Remarks

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

$$f(x) = (f_1(x), \dots, f_m(x))$$

f is differentiable at x_0

$$\Updownarrow$$

f_i, $1 \leq i \leq m$, is differentiable at x_0

And then

$$df(x_0) = \begin{pmatrix} df_1(x_0) \\ \vdots \\ df_m(x_0) \end{pmatrix}$$

(2) $\underline{m = 1}$

$u : \mathbb{R}^n \longrightarrow \mathbb{R}$ "linear form"

is given by

$$u(x_1, \ldots, x_n) = a_1 x_1 + \cdots + a_n x_n$$

for some vector $(a_1, \ldots, a_n) \in \mathbb{R}^n$.

One can also write

$$u(x) = a \cdot x$$

[where $x \cdot y = x_1 y_1 + \cdots + x_n y_n$ is a scalar product].

Prop. (3.4.4)

$f : \mathbb{R}^n \longrightarrow \mathbb{R}$ differentiable at x_0

(1) f is continuous at x_0

(2) f has partial derivatives at x_0,

and

$$\frac{\partial f}{\partial x_i}(x_0) = a_i$$

where $df(x_0) = a \cdot x$

with $a = (a_1, \ldots, a_n)$

<u>Proof</u> : (2) for $n = 2$

x, y as variables in $\mathbb{R}^2$

(x_0, y_0) $u =$ differential at (x_0, y_0)

Write $u(x, y) = ax + by$, a, b in $\mathbb{R}$

Compute $\frac{\partial f}{\partial x}$:

$$g(x) = f(x, y_0)$$

$$\frac{g(x) - g(x_0)}{x - x_0} = \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0}$$

Write

$$f(x,y) = f(x_0, y_0) + a(x-x_0) + b(y-y_0) + E(x,y)$$

$u(x-x_0, y-y_0)$

Then

$$\frac{g(x)-g(x_0)}{x-x_0} = \frac{a(x-x_0) + \cancel{b(y_0-y_0)}}{x-x_0} + \frac{E(x,y_0)}{x-x_0}$$

$$= a + \frac{E(x,y_0)}{\pm\|(x,y_0)-(x_0,y_0)\|}$$

0

f differentiable with $df(x_0, y_0) = u$

means

$$\lim_{(x,y) \to (x_0, y_0)} \frac{E(x,y)}{\|(x,y) - (x_0,y_0)\|} = 0$$

So we find that

$$\lim_{x \to x_0} \frac{g(x) - g(x_0)}{x - x_0} = a.$$

So $\frac{\partial f}{\partial x}(x_0, y_0) = a$.

<u>Conclusion:</u> to compute the differential, it is enough to compute partial derivatives.

<u>Examples:</u> (3.4.5)

(1) Let $f(x) = \underbrace{y_0}_{\in \mathbb{R}^m} + \underbrace{A}_{\substack{\text{matrix}\\ \text{with } m\\ \text{rows, } n\\ \text{columns}}} x$

$(x \in \mathbb{R}^n)$

(affine-linear)

Then f is differentiable on $\mathbb{R}^n$

with $\quad df(x_0) = u$

(where $u(x) = Ax$) for all

x_0 [constant differential]

Because:

$$\underbrace{f(x)}_{y_0 + Ax} - \underbrace{f(x_0)}_{y_0 + Ax_0} - \overbrace{u(x - x_0)}^{A(x - x_0)}$$

$$= 0$$

so $\displaystyle \lim_{x \to x_0} \frac{f(x) - f(x_0) - u(x - x_0)}{\|x - x_0\|} = 0.$

33

(2) $g(x,y) = \dfrac{xy}{\sqrt{x^2+y^2}}$, $g(0,0) = 0$

$\dfrac{\partial g}{\partial x} = 0$, $\dfrac{\partial g}{\partial y} = 0$

but is not differentiable

(3) <u>$m = 1$</u>

if f is differentiable at x_0

then $df(x_0)(x) = a \cdot x$

where

$$a = \begin{pmatrix} \partial_{x_1} f(x_0) \\ \vdots \\ \partial_{x_n} f(x_0) \end{pmatrix} \overset{\text{def.}}{\equiv} \nabla f(x_0)$$

"gradient" $\in \mathbb{R}^n$

\nabla

Nabla

Prop. [3.4.6, 3.4.7, 3.4.9]

(1) If f, g are differentiable, so is $f+g$, and fg (if $m=1$), f/g (if $m=1$, g never zero)

