Existence of partial derivatives is not good enough to approximate a function:

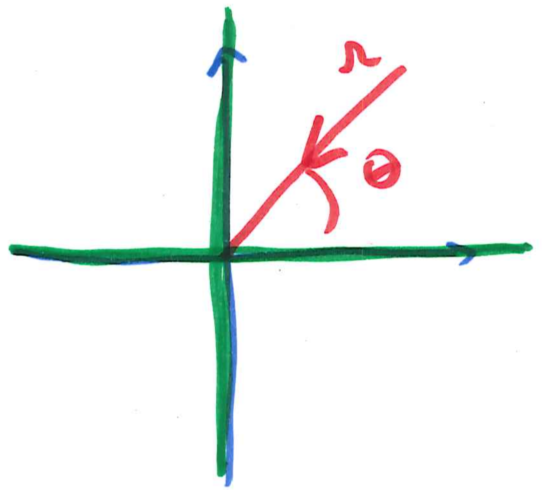
<u>Ex.</u> A function (of two variables) may have partial derivatives at $(0,0)$ ~~but~~ without being continuous at $(0,0)$!

For instance:

$$f:\mathbb{R}^2 \to \mathbb{R};\ f(x,y) = \frac{xy}{x^2+y^2},\quad f(0,0) = 0$$

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

① f has partial derivatives with respect to x and y:



$f(x,0) = 0$ for all x

$\Longrightarrow \partial_x f(0,0) = 0$

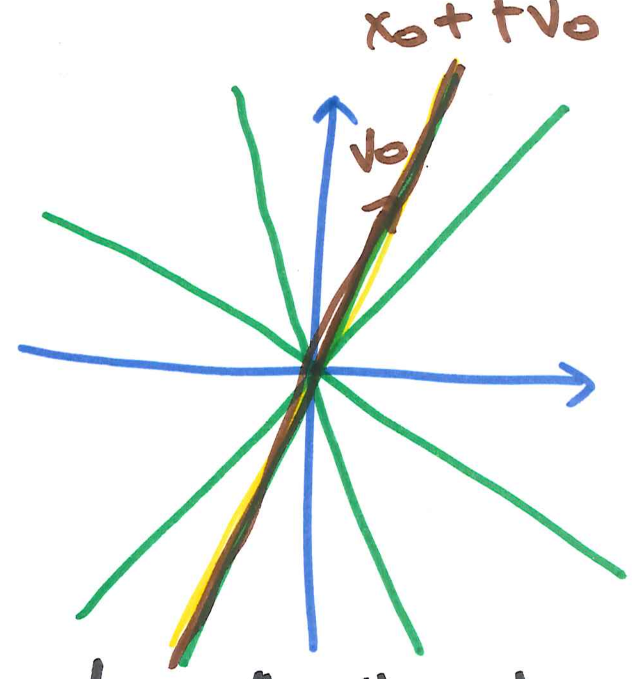
$f(0,y) = 0$ for all y

$\Longrightarrow \partial_y f(0,0) = 0$

② f is not continuous at $(0,0)$:

$(\theta = \frac{\pi}{4})$ $f(r\cos\theta, r\sin\theta) = \cos\theta \sin\theta \underset{r\to 0}{\not\longrightarrow} 0$

Pb. We do not look at the function "in all directions" when computing partial derivatives.



Def. (Directional derivative)

$X \subset \mathbb{R}^n$, open

$f: X \longrightarrow \mathbb{R}^m$

$x_0 \in X$, $v_0 \in \mathbb{R}^n - \{0\}$, maybe $\|v_0\| = 1$

Let $g(t) = f(x_0 + tv_0)$ $(t \in \mathbb{R})$

If g has derivative $y \in \mathbb{R}^m$ at $t=0$, we say that y is the <u>directional derivative</u> of f in direction v_0 at x_0.

Ex. $f(x,y) = \dfrac{xy}{x^2+y^2}$, $f(0,0) = 0$

$x_0 = (0,0)$, $v_0 = (a,b) \neq (0,0)$

$$g(t) = f(tv_0) = f(ta, tb)$$
$$= \frac{t^2 ab}{t^2(a^2+b^2)} = \frac{ab}{a^2+b^2}$$

so $g'(0) = 0$: f has all directional derivatives at $(0,0)$, but is not continuous.

