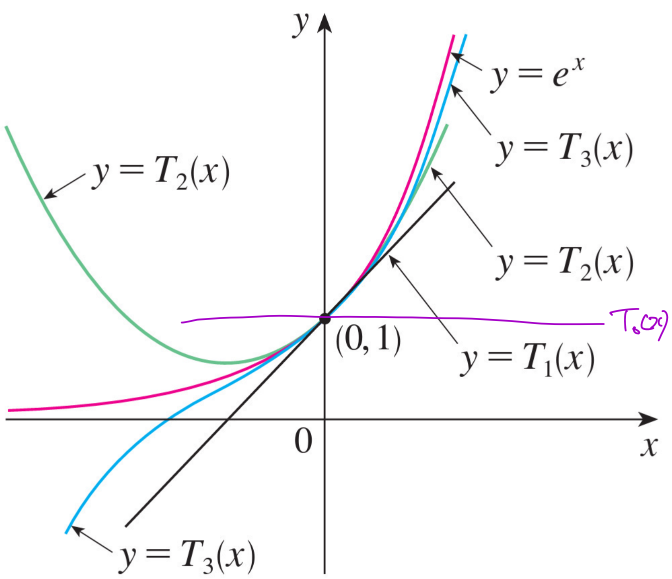
# 3.7 Taylor theorem for multivariate fn.

/* We will learn multi-dimensional (Ch. 4.2) version of Taylor theorem, which provides higher order approximation of a function than linear one. */

## Brief review of 1D Taylor formula




$f(x) = e^x$

$T_0(x) = 1$

$T_1(x) = 1 + x$

$T_2(x) = 1 + x + \frac{x^2}{2}$

$T_3(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6}$

$T_4(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{4!}$

$\vdots$

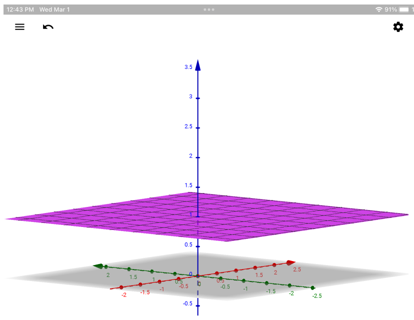
" $T_n(x) \longrightarrow f(x)$ "

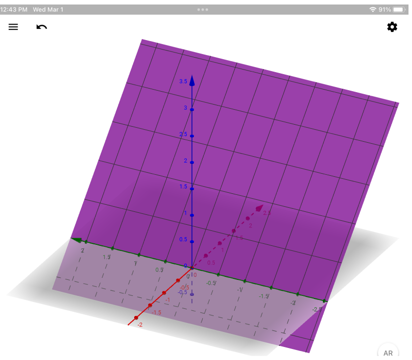
⊙ In 2D, things are more complicated.

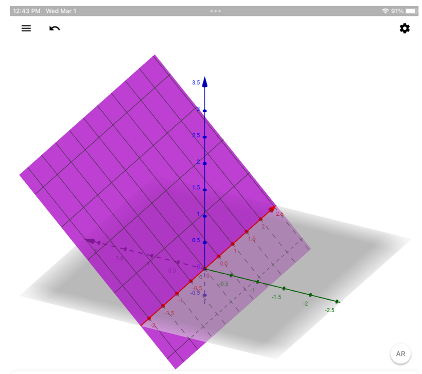
But the idea is the same

# 2D polynomials

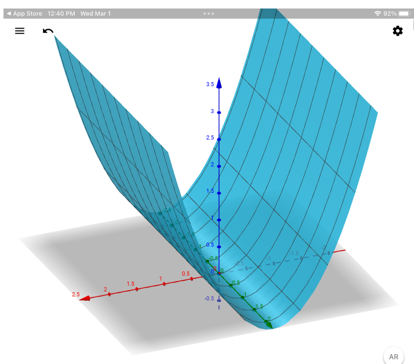
$z = 1$ $\qquad$ $z = x$ $\qquad$ $z = y$

