Recall $f : \mathbb{R}^n \longrightarrow \mathbb{R}$

f C^k (on X open in $\mathbb{R}^n$) ($k \geq 1$)

$$\Updownarrow$$

every partial derivative of order $\leq k$ exists on X and is a continous function on X

Then partial derivatives with respect to different variables commute:

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

77

$m = (m_1, \ldots, m_n)$, $\quad m_i \geq 0$ integers

$m_1 + \cdots + m_n = k$

$$\frac{\partial^m f}{\partial x_1^{m_1} \cdots \partial x_n^{m_n}} = \underbrace{\partial_{x_1} \cdots \partial_{x_1}}_{m_1 \text{ times}} \underbrace{\partial_{x_2} \cdots \partial_{x_2}}_{m_2 \text{ times}} \cdots \underbrace{\partial_{x_n} \cdots \partial_{x_n}}_{m_n \text{ times}} f$$

3.7. Taylor polynomials

Motivation: $n = 1$

(I open) $f : I \longrightarrow \mathbb{R}$, C^k

f is **well - approximated** close to $x_0 \in I$ by

$$f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2 + \dots$$
$$+ \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k$$

In ~~IR~~ n variables:

Def. (3.7.1) $k \geq 1$, $X \subset \mathbb{R}^n$ open

$f : X \longrightarrow \mathbb{R}$ class C^k

$x_0 \in X$

Taylor polynomial of degree k of f at x_0: functions of $y = (y_1, \ldots, y_n)$ given by:

$$T_k f(y; x_0) = f(x_0) + \underbrace{\sum_{i=1}^{n} \partial_{x_i} f(x_0)\, y_i}_{\nabla f(x_0) \cdot y}$$

$$+ \cdots\cdots +$$

$$\sum_{\substack{m_i \geq 0 \\ m_1 + \cdots + m_n = k}} \frac{1}{m_1! \cdots m_n!} \, \partial_m^k f(x_0) \, y_1^{m_1} \cdots y_n^{m_n}$$

$(m_1, \ldots, m_n)$

(Term of degree j, $0 \leq j \leq k$, is the same as ◯ with k replaced by j).

Using multi-index notation:

$$m = (m_1, \dots, m_n), \qquad m_i \geq 0$$

Write
$$\begin{cases} |m| = m_1 + \dots + m_n \\ m! = m_1! \cdot \ldots \cdot m_n! \\ y^m = y_1^{m_1} \cdot \ldots \cdot y_n^{m_n}, \quad y \in \mathbb{R}^n \end{cases}$$

$$\frac{1}{m_1! \cdot \ldots \cdot m_n!} \partial_m^{j} f(x_0)\, y_1^{m_1} \cdot \ldots \cdot y_n^{m_n}$$

$$= \frac{1}{m!} \partial_m^{|m|} f(x_0)\, y^m$$

With these notation:

$$T_k f(y; x_0) = \sum_{|m| \leq k} \frac{1}{m!} \partial_m^{|m|} f(x_0)\, y^m$$

sum over all $(m_1, \dots, m_n)$
$m_i \geq 0$, $|m| \leq k$

<u>$n = 1$:</u>

$$\sum_{m=0}^{k} \frac{1}{m!} f^{(m)}(x_0)\, y^m$$

Example:

$k=1$:

scalar product in $\mathbb{R}^n$

$$T_1 f(y; x_0) = f(x_0) + \nabla f(x_0) \cdot y$$
$$= f(x_0) + \sum \partial_{x_i} f(x_0) y_i$$

$k=2$: f C^2

$$T_2 f(y; x_0) = f(x_0) + \nabla f(x_0) \cdot y$$
$$+ \frac{1}{2} \sum_{i=1}^{n} \partial_{x_i}^2 f(x_0)\, y_i^2 + \sum_{1 \le i < j \le n} \partial_{x_i}\partial_{x_j} f(x_0)\, y_i y_j$$

