Ex: Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ differentiable

Define $\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}$ by $\varphi(x,y,z) = f(x^2 - yz, xyz)$

$$\frac{\partial \varphi}{\partial x}(x,y,z) = 2x \frac{\partial f}{\partial x}(x^2-yz, xyz) + yz \frac{\partial f}{\partial y}(x^2-yz, xyz)$$

$$\frac{\partial \varphi}{\partial y}(x,y,z) = -z \frac{\partial f}{\partial x}(x^2-yz, xyz) + xz \frac{\partial f}{\partial y}(x^2-yz, xyz)$$

$$\frac{\partial \varphi}{\partial z}(x,y,z) = -y \frac{\partial f}{\partial x}(x^2-yz, xyz) + xy \frac{\partial f}{\partial y}(x^2-yz, xyz)$$

$\frac{\partial f}{\partial x}$ is just a notation to say partial derivative w.r.t. first variable

You compute the partial derivative of f w.r.t the first variable and then evaluate it at (x^2-yz, xyz)

Be sure you understand the above computations before continuing...

Your worst enemy in calculus is going to be notation:

① There are as many notation as people

eg: $\partial_x, \partial_{e_1}, \partial^{e_1}, \frac{\partial}{\partial x}, f_x, D^{(1,0)}, \ldots$ In MAT237, we'll use $\partial_x f$ or $\frac{\partial f}{\partial x}$.

Be sure that you understand what you are computing

② Notation can be confusing

eg: $\frac{\partial f}{\partial x}(x^2-yz, xyz)$ means ⓐ compute $\frac{\partial f}{\partial x}$ first partial derivative of f

then ⓑ evaluate at (x^2-yz, xyz) → ✓

It does NOT mean: compute $f(x^2-yz, xyz)$ and then differentiate wrt to x → ✗

Ex: $f: \mathbb{R}^3 \longrightarrow \mathbb{R}$ differentiable

$S = \{(x,y) \in \mathbb{R}^2 : y \neq 0\}$ open

Define $\varphi(x,y) = f(x, xy, x/y)$

$$\frac{\partial \varphi}{\partial x}(x,y) = \frac{\partial f}{\partial x}(x, xy, x/y) + y\frac{\partial f}{\partial y}(x,xy,x/y) + \frac{1}{y}\frac{\partial f}{\partial z}(x,xy,x/y)$$

$$\frac{\partial \varphi}{\partial y}(x,y) = 0 + x\frac{\partial f}{\partial y}(x,xy,x/y) - \frac{x}{y^2}\frac{\partial f}{\partial z}(x,xy,x/y)$$

Comment:

$$D\varphi = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ y & x \\ 1/y & -\frac{x}{y^2} \end{pmatrix}$$

We recover the same result!

⚠ It's common to drop the variables during the computations as I did in the comment, in order to lighten the notation.

If you do so: ① Be careful to not forget to add them back at the end

② Keep track of them

Ex: Polar coordinates

$f: \mathbb{R}^2 \to \mathbb{R}$, $S = \{(r,\theta) \in \mathbb{R}^2 : r \geq 0\}$

$g: S \to \mathbb{R}^2$ defined by $g(r,\theta) = (r\cos\theta, r\sin\theta)$

$\varphi = f\circ g$ so that $\varphi(r,\theta) = f(r\cos\theta, r\sin\theta)$

(A) Chain rule for the Jacobian matrix:

$$D\varphi(r,\theta) = Df(g(r,\theta))\cdot Dg(r,\theta)$$
$$= Df(r\cos\theta, r\sin\theta)\cdot Dg(r,\theta)$$
$$= \begin{pmatrix} \frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) & \frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta) \end{pmatrix} \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

$$\begin{pmatrix} \frac{\partial \varphi}{\partial r}(r,\theta) & \frac{\partial \varphi}{\partial \theta}(r,\theta) \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta)\cos\theta + \frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta)\sin\theta, & -\frac{\partial f}{\partial x} r\sin\theta + \frac{\partial f}{\partial y} r\cos\theta \end{pmatrix}$$

(B) Chain rule for the partial derivatives

$$\frac{\partial \varphi}{\partial r}(r,\theta) = \frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta)\cos\theta + \frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta)\sin\theta$$

$$\frac{\partial \varphi}{\partial \theta}(r,\theta) = -\frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta)\, r\sin\theta + \frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta)\, r\cos\theta$$

We read them in the components of the Jacobian matrix

(C) Chain rule for the differentials:

$$d_{(r,\theta)}\varphi(h,k) = d_{g(r,\theta)}f \circ d_{(r,\theta)}g(h,k) = d_{g(r,\theta)}f(\cos\theta h - r\sin\theta k, \sin\theta h + r\cos\theta k)$$
$$= \frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta)(\cos\theta h - r\sin\theta k) + \frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta)(\sin\theta h + r\cos\theta k)$$

