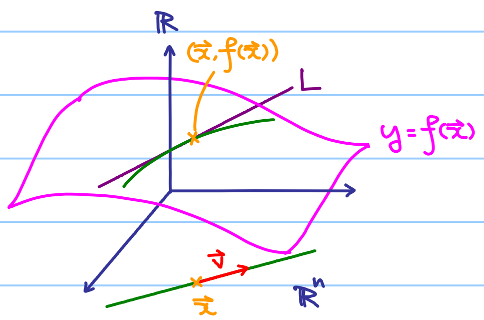
Definition 8.2

Let $D \subseteq \mathbb{R}^n$, $\vec{x} \in D$ and let $f: D \to \mathbb{R}$.

Given a nonzero vector $\vec{v} \in \mathbb{R}^n$, if $\lim_{h\to 0} \frac{f(\vec{x}+h\hat{v}) - f(\vec{x})}{h}$ exists,

where $\hat{v} = \frac{\vec{v}}{|\vec{v}|}$ is the unit vector of $\vec{v}$, then the limit is said to be

the directional derivative of f along $\vec{v}$, and it is denoted by $\nabla_{\vec{v}} f(\vec{x})$ or $D_{\vec{v}} f(\vec{x})$.



$\nabla_{\vec{v}} f(\vec{x}) = \text{slope of } L$

In particular, $\nabla_{\vec{e}_i} f(\vec{x}) = \frac{\partial f}{\partial x_i}$.

Example 8.4

Let $f(x,y) = xy$ and $\vec{v} = (1,2) \in \mathbb{R}^2$. Then $\hat{v} = \frac{\vec{v}}{|\vec{v}|} = \frac{1}{\sqrt{5}}(1,2)$.

$$\lim_{h\to 0} \frac{f(\vec{x}+h\hat{v}) - f(\vec{x})}{h} = \lim_{h\to 0} \frac{f(x+\frac{h}{\sqrt{5}}, y+\frac{2h}{\sqrt{5}}) - f(x,y)}{h}$$

$$= \lim_{h\to 0} \frac{(x+\frac{h}{\sqrt{5}})(y+\frac{2h}{\sqrt{5}}) - xy}{h}$$

$$= \lim_{h\to 0} \frac{2}{\sqrt{5}}xh + \frac{1}{\sqrt{5}}yh + \frac{2}{5}h^2$$

$$= \frac{2}{\sqrt{5}}x + \frac{1}{\sqrt{5}}y$$

$\therefore \nabla_{\hat{v}} f(\vec{x}) = \frac{2}{\sqrt{5}}x + \frac{1}{\sqrt{5}}y$

Higher Partial Derivatives

Example 8.5

Let $f(x,y) = x^3 + 2xy^2 + y^4$. Then, $\frac{\partial f}{\partial x} = 3x^2 + 2y^2$, $\frac{\partial f}{\partial y} = 4xy + 4y^3$.

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = 6x \qquad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = 4y$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = 4x + 12y^2 \qquad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = 4y$$

Are they always the same?

No, but it is true when f is "nice".

Example 8.6

$$\text{Let } f(x,y) = \begin{cases} \frac{xy(x^2-y^2)}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

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

$\therefore \frac{\partial f}{\partial x}(0,0) = 0$ $\qquad$ $\therefore \frac{\partial f}{\partial y}(0,0) = 0$

$$\frac{\partial f}{\partial x}(x,y) = \begin{cases} \frac{y(x^4+4x^2y^2-y^4)}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases} \qquad \frac{\partial f}{\partial y}(x,y) = \begin{cases} \frac{x(x^4-4x^2y^2-y^4)}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

$$\lim_{h\to 0} \frac{\frac{\partial f}{\partial x}(0,0+h) - \frac{\partial f}{\partial x}(0,0)}{h} = \lim_{h\to 0} -1 = -1 \qquad \lim_{h\to 0} \frac{\frac{\partial f}{\partial y}(0+h,0) - \frac{\partial f}{\partial y}(0,0)}{h} = \lim_{h\to 0} 1 = 1$$

$\therefore \frac{\partial^2 f}{\partial y \partial x}(0,0) = -1$ $\qquad$ $\therefore \frac{\partial^2 f}{\partial x \partial y}(0,0) = 1$

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) \neq \frac{\partial^2 f}{\partial x \partial y}(0,0)$$

Let $(a,b) \in \mathbb{R}^2$. Then,

$$\frac{\partial f}{\partial x}(a,y) = \lim_{h \to 0} \frac{f(a+h,y)-f(a,y)}{h} \qquad \frac{\partial f}{\partial y}(x,b) = \lim_{k \to 0} \frac{f(x,b+k)-f(x,b)}{k}$$

$$\frac{\partial^2 f}{\partial y \partial x}(a,b) = \lim_{k \to 0} \frac{\frac{\partial f}{\partial x}(a,b+k) - \frac{\partial f}{\partial x}(a,b)}{k} = \lim_{k \to 0} \lim_{h \to 0} \frac{[f(a+h,b+k)-f(a,b+k)]-[f(a+h,b)-f(a,b)]}{hk} \quad ①$$

$$\frac{\partial^2 f}{\partial x \partial y}(a,b) = \lim_{h \to 0} \frac{\frac{\partial f}{\partial y}(a+h,b) - \frac{\partial f}{\partial y}(a,b)}{h} = \lim_{h \to 0} \lim_{k \to 0} \frac{[f(a+h,b+k)-f(a+h,b)]-[f(a,b+k)-f(a,b)]}{hk} \quad ②$$

Main issue: ① and ② are not the same in general.

Definition 8.3

Let $D \subseteq \mathbb{R}^n$ be an open subset and let $f: D \to \mathbb{R}$. Let $r \geq 0$.

f is said to be a C^r function if all partial derivatives of f exist up to r-th order and are continuous on D.

In particular, a C^0 function is a continuous function;

a C^∞ function is a function such that partial derivatives exist for any order, which is also called a smooth function.

Theorem 8.1 (Mixed Derivative Theorem / Clairaut's Theorem)

Let $D \subseteq \mathbb{R}^n$ be an open subset and let $f: D \to \mathbb{R}$.

If f is a C^2 function on D, i.e. all second partial derivatives are continuous on D,

then for all $1 \le i, j \le n$, we have $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$.

proof (for $n=2$):

Let $(a,b) \in D$

$$\frac{\partial^2 f}{\partial x \partial y}(a,b) = \lim_{h \to 0} \lim_{k \to 0} \left(\frac{[f(a+h,b+k) - f(a+h,b)] - [f(a,b+k) - f(a,b)]}{k} \right) \cdot \frac{1}{h}$$

Apply MVT to the function $f(x,b+k) - f(x,b)$ on $[a,a+h]$ or $[a+h,a]$

there exists $c_1 \in (a,a+h)$ or $(a+h,a)$ such that

$$\frac{[f(a+h,b+k) - f(a+h,b)] - [f(a,b+k) - f(a,b)]}{k} = \frac{\partial f}{\partial x}(c_1,b+k) - \frac{\partial f}{\partial x}(c_1,b)$$

$$= \lim_{h \to 0} \lim_{k \to 0} \frac{\frac{\partial f}{\partial x}(c_1,b+k) - \frac{\partial f}{\partial x}(c_1,b)}{h}$$

Apply MVT to the function $\frac{\partial f}{\partial x}(c_1,y)$ on $[b,b+k]$ or $[b+k,b]$

there exists $c_2 \in (b,b+k)$ or $(b+k,b)$ such that $\frac{\frac{\partial f}{\partial x}(c_1,b+k) - \frac{\partial f}{\partial x}(c_1,b)}{h} = \frac{\partial^2 f}{\partial y \partial x}(c_1,c_2)$

$$= \lim_{h \to 0} \lim_{k \to 0} \frac{\partial^2 f}{\partial y \partial x}(c_1,c_2)$$

$= \lim_{h \to 0} \frac{\partial^2 f}{\partial y \partial x}(c_1,b)$ (As $k \to 0$, we have $c_2 \to b$. The equality hold by the continuity of $\frac{\partial^2 f}{\partial y \partial x}$.)

