MAT B41 — Week 4a

This week:

- Differentiation
- Chain Rule.

Recall from last week:

$$\frac{\partial f}{\partial x_i}(\vec{x}) = \lim_{h \to 0} \frac{f(\vec{x}+h\vec{e}_i)-f(\vec{x})}{h}$$

= "change of f in direction $\vec{e}_i$"

We need to be careful with this definition.

Ex: Find $\frac{\partial f}{\partial x}(0,0)$ and $\frac{\partial f}{\partial y}(0,0)$ for $f(x,y) = \sqrt[3]{xy}$

We get:
$$\frac{\partial f}{\partial x} = \lim_{h \to 0} \frac{f(h,0)-f(0,0)}{h}$$

$$= \lim_{h \to 0} \frac{0}{h} = 0$$

(!) $f(x,y)$ is zero along the axes.

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

$$= \lim_{h \to 0} \frac{0}{h} = 0$$

Thus, $\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0$.

Question: Does $\displaystyle\lim_{(x,y)\to(0,0)} \frac{f(x,y)-f(0,0)}{\|(x,y)\|}$ exist?

Can we call this the derivative?

Consider,
$$\lim_{h\to 0} \frac{f(h,h)-f(0,0)}{\|(h,h)\|}$$

$$=\lim_{h\to 0} \frac{\sqrt[3]{h\cdot h}-0}{\sqrt{2}\cdot h} = \lim_{h\to 0} \frac{1}{\sqrt{2}\cdot\sqrt[3]{h}}$$

This limit does not exist!

We get that $\lim\limits_{(x,y)\to(0,0)} \dfrac{f(x,y)-f(0,0)}{\|(x,y)\|}$ does not exist.

Defn (p136) The DIRECTIONAL DERIVATIVE of $f(\vec{x})$ in the direction $\vec{v}$ is:

$$\frac{\partial f}{\partial v} = \frac{d}{dt}\left[f(\vec{x}+t\vec{v})\right]\Big|_{t=0}$$
$$= \lim_{h\to 0} \frac{f(\vec{x}+h\vec{v})-f(\vec{x})}{h}$$

Ex: Find the direction derivative $\frac{\partial f}{\partial v}(1,2)$ for $\vec{v}=(1,-1)$ and $f(x,y)=x^2+y^2$.

$$\frac{\partial f}{\partial v} = \lim_{h\to 0} \frac{f((1,2)+h(1,-1))-f(1,2)}{h}$$

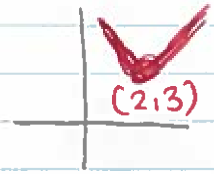
$$= \lim_{h\to 0} \frac{f(1+h,2-h)-5}{h} = \lim_{h\to 0} \frac{(1+h)^2+(2-h)^2-5}{h}$$

$$= \lim_{h\to 0} \frac{1+2h+h^2+4-4h+h^2-5}{h} = \lim_{h\to 0} \frac{-2h+2h^2}{h}$$

$$= \lim_{h\to 0} -2+2h = -2.$$

Differentiability

Recall from single variable calculus,



NOT all functions are differentiable. e.g. $y = |x-2|+3$ has no tangent at $(2,3)$

These issues are more subtle in $\mathbb{R}^n$.

Explore these in Geogebra!

CONES

FOLDS

JUMPS

$z = \|(x,y)\|$ $\quad$ $z = \min\{0, x-y\}$ $\quad$ $z = \begin{cases} 5 & (x,y) = 0 \\ 0 & (x,y) \neq 0 \end{cases}$

There are many reasons that $\mathbb{R}^n \to \mathbb{R}$ can fail to be differentiable.

<u>Defn</u>: We say $f: \mathbb{R}^n \to \mathbb{R}$ is DIFFERENTIABLE at $\vec{x} = \vec{p}$ if it has a tangent plane at that point.

We need:

$$\lim_{\vec{x} \to \vec{p}} \frac{\| f(\vec{x}) - f(\vec{p}) - \vec{n} \cdot (\vec{x} - \vec{p}) \|}{\| \vec{x} - \vec{p} \|} = 0$$

for some plane $\Pi = \{ \vec{x} : \vec{n}(\vec{x} - \vec{p}) = 0 \}$

<u>Question</u>: Show $f(x,y) = x + 2y$ if diff'able at $(x,y) = (2,3)$

<u>Ex</u>: Show $f(x,y) = xy$ is differentiable at $(1,2)$.

