June 4th

$f: \mathbb{R} \to \mathbb{R}$

$f(a+h) = f(a) + mh + E(h)$, $\lim_{h \to 0} \frac{E(h)}{h} = 0$

$f: \mathbb{R} \to \mathbb{R}^m$

$\vec{f}(a+h) = \vec{f}(a) + \vec{m} h + \vec{E}(h)$, $\lim_{h \to 0} \frac{|\vec{E}(h)|}{h} = 0$

$\Leftrightarrow$ each f_i is differentiable

Now we are considering $\mathbb{R}^n \to \mathbb{R}$.

does change things

$$f(\vec{a} + \vec{h}) = f(a_1+h_1, a_2+h_2, \cdots, a_n+h_n)$$

Def: $f_x(\vec{a}) = \partial_j f(\vec{a}) = \frac{\partial f(\vec{a})}{\partial x_j} = \lim_{h \to 0} \frac{f(a_1, a_2, \cdots, a_j+h, \cdots a_n) - f(a_1, a_2, \cdots, a_n)}{h}$

By holding all variables except the jth constant, get a function : $\mathbb{R} \to \mathbb{R}$

$\vec{h} = (0, \cdots, h, 0, \cdots, 0)$
↳ jth component

$|\vec{h}| = |h|$

Ex: $f(x,y) = \frac{xy^3}{x^2+y^6}$, $f(0,0) = 0$

Recall along paths $y = cx$ appeared continuous

But for $y = \sqrt[3]{cx}$, it was not cont.

$$\partial_x f(x,y) = \frac{y^3(x^2+y^6) - 2x(xy^3)}{(x^2+y^6)^2}$$

$$\partial_y f(x,y) = \frac{3xy^2(x^2+y^6) - 6y^5(xy^3)}{(x^2+y^6)^2}$$

These are defined away $(0,0)$

$f_x(0,0) = \lim_{h \to 0} \frac{f(h,0) - f(0,0)}{h} = 0 = \lim_{h \to 0} \frac{f(0,h) - f(0,0)}{h} = f_y(0,0)$

* partials exist, go along a path on axis, f is continuous.

But, very much not constant on other paths, not even continuous at $\vec{0}$.

So. partials don't reflect enough at $\vec{a}$.

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

$l(\vec{x}) = b + c_1 x_1 + \cdots + c_n x_n = b + \vec{c} \cdot \vec{x}$

$f(\vec{a}) = l(\vec{a}) = b + \vec{c} \cdot \vec{a} \Rightarrow b = f(\vec{a}) - \vec{c} \cdot \vec{a}$

set $\vec{h} = \vec{x} - \vec{a}$

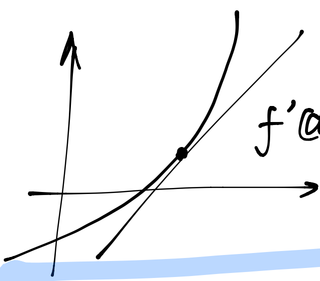
$* \; f(\vec{a}+\vec{h}) = \underbrace{f(\vec{a}) + \vec{c} \cdot \vec{h}}_{\text{linear}} + \underbrace{E(\vec{h})}_{\to \text{error}}$ This one makes sense!

Def'n : $f: \mathbb{R}^n \to \mathbb{R}$ is diff at $\vec{a}$ if $\exists \vec{c}$ s.t.

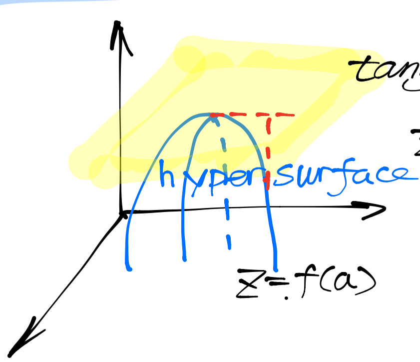
$*$ is true with $\lim_{\vec{h} \to \vec{0}} \frac{E(\vec{h})}{|\vec{h}|} = 0$.

$\vec{c} =: \nabla f(\vec{a})$, "gradient" of f

Note, ∇f is a vector, even though f is a scalar valued.