What is the right definition?

3.4 _ The differential

<u>Idea:</u> take the approximation property of the derivative as the main definition.

Recall: $n = 1$

$$f : \mathbb{R} \longrightarrow \mathbb{R}$$

$$x_0 \in \mathbb{R}$$

To say that f has derivative $a \in \mathbb{R}$ at x_0 means that

$$f(\overbrace{x_0 + \underbrace{t}_{\text{"small"}}}^{x}) \approx f(x_0) + \underbrace{at}_{= a(x - x_0)} + \overbrace{(\text{error})}^{E(x)}$$

where the error is much smaller than $\underbrace{|x - x_0|}_{= |t|}$

(i.e. $\lim\limits_{x \to x_0} \frac{E(x)}{|x - x_0|} = 0$)

We generalize this by replacing

$\underline{\underline{n = 1}}$ $\qquad a \, \underbrace{(x - x_0)}_{\text{real number}}$

number

by

$\underline{\underline{n \geq 2}}$ $\qquad A \cdot \underbrace{(x - x_0)}_{\text{vector of size } n}$

matrix

Def: $X \subset \mathbb{R}^n$ open

$f: X \longrightarrow \mathbb{R}^m$

$x_0 \in X$

f is differentiable at x_0 with differential $\left\{ \begin{array}{l} u: \mathbb{R}^n \longrightarrow \mathbb{R}^m \text{ linear} \\ A \text{ matrix with } m \text{ rows, } n \text{ columns} \end{array} \right.$

if

$$f(x) = f(x_0) + \underbrace{A\overbrace{(x - x_0)}^{\mathbb{R}^n}}_{\in \mathbb{R}^m} + E(x)$$

where

$$\lim_{x \to x_0} \frac{\|E(x)\|}{\|x - x_0\|} = 0.$$

(or $$f(x) = f(x_0) + \underbrace{u(x - x_0)}_{\text{evaluation of } u \text{ at } x - x_0} + \bar{E}(x)$$

with $\displaystyle \lim_{x \to x_0} \frac{\|\bar{E}(x)\|}{\|x - x_0\|} = 0$)

<u>Ex.</u> $m = n = 1$: same definition as f differentiable at $x_0 \in \mathbb{R}$ with derivative the only coefficient of A.

<u>Notation:</u> $\begin{cases} u = df(x_0) \\ A = J_f(x_0) \end{cases}$

If f is differentiable at every $x \in X$ then f is said to be <u>differentiable on X.</u> The differential is a function

$$df : X \longrightarrow M_{\substack{m \text{ rows} \\ n \text{ colums}}}(\mathbb{R})$$

or

$$X \longrightarrow (\text{linear maps } \mathbb{R}^n \to \mathbb{R}^m)$$

Properties $\quad f : X \longrightarrow \mathbb{R}^m$

(1) If f is differentiable at x_0 (on X), then it is continuous at x_0 (on X).

(2) If f is ______________, then it has all partial derivatives at x_0, and

$$A = \underbrace{J_f(x_0)}_{= \left(\frac{\partial f_j}{\partial x_i}(x_0)\right)}$$

is the Jacobi matrix.

(3) If $f, g : X \longrightarrow \mathbb{R}^m$ are differentiable then $f+g$ is and

$$d(f+g) = \underbrace{df + dg}_{\text{sum of matrices}}$$

(4) If $m = 1$ and f, g are differentiable, then fg is differentiable.

If g is never 0, then f/g is differentiable.

(5) If $f : X \longrightarrow \mathbb{R}^m$, $f = (f_1, \dots, f_m)$ has all partial derivatives

$$\frac{\partial f_j}{\partial x_i} : X \longrightarrow \mathbb{R}$$

and if these are all continuous on X, then f is differentiable on X, and

$$df(x) = J_f(x)$$

for every $x \in X$.