$= \frac{\partial^2 f}{\partial y \partial x}(a,b)$ (As $h \to 0$, we have $c_1 \to a$. The equality hold by the continuity of $\frac{\partial^2 f}{\partial y \partial x}$.)

Theorem 8.2 (Generalization of Mixed Derivative Theorem)

Let $D \subseteq \mathbb{R}^n$ be an open subset and let $f: D \to \mathbb{R}$ be a C^r function on D.

Then the order of differentiation does not matter for all partial derivatives up to r-th order.

§ 9 Differentiability

Definition 9.1

Let $D \subseteq \mathbb{R}^n$, let $\vec{x}_0 \in D$ and let $f : D \to \mathbb{R}$.

Assume that all first partial derivatives of f at $\vec{x}_0$ exist.

$\nabla f(\vec{x}_0) := \left(\frac{\partial f}{\partial x_1}(\vec{x}_0), \frac{\partial f}{\partial x_2}(\vec{x}_0), \dots, \frac{\partial f}{\partial x_n}(\vec{x}_0) \right)$ is said to be the gradient vector of f at $\vec{x}_0$.

Exercise 9.1

Let $D \subseteq \mathbb{R}^n$, let $\vec{x}_0 \in D$ and let $f, g : D \to \mathbb{R}$.

Assume that all first partial derivatives of f and g at $\vec{x}_0$ exist.

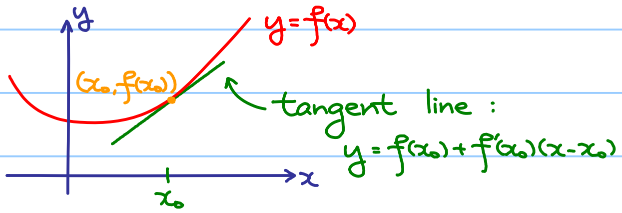
(1) $\nabla(f \pm g)(\vec{x}_0) = \nabla f(\vec{x}_0) \pm \nabla g(\vec{x}_0)$;

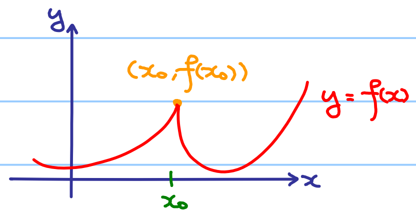
(2) $\nabla(f \cdot g)(\vec{x}_0) = g(\vec{x}_0) \nabla f(\vec{x}_0) + f(\vec{x}_0) \nabla g(\vec{x}_0)$;

(3) $\nabla\left(\frac{f}{g}\right)(\vec{x}_0) = \frac{g(\vec{x}_0) \nabla f(\vec{x}_0) + f(\vec{x}_0) \nabla g(\vec{x}_0)}{[g(\vec{x}_0)]^2}$.

Question: Is the condition that all first partial derivatives at $\vec{x} = \vec{x}_0$ exist, i.e. $\nabla f(\vec{x}_0)$ exists, a good generalization of differentiability for function $f : \mathbb{R}^n \to \mathbb{R}$ at $\vec{x} = \vec{x}_0$?

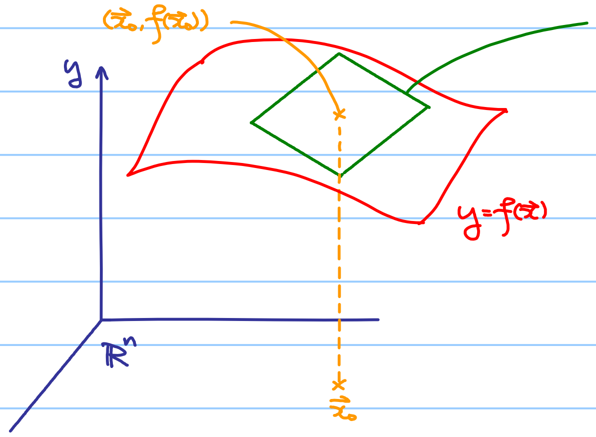
Idea: Differentiability of $y = f(x)$ at $x = x_0$ $\Leftrightarrow$ Construct a tangent line at $x = x_0$.





$f(x)$ is differentiable at $x = x_0$ $\qquad$ $f(x)$ is not differentiable at $x = x_0$

Differentiability of $y = f(\vec{x})$ at $\vec{x} = \vec{x}_0$ $\Leftrightarrow$ Construct a tangent plane at $\vec{x} = \vec{x}_0$



affine hyperplane which is tangent to the graph $y = f(\vec{x})$ at $\vec{x} = \vec{x}_0$.

Recall. Let $D \subseteq \mathbb{R}$ be an open subset, let $x_0 \in D$ and let $f: D \to \mathbb{R}$.

If f is differentiable at $x = x_0$, then f is continuous at $x = x_0$.

Question. In general, let $D \subseteq \mathbb{R}^n$ be an open subset, let $\vec{x}_0 \in D$ and let $f: D \to \mathbb{R}$.

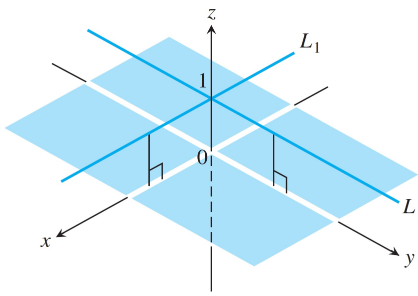
If all first partial derivatives of f at $\vec{x}_0$ exist,

does it imply f is continuous at $\vec{x}_0$?

Answer. Unfortunately, no!

Example 9.1

Let $f(x,y) = \begin{cases} 0 & \text{if } xy \neq 0 \quad \text{(i.e. both } x, y \neq 0) \\ 1 & \text{if } xy = 0 \quad \text{(i.e. } x = 0 \text{ or } y = 0) \end{cases}$



The graph of $z = f(x,y)$

$\lim_{h \to 0} \frac{f(0+h,0) - f(0,0)}{h} = \lim_{h \to 0} \frac{1-1}{h} = \lim_{h \to 0} 0 = 0 \quad \therefore \frac{\partial f}{\partial x}(0,0) = 0$

$\lim_{h \to 0} \frac{f(0,0+h) - f(0,0)}{h} = \lim_{h \to 0} \frac{1-1}{h} = \lim_{h \to 0} 0 = 0 \quad \therefore \frac{\partial f}{\partial y}(0,0) = 0$

All first partial derivatives of f at $(0,0)$ exist.

However, Let $\gamma_1(t) = (t,0)$ and $\gamma_2(t) = (t,t)$. Then, $\gamma_1(0) = \gamma_2(0) = (0,0)$, but

$\lim_{t \to 0} (f \circ \gamma_1)(t) = \lim_{t \to 0} f(\gamma_1(t)) = \lim_{t \to 0} f(t,0) = \lim_{t \to 0} 1 = 1$

$\lim_{t \to 0} (f \circ \gamma_2)(t) = \lim_{t \to 0} f(\gamma_2(t)) = \lim_{t \to 0} f(t,t) = \lim_{t \to 0} 0 = 0$

$\therefore \lim_{t \to 0} (f \circ \gamma_1)(t) \neq \lim_{t \to 0} (f \circ \gamma_2)(t)$ and $\lim_{(x,y) \to (0,0)} f(x,y)$ does not exist.