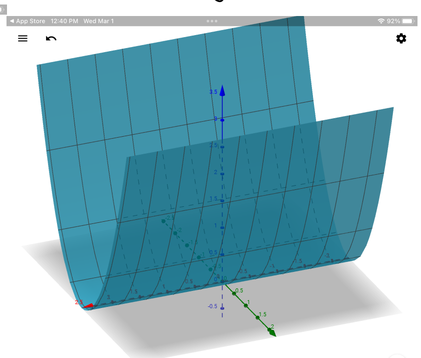


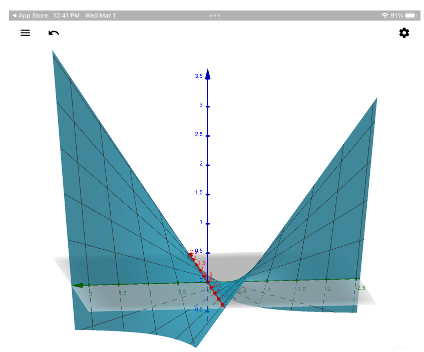




$z = x^2$ $\qquad$ $z = y^2$ $\qquad$ $z = xy = yx$

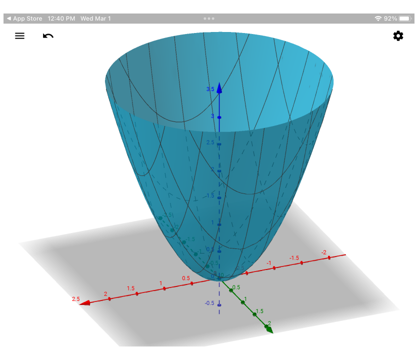


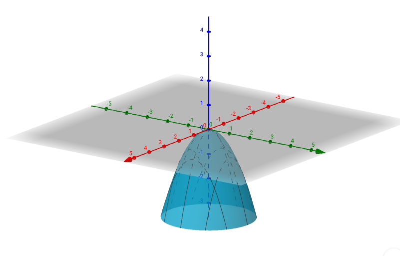




paraboloid

$z = x^2 + y^2$ $\qquad$ $z = -x^2 - y^2$





Given $f(x,y)$, let

$$T_1(x,y) = f(x_0,y_0) + \nabla f(x_0,y_0)\cdot(x-x_0,\, y-y_0)$$
$$= f + \frac{\partial f}{\partial x}\cdot\Delta x + \frac{\partial f}{\partial y}\Delta y$$

be the linear approximation near $\vec{a} = (x_0, y_0)$ (input omitted $\rightarrow$ evaluate at $(x_0,y_0)$)

What do you think should you add to get a better approximation?

$$T_2(x,y) = f \qquad \rightarrow 0^{th}\text{-order appr.}$$
$$+ \frac{\partial f}{\partial x}\cdot\Delta x + \frac{\partial f}{\partial y}\Delta y \qquad \rightarrow 1^{st}\text{-order appr.}$$

linear appr.

$$+ \underline{\frac{1}{2}\frac{\partial^2 f}{\partial x^2}}\,\underline{\Delta x^2} + \underline{\frac{1}{2}\frac{\partial^2 f}{\partial y\,\partial x}}\,\underline{\Delta x\,\Delta y}$$
$$+ \underline{\frac{1}{2}\frac{\partial^2 f}{\partial x\,\partial y}}\,\underline{\Delta y\,\Delta x} + \underline{\frac{1}{2}\frac{\partial^2 f}{\partial y^2}}\,\underline{\Delta y^2}$$

$2^{nd}$ order $\Rightarrow$ appr.

quadratic increments.

quadratic appr.

"right" coeffients

$$= f + \nabla f\cdot\Delta\vec{x} + \frac{1}{2}\Delta\vec{x}^{\,T} H_f\,\Delta\vec{x},$$

where $\Delta\vec{x} = \begin{bmatrix}\Delta x\\ \Delta y\end{bmatrix} = \begin{bmatrix}x-x_0\\ y-y_0\end{bmatrix}$ is the increment vector and $H_f$ is

$$H_f(x_0, y_0) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(x_0, y_0) & \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) \\ \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) & \frac{\partial^2 f}{\partial y^2}(x_0, y_0) \end{bmatrix}.$$

This is called **Hessian** matrix of $f$ at $\vec{a} = (x_0, y_0)$. corresponds to $f''$ in 1D.