$m = (0, \dots, 0, 2, 0, \dots 0)$ (i-th variable)

$m! = 2!$

$m = (0, \dots, 0, 1, 0, \dots, 1, 0, \dots, 0)$ (ith, j-th)

$m! = 1$

<u>Recall:</u> f C^2

$$\mathrm{Hess}_f(x_0) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0)\right)_{1 \le i,j \le n}$$

"Hessian"; <u>symmetric</u> matrix

The quadratic term in T_2 is

$$\frac{1}{2} y^t \underbrace{\mathrm{Hess}_f(x_0)}_{n \times n}\, y$$

transpose = row vector

$y \in \mathbb{R}^n$ column vector

[linear algebra]

So:

$$\boxed{T_2 f(y; x_0) = f(x_0) + Df(x_0) \cdot y + \frac{1}{2} y^t \, \mathrm{Hess}_f(x_0) \, y}$$

<u>Ex:</u> $n = 2$

$$\mathrm{Hess}_f(x_0) = \begin{pmatrix} a & b \\ b & d \end{pmatrix} = H$$

$$(y_1 \ \ y_2) H \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = (y_1 \ y_2) \begin{pmatrix} ay_1 + by_2 \\ by_1 + dy_2 \end{pmatrix} = $$

$$= ay_1^2 + by_1y_2 + by_1y_2 + dy_2^2$$

So

$$\frac{1}{2} y^t H y = \frac{1}{2} a y_1^2 + b y_1 y_2 + \frac{1}{2} d y_2^2$$

Compare with:

$$\frac{1}{2} \partial_{x_1}^2 f(x_0) y_1^2 + \frac{1}{2} \partial_{x_2}^2 f(x_0) y_2^2 + \partial_{x_i x_j}^2 f(x_0) y_1 y_2$$

$$H = \begin{pmatrix} \partial_{x_1}^2 f & \partial_{x_1 x_2} f \\ \partial_{x_1 x_2} f & \partial_{x_2}^2 f \end{pmatrix} = \mathrm{Hess}_f(x_1, x_2)$$

<u>Prop.</u> 3.7.3

$k \geq 1$, $X \subset \mathbb{R}^n$ open

$f : X \longrightarrow \mathbb{R}$ <u>class C^k</u>

Write

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

Then

$$\lim_{\substack{x \to x_0 \\ x \neq x_0}} \frac{E_k f(x; x_0)}{\|x - x_0\|^k} = 0$$

<u>Ex.</u> (3.7.4)