Hence, $f(x,y)$ is NOT continuous at $(0,0)$.

Explanation: $\frac{\partial f}{\partial x}(0,0)$ and $\frac{\partial f}{\partial y}(0,0)$ exist means that the function is "smooth" along the x-direction and y-direction. However, other directions are also considering for computing $\lim_{(x,y) \to (0,0)} f(x,y)$.

Question: What is the correct generalization of differentability?

Definition 9.2

Let $D \subseteq \mathbb{R}$ be an open subset, let $x_0 \in D$ and let $f: D \to \mathbb{R}$.

f is said to be differentiable at $x = x_0$ if $\lim_{h \to 0} \frac{f(x_0+h) - f(x_0)}{h}$ exists

(if it exists, denote it by $f'(x_0)$)

Rephrase : There exists $L \in \mathbb{R}$ such that $\lim_{h \to 0} \frac{f(x_0+h) - f(x_0)}{h} = L$ i.e. $\lim_{h \to 0} \frac{f(x_0+h) - (f(x_0) + Lh)}{h} = 0$.

Furthermore, if such a $L \in \mathbb{R}$ exists, $L = f'(x_0)$.

Idea: Let $\varepsilon(h) = f(x_0+h) - (f(x_0) + f'(x_0)h)$.

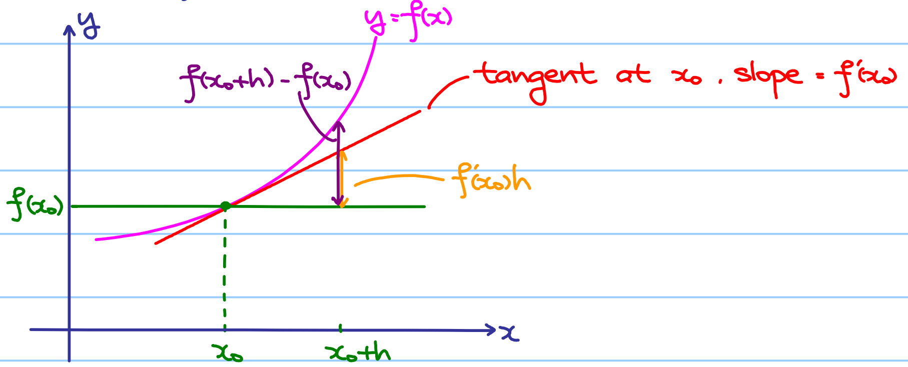
$\lim_{h \to 0} \frac{f(x_0+h) - (f(x_0) + f'(x_0)h)}{h} = \lim_{h \to 0} \frac{\varepsilon(h)}{h} = 0$ roughly means that when h is small,

$\varepsilon(h)$ is even smaller comparing to h.

Interpretation 1 : If we write $\underbrace{f(x_0+h) - f(x_0)}_{\text{change of } f} \sim f'(x_0)\underbrace{h}_{\text{change of } x}$

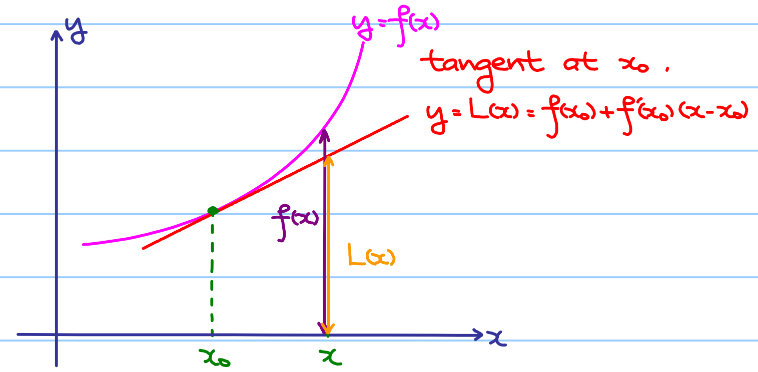
it means change of f can be approximated by a linear transformation on change of x

and $\varepsilon(h) = f(x_0+h) - (f(x_0) + f'(x_0)h)$ is the error of the approximation.



Interpretation 2 : If we let $x = x_0 + h$ (i.e. $h = x - x_0$) and we write $f(x) \sim \underbrace{f(x_0) + f'(x_0)(x - x_0)}_{\text{linear function } L(x)}$

it means $f(x)$ can be approximated by a linear function $L(x)$, called linearization of f at x_0



Example 9.2

What is the value of $\sqrt{101}$?

Let $f:[0,\infty)\to\mathbb{R}$ defined by $f(x)=\sqrt{x}$. Then $f'(x)=\frac{1}{2\sqrt{x}}$

(Put $x_0=100$, $x=101$, so $h=x-x_0=1$)

$$f(x_0+h)-f(x_0)\sim f'(x_0)h \qquad\qquad f(x)\sim L(x)=f(x_0)+f'(x_0)(x-x_0)$$

$$f(101)-f(100)\sim f'(100)\cdot 1 \qquad\qquad f(101)\sim L(101)=f(100)+f'(100)\cdot(101-100)$$

$$\sqrt{101}-10\sim\frac{1}{20}\cdot 1 \qquad\qquad =10+\frac{1}{20}\cdot 1$$

$$\sqrt{101}\sim 10.05 \qquad\qquad =10.05$$

Definition 9.2 (Restated)

Let $D\subseteq\mathbb{R}$ be an open subset, let $x_0\in D$ and let $f:D\to\mathbb{R}$.

f is said to be differentiable at $x=x_0$ if

there exists $L\in\mathbb{R}$ such that $\lim\limits_{h\to 0}\frac{f(x_0+h)-(f(x_0)+Lh)}{h}=0$.

Furthermore, if such a $L\in\mathbb{R}$ exists, $L=f'(x_0)$.

With the above, we are ready to generalize :

Differentiability of Real Valued Functions

Definition 9.3

Let $D\subseteq\mathbb{R}^n$ be an open subset, let $\vec{x}_0\in D$ and let $f:D\to\mathbb{R}$.

f is said to be differentiable at $\vec{x}_0$ if there exists $\vec{L}\in\mathbb{R}^n$ such that

$$\lim_{\vec{h}\to\vec{0}}\frac{f(\vec{x}_0+\vec{h})-(f(\vec{x}_0)+\vec{L}\cdot\vec{h})}{|\vec{h}|}=0$$

Furthermore, f is said to be a differentiable function if it is differentiable at every $x\in D$.

Remark: Note that $\lim\limits_{\vec{h}\to\vec{0}}\frac{f(\vec{x}_0+\vec{h})-(f(\vec{x}_0)+\vec{L}\cdot\vec{h})}{|\vec{h}|}=0 \Leftrightarrow \lim\limits_{\vec{h}\to\vec{0}}\frac{|f(\vec{x}_0+\vec{h})-(f(\vec{x}_0)+\vec{L}\cdot\vec{h})|}{|\vec{h}|}=0$

Therefore, some textbooks use the latter one in the definition of differentiability.

Idea: Now, change of $\vec{x}$ may not be a real number but a vector $\vec{h}\in\mathbb{R}^n$.

Therefore, change of $f=f(\vec{x}_0+\vec{h})-f(\vec{x}_0)\in\mathbb{R}$ should be approximated by a linear transformation on change of $\vec{x}=\vec{h}$, which is in form of $\vec{L}\cdot\vec{h}$.

Think: But, what is $\vec{L}$?

