5.11-20

Defn

- $f : X \to \mathbb{R}^m$, $X \subset \mathbb{R}^n$

  f is <u>differentiable in $x_0 \in X$</u>

  if $\exists$ a lin. map $u : \mathbb{R}^n \to \mathbb{R}^m$

  such that

  $$f(x) = f(x_0) + u(x - x_0) + E(f; x; x_0)$$

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

u is called the <u>differential of f</u>

Thm

- If f is differentiable at x_0 then f is continous at x_0.

- If f is diff. at x_0 then all of its partial derivatives

  $\dfrac{\partial f_i}{\partial x_j}$ exist at x_0 $\quad 1 \leq i \leq m$, $1 \leq j \leq n$.

  and the matrix that represents the differential of f is

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

$$(df)(x_0) : \mathbb{R}^n \to \mathbb{R}^m$$

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \longmapsto (J_f(x_0)) \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$$

- if $f : \mathbb{R}^n \to \mathbb{R}$ then

  $$df(x_0) : \mathbb{R}^n \to \mathbb{R}$$

  $$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \longmapsto \frac{\partial f}{\partial x_1}(x_0)\, x_1 + \dots + \frac{\partial f}{\partial x_n}(x_0)\, x_n.$$

Defn

- If $f : \mathbb{R}^n \to \mathbb{R}$, then the gradient ∇f of f is $\nabla f(x_0) = (J_f(x_0))^t$

Thm

- If $f : X \to \mathbb{R}^m$, $X \subset \mathbb{R}^m$ has all partial derivatives $\dfrac{\partial f_i}{\partial x_j}$ ~~on X~~ <u>and</u> if the partial derivatives are continous then f is differentiable.

Defn $f : X \to \mathbb{R}^m$ diff at x_0. Then the graph of the affine lin. transformation $g : \mathbb{R}^n \to \mathbb{R}^m$, $g(x) = f(x_0) + (df)(x_0)(x - x_0)$ is called the tangent space at x_0 ~~the~~ the graph of f. $G(f) = \{(x, f(x)) \in \mathbb{R}^{n+m}$

Special case

$f : \mathbb{X} \to \mathbb{R} \qquad \mathbb{X} \subset \mathbb{R}^2, \bar{x}_0 \in \mathbb{X}$

$$df(\bar{x}_0) : \mathbb{R}^2 \to \mathbb{R}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \left(\frac{\partial f}{\partial x}(x_0), \frac{\partial f}{\partial y}(x_0)\right)\begin{pmatrix} x \\ y \end{pmatrix}$$

$$= \langle \nabla f, \begin{pmatrix} x \\ y \end{pmatrix} \rangle$$

$$= \frac{\partial f}{\partial x}(\bar{x}_0) \cdot x + \frac{\partial f}{\partial y}(\bar{x}_0)\, y.$$

and $g(x,y) = f(\bar{x}_0) + \nabla f(\bar{x}_0) \cdot (\bar{x} - \bar{x}_0)$

and $z = f(\bar{x}_0) + \frac{\partial f}{\partial x}(\bar{x}_0)(x - x_0) + \frac{\partial f}{\partial y}(\bar{x}_0)(y - y_0)$

is the equation of the tangent plane.

Thm 1) $f, g : X \to \mathbb{R}^m$ diff in x_0 then so is $f+g$ and
$df(x_0) + dg(x_0) = d(f+g)(x_0)$

2) if $f, g : \mathbb{R}^n \to \mathbb{R}$ are diff. in x_0 then so is $f \cdot g$.
if $g \neq 0$ for any $x \in \mathbb{R}^n$, then f/g is also diff.

Thm (chain rule) let $X \subset \mathbb{R}^n$ open $Y \subset \mathbb{R}^m$ open. let $f : X \to Y$ $g : Y \to \mathbb{R}^p$ diff. functions. Then $g \circ f : X \to \mathbb{R}^p$ is diff on X. And for any $x_0 \in X$ its differential is given by the linear map which is the composition of $dg(f(x_0)) : \mathbb{R}^m \to \mathbb{R}^p$ and $df(x_0) : \mathbb{R}^n \to \mathbb{R}^m$

ie $d(g \circ f)(x_0) = dg(f(x_0)) \circ df(x_0)$ composition of maps.