$\boxed{\text{Clicker}}$ If $\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}$, $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, what is $\vec{x}^{T} A \vec{x}$? transpose

$\vec{x}^T = [x \ \ y]$

(A) $ax^2 + bx + cy + dy^2$

(B) $ax^2 + (b+c)xy + dy^2$ ✓

(C) $ax + by + cxy + d$

$$\vec{x}^T A \vec{x} = [x \ \ y] \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$
$$= [x \ \ y] \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}$$
$$= x(ax + by) + y(cx + dy)$$
$$= ax^2 + bxy + cxy + dy^2$$

$$\begin{bmatrix} ax^2 & bxy \\ bxy & cy^2 \end{bmatrix}$$
$$\downarrow$$
$$ax^2 + 2bxy + cy^2$$

⓪ $H_f$ is always symmetric if $f$ is smooth enough. (Recall $f_{xy} = f_{yx}$)

$\boxed{\text{Thm}}$ (2nd order Taylor formula for 2D)

Suppose $f: \mathbb{R}^2 \to \mathbb{R}$ is a $C^2$-fn, i.e. it has continuous partial's up to order 2.
Then, it is the case that

$$f(\vec{x}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a}) + \frac{1}{2} (\vec{x} - \vec{a})^T H_f(\vec{a}) (\vec{x} - \vec{a}) + R_2(\vec{x}, \vec{a}), \qquad (*)$$

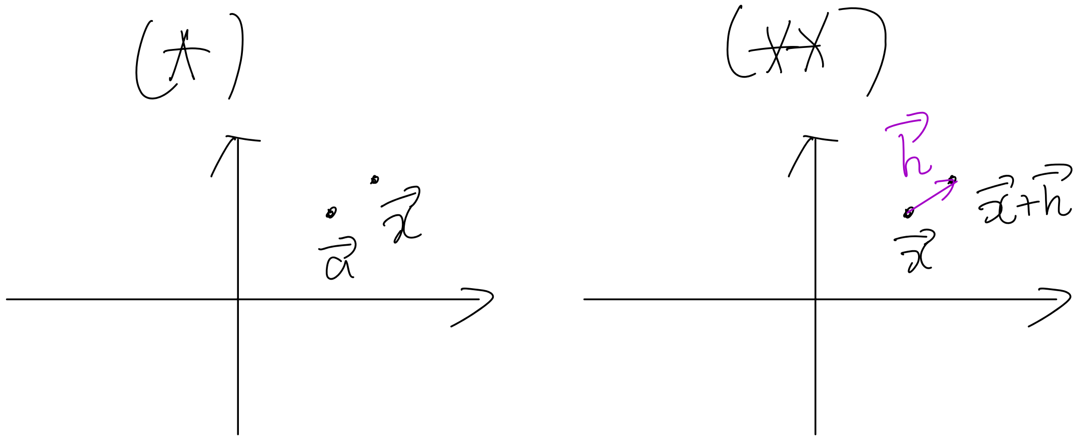
where

$$\lim_{\vec{x} \to \vec{a}} \frac{|R_2(\vec{x}, \vec{a})|}{\|\vec{x} - \vec{a}\|^2} = 0.$$

◎ The main equality can be written as follows by putting $\vec{a} \leftarrow \vec{x}$, $\vec{h} = \vec{x} - \vec{a}$.

$$f(\vec{x} + \vec{h}) = f(\vec{x}) + \nabla f(\vec{x}) \cdot \vec{h} + \frac{1}{2} \vec{h}^T H_f(\vec{x}) \vec{h} + R_2(\vec{x}, \vec{h}) \qquad (**)$$

This equivalent expression focus more on the increment $\vec{h}$ while the previous one focus more on the new input




⊙ 1st order Taylor formula is nothing but rewriting of differentiability.

$$f(\vec{x}) = \underbrace{f(\vec{a}) + \nabla f(\vec{a})\cdot(\vec{x}-\vec{a})}_{\text{lin. appr.}} + \underbrace{R_1(\vec{x},\vec{a})}_{\text{error}}$$

$$|R_1(\vec{x},\vec{a})| / \|\vec{x}-\vec{a}\| \longrightarrow 0 \quad \text{as} \quad \vec{x} \to \vec{a}.$$

Question: Does the following look natural?
(Substitute for the proof: see textbook pp. 235–237 for a more formal argument)

$$|R_1(\vec{x},\vec{a})| / \|\vec{x}-\vec{a}\| \longrightarrow 0 \quad \text{as} \quad \vec{x} \to \vec{a}$$

$$|R_2(\vec{x},\vec{a})| / \|\vec{x}-\vec{a}\|^2 \longrightarrow 0 \quad \text{as} \quad \vec{x} \to \vec{a}$$

**Example**: Compute $T_2(x,y)$ for

$$f(x,y) = \sin(x+y) + \cos(x-3y)$$

at $(x_0, y_0) = (0,0)$.