(2) [Chain rule]

$$\begin{array}{ccccc} X & \xrightarrow{f} & Y & \xrightarrow{g} & \mathbb{R}^p \\ \cap & & \cap & & \\ \mathbb{R}^n & & \mathbb{R}^m & & \end{array}$$

composition of linear maps

f, g differentiable on X, Y

$\Longrightarrow$ $g \circ f$ is differentiable on X with

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

(3) If $f : X \longrightarrow \mathbb{R}^m$ has partial derivatives on X **and** if all $\frac{\partial f_j}{\partial x_i}$ are continuous on X, then f is differentiable on X and the matrix of $df(x_0)$ [w.r.t canonical bases of $\mathbb{R}^n$ and $\mathbb{R}^m$] is the <u>Jacobi matrix</u>

$$J_f(x_0) = \left(\frac{\partial f_i}{\partial x_j} \right)_{\substack{1 \le i \le m \\ 1 \le j \le n}}$$

Examples -

(1) Any polynomial in n variables is differentiable (any partial derivative is a polynomial so it is continuous).

(2) $$f(x) = f_1(x_1) \cdots f_n(x_n)$$

If f_i is C^1 (as function of one variable)

then $\frac{\partial f}{\partial x_i} = f_1(x_1) \cdots \underline{f_i'(x_i)} \cdots f_n(x_n)$

so f is differentiable.

(3) [Chain rule examples]

In terms of Jacobi matrices the Chain Rule is

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

matrix product

$$\left(\begin{array}{ccccc} X & \xrightarrow{f} & Y & \xrightarrow{g} & \mathbb{R}^p \\ \cap & & \cap & & \\ \mathbb{R}^n & & \mathbb{R}^m & & \end{array} \right)$$

For instance:

$$n = m = p = 2$$

$$f(x,y) = \begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix}, \quad g(u,v) = \begin{pmatrix} g_1(u,v) \\ g_2(u,v) \end{pmatrix}$$

$$J_f(x,y) = \begin{pmatrix} \partial_x f_1 & \partial_y f_1 \\ \partial_x f_2 & \partial_y f_2 \end{pmatrix}, \quad J_g = \begin{pmatrix} \partial_u g_1 & \partial_v g_1 \\ \partial_u g_2 & \partial_v g_2 \end{pmatrix}$$

function of u,v ~~eval~~ evaluated at $(u,v) = (f_1(x,y), f_2(x,y))$

$$J_{g\circ f}(x,y) = \left(\begin{array}{c|c} \partial_u g_1 \, \partial_x f_1 + \partial_v g_1 \, \partial_x f_2 & \cdots \\ \hline \partial_u g_2 \, \partial_x f_1 + \partial_v g_2 \, \partial_x f_2 & \cdots \end{array}\right)$$

evaluate at (x,y)

<u>$p = 1$</u> $\quad f : X \longrightarrow \underset{\mathbb{R}^m}{Y} \xrightarrow{g} \mathbb{R}$

Chain rule gives $\dfrac{\partial(g \circ f)}{\partial x_i}$:

<u>$i = 1$</u> $\quad f(x) = (f_1(x), \dots, f_m(x))$

$$\frac{\partial(g \circ f)}{\partial x_1} = \frac{\partial g}{\partial y_1}\frac{\partial f_1}{\partial x_1} + \frac{\partial g}{\partial y_2}\frac{\partial f_2}{\partial x_1} + \dots + \frac{\partial g}{\partial y_m}\frac{\partial f_m}{\partial x_1}$$

40

(3) $f, g : \mathbb{R}^n \longrightarrow \mathbb{R}$

Let $h(x) = \begin{pmatrix} f(x) \\ g(x) \end{pmatrix}$, $h : \mathbb{R}^n \longrightarrow \mathbb{R}^2$

$m(u,v) = uv$, $\quad m : \mathbb{R}^2 \longrightarrow \mathbb{R}$

so $(m \circ h)(x) = f(x)\,g(x)$.