$n = 2$, $f(x,y) = e^{3x - \sin(xy)}$

$(x_0, y_0) = (0,0)$, $f(0,0) = 1$

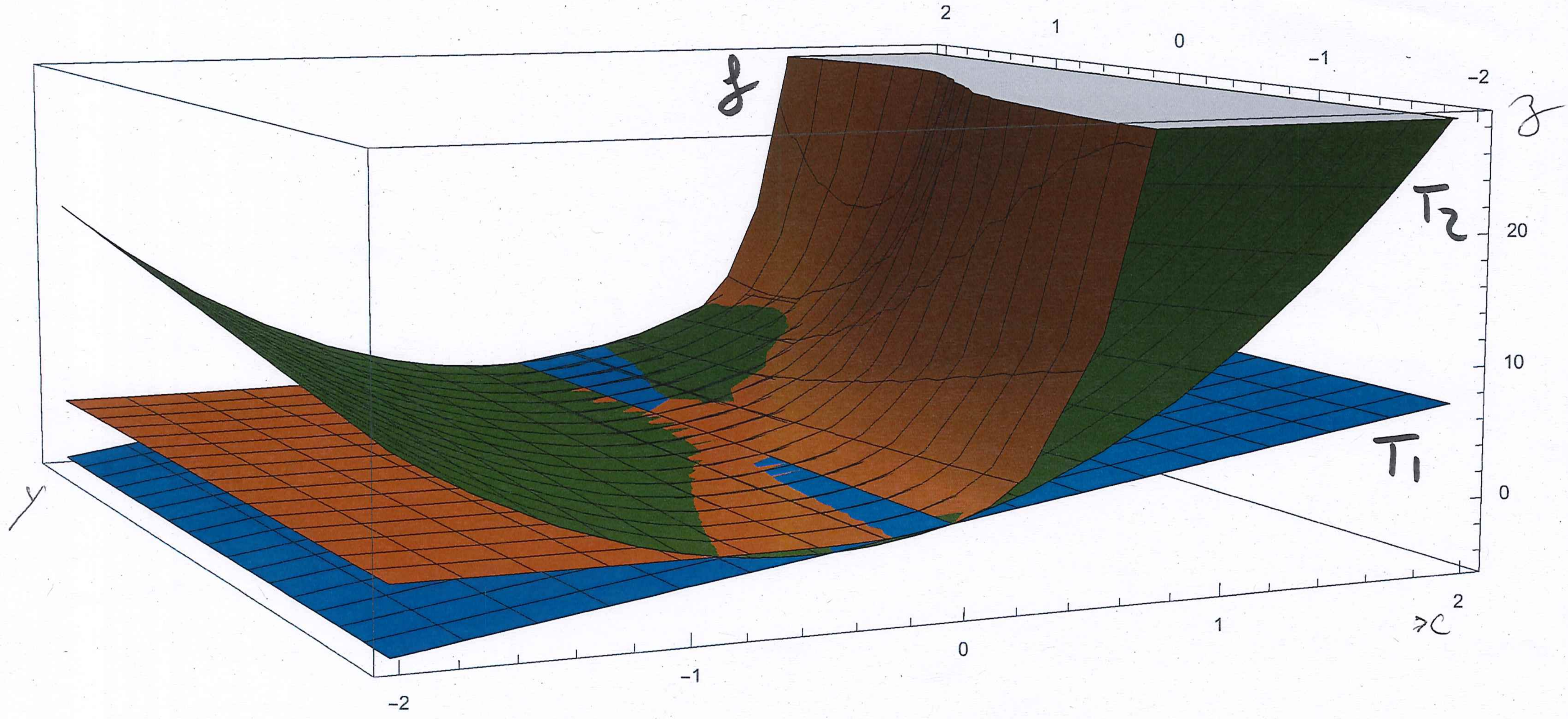
$$T_2 f(\cancel{x}, (0,0)) = 1 + \underbrace{3x}_{\text{gradient}} + \underbrace{\frac{9x^2}{2} - xy}_{\text{Hessian}}$$

(x_1, x_2)

$f(-0{,}0015 \,;\, 0{,}003) \simeq 0{,}9955145896\ldots$

$T_1 f(\;'' \;,\; ''\;) = 0{,}9955$

$T_2 f(\;'' \;,\; ''\;) = 0{,}995574625$


f
z
T2
T1
y
x
20
10
0
2
1
0
-1
-2
2
1
0
-1
-2

3.8 – Critical points

Goal: analogue of

(1) if $f : \underset{(\text{open})}{I} \longrightarrow \mathbb{R}$ is differentiable, and f max./min. at $x_0 \in I$ (local extremum) then $f'(x_0) = 0$

(2) if f is C^2, $f'(x_0) = 0$, $f''(x_0) > 0$, then f has a local minimum at x_0

Prop. 3.8.1. Let $X \subset \mathbb{R}^n$ $\boxed{\text{open}}$

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

If $x_0 \in X$ satisfies:

