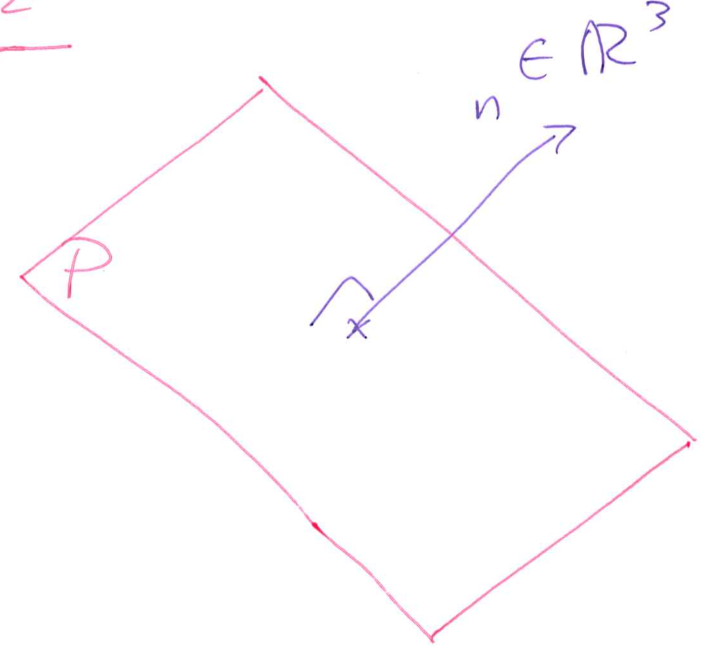
By definition, the tangent space to the graph of f is the graph of

$$g(x) = f(x_0) + \nabla f(x_0) \cdot x$$

so the set

$$f(x_0) + \underbrace{\{ (\overset{\mathbb{R}^n}{x}, \overset{\mathbb{R}}{y}) \mid y = \nabla f(x_0) \cdot x \}}_{\text{linear subspace of dimension } n}$$

Ex. $n = \cancel{3}\,2$



(A plane P in $\mathbb{R}^3$ is described by a vector n perpendicular to P, a "normal" vector)

$$y = \nabla f(x_0) \cdot x$$

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

$$\iff \begin{pmatrix} x \\ y \end{pmatrix} \cdot \underbrace{\begin{pmatrix} -\nabla f(x_0) \\ 1 \end{pmatrix}}_{n_0} = 0$$

in $\mathbb{R}^{n+1}$

Second geometric interpretation

$\nabla f(x_0)$ (if non-zero) gives the direction (in $\mathbb{R}^n$) in which f increases "fastest" when you move slightly.

Ex.

map of an area

colors represent height of a mountain

$h = f(x,y)$

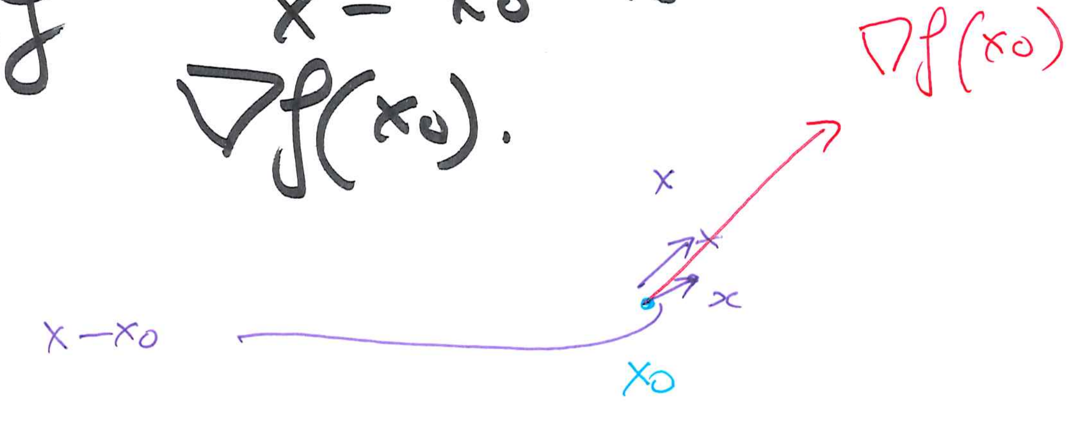
$$f(x) \approx f(x_0) + \nabla f(x_0) \cdot (x - x_0)$$