The Jacobi Matrices satisfy

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

matrix product

Thm If $f : X \to \mathbb{R}^m$ is diff at x_0, then the directional derivative of f at x_0 in the direction of $\vec{v} \in \mathbb{R}^n$ exists for every $\vec{v} \in \mathbb{R}^n$ and

$$\boxed{\frac{d}{dt} f(x_0 + t\vec{v})\Big|_{t=0} = df(x_0)(v) = J_f(x_0) \cdot v}$$

Rk. Geometric meaning of the gradient

$f : \mathbb{R}^n \to \mathbb{R}$. diff.

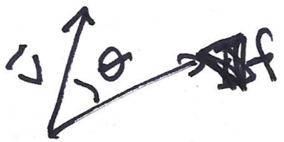
$$\nabla f(\bar{x}_0) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\bar{x}_0) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\bar{x}_0) \end{pmatrix}.$$

Suppose $\nabla f(x_0) \neq 0$.

Let $\vec{v}$ be a unit vector, $\|v\| = 1$.

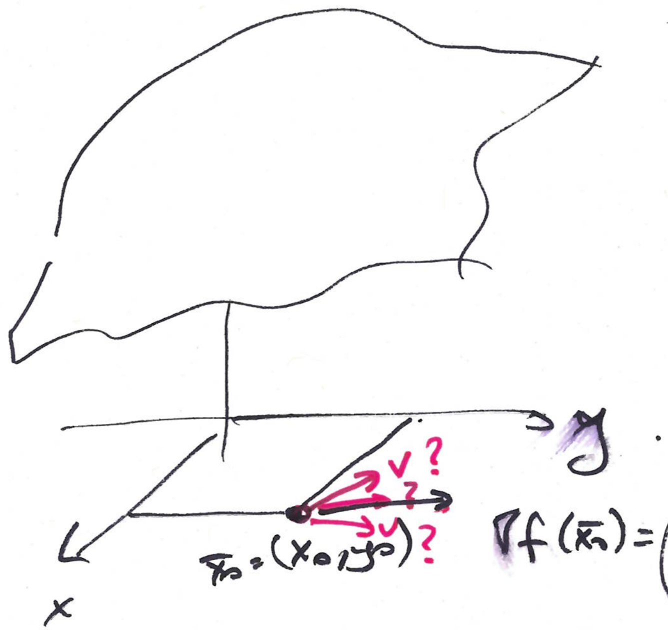
dir. der. of f in the direc. of $\vec{v}$ at the point $\bar{x}_0$ is given by

$$\langle \nabla f(x_0), \vec{v} \rangle = \underbrace{\|\nabla f(x_0)\| \overbrace{\|v\|}^{1} \cos\theta}_{\text{when is this largest?}}$$



We'll maximize the direc. der. at x_0, ie $\langle \nabla f(x_0), \vec{v} \rangle$ if we maximize $\cos\theta$

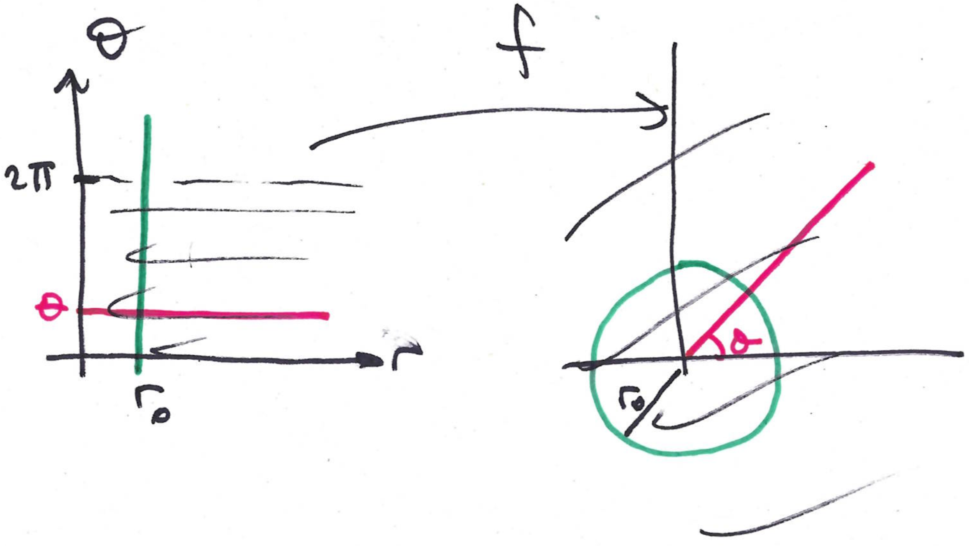
$\Rightarrow$ ie when $\theta = 0$.