order partials

$0^{th}$. $f(0,0) = \sin(0) + \cos(0) = 1$

$1^{st}$. $f_x(0,0) = \cos(x+y) - \sin(x-3y) \Big|_{\substack{x=0\\y=0}}$

$= \cos(0) - \sin(0)$

$= 1$

. $f_y(0,0) = \cos(x+y) - \sin(x-3y)\cdot(-3)$

$= \cos(x+y) + 3\sin(x-3y) \Big|_{\substack{x=0\\y=0}}$

$= \cos(0) + 3\cdot\sin(0)$

$= 1$

$2^{nd}$. $f_{xx}(0,0) = -\sin(x+y) - \cos(x-3y) \Big|_{\substack{x=0\\y=0}}$

$= -\sin(0) - \cos(0)$

$= -1$

- $f_{xy}(0,0) = -\sin(x+y) - \cos(x-3y)\cdot(-3)\Big|_{\substack{x=0\\y=0}}$

$$= -\sin(0) + 3\cos(0)$$

$$= 3$$

- $f_{yy}(0,0) = -\sin(x+y) + 3\cos(x-3y)\cdot(-3)\Big|_{\substack{x=0\\y=0}}$

$$= -\sin(0) - 9\cos(0)$$

$$= -9$$

$$T_2(x,y) = f(0,0) + \nabla f(0,0)\begin{bmatrix} x \\ y \end{bmatrix} + \frac{1}{2}\begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

$(x,y)$

$(0,0)$

$$= 1 + \begin{bmatrix} 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

$$+ \frac{1}{2}\begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} -1 & 3 \\ 3 & -9 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

$$= 1 + x + y - \frac{1}{2}x^2 + 3xy - \frac{9}{2}y^2$$

Exercise: graph $f(x,y)$ and $T_2(x,y)$

Exercise : Find $1^{st}$ and $2^{nd}$ order approximation of $f(x,y) = ye^{-x^2} + 2$ at $(1,0)$.

How to check? Plot them!

⓪ Though $H_f$ is symmetric, the correct one is $\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$, not $\begin{bmatrix} f_{xx} & f_{yx} \\ f_{xy} & f_{yy} \end{bmatrix}$

Think of this as follows. $\to [f_x \ f_y]$

① Find $\nabla f$. It is row vector $(1\times 2)$ as a derivative.

$D\begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$

② Change this to a column to see it as a vector field (output must be viewed as a column). This means we now see $\nabla f$ as a target to differentiate, not as a derivative.

③ Find $D(\underset{\text{(column)}}{\nabla f})$ as a derivative of a vector field. Then, everything works out.

# NT. 3.8 Max/Min (Extrema) of bivariate fn's.

/* Do I need to motivate the significance of max/min? */

## Basic definitions

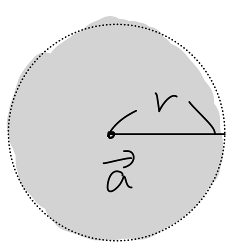
Given a function $f: D \to \mathbb{R}$, where $D \subseteq \mathbb{R}^m$ (mostly interested in $m=1$ or $2$)

- an **open ball** $B_r(\vec{a})$ of radius $r>0$ centered at $\vec{a} \in \mathbb{R}^m$ is defined

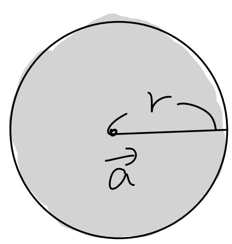
$$B_r(\vec{a}) = \{\vec{x} \in \mathbb{R}^m : \|\vec{x}-\vec{a}\| < r\}.$$

all points whose distance from $\vec{a}$ is strictly less than $r$

- a **closed ball** $\overline{B_r}(\vec{a})$ is given by $B_r(\vec{a}) = \{\vec{x} \in \mathbb{R}^m : \|\vec{x}-\vec{a}\| \leq r\}$.