$$f(x) - f(x_0) \approx \nabla f(x_0) \cdot (x - x_0)$$

Cauchy-Schwarz:

$$|\nabla f(x_0) \cdot (x - x_0)| \leq \|\nabla f(x_0)\| \cdot \|x - x_0\|$$

with equality if $x - x_0$ is proportional to $\nabla f(x_0)$.



so the ~~largest~~ fastest increase is in the direction of $\nabla f(x_0)$.

[Slowest increase if $x - x_0 \perp \nabla f(x_0)$]

<u>Third geometric interpretation</u>

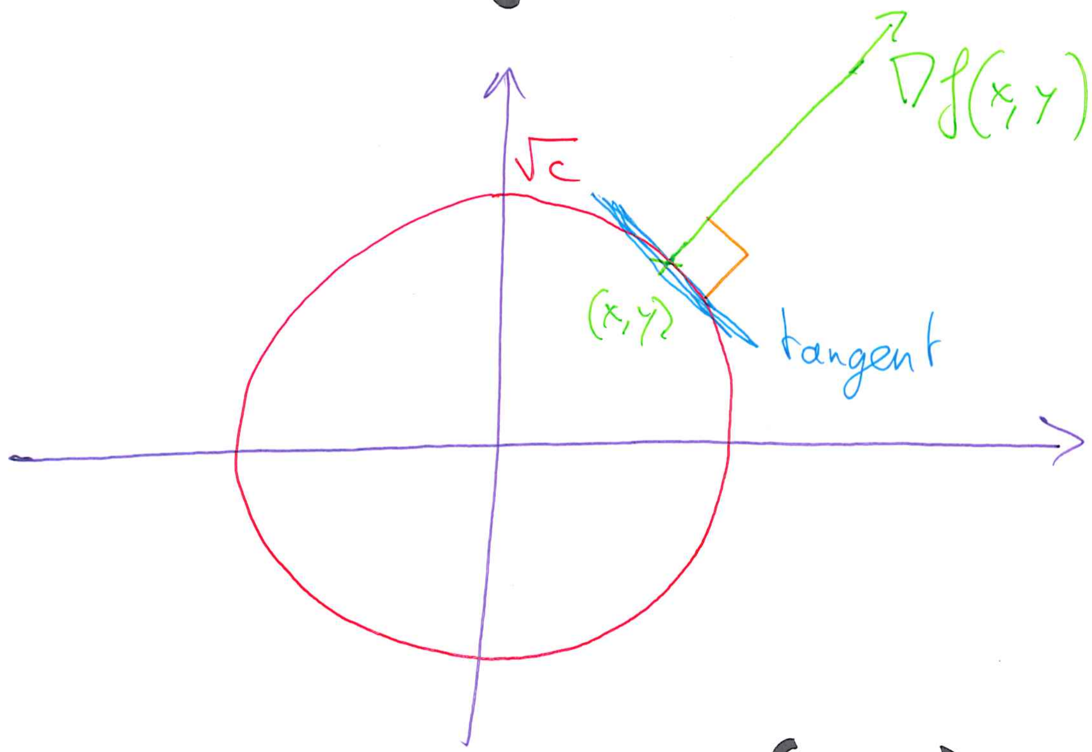
$\nabla f(x_0)$ is orthogonal to sets of the form $L_c = \{ x \in \mathbb{R}^n \mid f(x) = c \}$

(c fixed real number, assume $L_c \neq \emptyset$)

<u>Ex.</u>

$n = 2$

$f(x, y) = x^2 + y^2$ (L_c curve)



