Section 2.3

<u>Def</u> Let $a \in X$ and $f: X \subset \mathbb{R}^n \to \mathbb{R}$ be a real-valued func on X. The ~~the~~ partial derivative of f with respect to x_i ~~at the~~ at the point $\vec{x} = \vec{a}$ is the real number

$$\frac{\partial f}{\partial x_i}(\vec{a}) = \lim_{h \to 0} \frac{f(a_1, \ldots, a_{i-1}, a_i + h, a_{i+1}, \ldots, a_n) - f(\vec{a})}{h}$$

This is a Calc I limit!

and the partial derivative of f with respect to x_i is the real-valued function

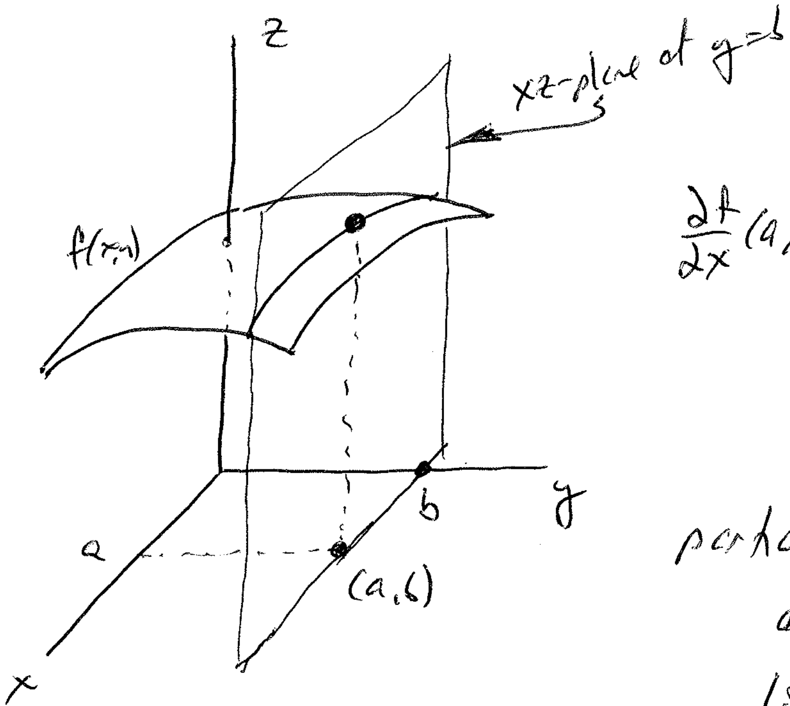
$$\frac{\partial f}{\partial x_i}(\vec{x}) = \lim_{h \to 0} \frac{f(x_1, \ldots, x_{i-1}, x_i + h, x_{i+1}, \ldots, x_n) - f(\vec{x})}{h}$$

<u>Notes</u> ① It is simply the ordinary derivative of f wrt x_i found by fixing all x_j, $j \neq i$ and ~~considering~~ varying <u>only</u> x_i.

② Alternate notation: $D_{x_i} f(\vec{x})$, $f_{x_i}(\vec{x})$

③ Geometrically, given $f(x, y)$ and $\vec{x} = \begin{bmatrix} a \\ b \end{bmatrix}$, the slice through graph(f) at $y = b$ is a 1-d curve inside the xz-plane at $y = b$. $\frac{\partial f}{\partial x}(a, b)$ is the slope of curve here @ (a, b):