If there exists $\vec{L}=(L_1,L_2,\cdots,L_n)\in\mathbb{R}^n$ such that $\lim_{\vec{h}\to\vec{0}}\frac{f(\vec{x}_0+\vec{h})-(f(\vec{x}_0)+\vec{L}\cdot\vec{h})}{|\vec{h}|}=0$.

In particular, consider $\vec{h}=h\vec{e}_i=(0,\cdots,0,\underset{\substack{\uparrow\\ i\text{-th}}}{h},0,\cdots,0)$.

$$\lim_{\vec{h}\to\vec{0}}\frac{f(\vec{x}_0+\vec{h})-(f(\vec{x}_0)+\vec{L}\cdot\vec{h})}{|\vec{h}|}=0$$

$$\lim_{h\to 0}\frac{f(\vec{x}_0+h\vec{e}_i)-f(\vec{x}_0)-L_ih}{|h|}=0$$

$$\Rightarrow \lim_{h\to 0}\frac{f(\vec{x}_0+h\vec{e}_i)-f(\vec{x}_0)-L_ih}{h}=0$$ (Why? Hint: Consider $h\to 0^+$ and $h\to 0^-$ separately.)

$$\lim_{h\to 0}\frac{f(\vec{x}_0+h\vec{e}_i)-f(\vec{x}_0)}{h}=L_i$$

$$L_i=\frac{\partial f}{\partial x_i}(\vec{x}_0).$$

$\therefore\ \vec{L}=\nabla f(\vec{x}_0)$

Theorem 9.1

Let $D\subseteq\mathbb{R}^n$ be an open subset, let $\vec{x}_0\in D$ and let $f:D\to\mathbb{R}$.

If f is differentiable at $\vec{x}_0$, then all first partial derivatives of f at $\vec{x}_0$ exist.

Also, $\nabla f(\vec{x}_0)=(\frac{\partial f}{\partial x_1}(\vec{x}_0),\frac{\partial f}{\partial x_2}(\vec{x}_0),\cdots,\frac{\partial f}{\partial x_n}(\vec{x}_0))\in\mathbb{R}^n$ which is called gradient vector, is the unique vector such that $\lim_{\vec{h}\to\vec{0}}\frac{f(\vec{x}_0+\vec{h})-(f(\vec{x}_0)+\nabla f(\vec{x}_0)\cdot\vec{h})}{|\vec{h}|}=0$.

Remark:

1) f is differentiable at $\vec{x}_0$, then all first partial derivatives of f at $\vec{x}_0$ exist.

2) Given a function f, to prove it is differentiable at $\vec{x}_0$,

how do we find $\vec{L}\in\mathbb{R}^n$ such that $\lim_{\vec{h}\to\vec{0}}\frac{f(\vec{x}_0+\vec{h})-(f(\vec{x}_0)+\vec{L}\cdot\vec{h})}{|\vec{h}|}=0$?

The above theorem says $\nabla f(\vec{x}_0)$ is the only possible candidate of $\vec{L}$.

Definition 9.3 (Restated)

Let $D\subseteq\mathbb{R}^n$ be an open subset, let $\vec{x}_0\in D$ and let $f:D\to\mathbb{R}$.

Suppose that all first partial derivatives of f at $\vec{x}_0$ exist.

Let $\varepsilon(\vec{h})=f(\vec{x}_0+\vec{h})-(f(\vec{x}_0)+\nabla f(\vec{x}_0)\cdot\vec{h})$, where $\nabla f(\vec{x}_0)=(\frac{\partial f}{\partial x_1}(\vec{x}_0),\frac{\partial f}{\partial x_2}(\vec{x}_0),\cdots,\frac{\partial f}{\partial x_n}(\vec{x}_0))\in\mathbb{R}^n$.

f is said to be differentiable at $\vec{x}_0$ if

$$\lim_{\vec{h}\to\vec{0}}\frac{f(\vec{x}_0+\vec{h})-(f(\vec{x}_0)+\nabla f(\vec{x}_0)\cdot\vec{h})}{|\vec{h}|}=\lim_{\vec{h}\to\vec{0}}\frac{\varepsilon(\vec{h})}{|\vec{h}|}=0.$$

Idea: Let $\varepsilon(\vec{h}) = f(\vec{x}_0+\vec{h}) - (f(\vec{x}_0) + \nabla f(\vec{x}_0)\cdot\vec{h})$.

$\lim_{\vec{h}\to\vec{0}} \frac{f(\vec{x}_0+\vec{h}) - (f(\vec{x}_0) + \nabla f(\vec{x}_0)\cdot\vec{h})}{|\vec{h}|} = \lim_{\vec{h}\to\vec{0}} \frac{\varepsilon(\vec{h})}{|\vec{h}|} = 0$ roughly means when $\vec{h}$ tends to $\vec{0}$,

$\varepsilon(\vec{h}) = f(\vec{x}_0+\vec{h}) - (f(\vec{x}_0) + \nabla f(\vec{x}_0)\cdot\vec{h})$ is even smaller comparing to $|\vec{h}|$.

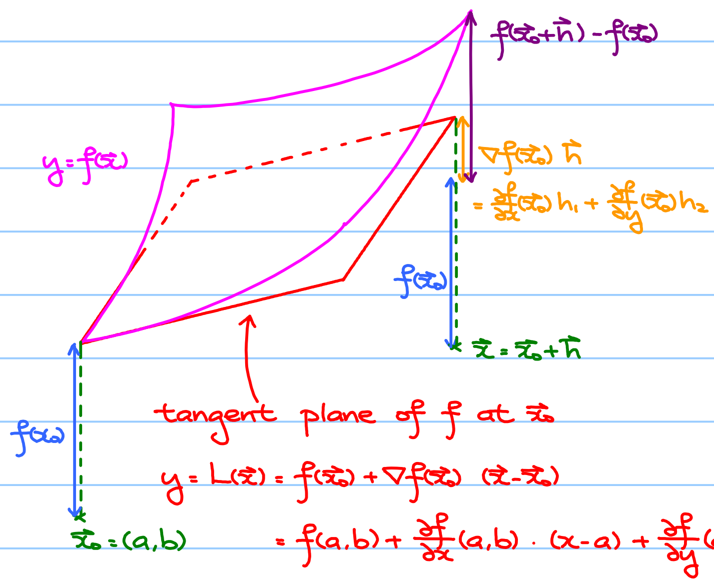
Interpretation 1: If we write $\underbrace{f(\vec{x}_0+\vec{h}) - f(\vec{x}_0)}_{\text{change of } f} \sim \nabla f(\vec{x}_0)\cdot \underbrace{\vec{h}}_{\text{change of } \vec{x}}$

it means change of f can be approximated by a linear transformation on change of $\vec{x}$ and $\varepsilon(\vec{h}) = f(\vec{x}_0+\vec{h}) - (f(\vec{x}_0) + \nabla f(\vec{x}_0)\cdot\vec{h})$ is the error of the approximation.

Interpretation 2: If we let $\vec{x} = \vec{x}_0 + \vec{h}$ (i.e. $\vec{h} = \vec{x} - \vec{x}_0$) and we write $f(\vec{x}) \sim \underbrace{f(\vec{x}_0) + \nabla f(\vec{x}_0)\cdot(\vec{x}-\vec{x}_0)}_{\text{linear function } L(\vec{x})}$

it means $f(\vec{x})$ can be approximated by a linear function $L(\vec{x})$, called linearization of f at $\vec{x}_0$

$n=2$ case:



tangent plane of f at $\vec{x}_0$

$y = L(\vec{x}) = f(\vec{x}_0) + \nabla f(\vec{x}_0)\ (\vec{x}-\vec{x}_0)$

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

Example 9.3

Let $f(x,y) = e^{\sin(x+2y)}$ How to find $f(0.05, 0.1)$?