$L_c = \emptyset, \quad c < 0$

$L_0 = \{(0, 0)\}, \quad c = 0$

$L_c =$ circle of radius $\sqrt{c}$

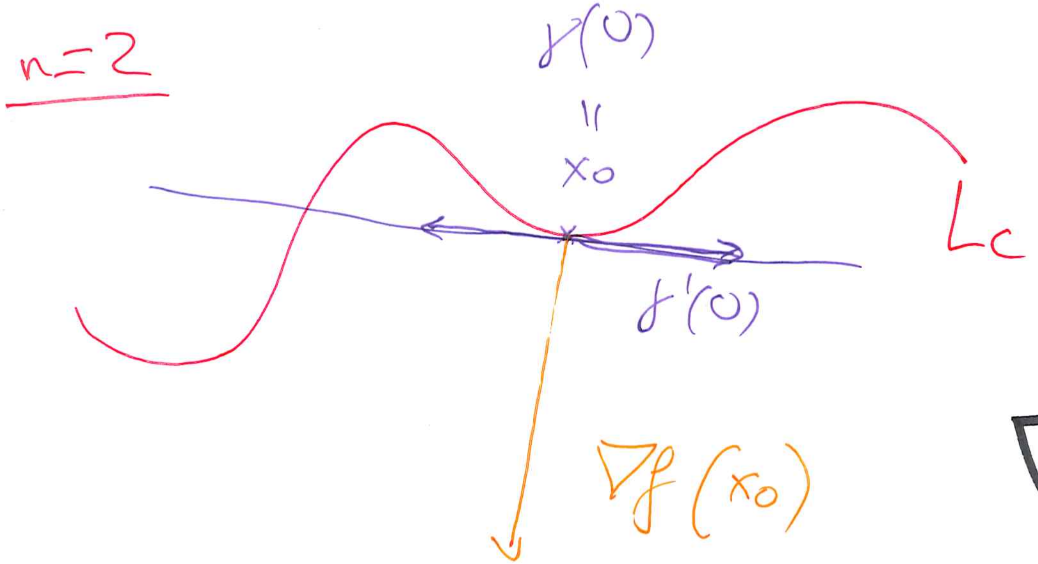
$$\nabla f(x, y) = \begin{pmatrix} 2x \\ 2y \end{pmatrix}$$

Why is that?

$x_0 \in L_c$

Take any $\gamma :]-1,1[\longrightarrow L_c$

differentiable, $\gamma(0) = x_0$.



Note that

$$(f \circ \gamma)'(0) = 0$$

$\| $ — (Chain rule)

$$\nabla f(x_0) \cdot \gamma'(0)$$

Higher derivatives [3.5]

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

f differentiable

$\hookrightarrow$ has $\partial_{x_i} f$, $1 \le i \le n$

$\hookrightarrow$ these may have partial derivatives or be differentiable

<u>Notation:</u>

$$\partial_{x_i}(\partial_{x_j} f) = \partial_{x_i x_j} f$$

$$= \frac{\partial^2 f}{\partial x_i \partial x_j}$$

<u>$i = j$:</u> $\partial_{x_i^2} f = \dfrac{\partial^2 f}{\partial x_i^2}$

Def. (3.5.1) $k \geq 1$

$f : X \longrightarrow \mathbb{R}^m$, $f = (f_1, \dots, f_m)$

(1) f C^1 if (f differentiable on X and) $\frac{\partial f_j}{\partial x_i}$ are continuous on X $(1 \leq i \leq n, 1 \leq j \leq m)$

(2) if $k \geq 2$, f is C^k if f is differentiable and all $\frac{\partial f_j}{\partial x_i}$ are C^{k-1}

(3) f is C^∞ if it is C^k for all k

Ex.