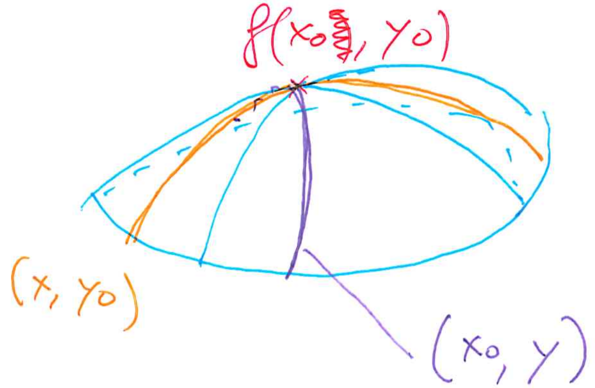
(1) $f(x) \leq f(x_0)$ if x close to x_0 (local maximum at x_0)

or

(2) $f(x) \geq f(x_0)$ ______ (local minimum)

Then $\nabla f(x_0) = 0$ $\left(\Leftrightarrow \dfrac{\partial f}{\partial x_i}(x_0) = 0 \text{ if } 1 \leq i \leq n \right)$

Proof. $\underline{n = 2}$



$$g(t) = f(x_0, y_0 + t)$$
(t small enough so that $(x_0, y_0 + t) \in X$)

$$g(0) = f(x_0, y_0)$$

g has local max. at $t = 0$

(Analysis I) $\Rightarrow$ $g'(0) = 0$

$$g'(0) = \frac{\partial f}{\partial y}(x_0, y_0) = 0$$

<u>Def.</u> 3.8.2.

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

A point x_0 such that $\nabla f(x_0) = 0$ is called a <u>critical point</u> of f.

So if f has a local extremum at a point x_0, this <u>must</u> be a critical point.

(Converse is false, already for $n=1$.)

How to find where $f: X \longrightarrow \mathbb{R}$ is maximal / minimal? (f C^1)

(1) f might not have a max./min (because X is ~~usually~~ not compact)

(2) Typically, $\overline{X}$ compact is given, closed + bounded

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

($\Rightarrow$ there is a max./min, Prop. 3.8.1)

(3) Write $\overline{X} = \underset{\text{open}}{X} \cup B$ "boundary"

(4) if f is differentiable on X, then the point(s) x_0 where $f(x_0)$ is maximal (or minimal) are either

(4i) a critical point of f on X

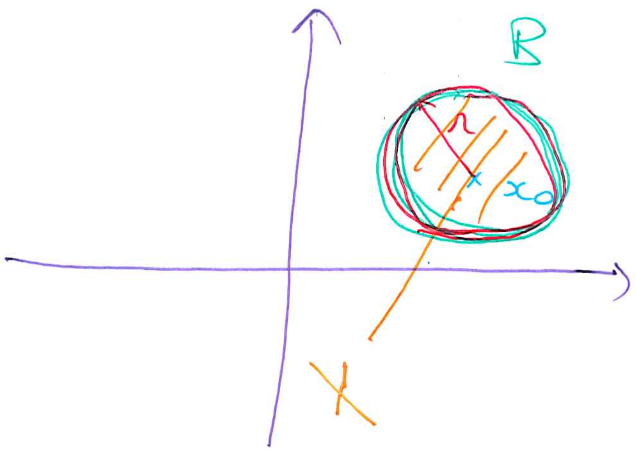
(4ii) a point of B

(5) So: we solve for $\nabla f(x_0) = 0$, evaluate f at those critical points, and evaluate f on B, and compare.

Ex. $\quad x_0 \in \mathbb{R}^n, \ r > 0$

(1) $\quad \overline{X} = \{ x \mid \|x - x_0\| \leq r \}$

(ball, **closed**, around x_0, of radius r)



$$\overline{X} = X \cup B$$

$$X = \{ x \mid \|x - x_0\| < r \}$$

$$B = \{ x \mid \|x - x_0\| = r \}$$

(2) [3.8.4]