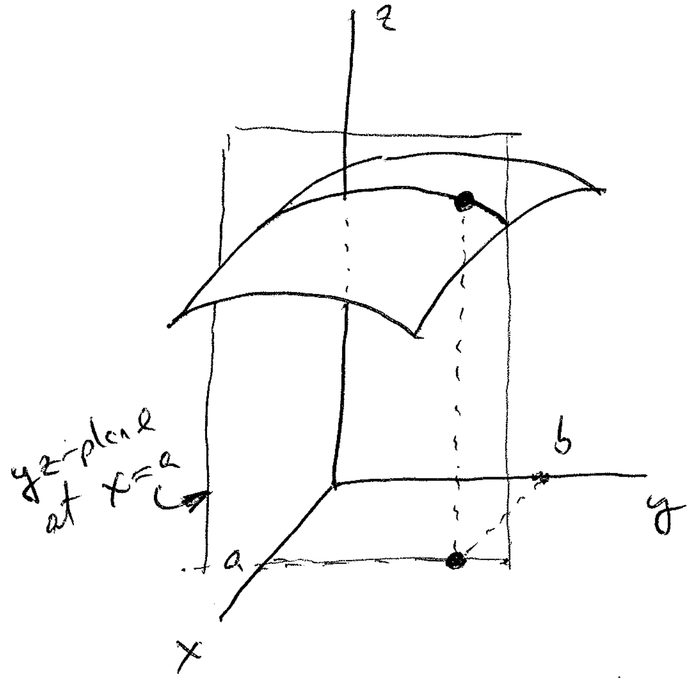
II



$$\frac{\partial f}{\partial x}(a,b) = \lim_{h\to 0} \frac{f(a+h,b) - f(a,b)}{h}$$

here a can vary but b does not.

partial derivative of f wrt x at (a,b) is slope of tangent line to graph(f) ∩ {xz-plane at y=b}.



$$\frac{\partial f}{\partial y}(a,b) = \lim_{h\to 0} \frac{f(a,b+h) - f(a,b)}{h}$$

is the ordinary derivative

And these ordinary derivatives satisfy all of the rules from Calculus I.

(IIa)

Now, for a nice pt of the graph(f), the 2 tangent lines, which meet at (a,b) and are perpendicular, will determine a plane:

Choose a vector from each line and find the set of all <u>linear combinations</u>

In higher dimensions, this setup is exactly the same:

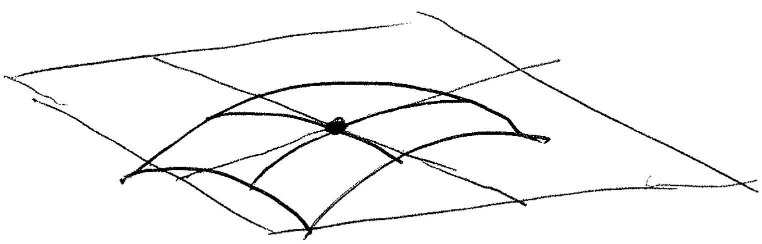
$\text{graph}(f) \subset \mathbb{R}^{n+1}$ for $f: X \subset \mathbb{R}^n \to \mathbb{R}$, and

$$\text{graph}(f) = \left\{ \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ z \end{bmatrix} \in \mathbb{R}^{n+1} \;\middle|\; z = f(x_1, \dots, x_n). \right\}$$

Fix all coordinates except x_i, and you get a 2-d subspace w/ coords (x_i, z).

Slice of graph is the graph of $z = f(c_1, \dots, c_{i-1}, x_i, c_{i+1}, \dots, c_n)$ in the $x_i z$-plane $\subset \mathbb{R}^{n+1}$

The partial derivative is the Calc I derivative of this restricted func.



This plane is called the tangent plane to graph of f at (a,b)

What is the equation of this 2-d plane in $\mathbb{R}^3$?

- it will be one linear eqn in 3 variables.
- The choice of vector in the xz-plane at $y=b$ will be based at (a,b) and have components $\begin{bmatrix} 1 \\ f_x(a,b) \end{bmatrix}$ why?