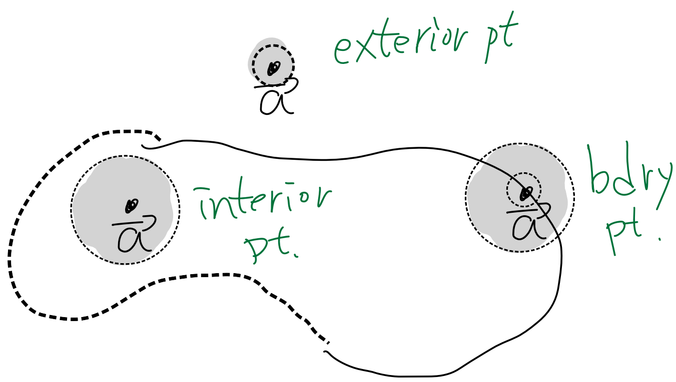


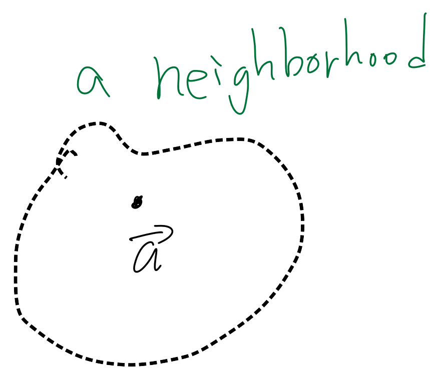

$B_r(\vec{a})$




$\overline{B_r}(\vec{a})$

⊙ $\vec{a} \in D$ is called an **interior point** if some open ball of $\vec{a}$ is contained in D: $B_r(\vec{a}) \subseteq D$, **exterior point** if some open ball is contained outside of D. If **any** open ball centered at $\vec{a}$ contains a point both inside and outside of D, $\vec{a}$ is called a **boundary point**.

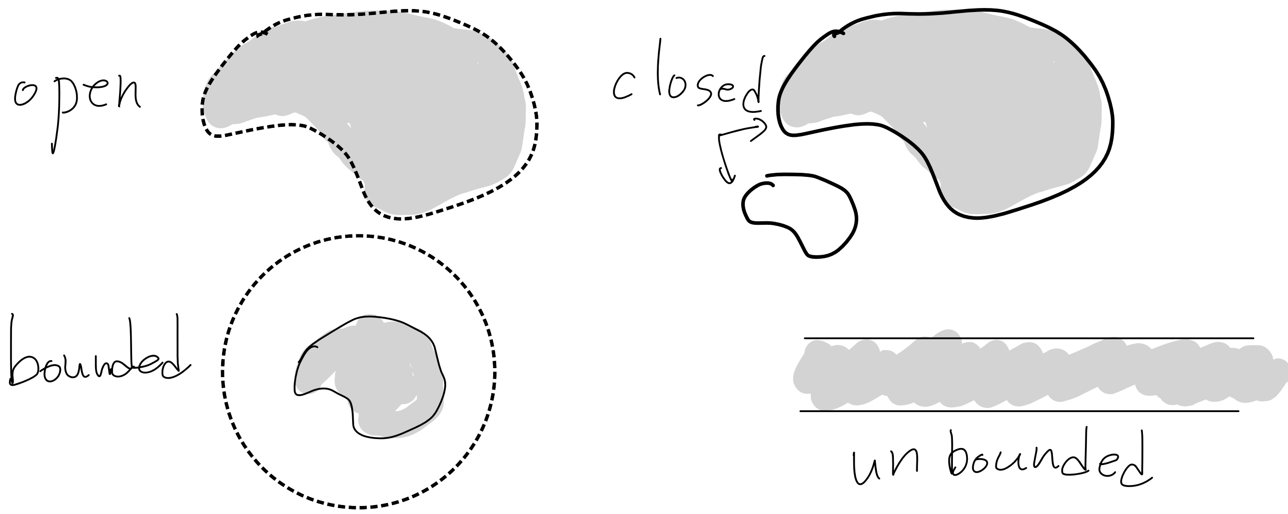



⊙ If $D \subseteq \mathbb{R}^m$ (mostly $m=2$ here) has only interior points, D is called **open**.

⊙ A **neighborhood** of $\vec{a} \in \mathbb{R}^m$, is an **open** region containing $\vec{a}$.

◎ If $D \subseteq \mathbb{R}^n$ (mostly $n=2$ here) has all boundary points, $D$ is called **closed**.




◎ In mathematics, "region" and "domain" have more specific meaning in this context. But we don't bother with those. We always think of them as a "nice connected area."

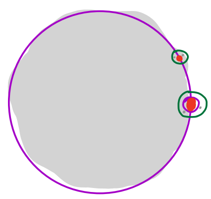
◎ If $D \subseteq \mathbb{R}^n$ (mostly $n=2$ here) is contained in some ball, it is called **bounded**. If not, it is **unbounded**.