97

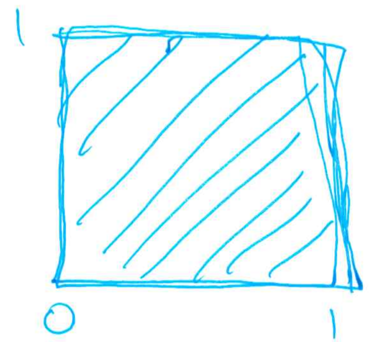
$\overline{X} = [0,1]^2$

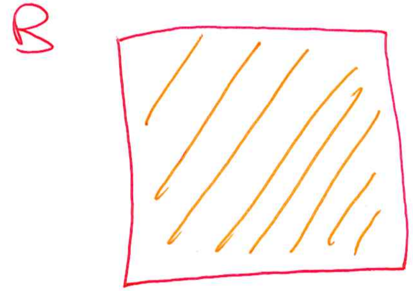
$X =]0,1[^2$

$B =$ union of four segments

$(x,0), (x,1)$ or $(0,y), (1,y)$

$0 \le x \le 1$ $\quad$ $0 \le y \le 1$





$$f(x,y) = x^2 - 2y^2$$

$$\nabla f(x,y) = \begin{pmatrix} 2x \\ -4y \end{pmatrix}$$

$(0,0)$ is the only critical point and $f(0,0) = 0$ (will not be max. or min)

Evaluate on B:

$f(x,0) = x^2$, $f(x,1) = x^2 - 2$ $\quad 0 \le x \le 1$

$f(0,y) = -2y^2$, $f(1,y) = 1 - 2y^2$ $\quad 0 \le y \le 1$

Max. of f on each of the line segments:

	1	-1
	0	1
Min:	0	-2
	-2	-1

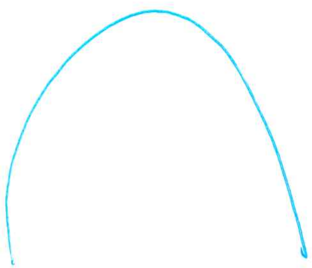
maximal value of f at $(1,0)$

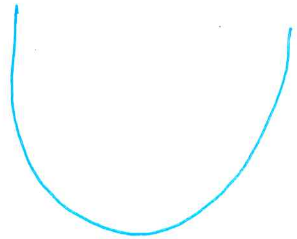
minimal value at $(0,1)$

Q. when does f, C^2, have a local extremum at a critical point x_0?

$n = 1$

$f''(x_0) < 0$

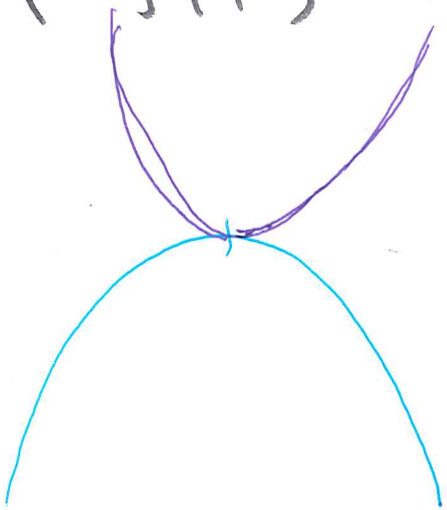




$f''(x_0) > 0$

Näive analogue: "if $\mathrm{Hess}_f(x_0)$ is $\neq 0$, we can say that f has a local extremum".

THIS IS FALSE!