  (as a vector in $\mathbb{R}^3$ it will have components $\begin{bmatrix} 1 \\ 0 \\ f_x(a,b) \end{bmatrix}$ and be based at $(a,b,f(a,b))$.
- The choice of vector in the yz-plane at $x=a$ will have components $\begin{bmatrix} 0 \\ 1 \\ f_y(a,b) \end{bmatrix}$ at $(a,b,f(a,b))$.
- The set of all linear combinations is then
  $$c_1 \begin{bmatrix} 1 \\ 0 \\ f_x(a,b) \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ f_y(a,b) \end{bmatrix}.$$

A better way: The vector

$$\vec{n} = \begin{bmatrix} 1 \\ 0 \\ f_x(a,b) \end{bmatrix} \times \begin{bmatrix} 0 \\ 1 \\ f_y(a,b) \end{bmatrix} = \begin{bmatrix} -f_x(a,b) \\ -f_y(a,b) \\ 1 \end{bmatrix}$$

is normal to both. Thus it is normal to all combinations. In fact, the solutions to

$$\underbrace{\begin{bmatrix} x-a \\ y-b \\ z-f(a,b) \end{bmatrix}}_{\substack{\text{generic vector} \\ \text{at } (a,b,f(a,b))}} \bullet \underbrace{\begin{bmatrix} -f_x(a,b) \\ -f_y(a,b) \\ 1 \end{bmatrix}}_{\substack{\text{normal vector} \\ \text{to all}}} = 0$$

$$-f_x(a,b)(x-a) - f_y(a,b)(y-b) + z - f(a,b) = 0$$

or

$$\boxed{z = f_x(a,b)(x-a) + f_y(a,b)(y-b) + f(a,b)}$$

This is the equation of the plane tangent to the graph of $f(x,y)$ at $(a,b,f(a,b)) \in \mathbb{R}^3$ when it is defined:

It is the best linear approximation to graph(f) at (a,b)

why? ① $z = h(x,y) =$ RHS is linear.

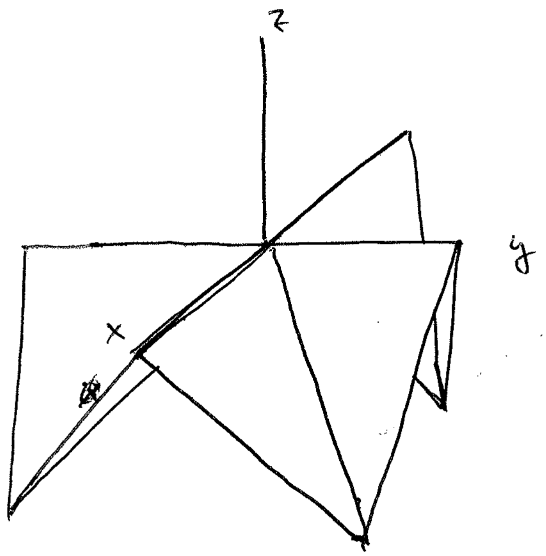
② $h(a,b) = f(a,b)$, $\frac{\partial h}{\partial x}(a,b) = \frac{\partial f}{\partial x}(a,b)$, $\frac{\partial h}{\partial y}(a,b) = \frac{\partial f}{\partial y}(a,b)$

Major Caveat!

Just because $\frac{\partial f}{\partial x}(a,b)$, $\frac{\partial f}{\partial y}(a,b)$ exist does not mean f is differentiable at (a,b).

Example 4, pg 121 is a great example!

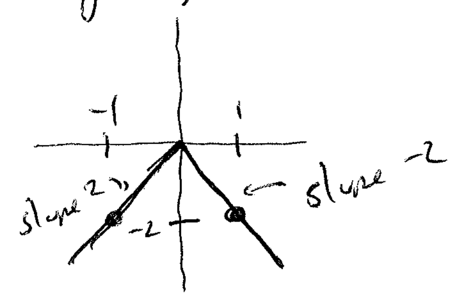
Roof top function: $f(x,y) = \big| |x| - |y| \big| - |x| - |y|$.



Here, both $\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0$ but step off the axes and one sees the sharp edges.

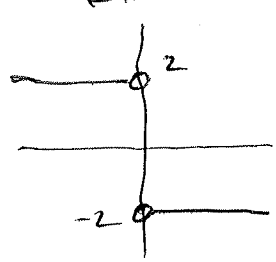
In fact, if one approached $(0,0)$ along the line $x=y$ in $\mathbb{R}^2$, then 2 side derivatives wouldn't match!