$\boxed{\text{Clicker}}$ For $\{(x,y): x^2+y^2 \le 1,\ (x,y) \ne (1,0)\}$,

the point $(1,0)$ is

(A) interior pt (B) exterior pt

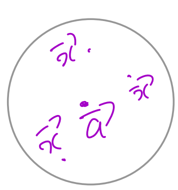
(C) boundary pt (D) none of these



$\boxed{\text{Clicker}}$ For $\{(x,y): x^2+y^2 \le 1,\ (x,y) \ne (1,0)\}$,

this is

(A) closed (B) open (C) neither



◎ If $f(\vec{x}) \geq f(\vec{a})$ (resp. $f(\vec{x}) \leq f(\vec{a})$) for all $\vec{x} \in B_r(\vec{a})$ on some neighborhood of $\vec{a}$, we say $f$ has a local / relative **minimum (resp. maximum)** at $\vec{a}$. $\vec{a}$ is called a **local minimizer (resp. maximizer)**

◎ If $f(\vec{x}) \geq f(\vec{a})$ (resp. $f(\vec{x}) \leq f(\vec{a})$) for all $\vec{x} \in D$, i.e. over the whole domain, we say $f$ has an **absolut / global minimum (resp. maximum)** at $\vec{a}$. $\vec{a}$ is called a **global minimizer (resp. maximizer)**

/* I will often call "minimum" (output) for "minimizer" (input) when there is no important reason to distinguish input and output. If you want to clarify, ask me. */

## Goal: Extend Cal1 to multivarite fn's

Review of Cal 1 on extreme values.

⊙ If $f'(a)=0$ or $f'(a)$ does not exist, $\vec{a}$ is called a **critical point**.

$\boxed{\text{Thm}}$ (Extreme value thm 1D)

If $f$ is continuous on $[a,b]$, then it has extrema : the maximum at $x_1 \in [a,b]$ and minimum at $x_2 \in [a,b]$. (Global)

$\boxed{\text{Thm}}$ (Fermat)

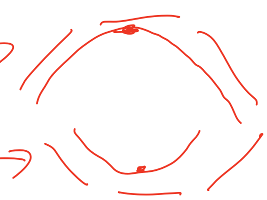
If $f$ has a local minimum or maximum at $x=a$ and $f$ is diff. there, then $f'(a)=0$.

$\boxed{\text{Thm}}$ (Second der. test)

Suppose $C^2$-fn $f$ has a critical point at $x=a$. Then, it has a

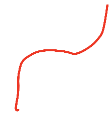
local max if $f''(a)<0 \rightarrow$

local min if $f''(a)>0 \rightarrow$



If $f''(a)=0$, we cannot decide.

# Max/min in 2D



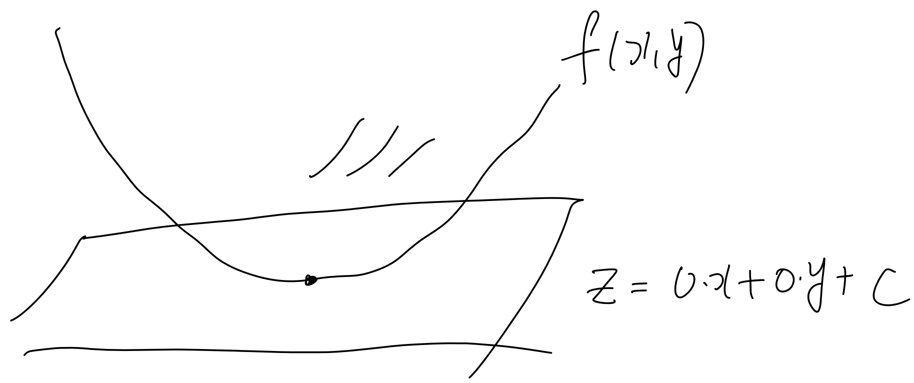
$x^3$

$\boxed{\text{Thm}}$ (Extreme value thm) (in 1D, $[a,b]$)

Suppose $D \subseteq \mathbb{R}^2$ is closed and bounded. Assume $f : D \to \mathbb{R}$ is continuous. Then, $f$ has a (global) maximum and a (global) minimum.

$\boxed{\text{Thm}}$ (Fermat 2D ver)

Suppose $f : D \to \mathbb{R}$ is differentiabe on a domain $D \subseteq \mathbb{R}^2$ and $f$ has a local max/min at interior point $\vec{a} \in D$. Then, $\nabla f(\vec{a}) = \vec{0}$.




⓪ If $\nabla f(\vec{a}) \neq \vec{0}$, there are always larger and smaller outputs nearby. (Why?)