$\nabla f(\bar{x}_0) = \left(\frac{\partial f}{\partial x}(x_0), \frac{\partial f}{\partial y}(x_0)\right)$

3 important examples of diff. functions.

① Polar coordinates.



let $f : [0,\infty) \times [0, 2\pi) \to \mathbb{R}^2$

$(r, \theta) \longmapsto (x,y) = (r\cos\theta, r\sin\theta)$

$f = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$ $\quad f_1(r,\theta) = r\cos\theta$

$\quad f_2(r,\theta) = r\sin\theta.$

$$J_f(r,\theta) = \begin{pmatrix} \frac{\partial f_1}{\partial r} & \frac{\partial f_1}{\partial \theta} \\ \frac{\partial f_2}{\partial r} & \frac{\partial f_2}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

$\det J_f(r,\theta) = \underline{r}.$

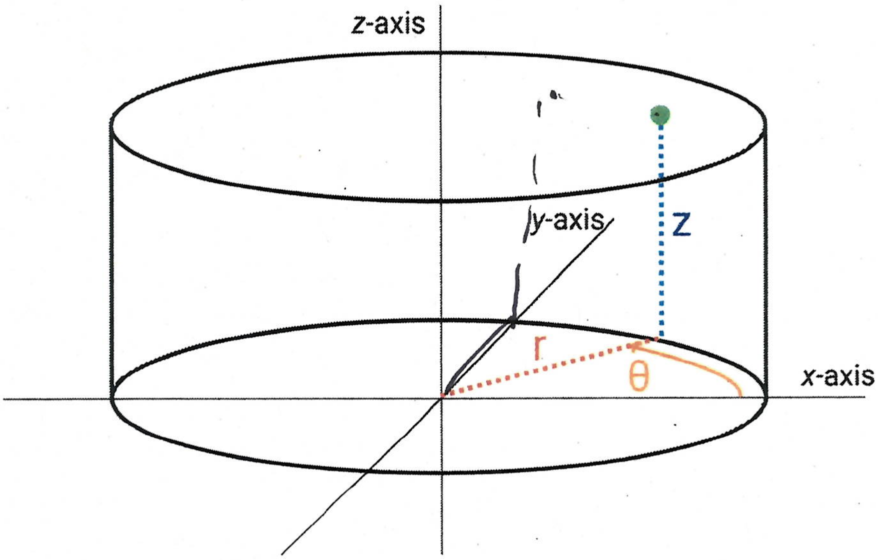


FIGURE 1. Cyclindrical Coordinates

$$f : (0,\infty) \times (0, 2\pi) \times \mathbb{R} \longrightarrow \mathbb{R}^3$$
$$(r, \theta, z) \longmapsto \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$J_f(r,\theta) = \begin{pmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \underline{\det J_f = r}$$

J

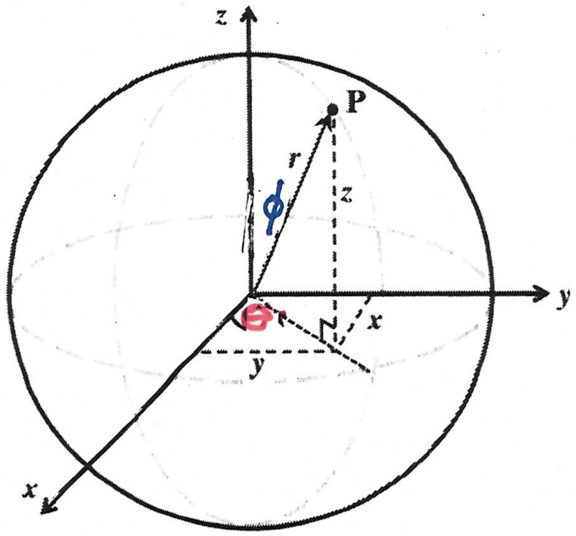


FIGURE 2. Spherical Coordinates

$$f : (0,\infty) \times (0, 2\pi) \times (0, \pi) \longrightarrow \mathbb{R}^3$$

$$(r, \theta, \phi) \longmapsto \begin{pmatrix} r\cos\theta\sin\phi \\ r\sin\theta\sin\phi \\ r\cos\phi \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$J_f(r,\theta,\phi) = \begin{pmatrix} \cos\theta\sin\phi & -r\sin\theta\sin\phi & r\cos\theta\cos\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\phi & 0 & -r\sin\phi \end{pmatrix}$$