$y=x$ plane



$$\lim_{(x,y)\to(0,0)} \frac{f(x,y) - \cancel{f(0,0)}^{\,0}}{x} \quad \text{along } y=x$$

$$\lim_{x\to 0} \frac{f(x,x)}{x} = \lim_{x\to 0} \frac{-2|x|}{x} \quad \text{DNE}$$



The existence of a proper tangent space relies on its ability to well-approximate the fnc from all directions.

So use the limit!

Notice how we can rewrite $f'(a) = \lim_{x\to a} \frac{f(x)-f(a)}{x-a}$

into $\lim_{x\to a} \frac{f(x) - (f(a) + f'(a)(x-a))}{x-a} = 0$

when (and only when) the limit exists (do this!)

This means that, for $h(x) = f(a) + f'(a)(x-a)$, we can say that f is differentiable at $x=a$ precisely when the tangent line $h(x)$ exists, precisely when

$$\lim_{x\to a} \frac{f(x)-h(x)}{x-a} = 0.$$

This is generalizable: For $X \subset \mathbb{R}^2$ open, with $f: X \to \mathbb{R}$, f is diff at $(a,b) \in X$ if

① Both $\frac{\partial f}{\partial x}(a,b)$ and $\frac{\partial f}{\partial y}(a,b)$ exist, and

② If $h(x,y) = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b)$ satisfies

$$\lim_{(x,y)\to(a,b)} \frac{f(x,y)-h(x,y)}{\|(x,y)-(a,b)\|} = 0.$$

Note: When $h(x,y)$ exists and satisfies the limit, then $z = h(x,y)$ is the tangent plane to graph(f) at $(a,b,f(a,b)) \in$ graph$(f) \subset \mathbb{R}^3$.

More notes: ① Alternate notion of differentiable:
For $X \subset \mathbb{R}^2$ open, and $f: X \to \mathbb{R}$, f is diff. at (a,b) if $f_x(x,y)$ and $f_y(x,y)$ are continuous in a neighborhood of (a,b).

② Like in Calculus I, differentiability always implies continuity.

③ True in n-dimensions: Given $X \subset \mathbb{R}^n$ open, with $f: X \to \mathbb{R}$, f is diff. at $\vec{a} \in \mathbb{R}^n$ if ① $\frac{\partial f}{\partial x_i}(\vec{a})$ exists for $i = 1, \dots, n$, and

② if $h(\vec{x}) = f(\vec{a}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{a})(x_i - a_i)$ satisfies $\lim_{\vec{x} \to \vec{a}} \frac{f(\vec{x}) - h(\vec{x})}{\|\vec{x} - \vec{a}\|} = 0$

There is an easier way to write this:

Def For $f: X \subset \mathbb{R}^n \to \mathbb{R}$ differentiable at $\vec{a} \in X$, the derivative of f at $\vec{a}$ is the $1 \times n$ matrix

$$Df(\vec{a}) = [f_{x_1}(\vec{a}) \; f_{x_2}(\vec{a}) \; \cdots \; f_{x_n}(\vec{a})]$$

And the derivative function is the $1\times n$ matrix of functions $Df(\vec{x}) = [f_{x_1}(\vec{x}) \quad \cdots \quad f_{x_n}(\vec{x})]$.

Knowing this, the tangent linear function in the defs above can be written

$$h(\vec{x}) = f(\vec{c}) + \sum_{i=1}^{n} f_{x_i}(\vec{c})(x_i - c_i)$$
$$= f(\vec{c}) + \begin{bmatrix} f_{x_1}(\vec{c}) & \cdots & f_{x_n}(\vec{c}) \end{bmatrix} \begin{bmatrix} x_1 - c_1 \\ x_2 - c_2 \\ \vdots \\ x_n - c_n \end{bmatrix}$$
$$= f(\vec{c}) + \underbrace{Df(\vec{c})}_{\substack{1\times n \\ \text{row matrix}}} \underbrace{(\vec{x}-\vec{c})}_{\substack{n\times 1 \\ \text{column vector}}}$$