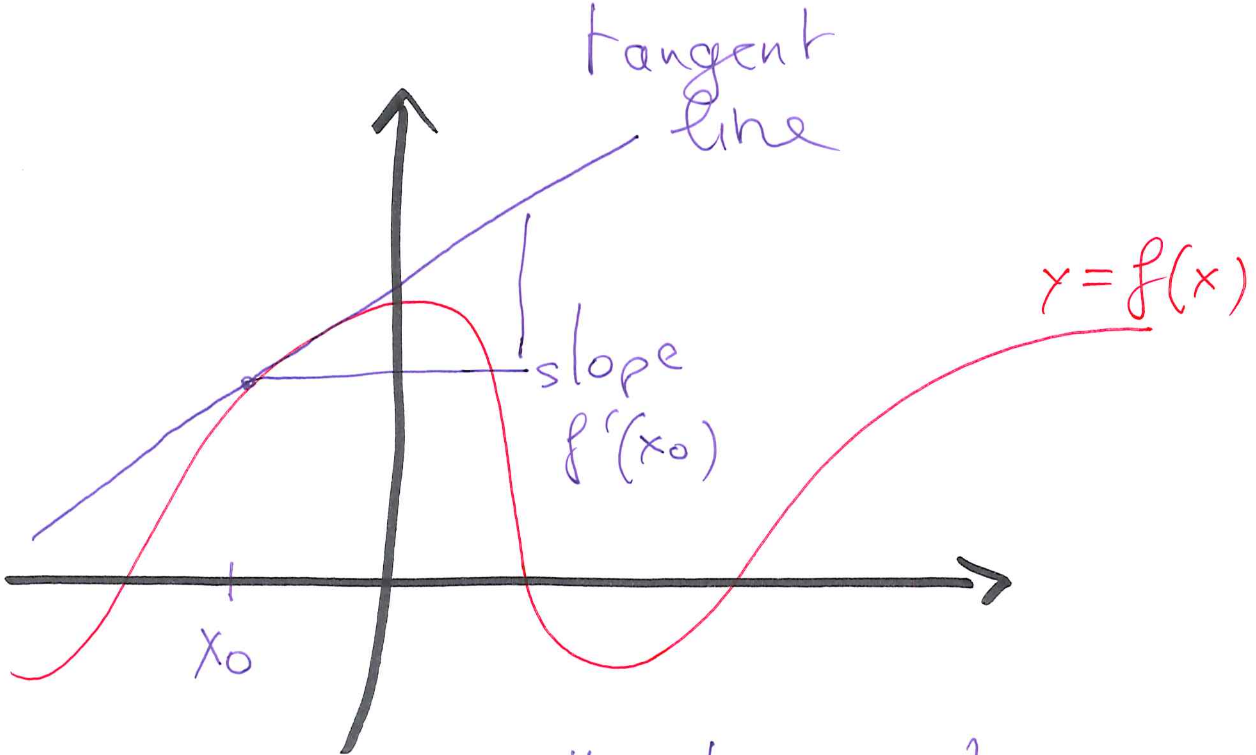
$$J_m(u,v) = \begin{pmatrix} v, & u \end{pmatrix}$$

so $$\frac{\partial (m \circ h)}{\partial x_1} = \frac{\partial (fg)}{\partial x_1} = v\,\frac{\partial f}{\partial x_1} + u\,\frac{\partial g}{\partial x_1}$$

$$= g(x)\frac{\partial f}{\partial x_1} + f(x)\frac{\partial g}{\partial x_1}$$

Geometric interpretation

(1) Tangent space to a graph



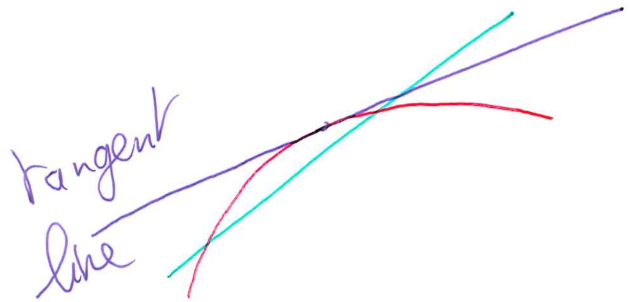
$n=1$

$$\Gamma = \{(x,y) \in \mathbb{R}^2 \mid y = f(x)\}$$

(graph of f)

" the graph is all to one side of the tangent line "

Close enough:



General case: $n \geq 2$

If $f : X \longrightarrow \mathbb{R}^m$ is differentiable, then we obtain an affine-linear subspace $H \subset \mathbb{R}^n \times \mathbb{R}^m$ s.t. "the graph of f is best approximated by H"

Def. (3.4.11)

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

The tangent space at x_0 to the graph $\Gamma = \{(x,y) \in \mathbb{R}^{n+m} \mid y = f(x)\}$ is the graph of the affine-linear map

$$g(x) = f(x_0) + u(x - x_0)$$
$$= f(x_0) - u(x_0) + u(x)$$

where $u = df(x_0)$.

The tangent space is all

$$\begin{matrix} n\{ \\ m\{ \end{matrix}\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{n+m}$$ such that

$$y = f(x_0) - u(x_0) + u(x)$$

or

$$\begin{pmatrix} x \\ y \end{pmatrix} = \underbrace{\begin{pmatrix} x_0 \\ f(x_0) \end{pmatrix}}_{\text{fixed}} + \underbrace{\begin{pmatrix} x - x_0 \\ u(x - x_0) \end{pmatrix}}_{\text{gives an } n\text{-dim. linear subspace}}$$

It is an <u>affine</u> subspace of dim. n.