We need $\vec{n}$ and $\vec{p}$ so that:

$$\lim_{(x,y) \to (1,2)} \frac{\| f(x,y) - f(1,2) - \vec{n}((x,y) - (1,2)) \|}{\| (x,y) - (1,2) \|} = 0$$

Ex (con't)

$$\lim_{(x,y)\to(1,2)} \frac{\| xy - 2 - \vec{n}((x-1, y-2))\|}{\|(x-1, y-2)\|}$$

Question: How do we pick $\vec{n}$?
→ We want the tangent plane.
→ We want same slopes as $f(x,y)$.

$$\vec{n} = \begin{bmatrix} \frac{\partial f}{\partial x}(1,2) \\ \frac{\partial f}{\partial y}(1,2) \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

We get:

$$\lim_{(x,y)\to(1,2)} \frac{\| xy - 2 - \begin{bmatrix} 2 \\ 1 \end{bmatrix}\cdot(x-1, y-2)\|}{\|(x-1, y-2)\|}$$

$$= \lim_{(x,y)\to(1,2)} \frac{\| xy - 2 - 2(x-1) - 1(y-2)\|}{\|(x-1, y-2)\|}$$

$$= \lim_{(x,y)\to(1,2)} \frac{\| xy - 2 - 2x + 2 - y + 2\|}{\|(x-1, y-2)\|}$$

$$= \lim_{(x,y)\to(1,2)} \frac{\| xy - 2x - y + 2\|}{\|(x-1, y-2)\|}$$

$(x-1)(y-2)$
$xy - 2x - y + 2$

$$= \lim_{(x,y)\to(1,2)} \frac{\|(x-1)(y-2)\|}{\|(x-1, y-2)\|} = 0$$

By Circle-Box argument. $\boxed{\text{Check this in Geogebra.}}$

Differentiability

Fact: If $f(\vec{x})$ is differentiable at $\vec{x}=\vec{p}$ then it has tangent plane with

$$\vec{n} = \left[\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \cdots \quad \frac{\partial f}{\partial x_n}\right]$$

Fact: If $f(\vec{x})$ is differentiable at $\vec{x}=\vec{p}$ then f is continuous at $\vec{x}=\vec{p}$.

Fact: If $f(\vec{x})$ is differentiable at $\vec{x}=\vec{p}$ then

$$\frac{\partial f}{\partial v}(\vec{p}) \text{ exists for all } \vec{v}.$$

Moreover,

$$\frac{\partial f}{\partial v} = \vec{v} \cdot \left[\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \cdots \quad \frac{\partial f}{\partial x_n}\right]$$

! To make $\frac{\partial f}{\partial v}$ as large as possible we pick:

$$\vec{v} = \left[\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \cdots \quad \frac{\partial f}{\partial x_n}\right]$$

Def^n: For $f: \mathbb{R}^n \to \mathbb{R}$ the GRADIENT VECTOR

$$\nabla f = \left[\frac{\partial f}{\partial x_1} \quad \cdots \quad \frac{\partial f}{\partial x_n}\right]$$

It is the direction which maximizes

$$\frac{\partial f}{\partial v}$$

This week:
- Differentiation
- The chain Rule

Recall, from last lecture, for $f: \mathbb{R}^n \to \mathbb{R}$:

$$\frac{\partial f}{\partial v} = \lim_{h \to 0} \frac{f(\vec{x} + h\vec{v}) - f(\vec{x})}{h}$$

= "change in direction $\vec{v}$"

$$\nabla f = \left[\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \cdots \quad \frac{\partial f}{\partial x_n} \right]$$

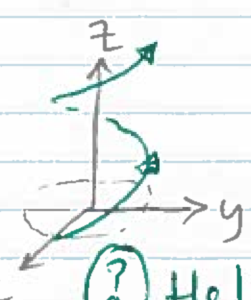
= direction of fastest increase"

In general, $\boxed{\frac{\partial f}{\partial v} = \vec{v} \cdot \nabla f}$

These are both notions of:

"the derivative of $f: \mathbb{R}^n \to \mathbb{R}$"

Question: What about $\mathbb{R} \to \mathbb{R}^n$?



Ex: Consider $c(t) = \begin{bmatrix} \cos(t) \\ \sin(t) \\ t \end{bmatrix}$

(?) Helix.