At least, we can approximate it by the following: Let $\vec{x}_0 = (0,0)$, $\vec{x} = (0.05, 0.1)$,

then $f(0.05,0.1) = f(\vec{x}) \approx L(\vec{x}) = f(\vec{x}_0) + \nabla f(\vec{x}_0)\cdot(\vec{x}-\vec{x}_0) = 1 + (1,2)\cdot(0.05, 0.1) = 1.25$

Remark:

1) How good is the above approximation? (Later!)

2) In general, we expect the linearization $L(\vec{x})$ of f at $\vec{x}_0$ is a "good" approximation "near" $\vec{x}_0$

Example 94

Let $f(x,y) = xy$. Show that f is differentiable at $(1,2)$

Furthermore, find the tangent plane of f at $(1,2)$

Step 1: Find $\frac{\partial f}{\partial x}(1,2)$ and $\frac{\partial f}{\partial y}(1,2)$.

$$\frac{\partial f}{\partial x}(1,2) = 2 \qquad \frac{\partial f}{\partial y}(1,2) = 1$$

$$\therefore \nabla f(1,2) = \left(\frac{\partial f}{\partial x}(1,2), \frac{\partial f}{\partial y}(1,2)\right) = (2,1)$$

Step 2: Check the definition.

$$\lim_{\vec{h}\to\vec{0}} \frac{f(\vec{x}_0+\vec{h}) - (f(\vec{x}_0) + \nabla f(\vec{x}_0)\cdot\vec{h})}{|h|} \qquad \text{where } \vec{x}_0 = (1,2),\ \vec{h} = (h_1,h_2)$$

$$= \lim_{(h_1,h_2)\to(0,0)} \frac{f(1+h_1, 2+h_2) - [f(1,2) + (2h_1+h_2)]}{\sqrt{h_1^2+h_2^2}}$$

$$= \lim_{(h_1,h_2)\to(0,0)} \frac{(1+h_1)(2+h_2) - (2+2h_1+h_2)}{\sqrt{h_1^2+h_2^2}}$$

$$= \lim_{(h_1,h_2)\to(0,0)} \frac{h_1h_2}{\sqrt{h_1^2+h_2^2}}$$

Note: $-\frac{|h_1h_2|}{\sqrt{h_1^2+h_2^2}} \le \frac{h_1h_2}{\sqrt{h_1^2+h_2^2}} \le \frac{|h_1h_2|}{\sqrt{h_1^2+h_2^2}}$

$$\lim_{(h_1,h_2)\to(0,0)} \frac{|h_1h_2|}{\sqrt{h_1^2+h_2^2}} = \lim_{(h_1,h_2)\to(0,0)} \frac{1}{\sqrt{\frac{1}{h_2^2}+\frac{1}{h_1^2}}} = 0 \qquad \text{so} \qquad \lim_{(h_1,h_2)\to(0,0)} -\frac{|h_1h_2|}{\sqrt{h_1^2+h_2^2}} = 0$$

By sandwich theorem, $\lim_{(h_1,h_2)\to(0,0)} \frac{h_1h_2}{\sqrt{h_1^2+h_2^2}} = 0$

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

(You may also use polar coordinates to deduce the limit to be 0.)

Furthermore, the tangent plane of f at $(1,2)$ is

$$z = f(1,2) + \nabla f(1,2) \quad [(x,y) - (1,2)]$$

$$= f(1,2) + \frac{\partial f}{\partial x}(1,2)\cdot(x-1) + \frac{\partial f}{\partial y}(1,2)\cdot(y-2)$$

$$= 2 + 2(x-1) + (y-2)$$

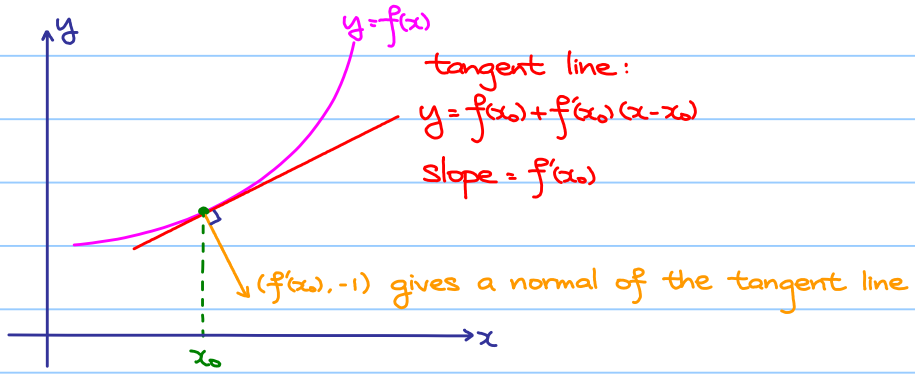
$$2x + y - z = 2$$

Remark. $\left(\frac{\partial f}{\partial x}(1,2), \frac{\partial f}{\partial y}(1,2), -1\right)$ gives a normal of the tangent plane.

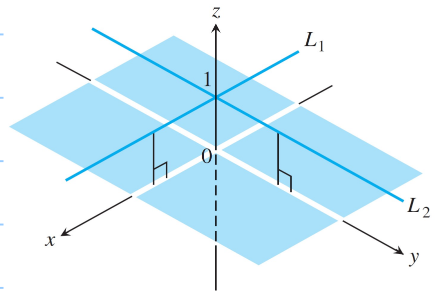
Exercise 9.3

Let $D \subseteq \mathbb{R}^n$ be an open subset, let $\vec{x}_0 \in D$ and let $f: D \to \mathbb{R}$ be a differentiable function.
Show that $\left(\frac{\partial f}{\partial x_1}(\vec{x}_0), \frac{\partial f}{\partial x_2}(\vec{x}_0), \dots, \frac{\partial f}{\partial x_n}(\vec{x}_0), -1\right) \in \mathbb{R}^{n+1}$ gives a normal of the tangent plane of f at $\vec{x}_0$.

In particular, when $n=1$,



Example 9.5



Let $f(x,y) = \begin{cases} 0 & \text{if } xy \neq 0 \quad (\text{i.e. both } x,y \neq 0) \\ 1 & \text{if } xy \neq 0 \quad (\text{i.e. } x=0 \text{ or } y=0) \end{cases}$

From example 9.1, we have $\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0$

$$\lim_{\vec{h} \to \vec{0}} \frac{f(\vec{x}_0+\vec{h}) - (f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot \vec{h})}{|\vec{h}|} \qquad \text{where } \vec{x}_0 = (0,0),\ \vec{h} = (h_1, h_2)$$

$$= \lim_{(h_1,h_2)\to(0,0)} \frac{f(0+h_1, 0+h_2) - [f(0,0) + (0h_1 + 0h_2)]}{\sqrt{h_1^2+h_2^2}}$$

$$= \lim_{(h_1,h_2)\to(0,0)} \frac{f(h_1,h_2) - 1}{\sqrt{h_1^2+h_2^2}}$$

Exercise: By considering the curve $(h_1(t), h_2(t)) = (t,t)$, show the above limit does not exist.

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

f is differentiable at $\vec{x}_0 \Rightarrow \frac{\partial f}{\partial x_i}(\vec{x}_0)$ exists for $i=1,2,\cdots,n$

However, example 9.5 shows that the converse is not true. Fortunately, we have:

Theorem 9.2

Let $D \subseteq \mathbb{R}^n$ be an open subset and let $f: D \to \mathbb{R}$.

If f is a C^1 function on D, then f is differentiable on D.