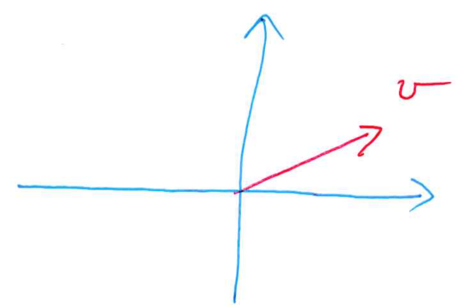
(2) Directional derivatives

Def. (3.4.13)

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

$0 \neq v \in \mathbb{R}^n$

$x_0 \in X$



We say that f has directional derivative a vector $w \in \mathbb{R}^m$ at x_0 in direction v if the function

$$g(t) = f(x_0 + t\,v), \quad t \in \mathbb{R}$$

is differentiable at $t = 0$ with $g'(0) = w$.

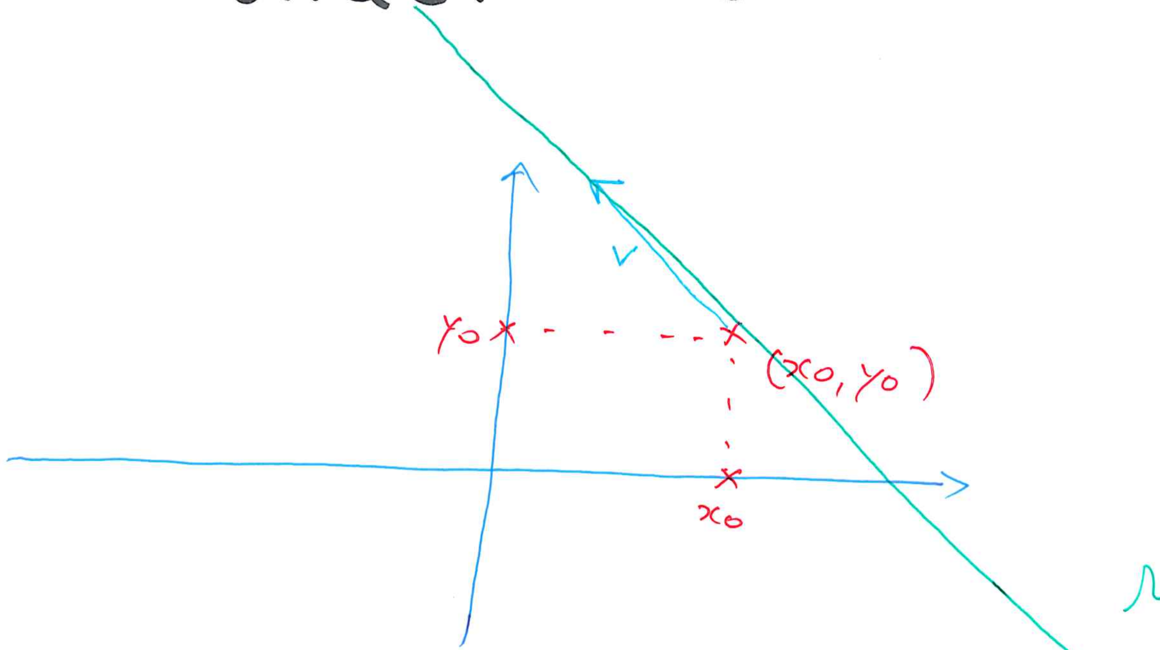
Ex.

(1) $v = e_i$, canonical basis vector

$$\Big(= (0, \dots, \underset{\substack{\uparrow \\ i\text{-th}}}{1}, 0, \dots, 0)\Big)$$

$\rightsquigarrow$ directional derivative $= \dfrac{\partial f}{\partial x_i}$

(2)



$\begin{pmatrix} n = 2 \\ m = 1 \end{pmatrix}$

restrict f to that line

and check if there is a tangent line

<u>Prop.</u> (3.4.15)

$f : X \longrightarrow \mathbb{R}^m$ is differentiable at x_0, with $df(x_0) = u$, then f has directional ~~derivati~~ derivative in all directions $v \neq 0$, given by $u(\underset{\mathbb{R}^n}{\overset{\Uparrow}{v}})$.

In particular, directional derivative is <u>linear</u> with respect to v.

(3) $\underline{m = 1}$ $\qquad X \subset \mathbb{R}^n$

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

differentiable at x_0

so $df(x_0)(x) = \nabla f(x_0) \cdot x$

$\nabla f(x_0) \longrightarrow \begin{pmatrix} \partial_{x_1} f \\ \vdots \\ \partial_{x_n} f \end{pmatrix}$

$\longrightarrow$ tangent space in $\mathbb{R}^{n+1}$

is the ~~[illegible]~~ affine subspace $f(x_0) + n_0^{\perp}$

vectors perpendicular to n_0

where

$$n_0 = \begin{pmatrix} -\nabla f(x_0) \\ 1 \end{pmatrix}$$

$\hookrightarrow$ $\nabla f(x_0) \in \mathbb{R}^n$ is the direction of greatest increase of f.