Find $c'(\pi/2)$

$$c'\left(\tfrac{\pi}{2}\right) = \lim_{h \to 0} \frac{c(\pi/2 + h) - c(\pi/2)}{h}$$

Take derivatives

We can use algebra! $= \lim_{h \to 0} \frac{1}{h} \begin{bmatrix} \cos(\frac{\pi}{2} + h) - \cos(\frac{\pi}{2}) \\ \sin(\frac{\pi}{2} + h) - \sin(\frac{\pi}{2}) \\ \frac{\pi}{2} + h \quad - \frac{\pi}{2} \end{bmatrix} = \begin{bmatrix} -\sin(\frac{\pi}{2}) \\ \cos(\frac{\pi}{2}) \\ 1 \end{bmatrix}$

? The vector $c'(t)$ represents the direction (velocity) that the curve is travelling.

Fact: If $f(t) = \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix}$ is differentiable

$$f'(t) = \begin{bmatrix} \frac{d}{dt} f_1(t) \\ \vdots \\ \frac{d}{dt} f_n(t) \end{bmatrix} = \begin{bmatrix} f_1'(t) \\ \vdots \\ f_n'(t) \end{bmatrix}$$

Ex: Find $f'(x)$ for $f(x) = \begin{bmatrix} x^2 \\ \cos(x) \end{bmatrix} \Rightarrow f'(x) = \begin{bmatrix} 2x \\ -\sin(x) \end{bmatrix}$

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

Def^n: We say $f(\vec{x}) = \vec{y}$ is DIFFERENTIABLE at $\vec{x} = \vec{p}$ if there is a matrix M such that:

$$\lim_{\vec{x} \to \vec{p}} \frac{\| f(\vec{x}) - f(\vec{p}) - M(\vec{x} - \vec{p}) \|}{\| \vec{x} - \vec{p} \|} = 0.$$

? Recall from last lecture we needed a tangent plane for $f: \mathbb{R}^n \to \mathbb{R}$ to be diff'able.

Ex: Show that $f(\vec{x}) = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \vec{x}$ is differentiable.

We pick $M = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$ and obtain:

$$\lim_{\vec{x} \to \vec{p}} \frac{\| \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \vec{x} - \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \vec{p} - M(\vec{x} - \vec{p}) \|}{\| \vec{x} - \vec{p} \|} = \lim_{\vec{x} \to \vec{p}} \frac{0}{\| \vec{x} - \vec{p} \|}$$

$$= 0.$$

Defn: If $f: \mathbb{R}^n \to \mathbb{R}^k$ is differentiable then $M = Df$ is the TOTAL DERIVATIVE

Fact: If $f(\vec{x}) = \begin{bmatrix} f_1(\vec{x}) \\ \vdots \\ f_k(\vec{x}) \end{bmatrix}$ then

(!) In general Df is not symmetric

Order matters!

$$Df = \left[\frac{\partial f_i}{\partial x_j}\right] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_1} \end{bmatrix}$$

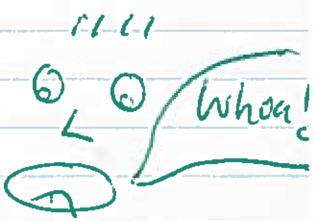
(?) Df transforms vectors according to how f would map them.

Ex: For $f(x,y) = \begin{bmatrix} x+y \\ xe^y \end{bmatrix}$ find Df.

$$Df = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ e^y & xe^y \end{bmatrix}$$

Ex: For $f(x,y) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+2y \\ 3x+4y \end{bmatrix}$ find Df.

$$Df = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$



(?) The total derivative of a linear map is the linear map itself.

<u>Ex</u>: If $f: \mathbb{R} \to \mathbb{R}^3$ is $f(t) = \begin{bmatrix} \cos(t) \\ \sin(t) \\ t \end{bmatrix}$

$g: \mathbb{R}^3 \to \mathbb{R}$ is $g(x,y,z) = x^2 + y^2 + z$

then find $\frac{d}{dt}(g \circ f)$, $f'(t)$, and ∇g.

Calculate $\frac{d}{dt}(g \circ f)$.

$$g \circ f(t) = g(\cos(t), \sin(t), t)$$
$$= \cos^2(t) + \sin^2(t) + t = 1 + t$$
$$\Rightarrow \frac{d}{dt}(g \circ f) = 1$$

Calculate $f'(t)$

$$f'(t) = \begin{bmatrix} \frac{d}{dt}\cos(t) \\ \frac{d}{dt}\sin(t) \\ \frac{d}{dt}(t) \end{bmatrix} = \begin{bmatrix} -\sin(t) \\ \cos(t) \\ 1 \end{bmatrix}$$

Calculate ∇g

$$\nabla g = \left[\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}, \frac{\partial g}{\partial z}\right] = \left[2x \quad 2y \quad 1\right]$$

<u>Question</u>: How do these relate to each other?