$$\det J_f = -r^2\sin\phi$$

Defn. Let $X \subset \mathbb{R}^n$ an open set and $f : X \longrightarrow \mathbb{R}^n$ diff. We say f is a change of variables around x_0 if there is a radius $\rho > 0$ such that the restriction of f to the Ball around x_0 of radius ρ
$$B = \{ x \in \mathbb{R}^n \mid \|x - x_0\| < \rho \}$$
so that the image $Y = f(B)$ is open in $\mathbb{R}^n$ and $\exists$ a diff map $g : Y \longrightarrow B$ such that $f \circ g = id_Y$, $g \circ f = id_B$

ie. $f|_{B_\rho(x_0)}$ is a bijection to the image with a inverse g which is also differentiable.

Thm (Inverse function thm).

Let $X \subseteq \mathbb{R}^n$ open, $f : X \to \mathbb{R}^n$ diff. If $x_0 \in X$ is such that $\det(J_f(x_0)) \neq 0$ ie $J_f(x_0)$ is invertible then f is a change of variables around x_0.

Moreover the Jacobian of g at x_0 is defined by
$$J_g(f(x_0)) = J_f(x_0)^{-1}$$

Higher derivatives.

Defn. $X \subset \mathbb{R}^n$, $f : X \to \mathbb{R}^m$
we say of ~~of~~ class C^1
if f is diff on X and
all of its partial derivatives
are continuous.

$C^1(X; \mathbb{R}^m)$.

let $k \geq 2$, we say $f \in C^k$
if ~~~~ it is diff and
each $\partial_{x_i} f : X \to \mathbb{R}^m$ is
of class C^{k-1}.

$C^k(X, \mathbb{R}^m)$.

f is <u>smooth</u> or C^∞ if
$f \in C^k \ \forall k$

Rk all polys, trig functions,
exponentials are smooth.