$f'(a)$ = slope of tangent line

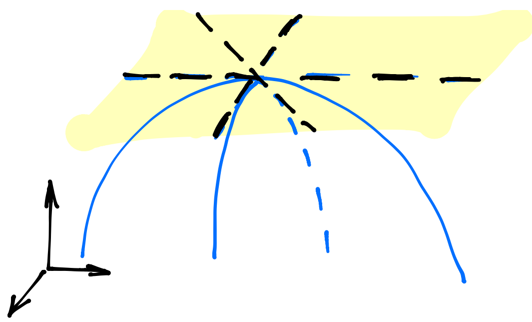


equation of a plane
$z = f(\vec{a}) + \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a})$

$a(x-x_0) + b(y-y_0) + c(z-z_0) = d \leftarrow$ familiar eq'n of plane

plane & surface intersect at $(\vec{a}, f(\vec{a}))$

$z_{surface} - z_{plane} = f(\vec{a}+\vec{h}) - f(\vec{a}) - \nabla f(\vec{a}) \cdot \vec{h}$



namely, partial derivative is restriction to $\vec{h} = (0, \cdots, h, \cdots, 0)$
$\hookrightarrow$ jth component

Differentiability is equivalent to $\lim \dfrac{f(\vec{a}+\vec{h}) - f(\vec{a}) - \nabla f(\vec{a}) \cdot \vec{h}}{|\vec{h}|} = 0$ ⊛

If f is diff. at $\vec{a}$ then all partials exist at $\vec{a}$.
If f is diff at $\vec{a}$ then f is cont. at $\vec{a}$.

Pf: multiply ⊛ by $|\vec{h}|$

$$f(\vec{a}+\vec{h}) - f(\vec{a}) - \underbrace{\nabla f(a) \cdot \vec{h}}_{\to 0 \text{ as } \vec{h} \to \vec{0}}$$

$$\Rightarrow f(\vec{a}+\vec{h}) - f(\vec{a}) = 0 \quad \text{as } \vec{h} \to \vec{0}$$

$\Rightarrow f$ is continuous.

■

Note: Continuity $\not\Rightarrow$ Differentiability

Ex: $f(x) = |x|$ $\leftarrow$ conti. but not diff.

Ex: $\dfrac{xy^3}{x^2+y^6}$ $\leftarrow$ partials exist, but not cont. $\Rightarrow$ not diff.

$\vec{h} = (0, \cdots, h, \cdots, 0)$

$$\lim_{\vec{h} \to \vec{0}} \frac{f(x_1, \cdots, x_j + h, \cdots, x_n) - f(x_1, \cdots, x_n) - c_j h}{h}$$

$\nabla f(\vec{x}) \cdot \vec{h} = \vec{c} \cdot \vec{h}$
for our $\vec{h}$

in other words, $\partial_j f(\vec{a}) = c_j$

$\nabla f(\vec{a})$ has partials for components.

Thm: let f be defined on an open set S and for $\vec{a} \in S$, all partials $\partial_j f$ exist in a neighbourhood of $\vec{a}$, and partials are cont. at $\vec{a}$.
$\hookrightarrow$ $\vec{a}$ is an interior point
Then f is diff.

Prove for n=2.

$$f(\vec{a}+\vec{h})-f(\vec{a}) = [f(a_1+h_1, a_2+h_2) - f(a_1, a_2+h_2)] + [f(a_1, a_2+h_2) - f(a_1, a_2)]$$

Define $g(t) = f(t, a_2+h_2)$ $\quad g: \mathbb{R} \to \mathbb{R}$

$f(a_1+h_1, a_2+h_2) - f(a_1, a_2+h_2) = g(a_1+h_1) - g(a_1)$

By MVT, $\exists c_1$ s.t. difference $= g'(a_1+c_1)h_1 = \partial_1 f(a_1+c_1, a_2+h_2)h_1$

Define $k(t) = f(a_1, t)$ $\quad k: \mathbb{R} \to \mathbb{R}$

$f(a_1, a_2+h_2) - f(a_1, a_2) = \partial_2 f(a_1, a_2+c_2)h_2$

$\vec{c} = (\partial_1 f(\vec{a}), \partial_2 f(\vec{a}))$

$$\frac{f(\vec{a}+\vec{h})-f(\vec{a})-\vec{c}\cdot\vec{h}}{|\vec{h}|} = \frac{\partial_1 f(a_1+c_1, a_2+h_2)h_1 - \partial_1 f(a_1, a_2)h_1}{|\vec{h}|} + \frac{\partial_2 f(a_1, a_2+c_2)h_2 - \partial_2 f(\vec{a})h_2}{|\vec{h}|}$$

! should be h_1, Trefor's wrong.

$\frac{h_1}{|\vec{h}|} \leq 1$ $\quad$ as have continuity of ∂_1 & ∂_2

$\frac{h_2}{|\vec{h}|} \leq 1$ $\quad$ top expressions $\to 0$

$\to$ generalized for general n.

Common trick: – compute partials
- where they are diff. $\Rightarrow$ conti.
- where f is not conti. $\Rightarrow$ not diff. see earlier
- "investigate problem spots"

! NOTE: I think Trefor made an error here as well.