Ex: changing the names of the variables

$$f : \begin{array}{l} \mathbb{R}^2 \longrightarrow \mathbb{R} \\ (x,y) \longmapsto f(x,y) \end{array}$$

$\varphi : \mathbb{R}^2 \longrightarrow \mathbb{R}$ defined by $\varphi(r,s) = f(re^s, rs)$

$$\frac{\partial \varphi}{\partial r}(r,s) = e^s \frac{\partial f}{\partial x}(re^s, rs) + s \frac{\partial f}{\partial y}(re^s, rs)$$

$$\frac{\partial \varphi}{\partial s}(r,s) = re^s \frac{\partial f}{\partial x}(re^s, rs) + r \frac{\partial f}{\partial y}(re^s, rs)$$

Level sets and the gradient

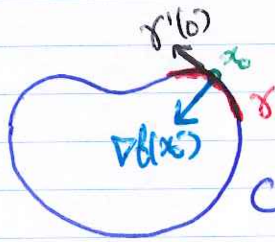
Setup: $\mathcal{U} \subset \mathbb{R}^n$ open, $f : \mathcal{U} \to \mathbb{R}$ differentiable at $x_0 \in \mathcal{U}$

$$C = \{x \in \mathcal{U} : f(x) = f(x_0)\}$$

C is the level set of f at $f(x_0)$

Def: We say that $v \in \mathbb{R}^n$ is $\boxed{\text{tangent}}$ to C at x_0 if there exists $\gamma : I \to \mathbb{R}^n$, $I \subset \mathbb{R}$ open interval, $0 \in I$, such that $\forall t \in I$, $\gamma(t) \in C$, $\gamma(0) = x_0$, $\gamma'(0) = v$

↳ particularly, $\gamma'(0)$ exists



$C = \{x \in \mathcal{U} \text{ st. } f(x) = f(x_0)\} \subset \mathcal{U}$ ⚠

Claim: If v is tangent to C at x_0 then $v \cdot \nabla f(x_0) = 0$

i.e.: $\nabla f(x_0)$ is orthogonal to the level set

Δ Take γ as in the above definition

Define $h : I \to \mathbb{R}$ by $h(t) = f(\gamma(t))$, then

① $\forall t \in I$, $h(t) = f(\gamma(t)) = f(x_0)$ since $\gamma(t) \in C$

$\Rightarrow h'(0) = 0$

② By the chain rule

$$\underset{\text{by ①}}{0 =} h'(0) = (f \circ \gamma)'(0)$$

$= d_0(f \circ \gamma)(1)$ → Recall that for $g : \underset{\mathbb{R} \text{ open interval}}{I} \to \mathbb{R}$, $d_{t_0} g(h) = g'(t_0) h$

Hence $g'(t_0) = d_{t_0} g(1)$

$= d_{\gamma(0)} f \circ d_0 \gamma(1)$ ↳ chain rule for the differentials

$= d_{\gamma(0)} f(d_0 \gamma(1))$

$= d_{x_0} f(v)$ → $\gamma(0) = x_0$, $d_0 \gamma(1) = \gamma'(0) \cdot 1 = v$

$= \nabla f(x_0) \cdot v$

$\square$

<u>Ex</u>: Find the tangent plane of

$$C = \{(x,y,z) \in \mathbb{R}^3 : x^2 - 2xy + 4yz - z^2 = 2\}$$

at $a = (1,1,1)$

Δ C is the level set $f(x,y,z) = 2$ for $f(x,y,z) = x^2 - 2xy + 4yz - z^2$

$\nabla f(a) = (0,2,2)$

So the tangent plance of C at a is

$$\{(x,y,z) \in \mathbb{R}^3 : (x-1, y-1, z-1) \cdot (0,2,2) = 0\}$$

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

It has a for equation $y+z=2$.

$\blacksquare$

<u>TODO</u>: Recap slides

<u>Homework</u>: Questions from § 2.3

DIFFERENTIABILITY

The Mean Value Theorem

Theorem: (MVT) $U \subset \mathbb{R}^m$ open, $f: U \to \mathbb{R}$ differentiable on U

Let $a, b \in U$, assume that $L_{a,b} = \{(1-t)a + tb : t \in [0,1]\} \subset U$

Then $\exists c \in L_{a,b}$ such that $f(b) - f(a) = \nabla f(c) \cdot (b-a)$

Δ Set $\gamma(t) = (1-t)a + tb$, $t \in [0,1]$

and $\varphi(t) = f(\gamma(t))$, $\varphi: [0,1] \to \mathbb{R}$ is differentiable by composit°

By 1st year MVT, $\exists t_0 \in (0,1)$ such that

$$\varphi'(t_0) = \frac{\varphi(1) - \varphi(0)}{1-0} = f(b) - f(a)$$

By the Chain Rule:
$$\begin{aligned}\varphi'(t_0) &= d_{t_0}\varphi(1) = d_{t_0} f\circ\gamma(1) \\ &= d_{\gamma(t_0)} f \circ d_{t_0}\gamma(1) \\ &= d_{\gamma(t_0)} f(b-a) \\ &= \nabla f(\gamma(t_0)) \cdot (b-a)\end{aligned}$$
$\square$