Thm For $f \in C^k$, $k \geq 2$
~~~~ then the partial
derivatives of order $\leq k$
are indep of order of
differentiation.

(ie. Mixed partial derivatives
(up to order k) all commute.

eg: 1) If $k=2$, $f \in C^2$ then

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

2) if $f \in C^4$

$$\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k} = \frac{\partial^3 f}{\partial x_j \partial x_i \partial x_k} = \cdots$$

$$\frac{\partial^4 f}{\partial x_1 \partial x_3^3} = \frac{\partial^4 f}{\partial x_3 \partial x_1 \partial^2 x_3} = \cdots$$

<u>Warning!</u>

$$f(x,y) = \begin{cases} \dfrac{xy(x^2-y^2)}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & \text{o.w.} \end{cases}$$

$$\partial_x \partial_y f(0,0) = 1$$

$$\partial_y \partial_x f(0,0) = -1$$

<u>Defn.</u> if $f \in C^2(X \to \mathbb{R})$. $X \subset \mathbb{R}^n$; Then the $n \times n$ matrix $\left(\dfrac{\partial^2 f}{\partial x_i \partial x_j}(x_0)\right) =: \mathrm{Hess}_f(x_0)$ is called the <u>Hessian of f at x_0</u>.

$H = \mathrm{Hess}_f(x_0)$ is a sym. matrix.
$H^t = H$.

eg: $f: \mathbb{R}^3 \to \mathbb{R}$
$(x,y,z) \mapsto x^2y + y^2$

$\frac{\partial f}{\partial x} = 2xy$ $\quad \frac{\partial f}{\partial y} = x^2 + z$ $\quad \frac{\partial f}{\partial z} = y$

$\frac{\partial^2 f}{\partial x^2} = 2y$ $\quad \frac{\partial^2 f}{\partial y^2} = 0$ $\quad \frac{\partial^2 f}{\partial z^2} = 0.$

$$\frac{\partial^2 f}{\partial x \partial y} = 2x = \frac{\partial^2 f}{\partial y \partial x}$$

$$\frac{\partial^2 f}{\partial x \partial z} = 0 = \frac{\partial^2 f}{\partial z \partial x}$$

$$\frac{\partial^2 f}{\partial y \partial z} = 1 = \frac{\partial^2 f}{\partial z \partial y}.$$

$$\mathrm{Hess}_f(x,y,z) = \begin{pmatrix} 2y & 2x & 0 \\ 2x & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Notation: When we are dealing w/ partial derivatives of higher order we use multi index notation.

$f: \mathbb{R}^n \to \mathbb{R}$

let $m = (m_1, m_2, \dots m_n)$

let $|m| = m_1 + m_2 + \dots + m_n$

For the partial derivative for

$$\frac{\partial^{|m|} f}{\partial x_1^{m_1} \partial x_2^{m_2} \dots \partial x_n^{m_n}}$$

we write $\partial_{x^m}^{|m|} f$

x^m means $(x_1^{m_1}, x_2^{m_2}, \dots x_n^{m_n})$

$m! := m_1! m_2! \dots m_n!$

We've already seen a first order approximation to $f: \mathbb{R}^n \to \mathbb{R}$ at $\bar{x}_0$

It is given by

$$f(x) = \overbrace{f(\bar{x}_0) + \nabla f(x_0) \cdot (\bar{x} - \bar{x}_0)}^{\text{affine lin. approx. to } f} + E_1(f : \bar{x}, x_0).$$

Ex: Find an approx. value for the number

$$\alpha := \sqrt{(3.03)^2 + (3.95)^2}$$

$$f: \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto \sqrt{x^2+y^2} \qquad \bar{x}_0 = (3,4)$$

$$f(x_0) = 5$$

$$f(3.03,\ 3.95) \cong f(3,4) + (\nabla f(3,4)) \cdot \underbrace{(0.03,\ -0.05)}_{x - x_0}$$

$x - x_0$: $(3.03, 3.95) \leftarrow \quad \rightarrow (3,4)$

$$\nabla f(3,4) = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix}\Bigg|_{(x,y)=(3,4)} = \begin{pmatrix} x/\sqrt{x^2+y^2} \\ y/\sqrt{x^2+y^2} \end{pmatrix} = \begin{pmatrix} 3/5 \\ 4/5 \end{pmatrix}$$

$$\alpha \cong 5 + \begin{pmatrix} 3/5 \\ 4/5 \end{pmatrix} \cdot \begin{pmatrix} 0.03 \\ -0.05 \end{pmatrix} \cong 4.978$$

actual value $= 4.97829\ldots$

For $f : \mathbb{R}^n \to \mathbb{R}$, $y = (y_1, \dots, y_n)$

let $T_1 f(y; x_0) :=$

$$f(x_0) + \nabla f(x_0) \cdot y$$

$T_1 f$ is called the Taylor poly of f at x_0 of order 1

$T_1 f(x - x_0; x_0)$ gives the first order approx to f at x_0. (This is a poly in n variables).

What about higher order approximations?

Defn. Let $f : X \to \mathbb{R}$, $f \in C^k$, $X \subset \mathbb{R}^n$, $x_0 \in X$. The k-th Taylor polynomial of f at x_0 is a poly in n-variables of degree $\leq k$ given by