Observe,

$$\frac{d}{dt}(g \circ f(t)) = \nabla g(f(t)) \cdot f'(t)$$

$$= \left[2\cos(t) \quad 2\sin(t) \quad 1\right] \begin{bmatrix} -\sin(t) \\ \cos(t) \\ 1 \end{bmatrix}$$

$$= 2\cos(t)(-\sin t) + 2\cos(t)\sin(t) + 1$$

$$= 1$$

Next up:

The Chain Rule!

MAT B41 — Week 4c

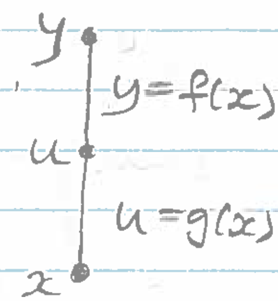
The Chain Rule $\mathbb{R}^n \to \mathbb{R}$

Recall, the chain rule in one-variable:

If $y = f(u)$ and $u = g(x)$

then $$\frac{dy}{dx} = \frac{dy}{du}\cdot\frac{du}{dx}$$

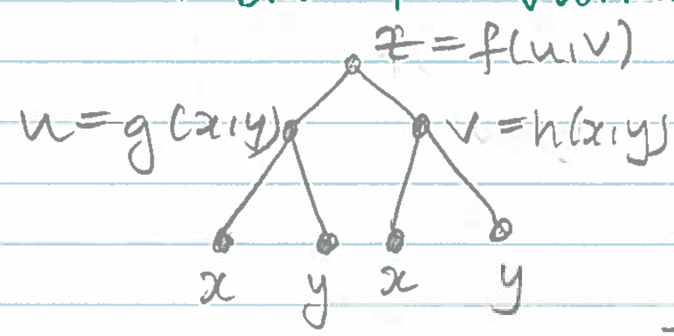
$$= f'(g(x))\cdot g'(x)$$



"A change in x produces a change in u which produces a change in y."

Ex: If $z = f(u,v)$ $u = g(x,y)$ $v = h(x,y)$ then how does a change in x effect a change in z?

Draw the variable dependence diagram



Apply chain rule along each path from x to z.

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial u}\cdot\frac{\partial u}{\partial x} + \frac{\partial z}{\partial v}\cdot\frac{\partial v}{\partial x}$$

Ex: Find $\frac{\partial z}{\partial x}(x,y)$ if $z = u^2+v^2$ $u = x+2y$ $v = 3x+4y$

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial u}\cdot\frac{\partial u}{\partial x} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial x}$$

$$= 2u\cdot 1 + 2v\cdot 3 = 2u+6v = 20x+28y$$

Chain Rule $\mathbb{R}^n \to \mathbb{R}^n$

We know that if $f(\vec{x}) = A\vec{x}$ then $Df = A$.

Suppose $f: \mathbb{R}^n \to \mathbb{R}^n$ is $f(\vec{x}) = A\vec{x}$
$g: \mathbb{R}^n \to \mathbb{R}^n$ is $g(\vec{x}) = B\vec{x}$

We get $f \circ g(\vec{x}) = A(B\vec{x}) = (AB)\vec{x}$

Moreover, $D(f \circ g) = (Df)(Dg)$? Matrix multiplication is designed so that this works out!

Fact: If $f: \mathbb{R}^n \to \mathbb{R}^k$ is differentiable at $\vec{x} = \vec{p}$ and $g: \mathbb{R}^k \to \mathbb{R}^l$ is differentiable at $\vec{x} = f(\vec{p})$ then

$$D(g \circ f)(\vec{p}) = (Dg)(f(\vec{p})) \cdot (Df)(\vec{p})$$

Ex: Suppose $f(x,y) = (xy, y^2)$ and $g(x,y) = (3x+2y, 10y)$
Find $D(g \circ f)(1,1)$. Question: What does this measure?

We calculate: $D(g \circ f)(1,1)$

$$= (Dg)(f(1,1))\,(Df)(1,1)$$

$$= \begin{bmatrix} 3 & 2 \\ 10 & 0 \end{bmatrix}(f(1,1)) \begin{bmatrix} y & x \\ 0 & 2y \end{bmatrix}(1,1)$$

$$= \begin{bmatrix} 3 & 2 \\ 10 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 7 \\ 10 & 10 \end{bmatrix}$$

Thus, the total derivative at $(1,1)$ is $\begin{bmatrix} 3 & 7 \\ 10 & 10 \end{bmatrix}$