Take $c = \gamma(t_0) \in L_{a,b}$
$= (1-t_0)a + t_0 b, t_0 \in (0,1)$

Def: A subset $S \subset \mathbb{R}^m$ is convex if:

$$\forall a, b \in S, \forall t \in [0,1], \quad (1-t)a + tb \in S$$

ie: given 2 points in S, the line segment between them is in S

Prop: A convex subset is path-connected

Δ For $a, b \in S$, take $\gamma(t) = (1-t)a + tb$ $\square$

Theorem: $\mathcal{U} \subset \mathbb{R}^m$ open and convex, $f : \mathcal{U} \longrightarrow \mathbb{R}$ differentiable on $\mathcal{U}$

If there exists $M > 0$ s.t. $\forall x \in \mathcal{U}$, $\|\nabla f(x)\| \leq M$

then $\forall a, b \in \mathcal{U}$, $|f(b) - f(a)| \leq M \cdot \|b-a\|$

Δ Let $a, b \in \mathcal{U}$, since $\mathcal{U}$ is convex $L_{a,b} = \{(1-t)a + tb : t \in [0,1]\} \subset \mathcal{U}$

so by the MVT, $\exists c \in L_{a,b}$ s.t.

$$f(b) - f(a) = (b-a) \cdot \nabla f(c)$$

$$\Rightarrow |f(b) - f(a)| = |(b-a) \cdot \nabla f(c)|$$

$\leq \|b-a\| \cdot \|\nabla f(c)\|$ by Cauchy-Schwarz

$\leq M \cdot \|b-a\|$ by assumption

$\square$

Theorem: $\mathcal{U} \subset \mathbb{R}^m$ open and convex, $f : \mathcal{U} \longrightarrow \mathbb{R}$ differentiable on $\mathcal{U}$

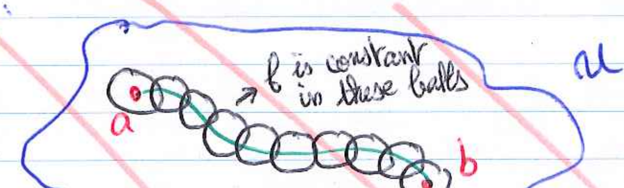
If $\forall x \in \mathcal{U}$, $\nabla f(x) = \vec{0}$ then f is constant on $\mathcal{U}$

Δ Let $a, b \in \mathcal{U}$, then
$|f(a) - f(b)| \leq 0 \cdot \|b-a\| = 0$, ie: $\forall a, b \in \mathcal{U}$, $f(a) = f(b)$ $\square$

Theorem: $\mathcal{U} \subset \mathbb{R}^m$ open and path-connected, $f : \mathcal{U} \to \mathbb{R}$ differentiable on $\mathcal{U}$

If $\forall x \in \mathcal{U}$, $\nabla f(x) = \vec{0}$ then f is constant on $\mathcal{U}$

Δ Idea:



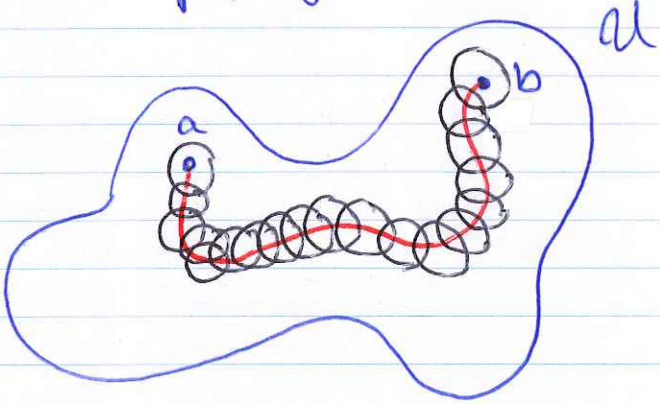
take $a, b \in \mathcal{U}$

take γ a C^0 path from a to b in $\mathcal{U}$

⚠ we don't know if γ is differentiable $\Rightarrow$ don't try to apply the chain rule

$\square$

△ Idea of proof



take $a, b \in \mathcal{U}$

By path-connectedness,

$\exists \gamma : [0,1] \longrightarrow \mathbb{R}^m$ s.t.

$$\begin{cases} \gamma \text{ is continuous} \\ \gamma(0) = a \\ \gamma(1) = b \\ \forall t \in [0,1], \gamma(t) \in \mathcal{U} \end{cases}$$

⚠ We don't know if γ is differentiable, so we can't mimick the proof of the MVT and compute $(f \circ \gamma)'(t)$

FACT: $\gamma([0,1])$ is compact as the continuous image of a compact set

Hence we may cover $\gamma([0,1])$ by finitely many open balls included in $\mathcal{U}$ which overlap as in the above drawing

Each of the balls are convex, hence f is ~~contin~~ constant on the balls and hence along γ.

Therefore $f(a) = f(b)$ □