Remark: Let $\vec{x}_0 \in D$. Since D is open, $\vec{x}_0$ is an interior point of D This theorem says if all partial derivatives "behave well" around $\vec{x}_0$, then differentiability of f at $\vec{x}_0$ can be guaranteed.

proof (for $n=2$):

Let $(a,b) \in D$. By assumption, $\frac{\partial f}{\partial x}(a,b)$ and $\frac{\partial f}{\partial y}(a,b)$ exists.

$$0 \leq \left| \frac{f(a+h_1, b+h_2) - f(a,b) - \frac{\partial f}{\partial x}(a,b)\, h_1 - \frac{\partial f}{\partial y}(a,b)\, h_2}{\sqrt{h_1^2+h_2^2}} \right|$$

$$= \left| \frac{[f(a+h_1,b) - f(a,b) - \frac{\partial f}{\partial x}(a,b)\cdot h_1] + [f(a+h_1,b+h_2) - f(a+h_1,b) - \frac{\partial f}{\partial y}(a,b)\cdot h_2]}{\sqrt{h_1^2+h_2^2}} \right|$$

Apply MVT,

there exist $c_1 \in (a, a+h_1)$ or $(a+h_1, a)$ such that $f(a+h_1,b) - f(a,b) = h_1 \frac{\partial f}{\partial x}(c_1,b)$

there exist $c_2 \in (b, b+h_2)$ or $(b+h_2, b)$ such that $f(a+h_1,b+h_2) - f(a+h_1,b) = h_2 \frac{\partial f}{\partial y}(a+h_1,c_2)$

$$= \left| \frac{(\frac{\partial f}{\partial x}(c_1,b) - \frac{\partial f}{\partial x}(a,b))\cdot h_1 + (\frac{\partial f}{\partial y}(a+h_1,c_2) - \frac{\partial f}{\partial y}(a,b))\cdot h_2}{\sqrt{h_1^2+h_2^2}} \right|$$

$$\leq \left| \frac{\partial f}{\partial x}(c_1,b) - \frac{\partial f}{\partial x}(a,b) \right| \frac{|h_1|}{\sqrt{h_1^2+h_2^2}} + \left| \frac{\partial f}{\partial y}(a+h_1,c_2) - \frac{\partial f}{\partial y}(a,b) \right| \frac{|h_2|}{\sqrt{h_1^2+h_2^2}} \quad (\because \text{triangle inequality})$$

$$\leq \left| \frac{\partial f}{\partial x}(c_1,b) - \frac{\partial f}{\partial x}(a,b) \right| + \left| \frac{\partial f}{\partial y}(a+h_1,c_2) - \frac{\partial f}{\partial y}(a,b) \right|$$

As $(h_1,h_2) \to (0,0)$, we have $c_1 \to a$ and $c_2 \to b$ By the continuity of $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ at (a,b),

we have $\lim_{(h_1,h_2)\to(0,0)} \left| \frac{\partial f}{\partial x}(c_1,b) - \frac{\partial f}{\partial x}(a,b) \right| + \left| \frac{\partial f}{\partial y}(a+h_1,c_2) - \frac{\partial f}{\partial y}(a,b) \right|$

By sandwich theorem, $\lim_{(h_1,h_2)\to(0,0)} \left| \frac{f(a+h_1,b+h_2) - f(a,b) - \frac{\partial f}{\partial x}(a,b)\, h_1 - \frac{\partial f}{\partial y}(a,b)\cdot h_2}{\sqrt{h_1^2+h_2^2}} \right| = 0$

It implies $\lim_{(h_1,h_2)\to(0,0)} \frac{f(a+h_1,b+h_2) - f(a,b) - \frac{\partial f}{\partial x}(a,b)\, h_1 - \frac{\partial f}{\partial y}(a,b)\cdot h_2}{\sqrt{h_1^2+h_2^2}} = 0$

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

Example 9.6

Let $f(x,y) = \begin{cases} (x^2+y^2)^2 \sin\left(\frac{1}{\sqrt{x^2+y^2}}\right) & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$

$$\lim_{h\to 0} \frac{f(0+h,0)-f(0,0)}{h} = \lim_{h\to 0} \frac{(h^2+0^2)^2 \sin\left(\frac{1}{\sqrt{0^2+h^2}}\right) - 0}{h} = \lim_{h\to 0} h^3 \sin\left(\frac{1}{\sqrt{h^2}}\right) \overset{\substack{\text{Sandwich theorem} \\ \downarrow}}{=} 0 \qquad \therefore \frac{\partial f}{\partial x}(0,0) = 0.$$

$$\frac{\partial f}{\partial x}(x,y) = \begin{cases} 2x(x^2+y^2)\sin\left(\frac{1}{\sqrt{x^2+y^2}}\right) - x\sqrt{x^2+y^2}\cos\left(\frac{1}{\sqrt{x^2+y^2}}\right) & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

$$\lim_{(x,y)\to(0,0)} \frac{\partial f}{\partial x}(x,y) \overset{\substack{\text{Sandwich theorem} \\ \downarrow}}{=} 0 = \frac{\partial f}{\partial x}(0,0)$$

$\therefore \frac{\partial f}{\partial x}$ is continuous at $(0,0)$

Furthermore, when $(x,y) \neq (0,0)$, $\frac{\partial f}{\partial x}(x,y) = 2x(x^2+y^2)\sin\left(\frac{1}{\sqrt{x^2+y^2}}\right) - x\sqrt{x^2+y^2}\cos\left(\frac{1}{\sqrt{x^2+y^2}}\right)$ which is continuous

Therefore $\frac{\partial f}{\partial x}$ is continuous everywhere.

Similarly, we can show that $\frac{\partial f}{\partial y}$ is continuous everywhere.

By theorem 9.2, f is differentiable everywhere.

Exercise 9.4

Show that f is differentiable at $(0,0)$ by using definition 9.3 and theorem 9.1.

As expected, we have:

Theorem 9.3

Let $D \subseteq \mathbb{R}^n$ be an open subset, let $\vec{x}_0 \in D$ and let $f: D \to \mathbb{R}$.

If f is differentiable at $\vec{x}_0$, then f is continuous at $\vec{x}_0$.

proof.

f is differentiable at $\vec{x}_0 \Rightarrow$ there exists $\vec{L} \in \mathbb{R}^n$ such that $\lim_{\vec{h}\to\vec{0}} \frac{f(\vec{x}_0+\vec{h}) - (f(\vec{x}_0) + \vec{L}\cdot\vec{h})}{|\vec{h}|} = 0$

$$\lim_{\vec{h}\to\vec{0}} f(\vec{x}_0+\vec{h}) - (f(\vec{x}_0) + \vec{L}\cdot\vec{h}) = \lim_{\vec{h}\to\vec{0}} \frac{f(\vec{x}_0+\vec{h}) - (f(\vec{x}_0) + \vec{L}\cdot\vec{h})}{|\vec{h}|} |\vec{h}| = 0$$

$$\therefore \lim_{\vec{h}\to\vec{0}} f(\vec{x}_0+\vec{h}) = \lim_{\vec{h}\to\vec{0}} f(\vec{x}_0) + \vec{L}\cdot\vec{h} = f(\vec{x}_0)$$ i.e. f is continuous at $\vec{x}_0$

Summary:

Let $D \subseteq \mathbb{R}^n$ be an open subset, let $\vec{x}_0 \in D$ and let $f: D \to \mathbb{R}$.