Property (5) means in practice that any function given by a simple concrete formula is differentiable:

↳ such a function has partial derivatives $\frac{\partial f_j}{\partial x_i}$

↳ these are also simple functions

↳ so they are continuous

↳ so (5) applies

Ex.

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

$(x,y) \neq (0,0)$

(1) f is differentiable on $X = \underbrace{\mathbb{R}^2 - \{0\}}_{\text{open set}}$

(2) f is not differentiable at $(0,0)$

(because f is not even continuous at $(0,0)$!)

[Property (1)]

application of (5)

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

for $(x,y) \neq (0,0)$, continuous on $\mathbb{R}^2 - \{(0,0)\}$

Similarly for $\frac{\partial f}{\partial y}$

$\overset{(5)}{\Longrightarrow}$ f is differentiable on $\mathbb{R}^2 - \{(0,0)\}$

and

$$J_f(x,y) = \left(\frac{-x^2y+y^3}{(x^2+y^2)^2}, \frac{-y^2x+x^3}{(x^2+y^2)^2} \right)$$

Recall : we computed

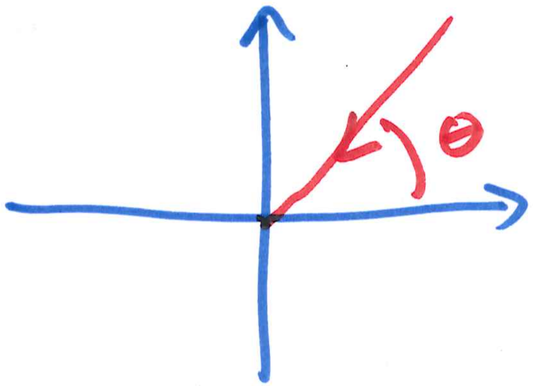
$$\frac{\partial f}{\partial x}(0,0) = 0 = \frac{\partial f}{\partial y}(0,0)$$

Note

$$\frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) = r^3 \frac{(\sin^3\theta - \sin\theta\cos^2\theta)}{r^4}$$

$$= \frac{\sin^3\theta - \sin\theta\cos^2\theta}{r}$$

$$\xrightarrow[r \to 0]{} \pm\infty \quad \text{if} \quad \ne 0$$



Examples:

(1) Polynomials are differentiable (and the coefficients of Jf are also polynomials).

(2) $f(x) = y_0 + A x$, affine-linear

matrix, m rows
n colums

$\Longrightarrow$ f differentiable

and df = matrix A at every point.

Notation:

(1) (Higher partial derivatives)

If $\frac{\partial f}{\partial x_i}$ exists and has partial derivative with respect to x_j, we write

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)$$

$j=i$:

$$\frac{\partial^2 f}{\partial x_j^2} = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_j}\right)$$

Sometimes $\partial^2_{x_j x_i} f = \dfrac{\partial^2 f}{\partial x_j \partial x_i}$, ... 109

(2) <u>**Gradient**</u> : $f : X \longrightarrow \mathbb{R}$ $(m = 1)$

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix} = \text{transpose of Jacobi matrix}$$

("gradient of f"
"nabla")

The ~~...~~ differential is a linear map ~~...~~ $\mathbb{R}^n \longrightarrow \mathbb{R}$

It's value at $x \in \mathbb{R}^n$ is the scalar product $\nabla f \cdot x$: scalar product on $\mathbb{R}^n$

$$\underbrace{\nabla f}_{\mathbb{R}^n} \cdot \underbrace{x}_{\mathbb{R}^n} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \cdot x_i$$

The approximation to f is then:

$$f(x) \cong f(x_0) + (\nabla f)(x_0) \underset{\text{scalar product}}{\bullet} (x - x_0)$$

