Recall for a fcn $f: D \to \mathbb{R}^p$, $\underbrace{D \subseteq \mathbb{R}^n}_{\text{open}}$ and $\vec{x}_0 \in D$

We say f is diff'able at $\vec{x}_0$ if $\exists$ a linear map: $f'(\vec{x}_0): \mathbb{R}^n \to \mathbb{R}^p$ such that:

$$\lim_{\vec{x} \to \vec{x}_0} \frac{\| f(\vec{x}) - f'(\vec{x}_0)(\vec{x} - \vec{x}_0) - f(\vec{x}_0) \|}{\| \vec{x} - \vec{x}_0 \|} = 0$$

aka $f(\vec{x}) \approx f'(\vec{x}_0)(\vec{x} - \vec{x}_0) + f(\vec{x}_0)$ for $\vec{x} \approx \vec{x}_0$

Let $e(x) := \dfrac{f(\vec{x}) - f'(\vec{x}_0)(\vec{x} - \vec{x}_0) - f(\vec{x}_0)}{\| \vec{x} - \vec{x}_0 \|}$

Note: $e(\vec{x})$ is a vector in $\mathbb{R}^p$

so:

$f(\vec{x}) = f'(\vec{x}_0)(\vec{x} - \vec{x}_0) + f(\vec{x}_0) + \| \vec{x} - \vec{x}_0 \| e(\vec{x})$

To say $f'(\vec{x}_0)$ is the derivative is equivalent to:

$\lim_{\vec{x} \to \vec{x}_0} e(\vec{x}) = \vec{0} \in \mathbb{R}^p$

This makes precise what we mean by "good linear approximation to f near $\vec{x}_0$"

Recall

- If f differentiable a $\vec{x}_0$ then all np partials exist then all np partials exist and $f'(x_0)$ is represented by the matrix of partials
- If all partials exist and are continuous in a nbhd of $\vec{x}_0$ then f is diff'able at $\vec{x}_0$
- In anomolous cases, the partials might exist but f is not diff'able at $\vec{x}_0$

Thm Chain Rule:

If $f: D_1 \to \mathbb{R}^p$, $g: D_2 \to \mathbb{R}^q$, $D_1 \subseteq \mathbb{R}^n$, $D_2 \subseteq \mathbb{R}^p$, $\vec{x}_0 \in D_1$ and $f(\vec{x}_0) \in D_2$ and f diff'able at $\vec{x}_0$ and g diff'able at $\vec{x}_0$ and

$$(g \circ f)'(\vec{x}_0) = g'(f(\vec{x}_0)) \, f'(\vec{x}_0)$$

$\uparrow$ mat mul

Proof sketch

say $f(\vec{x}) = f'(\vec{x}_0)(\vec{x} - \vec{x}_0) + f(\vec{x}_0) + \| \vec{x} - \vec{x}_0 \| e_1(\vec{x}_0)$

($\bigstar$) $g(\vec{y}) = g'(f(\vec{x}_0))(\vec{y} - f(\vec{x}_0)) + g(f(\vec{x}_0)) + \| \vec{y} - f(\vec{x}_0) \| e_2(\vec{y})$

think: $\vec{y}_0 = f(\vec{x}_0)$

Plug in $\vec{y} = f(\vec{x})$,

then $\underbrace{f(\vec{x})}_{=y} = f'(\vec{x}_0)(\vec{x} - \vec{x}_0) + f(\vec{x}_0) + \| \vec{x} - \vec{x}_0 \| e(x)$

and then plug this into ($\bigstar$)

$g(f(\vec{x})) = g(\vec{y}) = g'(\vec{y}_0)(f'(\vec{x}_0)(\vec{x} - \vec{x}_0)) + g(\vec{y}_0) +$ error terms involving e_1 and e_2

you'll get a term of the form:

$g'(\vec{y}_0)(\| \vec{x} - \vec{x}_0 \| e_1(\vec{x})) = \| \vec{x} - \vec{x}_0 \| \underbrace{g'(\vec{y}_0)(e_1(\vec{x}))}_{\text{goes to } 0 \text{ as } \vec{x} \to \vec{x}_0}$

Conclusion

— To compute, you use matmul but proof of chain rule just uses composition of linear fcns

Note Key case is $n=q=1$
then f is a vector-valued fcn of one input
g is a scalar-valued fcn
so f' is same as in first few weeks
g' is ∇g viewed as a row vector
and $\underbrace{g'(f(\vec{x_0}))}_{\text{row vector}}\underbrace{f'(\vec{x_0})}_{\text{column vector}}$

$= \nabla g(f(\vec{x_0})) \underset{\text{dot prod}}{\cdot} f'(\vec{x_0})$

"Key case" bc you can prove multidim chain rule using this case (once for each of nq coeffs of $(g \circ f)'$)

Note formula for direction derivative
$D_{\vec{u}} g = \nabla g \cdot \vec{u}$ is a special case of chain rule
(where $f(t) = t\vec{u} + \vec{x_0}$)

Maxima & Minima

Suppose $f: D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$ and $\vec{x_0} \in D$

Def We say that f has a:

① Local max at $\vec{x_0}$ if
$f(\vec{x}) \leq f(\vec{x_0})$ $\forall \vec{x}$ near $\vec{x_0}$
(ie $\exists \varepsilon > 0$ s.t. it's true for all $\|x - \vec{x_0}\| < \varepsilon$)

② Local min at $\vec{x_0}$ if
$f(\vec{x}) \geq f(\vec{x_0})$ $\forall \vec{x}$ near $\vec{x_0}$

③ Global max at $\vec{x_0}$ if
$f(\vec{x}) \leq f(\vec{x_0})$ $\forall \vec{x} \in D$