$\overset{\times}{\Longleftarrow}$ True only when $\frac{\partial f}{\partial x_i}$ are continuous on D for all i (Theorem 9.2)

1) f is differentiable at $\vec{x}_0 \Rightarrow \frac{\partial f}{\partial x_i}(\vec{x}_0)$ exists for all i (Theorem 9.1)

Exercise $\times \Uparrow$ $\qquad\qquad$ $\Downarrow \times$ See example 9.1

2) f is differentiable at $\vec{x}_0 \Rightarrow f$ is continuous at $\vec{x}_0$ (Theorem 9.3)

Proposition 9.1

Let $D \subseteq \mathbb{R}^n$ be an open subset, $\vec{x}_0 \in D$ and let $f, g: D \to \mathbb{R}$ such that f and g are differentiable at $\vec{x}_0$. Then,

(1) $f \pm g$ is differentiable at $\vec{x}_0$,

(2) $f \cdot g$ is differentiable at $\vec{x}_0$,

(3) $\frac{f}{g}$ is differentiable at $\vec{x}_0$ if $g(\vec{x}_0) \neq 0$.

proof of (2):

From exercise 9.1, $\nabla(f \cdot g)(\vec{x}_0) = g(\vec{x}_0)\nabla f(\vec{x}_0) + f(\vec{x}_0)\nabla g(\vec{x}_0)$.

$$\lim_{\vec{h} \to \vec{0}} \frac{(f \cdot g)(\vec{x}_0 + \vec{h}) - \left((f \cdot g)(\vec{x}_0) + \nabla(f \cdot g)(\vec{x}_0) \cdot \vec{h}\right)}{|\vec{h}|}$$

$$= \lim_{\vec{h} \to \vec{0}} \frac{f(\vec{x}_0+\vec{h})g(\vec{x}_0+\vec{h}) - \left(f(\vec{x}_0)g(\vec{x}_0) + [g(\vec{x}_0)\nabla f(\vec{x}_0) + f(\vec{x}_0)\nabla g(\vec{x}_0)] \cdot \vec{h}\right)}{|\vec{h}|}$$

$$= \lim_{\vec{h} \to \vec{0}} \frac{f(\vec{x}_0+\vec{h})g(\vec{x}_0+\vec{h}) - [f(\vec{x}_0)g(\vec{x}_0+\vec{h}) + g(\vec{x}_0+\vec{h})\nabla f(\vec{x}_0)\cdot\vec{h}] + f(\vec{x}_0)g(\vec{x}_0+\vec{h}) - [f(\vec{x}_0)g(\vec{x}_0) + f(\vec{x}_0)\nabla g(\vec{x}_0)\cdot\vec{h}] - g(\vec{x}_0)\nabla f(\vec{x}_0)\cdot\vec{h} + g(\vec{x}_0+\vec{h})\nabla f(\vec{x}_0)\cdot\vec{h}}{|\vec{h}|}$$

$$= \lim_{\vec{h} \to \vec{0}} \underbrace{\frac{f(\vec{x}_0+\vec{h}) - (f(\vec{x}_0) + \nabla f(\vec{x}_0)\cdot\vec{h})}{|\vec{h}|}}_{\substack{\downarrow \text{ diff. of } f \\ 0 \quad \text{at } \vec{x}_0}} \underbrace{g(\vec{x}_0+\vec{h})}_{\substack{\text{cont. of } g \\ g(\vec{x}_0) \quad \text{at } \vec{x}_0}} + f(x_0)\underbrace{\frac{g(\vec{x}_0+\vec{h}) - (g(\vec{x}_0) + \nabla g(\vec{x}_0)\cdot\vec{h})}{|\vec{h}|}}_{\substack{\downarrow \text{ diff of } g \\ 0 \quad \text{at } \vec{x}_0}} + \underbrace{(g(\vec{x}_0+\vec{h}) - g(\vec{x}_0))}_{\substack{\downarrow \text{ cont. of } g \\ 0 \quad \text{at } \vec{x}_0}} \underbrace{\nabla f(\vec{x}_0)\cdot\vec{h}}_{\substack{\downarrow \\ 0}}$$

$= 0$

Total Differential

Furthermore, $f(\vec{x}) \sim f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)$

$$\underbrace{f(\vec{x}) - f(\vec{x}_0)}_{\text{change of } f} \sim \nabla f(\vec{x}_0) \cdot \underbrace{(\vec{x} - \vec{x}_0)}_{\text{change of } \vec{x}.} = \frac{\partial f}{\partial x_1}(\vec{x}_0)\, \Delta x_1 + \cdots + \frac{\partial f}{\partial x_n}(\vec{x}_0)\, \Delta x_n$$, here we write $\vec{x} - \vec{x}_0 = (\Delta x_1, \cdots, \Delta x_n)$

$$\Delta f \sim \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{x}_0)\, \Delta x_i$$

Classically, we write $df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{x}_0)\, dx_i$ which is said to be the total differential of f at $\vec{x}_0$

df, dx_i can be regarded as variables, and if $dx_i = \Delta x_i$, then $\Delta f \approx df$.

In more advanced level, df and dx_i are treated as linear maps from $\mathbb{R}^n \to \mathbb{R}$, called differential forms.

They are particularly useful in discussion of multiple integrals.

Exercise 9.5

Let $D \subseteq \mathbb{R}^n$ be an open subset, let $f, g : D \to \mathbb{R}$ be differentiable functions and let $c \in \mathbb{R}$

Show that

(1) $d(c) = 0$

(2) $d(f+g) = df + dg$

(3) $d(fg) = g\,df + f\,dg$, so in particular $d(cf) = c\,df$

(4) $d(\frac{f}{g}) = \dfrac{g\,df - f\,dg}{g^2}$

Example 9.7

The volume of a cone V is related to the radius r and height h by the equation $V = \frac{1}{3}\pi r^2 h$.

Then we have $dV = d(\frac{1}{3}\pi r^2 h)$

$$= \frac{1}{3}\pi(2rh\,dr + r^2\,dh)$$

Suppose that $\mathcal{C}$ is a metallic cone with radius $r = 5$ and height $h = 10$.

If $\mathcal{C}$ is heated so that the radius and height are increased by 0.2 and 0.4 respectively, i.e. $dr = \Delta r = 0.2$ and $dh = \Delta h = 0.4$, then

$$\Delta V \sim dV = \frac{1}{3}\pi(2rh\,dr + r^2\,dh)$$

$$= 10\pi$$

Differentiability of Vector Valued Functions

Definition 94

Let $D \subseteq \mathbb{R}^n$ be an open subset, let $\vec{x}_0 \in D$ and let $f: D \to \mathbb{R}^m$.

f is said to be differentiable at $\vec{x}_0$ if there exists $P \in M_{m\times n}(\mathbb{R})$ such that

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

Furthermore, f is said to be a differentiable function if it is differentiable at every $x \in D$.

Idea: Now, change of $\vec{x}$ may not be a real number but a vector $\vec{h} \in \mathbb{R}^n$.

Therefore, change of $f = f(\vec{x}_0+\vec{h}) - f(\vec{x}_0) \in \mathbb{R}^m$ should be approximated by a linear transformation on change of $\vec{x} = \vec{h}$, which is in form of $P\vec{h} \in \mathbb{R}^m$.

Think: But, what is P?