- Polynomials are C^∞
- $f+g$ is C^k if f, g are C^k
- (m=1) fg ———— ————
- f/g ———— ———— if $g(x) \neq 0$ on X
- $f \circ g$ is C^k if f and g are (chain rule)

<u>$k=2$:</u> a priori n^2

$(m=1)$ 2nd-order derivatives

$$\partial_{x_i}\partial_{x_j} f$$

<u>Ex.</u> take $n=3$

$$f(x,y,z) = x^a y^b z^c$$

$$\partial_x f = a\, x^{a-1} y^b z^c$$

$$\partial_y \partial_x f = ab\, x^{a-1} y^{b-1} z^c = \partial_x \partial_y f$$

<u>symmetry</u>

<u>Prop.</u> (3.5.4)

$k \geq 2$; $f : X \longrightarrow \mathbb{R}^m$ class C^k

The partial derivatives with respect to different variables "commute"

$$\partial_{x_i} \partial_{x_j} f = \partial_{x_j} \partial_{x_i} f$$

$$\partial_{x_i} \partial_{x_j} \partial_{x_k} f = \partial_{x_i} \partial_{x_k} \partial_{x_j} f = \partial_{x_j} \partial_{x_k} \partial_{x_i} f = \ldots$$

$$\vdots$$

<u>Important:</u> all partial derivatives considered must be continuous!

<u>Ex.</u> $f(x,y) = \dfrac{xy(x^2-y^2)}{x^2+y^2}$

$f(0,0) = 0$

f differentiable at $(0,0)$ with

$\nabla f(0,0) = 0, \quad \dfrac{\partial^2 f}{\partial x \partial y}(0) = 1 \neq -1 = \dfrac{\partial^2 f}{\partial y \partial x}(0)$

<u>Notation:</u> ("multi-index notation")

$k \geq 2$

Want to "parameterize" all possible <u>different</u> partial derivatives of order k.

Prop. $\Longrightarrow$ a derivative like this only depends on

m_1 = nb. of times we differentiate w.r.t. x_1

m_2 = ______________ x_2

⋮

m_n = ______________ x_n

which are non-negative integers

and $m_1 + m_2 + \dots + m_n = k$

We write

$$\frac{\partial^k f}{\partial x_1^{m_1} \cdots \partial x_n^{m_n}} = \partial^m f = \frac{\partial^k f}{\partial x^m} \quad \text{for this}$$

$(m = (m_1, \ldots, m_n))$

derivative:

$$\underbrace{\partial_{x_1} \cdots \partial_{x_1}}_{m_1} \underbrace{\partial_{x_2} \cdots \partial_{x_2}}_{m_2} \cdots \underbrace{\partial_{x_n} \cdots \partial_{x_n}}_{m_n} f$$

$m = (m_1, \ldots, m_n)$ is called a <u>multi-index</u>

Ex. $k = 4$

$m = (1, 1, 2)$

$$\partial_{x_1} \partial_{x_2} \partial_{x_3} \partial_{x_3} f = \partial^{(1,1,2)} f$$

Def. (Hessian)

$k=2$: f class C^2

$$\left(\frac{\partial^2 f}{\partial x_i \, \partial x_j} \right)_{1 \le i,j \le n} = \mathrm{Hess}_f(x)$$

(Hessian matrix; symmetric matrix)

Change of variable / coordinates

(3.6)

Idea

$(x_1, \dots, x_n)$

$$\begin{array}{c} f : X \longrightarrow \mathbb{R} \\ \big\uparrow g \\ U \end{array}$$

$(y_1, \dots, y_n)$

$y \longleftrightarrow$ "new" variables

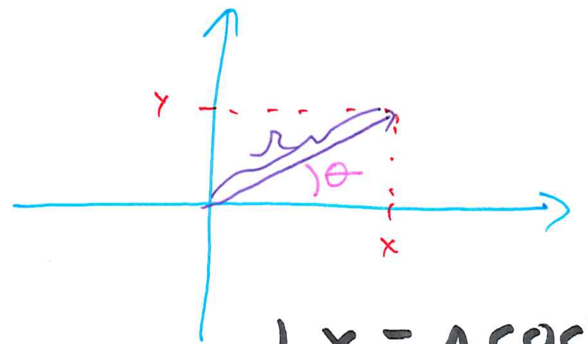
$x \longleftrightarrow$ "old" variables

Ex. polar coordinates

$f(x,y)$

$(x,y) \in \mathbb{R}^2$

r, θ "new" variable



$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases}$$

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

$f \circ g : U \longrightarrow \mathbb{R}$

"new" function

<u>Goal:</u> express ∇f, other partial derivatives in terms of $\nabla(f \circ g)$ or conversely