Differentials

$$f(\vec{a}+\vec{h}) - f(\vec{a}) = \nabla f(\vec{a})\cdot\vec{h} + E(\vec{h})$$

Differentials forget error.

$$d_{\vec{a}} f(\vec{h}) := \nabla f(\vec{a})\cdot\vec{h} = \sum_{i=1}^{n} \partial_i f(\vec{a})\cdot h_i$$

$\vec{h} = d\vec{x} = (dx_1, \cdots, dx_n)$ $\quad \leftarrow$ just symbols

$f(\vec{x}) = u$, $\boxed{du = \sum_{i=1}^{n} \partial_i f\, dx_i}$ $\quad$ when we use "d" we usually mean an infitessimal change.

Product Rule

$f(x)=u\ ,\ g(x)=w$

$$d(uw)=\sum_{i=1}^{n}\partial_i(fg)\,dx_i=\sum_{i=1}^{n}\left[(\partial_i f)g+f\partial_i g\right]dx_i=(du)w+u(dw)$$

- linearity
- quotient rule

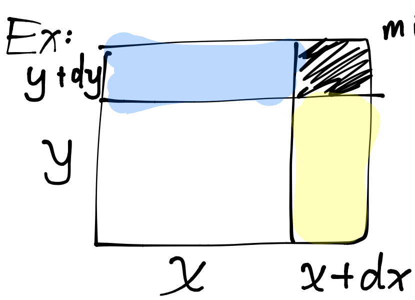
Ex: $V=xyz\ ,\ (x,y,z)=(1,2,3)\longrightarrow(1.1,2.1,3.1)$

$$dV=\frac{\partial V}{\partial x}dx+\frac{\partial V}{\partial y}\cdot dy+\frac{\partial V}{\partial z}\cdot dz \qquad d\vec{x}=(0.1,0.1,0.1)$$

$$=yz\,dx+xz\,dy+xy\,dz$$
$$=6\cdot0.1+3\cdot0.1+2\cdot0.1$$
$$=1.1$$

So dV measures change in volume from $(1,2,3)\longrightarrow(1.1,2.1,3.1)$

Ex:

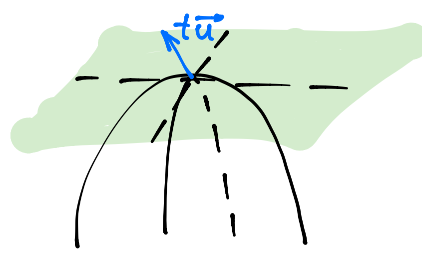


$V=xy$

$dV=x\,dy+y\,dx$

Directional Derivatives

$g(t)=\vec{a}+t\vec{u}\ ,\ |\vec{u}|=1\ ,\ \vec{u}$ is normalized



$$\partial_u f(\vec{a})=\frac{d}{dt}f(\vec{a}+t\vec{u})\Big|_{t=0}$$

$$=\lim_{t\to0}\frac{f(\vec{a}+t\vec{u})-f(\vec{a})}{t}$$

Partials are $\vec{u}=(0,0,\cdots,1,\cdots,0)$

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad \hookrightarrow$ jth component

Thm: If f is differentiable at a, then all directional derivatives exist. and $\partial_u f(\vec{a})=\nabla f(\vec{a})\cdot\vec{u}$

Proof: As f is diff, $\dfrac{E(\vec{h})}{|\vec{h}|} = \dfrac{f(\vec{a}+\vec{h})-f(\vec{a})-\nabla f(\vec{a})\cdot\vec{h}}{|\vec{h}|}$

with $\vec{h} = t\vec{u}$

For $t>0$, $|\vec{h}| = |t\vec{u}| = t$

$\dfrac{f(\vec{a}+\vec{h})-f(\vec{a})-(\nabla f(\vec{a})\cdot\vec{u})}{t}$

For $t<0$, $\dfrac{f(\vec{a}+\vec{h})-f(\vec{a})+\nabla f(\vec{a})\cdot\vec{u}}{-t}$

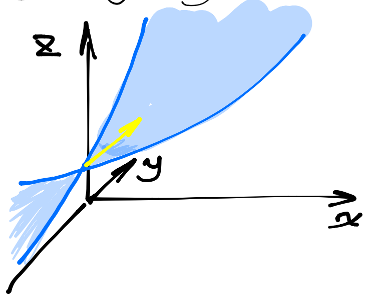
Hence as $t \to 0$ in either case has $\partial_u f(\vec{a}) = \nabla f(\vec{a})\cdot\vec{u}$ $\blacksquare$

$\partial_u f(\vec{a}) \le |\nabla f(\vec{a})||\vec{u}|$ get equality when $\nabla f(\vec{a})$ and $\vec{u}$ are parallel.