④ Global min at $\vec{x_0}$ if
$f(\vec{x}) \geq f(\vec{x_0})$ $\forall \vec{x} \in D$

Note global max/min $\Rightarrow$ local max/min
$\not\Leftarrow$

Note for $n=1$, if local max/min then $f'(x_0) = 0$
Similarly for general n, if f has a local max/min at $\vec{x_0}$ then
$\nabla f(\vec{x_0}) = 0$

Also for $n=1$ sometimes $f'(x_0) = 0$ but f doesn't have a local max
Similarly can have $\nabla f(\vec{x_0}) = 0$ but no local max or min

eg

① $f(x,y) = x^3 + y^3$ $\quad \nabla f = (3x^2, 3y^2)$ $\quad x_0 = (0,0)$
but not local max/min $\quad \nabla f(x_0) = (0,0)$
(2D version of $f(x) = x^3$)

② $f(x,y) = x^2 - y^2$
$x_0 = (0,0)$ $\quad \nabla f = (2x, -2y)$ $\quad \nabla f(x_0) = (0,0)$
$\nabla f = \vec{0}$ but not a local min/max
↳ called a saddle point (fundamentally multidim)

Defn If $\nabla f(\vec{x_0}) = \vec{0}$ then we say that $\vec{x_0}$ is a critical point of f

Thus
- any local max/min is a critical point
- above we gave ex of critical pts that weren't max/min

Recall in 1 var, if:
$f''(x_0) > 0 \longrightarrow$ local min
$f''(x_0) < 0 \longrightarrow$ local max
$f''(x_0) = 0 \longrightarrow$ unclear

in multivar:
Define Hessian:

If $\vec{x_0}$ is a critical pt of F,

$F: D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$

Then define an $n \times n$ matrix of 2^{nd} derivatives, whose ij – coeff is $\dfrac{\partial^2 F}{\partial x_i x_j}$

<u>Notice</u> ij – coeff equals the ji – coeff

$\Rightarrow$ it's a symmetric matrix

<u>eg</u> $n = 2$, $x_1 = x$, $x_2 = y$ $\rightarrow$ $\vec{x}_0 = (x, y)$

$$\text{Hess}_{\vec{x_0}}(F) = \begin{bmatrix} \dfrac{\partial^2 F}{\partial x^2} & \dfrac{\partial^2 F}{\partial x \partial y} \\[2ex] \dfrac{\partial^2 F}{\partial y \partial x} & \dfrac{\partial^2 F}{\partial y^2} \end{bmatrix}$$

$$D = \det(\text{Hess}_{\vec{x_0}}(F))$$

$$= D_{xx} F \quad D_{yy} F \longleftarrow (D_{xy} F)^2$$

↳ can ask – is this scalar positive/negative

<u>Two-variable 2^{nd} derivative Test</u>

① If $D > 0$, then F has a local max or local min at $\vec{x_0}$

② If $D < 0$, then F has a saddle point

③ If $D = 0$, then the test doesn't determine what happens.

<u>Remark</u>

– In case $D > 0$, you can tell if local max/min by finding the eigenvalues of the Hessian

→ positive eigenvalues: local min

→ negative eigenvalues: local max

<u>Comment on saddlepoints</u>

- when F has a local max in one direction and a local min in the other
- NOT like 1-D critical pts that aren't a local max/min

→ rather, you have a local max in 1-D and a local min in an orthogonal direction (fundamentally multidim)

<u>eg</u> $F(x, y) = x^2 - y^2$ at $(0, 0)$

<u>then</u> if you fix $y = 0$ and let x vary, then F has a local min at x_0

if you fix $x = 0$ and let y vary then you get a local max at x_0

<u>eg</u> $F(x, y) = xy$

<u>then</u> it's a local min in the direction $\vec{u} = (1,1)$

ie, $D_{\vec{u}} D_{\vec{u}} F > 0$

<u>but</u> local max in direction $\vec{u} = (1, -1)$

<u>Intro to Inverse Fcn Thm</u>

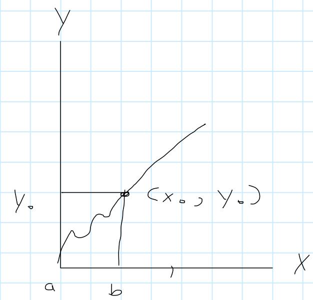
Let $f: [a, b] \to \mathbb{R}$ (... say ...

Intro to Inverse Fcn Thm

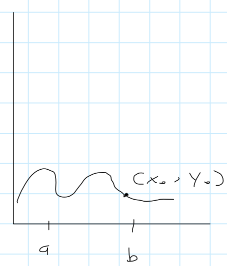
Let $f: [a,b] \to \mathbb{R}$ (one-var fcn). say $x_0 \in (a,b)$, say f is diffable
$y_0 = f(x_0)$

Consider 3 cases for $f'(x_0)$. If

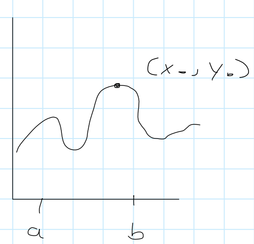
(1) $f'(x_0) > 0$



(2) $f'(x_0) < 0$



(3) $f(x_0) = 0 \to$ suppose $f''(x_0) < 0$



Suppose we want to inverse the function $f^{-1}(y)$

$$x = f^{-1}(y) \iff y = f(x)$$

Case (1)

- say we want $f^{-1}(y_0)$ that should be x_0
- say we want $f^{-1}(y_1)$ have ≥ 2 possibilities for its value

but if y near y_0 and $f'(x_0) \neq 0$, can choose $f^{-1}(y)$ consistently for y near y_0

but not necessarily if $f'(x_0) = 0$