If we write $f(\vec{x}) = f(x_1, x_2, \cdots, x_n) = (f_1(x_1, x_2, \cdots, x_n), f_2(x_1, x_2, \cdots, x_n), \cdots, f_m(x_1, x_2, \cdots, x_n))$,

and $P = \begin{bmatrix} \vec{P}_1 \\ \vdots \\ \vec{P}_m \end{bmatrix} \in M_{m\times n}(\mathbb{R})$, where each $\vec{P}_i$ is a row vector in $\mathbb{R}^n$, we have:

f is differentiable at $\vec{x}_0$

$\Leftrightarrow$ There exists $P \in M_{m\times n}(\mathbb{R})$ such that $\displaystyle\lim_{\vec{h}\to\vec{0}} \frac{|f(\vec{x}_0+\vec{h}) - (f(\vec{x}_0) + P\vec{h})|}{|\vec{h}|} = 0$

row vector in $\mathbb{R}^n$

$$\text{i.e. } \lim_{\vec{h}\to\vec{0}} \frac{\left\| \begin{bmatrix} \underline{f_1(\vec{x}_0+\vec{h}) - (f_1(\vec{x}_0) + \vec{P}_1\cdot\vec{h})} \\ \vdots \\ f_m(\vec{x}_0+\vec{h}) - (f_m(\vec{x}_0) + \vec{P}_m\cdot\vec{h}) \end{bmatrix} \right\|}{|\vec{h}|} = 0$$

$\Leftrightarrow$ There exists $\vec{P}_1, \cdots, \vec{P}_m \in \mathbb{R}^n$ such that $\displaystyle\lim_{\vec{h}\to\vec{0}} \frac{f_i(\vec{x}_0+\vec{h}) - (f_i(\vec{x}_0) + \vec{P}_i\cdot\vec{h})}{|\vec{h}|} = 0$ for $i = 1, 2, \cdots, m$

(In this case $\vec{P}_i = \nabla f_i(\vec{x}_0)$.)

$\Leftrightarrow$ f_i is differentiable at $\vec{x}_0$ for $i = 1, 2, \cdots, m$

$$P = \begin{bmatrix} \nabla f_1(\vec{x}_0) \\ \vdots \\ \nabla f_m(\vec{x}_0) \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\vec{x}_0) & \cdots & \frac{\partial f_1}{\partial x_n}(\vec{x}_0) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\vec{x}_0) & \cdots & \frac{\partial f_m}{\partial x_n}(\vec{x}_0) \end{bmatrix} \in M_{m\times n}(\mathbb{R})$$

Theorem 9.4

Let $D \subseteq \mathbb{R}^n$ be an open subset, let $\vec{x}_0 \in D$ and let $f: D \to \mathbb{R}^m$.

f is differentiable at $\vec{x}_0$ if and only if each f_i is differentiable.

Also, $Df(\vec{x}_0) = \begin{bmatrix} \nabla f_1(\vec{x}_0) \\ \vdots \\ \nabla f_m(\vec{x}_0) \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\vec{x}_0) & \cdots & \frac{\partial f_1}{\partial x_n}(\vec{x}_0) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\vec{x}_0) & \cdots & \frac{\partial f_m}{\partial x_n}(\vec{x}_0) \end{bmatrix} \in M_{m\times n}(\mathbb{R})$ which is called total derivative,

is the unique matrix such that $\lim_{\vec{h}\to\vec{0}} \frac{|f(\vec{x}_0+\vec{h}) - (f(\vec{x}_0) + Df(\vec{x}_0)\cdot\vec{h})|}{|\vec{h}|} = 0$

Remark. $Df(\vec{x}_0)$ is also called Jacobi matrix and denoted by $J_f(\vec{x}_0)$ or $\frac{\partial(f_1,\cdots,f_m)}{\partial(x_1,\cdots,x_n)}(\vec{x}_0)$.

Definition 9.4 (Restated)

Let $D \subseteq \mathbb{R}^n$ be an open subset, let $\vec{x}_0 \in D$ and let $f: D \to \mathbb{R}^m$.

Suppose that $Df(\vec{x}_0) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\vec{x}_0) & \cdots & \frac{\partial f_1}{\partial x_n}(\vec{x}_0) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\vec{x}_0) & \cdots & \frac{\partial f_m}{\partial x_n}(\vec{x}_0) \end{bmatrix} \in M_{m\times n}(\mathbb{R})$ exists.

Let $\varepsilon(\vec{h}) = f(\vec{x}_0+\vec{h}) - (f(\vec{x}_0) + Df(\vec{x}_0)\,\vec{h})$.

f is said to be differentiable at $\vec{x}_0$ if

$$\lim_{\vec{h}\to\vec{0}} \frac{|f(\vec{x}_0+\vec{h}) - (f(\vec{x}_0) + Df(\vec{x}_0)\cdot\vec{h})|}{|\vec{h}|} = \lim_{\vec{h}\to\vec{0}} \frac{|\varepsilon(\vec{h})|}{|\vec{h}|} = 0 .$$

Idea: Let $\varepsilon(\vec{h}) = f(\vec{x}_0+\vec{h}) - (f(\vec{x}_0) + Df(\vec{x}_0)\cdot\vec{h})$

$\lim\limits_{\vec{h}\to\vec{0}} \dfrac{|f(\vec{x}_0+\vec{h}) - (f(\vec{x}_0)+Df(\vec{x}_0)\cdot\vec{h})|}{|\vec{h}|} = \lim\limits_{\vec{h}\to\vec{0}} \dfrac{|\varepsilon(\vec{h})|}{|\vec{h}|} = 0$ roughly means when $\vec{h}$ tends to $\vec{0}$,

$|\varepsilon(\vec{h})| = |f(\vec{x}_0+\vec{h}) - (f(\vec{x}_0)+Df(\vec{x}_0)\,\vec{h})|$ is even smaller comparing to $|\vec{h}|$.

Interpretation 1: If we write $\underbrace{f(\vec{x}_0+\vec{h}) - f(\vec{x}_0)}_{\text{change of } f} \sim Df(\vec{x}_0)\cdot \underbrace{\vec{h}}_{\text{change of } \vec{x}}$

it means change of f can be approximated by a linear transformation on change of $\vec{x}$,

and $\varepsilon(\vec{h}) = f(\vec{x}_0+\vec{h}) - (f(\vec{x}_0)+Df(\vec{x}_0)\cdot\vec{h})$ is the error of the approximation.

Interpretation 2: If we let $\vec{x} = \vec{x}_0+\vec{h}$ (i.e. $\vec{h} = \vec{x}-\vec{x}_0$) and we write $f(\vec{x}) \sim \underbrace{f(\vec{x}_0) + Df(\vec{x}_0)\cdot(\vec{x}-\vec{x}_0)}_{\text{linear function } L(x)}$

it means $f(\vec{x})$ can be approximated by a linear function $L(\vec{x})$, called linearization of f at $\vec{x}_0$

Example 9.8

Let $f: \mathbb{R}^3 \to \mathbb{R}^2$ defined by $f(x,y,z) = ((2x-y+1)\cos z,\ e^x \sin(2y+z))$

Clearly, $f(\vec{0}) = (1,0)$, but how about $f(\vec{x})$ for those $\vec{x}$ "near" $\vec{0}$?

$Df = J_f = \dfrac{\partial(f_1,f_2)}{\partial(x,y,z)} = \begin{bmatrix} 2\cos z & -\cos y & -(2x-y+1)\sin z \\ e^x\sin(2y+z) & 2e^x\cos(2y+z) & e^x\cos(2y+z) \end{bmatrix} \in M_{2\times 3}(\mathbb{R})$ and so

$Df(\vec{0}) = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 2 & 1 \end{bmatrix} \in M_{2\times 3}(\mathbb{R})$

Let $\vec{x} = (0.1, 0.1, 0.1)$,

$f(\vec{x}) - f(\vec{0}) \sim Df(\vec{0})\ (\vec{x}-\vec{0}) = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} 0.1 \\ 0.1 \\ 0.1 \end{bmatrix} = \begin{bmatrix} 0.1 \\ 0.3 \end{bmatrix}$

Therefore $f(\vec{x}) - f(\vec{0}) \sim (0.1, 0.3)$

$f(0.1, 0.1, 0.1) \sim f(\vec{0}) + (0.1, 0.3) = (1.1, 0.3)$