Cauchy Ineq.

* Namely, $\nabla f(\vec{a})$ points in the direction $\vec{u}$ of maximal slope.

Ex: $f(x,y) = e^{x+y} = z$



$f(0,0) = 1$

$\nabla f(0,0) = (e^{x+y}, e^{x+y})|_{(0,0)} = (1,1)$

$\partial_u f(0,0) = \nabla f(0,0)\cdot\vec{u} = (1,1)\cdot(1,0) = 1$

$\vec{u} = (1,0)$

$1 = |\partial_u f(0,0)| < |\nabla f(0,0)| = \sqrt{2}$

directional derivative in $(1,0)$ direction.

---

§ 2.3 Chain Rule

for $f:\mathbb{R}\to\mathbb{R}$, $g:\mathbb{R}\to\mathbb{R}$

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

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

$g:\mathbb{R}^m \to \mathbb{R}^n \longrightarrow f\circ g:\mathbb{R}^m\to\mathbb{R}^n\to\mathbb{R}^k$

MVT (I) $:\mathbb{R}\to\mathbb{R}^n\to\mathbb{R}$

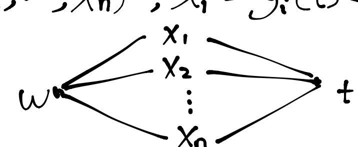
MVT (II) $:\mathbb{R}^m\to\mathbb{R}^n\to\mathbb{R}$

MVT (III) $:\mathbb{R}^m\to\mathbb{R}^n\to\mathbb{R}^k$ (Talk about these in later lecs.)

MVT(I)

$f(\vec{x}) = f(x_1,\dots,x_n)$, $x_i = g_i(t)$ ← independent variable

$w = f(\vec{x})$



Consider $\varphi:\mathbb{R}\to\mathbb{R}$, $\varphi = f(g(x))$

MVT #1

Sps g_i are diff at $t=a$
f is diff. at $\vec{x}=\vec{b}=\vec{g}(\vec{a})$

then $\varphi(t)$ is diff at $t=a$ and $\varphi'(a)=\nabla f(b)\cdot\vec{g}'(a)$.

Proof: $f(\vec{b}+\vec{h})=f(\vec{b})+\nabla f(\vec{b})\cdot\vec{h}+E(\vec{h})$, where $\lim_{\vec{h}\to\vec{0}}\frac{E(\vec{h})}{|\vec{h}|}=0$

$\vec{g}(a+u)=\vec{g}(a)+\vec{g}'(a)u+\vec{E}(u)$, where $\lim_{u\to 0}\frac{|\vec{E}(u)|}{u}=0$

$$\vec{h}=\vec{g}(a+u)-\vec{g}(a)=\vec{g}(a+u)-b$$
$$=\vec{g}'(a)u+\vec{E}(u)$$

?

$$\varphi(a+u)=f(g(a+u))=f(\vec{h}+\vec{b})=f(\vec{b})+\nabla f(\vec{b})\cdot\vec{h}+E(\vec{h})$$
$$=\varphi(a)+\nabla f(\vec{b})[\underbrace{g'(a)u}_{\text{linear}}+\underbrace{\vec{E}(u)]+E(\vec{h})}_{\text{Error } E_\varphi}$$

namely this shows, $\varphi'(a)=\nabla f(\vec{b})\cdot\vec{g}'(a)$ if E_φ is small.

Now we ~~finally need~~ to show E_φ small

$$E_\varphi(u)=\underbrace{\nabla f(\vec{b})\cdot\vec{E}_g(u)}_{\leq|\nabla f(\vec{b})|\,|\vec{E}_g(u)|}+E_f(\vec{h})$$

both sides divided by u

$$\frac{|\nabla f(\vec{b})\cdot\vec{E}_g(u)|}{u}\leq\frac{|\nabla f(\vec{b})|\,|\vec{E}_g(u)|}{u}\to 0 \text{ as } u\to 0 \quad \text{as } \frac{|\vec{E}_g(u)|}{u}\to 0$$

So $|\vec{E}_g(u)|\leq|\vec{u}|$ for small $\vec{u}$

$$|\vec{h}|=|\vec{g}'(a)u+\vec{E}_g(u)|\leq|u|\underbrace{(|\vec{g}'(a)|+1)}_{\text{number}}$$

so $\vec{h}\to\vec{0}$ as $u\to 0\Rightarrow\frac{E_f(h)}{u}$ gets small as $u\to$ small. ∎