$\partial^2_{y^2} f(x_0) > 0$

$\partial^2_{x^2} f(x_0) < 0$

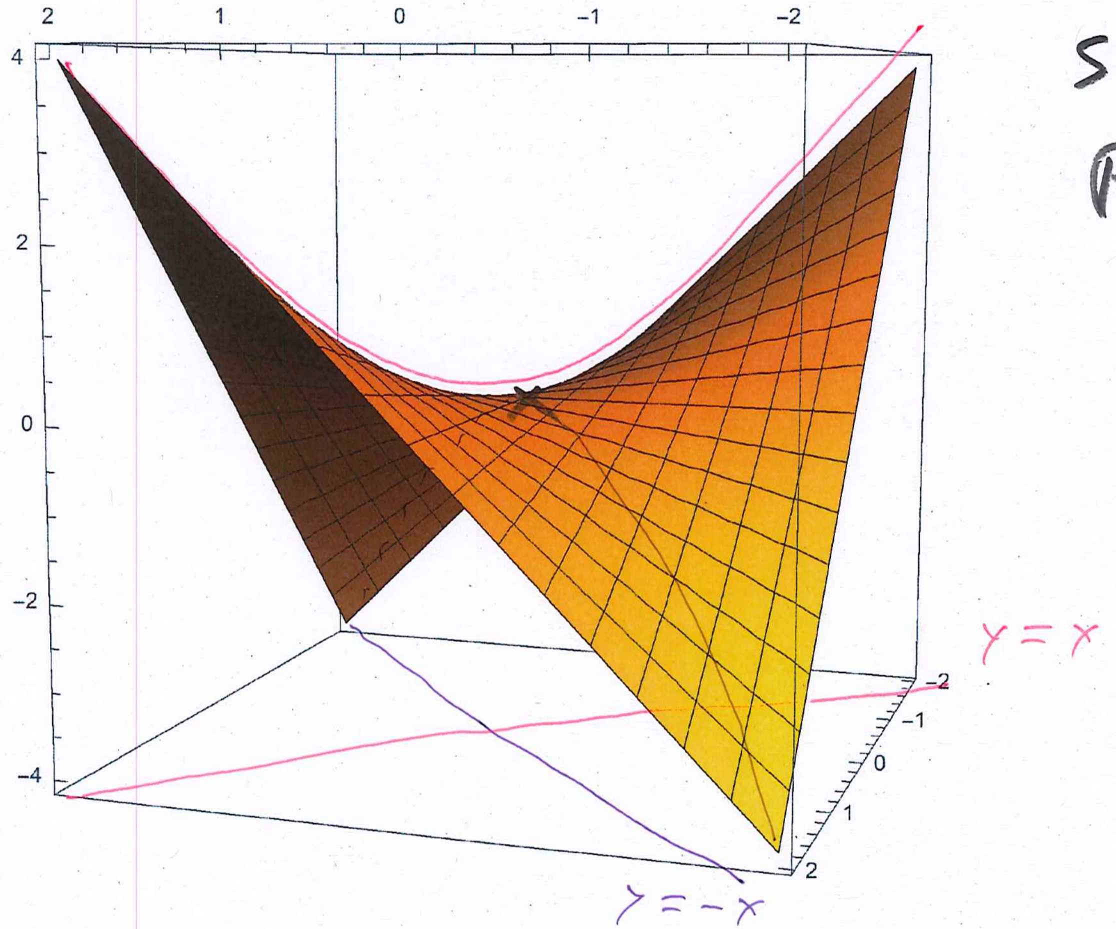
$x_0 = (0,0)$ is a critical point

"saddle point"

$f(x,y) = xy$

$f(x,x) = x^2$

$f(x,-x) = -x^2$



Def. (3.8.6) f C^2

A critical point of f is **non-degenerate** if $\det \mathrm{Hess}_f(x_0) \neq 0$.

Remark: for degenerate critical points, the problem is much more complicated.

Ex: $f_1(x,y) = x^4 + y^4$, $f_2(x,y) = x^4 - y^4$
$f_3(x,y) = -x^4 - y^4$
have $(0,0)$ as degenerate critical point

Suppose x_0 is a non-degenerate critical point of f.

$H = \mathrm{Hess}_f(x_0)$ is a symmetric invertible matrix

$\hookrightarrow$ (linear algebra) H can be diagonalized in an orthonormal basis $(v_1, \dots, v_n)$

$y = t_1 v_1 + \dots + t_n v_n$

$\hookrightarrow$ $y^t \, \mathrm{Hess}_f(x_0) \, y = \lambda_1 t_1^2 + \dots + \lambda_n t_n^2$

where λ_i is the eig. of v_i