Here the limit in the def becomes

$$\lim_{\vec{x}\to\vec{c}} \frac{f(\vec{x}) - h(\vec{x})}{\|\vec{x}-\vec{c}\|} = \lim_{\vec{x}\to\vec{c}} \frac{f(\vec{x}) - \left(f(\vec{c}) + Df(\vec{c})(\vec{x}-\vec{c})\right)}{\|\vec{x}-\vec{c}\|}$$

Now what about $\vec{f}: X \subset \mathbb{R}^n \to \mathbb{R}^m$? Here, ~~each~~ for $\vec{x} \in X$, $\vec{f}(\vec{x}) = \begin{bmatrix} f_1(\vec{x}) \\ \vdots \\ f_m(\vec{x}) \end{bmatrix}$. w/ n inputs and m outputs.

If the derivative exists, then each component $f_i(\vec{x})$ has a derivative w/ n partials. We have

$$D\vec{f}(\vec{a}) = \begin{bmatrix} Df_1(\vec{a}) \\ \vdots \\ Df_m(\vec{a}) \end{bmatrix}$$ but each is a $1 \times n$ matrix.

Hence

$$D\vec{f}(\vec{a}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\vec{a}) & \frac{\partial f_1}{\partial x_2}(\vec{a}) & \cdots & \frac{\partial f_1}{\partial x_n}(\vec{a}) \\ \vdots & & & \\ \frac{\partial f_m}{\partial x_1}(\vec{a}) & \cdots & & \frac{\partial f_m}{\partial x_n}(\vec{a}) \end{bmatrix}$$ an <u>$m \times n$ matrix</u>

where the ijth entry is $\frac{\partial f_i}{\partial x_j}(\vec{a})$.

Back to our definition:

Let $\vec{f}: X \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a vector-valued fnc on an open X, and $\vec{a} \in X$. $\vec{f}$ is diff. at $\vec{x} = \vec{a}$

if (1) $\frac{\partial f_i}{\partial x_j}(\vec{a})$ all exist for $i = 1, \ldots, m$, $j = 1, \ldots, n$,

and (2) The linear space $\vec{h}(\vec{x}) = \vec{f}(\vec{a}) + D\vec{f}(\vec{a})(\vec{x} - \vec{a})$ satisfies

$$\lim_{\vec{x} \to \vec{a}} \frac{\| \vec{f}(\vec{x}) - \vec{h}(\vec{x}) \|}{\| \vec{x} - \vec{a} \|} = 0$$

Notes (1) In this last definition $\|\vec{f}(\vec{x})-\vec{h}(\vec{x})\|$ measures the distance between $\vec{f}(\vec{x})$ and $\vec{h}(\vec{x})$ as vector-valued fncs.

~~(4) $\vec{h}(\vec{x})$ here is the fnc $\vec{h}: \mathbb{R}^n \to \mathbb{R}^m$ that is the best~~

(2) $Df(\vec{c})$, as a matrix of numbers, represents a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$. It's entries vary as $\vec{c}$ varies, but it represents the <u>best</u> linear map to approximate $f(\vec{x})$ at $\vec{x}=\vec{c}$.

(3) $Df(\vec{c})(\vec{x}-\vec{c}) \in \mathbb{R}^m$ for each $\vec{x}$ and represents a catalog of ways moving around near $\vec{c}$ in the domain affects function values in the codomain.

(4) $\vec{h}(\vec{x}) = \vec{f}(\vec{c}) + D\vec{f}(\vec{c})(\vec{x}-\vec{c})$ defines an affine map $\vec{h}: \mathbb{R}^n \to \mathbb{R}^m$ (linear map with a transl.)

for $m=1$, $z=h(\vec{x})$ has a graph in $\mathbb{R}^{n+1}$ called the tangent space to graph(f) at $(\vec{c}, f(\vec{c}))$.

It is the same for $m>1$, but geometrically this is less easy to see.