$$g(r, \theta) = (r\cos\theta, r\sin\theta)$$

$$J_g(r, \theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

so for

$f : \mathbb{R}^2 \longrightarrow \mathbb{R}$

we get

$$\frac{\partial f}{\partial r} = \cos\theta \frac{\partial f}{\partial x} + \sin\theta \frac{\partial f}{\partial y}$$

$$\frac{\partial f}{\partial \theta} = -r\sin\theta \frac{\partial f}{\partial x} + r\cos\theta \frac{\partial f}{\partial y}$$

Chain rule: $\quad h = f \circ g$

$$\partial_{y_1} h(y) = \frac{\partial f}{\partial x_1} \frac{\partial g_1}{\partial y_1} + \dots + \frac{\partial f}{\partial x_n} \frac{\partial g_n}{\partial y_1}$$

$$\left[g(y) = (g_1(y), \dots, g_n(y)) \right]$$

evaluated at $g(y)$

Abuse of notation!

(1) "$h = f$"

(2) "$x_1 = g_1, \dots, x_n = g_n$"

$$\rightarrow \frac{\partial f}{\partial y_1} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial y_1} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial y_1}$$

Similarly for $\frac{\partial f}{\partial y_2}, \ldots, \frac{\partial f}{\partial y_n}$.

given

One can solve for

$\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}$:

express y_1 as function of $(x_1, \ldots, x_n)$, and differentiate

$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial y_1}\frac{\partial y_1}{\partial x_1} + \cdots + \frac{\partial f}{\partial y_n}\frac{\partial x_n}{\partial x_1}$$

$\vdots$

Ex. $f(x,y) = \exp(x^2+y^2)$

$$\updownarrow$$

$f(r,\theta) = \exp(r^2)$

$\cancel{\frac{\partial f}{\partial r}}$ Solving for $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$,

we get

$$\frac{\partial f}{\partial x} = \cos(\theta)\frac{\partial f}{\partial r} - \frac{1}{r}\sin(\theta)\frac{\partial f}{\partial \theta}$$

$$\frac{\partial f}{\partial y} = \sin(\theta)\frac{\partial f}{\partial r} + \frac{1}{r}\cos(\theta)\frac{\partial f}{\partial \theta}$$

$$\nabla f(x,y) = \begin{pmatrix} \partial_x f \\ \partial_y f \end{pmatrix}$$

polar coordinates

$$= \begin{pmatrix} \cos\theta \, 2r e^{r^2} + 0 \\ \sin\theta \, 2r e^{r^2} + 0 \end{pmatrix}$$

f does not depend on θ

$$= \begin{pmatrix} 2x \exp(x^2+y^2) \\ 2y \exp(x^2+y^2) \end{pmatrix}$$

Higher derivatives?

Just iterate...

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \qquad \text{"Laplacian"}$$

$$\begin{aligned}\partial_{x^2} f &= \cos\theta \, \partial_r(\partial_x f) \\ &\quad - \frac{1}{r} \sin\theta \, \partial_\theta(\partial_x f) \\ &= \cos\theta \, \partial_r\left(\cos\theta \, \partial_r f - \frac{1}{r} \sin\theta \, \partial_\theta f\right) \\ &\quad + (\text{rest})\end{aligned}$$

One obtains:

$$\Delta f = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2}$$

Ex. $f = \exp(r^2)$

$$\Delta f = \partial_{r^2} f + \frac{1}{r}\partial_r f$$

$$= (2 + 4r^2)e^{r^2} + 2e^{r^2}$$

$$= 4(1 + x^2 + y^2)e^{x^2+y^2}$$