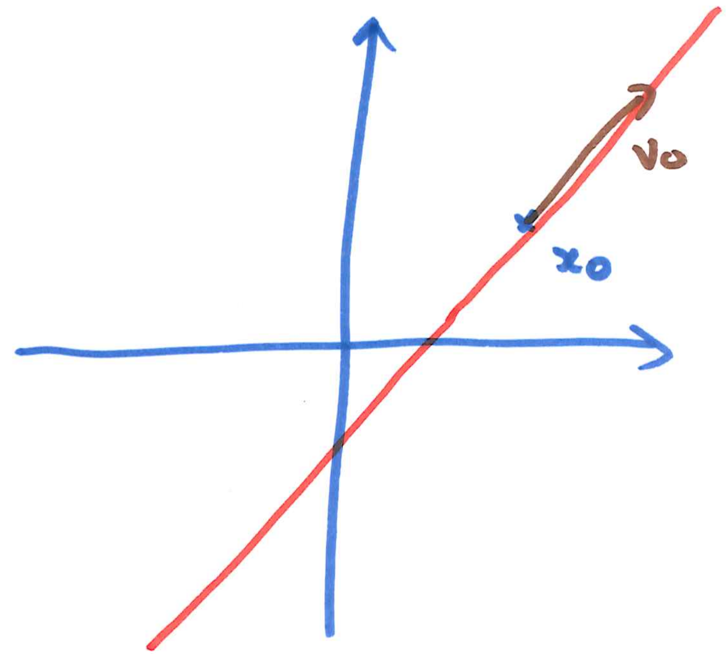
Further properties :

(6) $f : X \longrightarrow \mathbb{R}^m$ differentiable

$\Longrightarrow$ for any $x_0 \in X$, any $v_0 \neq 0$ in $\mathbb{R}^n$, the directional derivative at x_0 in direction v_0 exists and is equal to

$$\boxed{J_f(x_0) \cdot v_0}$$

matrix product



<u>In particular:</u> if one has

directional deriv (at x_0)

y_0 in direction v_0

y_1 ———————— v_1

$\Longrightarrow$ ————————————

$y_0 + y_1$ ———————— $v_0 + v_1$

$y_0 - y_1$ ———————— $v_0 - v_1$

$\vdots$

if non-zero

<u>Special case:</u> $m=1$

Directional derivative is

$$\nabla f(x_0) \cdot v_0$$

scalar product

(7) [<u>Chain rule for many variables</u>]

$$X \xrightarrow{f} \mathbb{R}^m \xrightarrow{g} \mathbb{R}^p$$

If f, g are differentiable then $g \circ f$ is also differentiable and

$$d(g \circ f)(x_0) = dg(f(x_0)) \circ df(x_0)$$

composition of linear maps

<u>or</u>

$$J_{g\circ f}(x_0) = J_g(f(x_0)) \cdot J_f(x_0)$$

matrix product

<u>Warning!</u> careful with order in the product!

<u>Ex.</u> (1) <u>$m = n = 1$</u>

$$\mathbb{R} \xrightarrow{f} \mathbb{R} \xrightarrow{g} \mathbb{R}$$

$$(g\circ f)'(x_0) = g'(f(x_0))\, f'(x_0)$$

is the usual chain-rule.

(2) $\mathbb{R}^n \xrightarrow{f} \mathbb{R}^2 \xrightarrow{m} \mathbb{R}$

$\qquad\qquad\qquad\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto xy$

polynomial

$f(x) = (f_1(x), f_2(x))$

so $m \circ f = f_1 f_2$ is the product.

If f is differentiable then so is $f_1 f_2$

since m is a polynomial, so is differentiable. We can compute $d(f_1 f_2)$ using the chain rule.

$$J_f(x) = \left(\begin{array}{ccccc} \frac{\partial f_1}{\partial x_1} & \cdot & \cdots & \cdot & \frac{\partial f_1}{\partial x_n} \\ \hline \frac{\partial f_2}{\partial x_1} & \cdot & \cdots & \cdot & \frac{\partial f_2}{\partial n} \end{array}\right)$$

$$J_m = \left(\begin{array}{c|c} y & x \end{array}\right)$$

$$\begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ \begin{pmatrix} x \\ y \end{pmatrix} & \longmapsto & xy \end{array}$$

$$J_{f_1 f_2} = J_{mof} = (y \mid x) \cdot \begin{pmatrix} \frac{\partial f_1}{\partial x_i} \\ \hline \frac{\partial f_2}{\partial x_i} \end{pmatrix}$$