$$T_k f(y; x_0) = f(x_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0) y_i + \dots + \sum_{m_1 + \dots + m_n = k} \frac{1}{m_1! \dots m_n!} \frac{\partial^k f}{\partial x_1^{m_1} \dots \partial x_n^{m_n}} y_1^{m_1} \dots y_n^{m_n}$$

$$= \sum_{\substack{m \\ |m| \leq k}} \frac{1}{m!} \partial_x^m f(x_0)\, y^m .$$

eg. $f: \mathbb{R}^2 \to \mathbb{R}$, $f \in C^2$

$x_0 \in \mathbb{R}^2$ $\quad y = (y_1, y_2)$

$$T_2 f(y, \overline{x_0}) = f(x_0) + \nabla f(x_0) \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \frac{1}{2!} (y_1 \ \ y_2) \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

$$= f(x_0) + \partial_{x_1} f(x_0) \cdot y_1 + \partial_y f(x_0) y_2 + \frac{1}{2!} \left(\frac{\partial^2 f(x_0)}{\partial x^2} y_1^2 + 2 \frac{\partial^2 f(x_0)}{\partial x \partial y} y_1 y_2 + \frac{\partial^2 f}{\partial y^2} y_2^2 \right).$$

Thm. Let $f \in C^k(X, \mathbb{R})$, $x_0 \in X$. Then we have

$$f(x) = T_k f(x - x_0; x_0) + E_k(f, x; x_0)$$

where $\frac{E_k(f, x, x_0)}{\|x - x_0\|^k} \xrightarrow[x \to x_0]{} 0$.

Ex: $f(x,y) = e^{x+y} \cos x \quad x_0 = (0,0)$

$f(0,0) = 1 \quad \left.\frac{\partial f}{\partial x}\right|_{(0,0)} = e^x \cdot e^y \cos x - e^{x+y} \sin x = 1.$

$\left.\frac{\partial f}{\partial y}\right|_{0,0} = 1.$

$$\left.\frac{\partial^2 f}{\partial x^2}\right|_{(0,0)} = 0$$

$$\left.\frac{\partial^2 f}{\partial y^2}\right|_{(0,0)} = 1$$

$$\left.\frac{\partial^2 f}{\partial x \partial y}\right|_{(0,0)} = 1$$

$$T_1 f = f(0,0) + \nabla f(0,0)\big((x,y)-(0,0)\big)$$

$$T_1 f = 1 + x + y$$

$$T_2 f = 1 + x + y + \frac{1}{2}(x\ y)\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

$$T_2 f = 1 + x + y + xy + \frac{1}{2}y^2$$

$$\left.\frac{\partial^3 f}{\partial x^3}\right|_{(0,0)} = -2 \qquad \frac{\partial^3 f}{\partial y^3} = 1 \qquad \frac{\partial^3 f}{\partial^2 x \partial y} = 0$$

$$\frac{\partial^3}{\partial^2 y \partial x} = 1$$

$$T_3 f(x,y) = 1 + x + y + xy + \frac{1}{2}y^2$$
$$+ \frac{1}{6}(-2)\, x^3$$
$$+ \frac{1}{6} \cdot 1 \cdot y^3$$
$$+ 0 \cdot x^2 y$$
$$+ \frac{1}{2!\,1!} 1 \cdot x y^2$$

$$T_3 f = 1 + x + y + xy + \frac{1}{2}y^2 - \frac{1}{3}x^3 + \frac{1}{6}y^3 + \frac{1}{2}xy^2.$$

$$C^k(X,Y) = \{ f : X \to Y \mid f \text{ is of class } C^k \}$$

/ik.

$1^5 \cdot t$

$1^3 \cdot t$

$1' \cdot t$

§ 3.8. Critical points and extrema of functions $f : \mathbb{R}^n \to \mathbb{R}$.

Def^n $f : X \subset \mathbb{R}^n \to \mathbb{R}$ diff. We say $x_0 \in X$ is a local maximum (min) if we can find a nbhd $B_r(x_0) = \{x \in \mathbb{R}^n \mid \|x - x_0\| < r\}$ st $B_r(x_0) \subset X$ and st $\forall x \in B_r(x_0)$

$$f(x) \leq f(x_0)$$

(resp $f(x) \geq f(x_0)$).

Recall $f : \mathbb{R} \to \mathbb{R}$. If f has a loc min or max at x_0 then $f'(x_0) = 0$.

Thm. Let $X \subset \mathbb{R}^n$ open, $f : X \to \mathbb{R}$ is diff. If $x_0 \in X$ is a local extrema (ie a min or max) then $\nabla f(x_0) = 0$

ie $$\frac{\partial f}{\partial x_1}(x_0) = \frac{\partial f}{\partial x_2}(x_0) = \cdots = \frac{\partial f}{\partial x_n}(x_0) = 0.$$

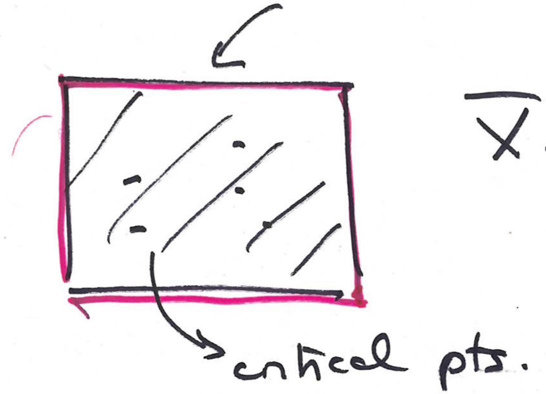
Defn A pt $x_0 \in X$ is called a **critical point** of f if $\nabla f(x_0) = 0$.

- Critical points are candidates for loc. extrema.

Defn A critical pt which is not a loc. min or max is called a **saddle point**.

Recall. If $f : [a,b] \to \mathbb{R}$ global extreme of f is either at an interior pt $x_0 \in (a,b)$ for which $f'(x_0) = 0$ or at $x=a$, $x=b$

Thm. If $f : \overline{X} \to \mathbb{R}$ diff on the interior of $\overline{X}$ where $\overline{X}$ is closed and bounded. Then ~~~~ global extrema of f exists and it is either at a critical point of f or on the boundary of $\overline{X}$.



$\overline{X} = \mathrm{int}(X) \cup \mathrm{bd}(X)$