Lecture 28: Chain rule (concl.)

Suppose you are interested in a composition

$$\begin{tikzcd} \mathcal{J} \arrow[r, "G"] \arrow[rr, bend left, "H"] \arrow[d, phantom, "\cap" rotate=90] & \mathcal{S} \arrow[r, "F"] \arrow[d, phantom, "\cap" rotate=90] & \mathbb{R}^m \\ \mathbb{R}^p & \mathbb{R}^n & \end{tikzcd}$$

$$\mathcal{J} \ni \vec{a} \longmapsto \vec{b} \in \mathcal{S}$$

coords: $\vec{x}$ $\quad \vec{y}$ $\quad \vec{z}$

and want the total derivative $H'(\vec{a}) : \mathbb{R}^p \to \mathbb{R}^m$ at $\vec{x} = \vec{a}$.

Apostol's notation

i.e. H is differentiable at $\vec{a}$

If it exists, this is a linear transformation with the property that $H(\vec{a}+\vec{u}) - H(\vec{a}) = H'(\vec{a})\,\vec{u} + \|\vec{u}\|\, E_H(\vec{a};\vec{u})$ where $E_H(\vec{a};\vec{u}) \to 0$ with $\vec{u}$. (This is the <u>definition</u>.) As we saw last time, when $H'(\vec{a})$ exists it is represented by the "Jacobian matrix" $J_H(\vec{a}) = \left[\frac{\partial h_i}{\partial x_j}(\vec{a})\right]$.

<u>Theorem</u>: If F is differentiable at $\vec{b} = G(\vec{a})$ and G is differentiable at $\vec{a}$, then H is differentiable at $\vec{a}$ and

$$\begin{cases} H'(\vec{a}) = F'(\vec{b}) \circ G'(\vec{a}) & \text{(composition of L.T.s)} \\ J_H(\vec{a}) = J_F(\vec{b}) \cdot J_G(\vec{a}) & \text{(matrix multiplication).} \end{cases}$$

<u>Proof</u>: $H(\vec a+\vec u) - H(\vec a) = F(\underbrace{G(\vec a+\vec u)}) - F(G(\vec a))$

$$\left(\begin{array}{l}\text{since}\\ G \text{ diff.}\end{array}\right) = G(\vec a) + \underbrace{G'(\vec a)\vec u + \|\vec u\|\, E_G(\vec a;\vec u)}_{\text{call this } \vec v}$$

$$= F(\vec b+\vec v) - F(\vec b)$$

$$= F'(\vec b)\vec v + \|\vec v\|\, E_F(\vec b;\vec v)$$

$$= F'(\vec b)\,G'(\vec a)\,\vec u + \|\vec u\|\left(F'(\vec b)\overbrace{E_G(\vec a;\vec u)}^{\to 0} + \underbrace{\frac{\|\vec v\|}{\|\vec u\|}}_{\text{bounded}}\underbrace{E_F(\vec b;\vec v)}_{\to 0}\right)$$

0 as $\vec u \to \vec 0$

$\square$

ASIDE:

$G'(\vec a)\,\vec u = \sum_{k=1}^{n} (\vec\nabla g_k(\vec a)\cdot\vec u)\,e_k$

$\Rightarrow \|G'(\vec a)\vec u\| \le \left(\sum \|\vec\nabla g_k(\vec a)\|\right)\|\vec u\|$

Cauchy-Schwarz + Δ ineq.

$\Rightarrow \frac{\|\vec v\|}{\|\vec u\|} \le \sum \|\vec\nabla g_k(\vec a)\| + E_G(\vec a;\vec u)$ is bounded.

So in matrix form we get

$$\left[\frac{\partial h_i}{\partial x_j}(\vec a)\right] = \left[\frac{\partial f_i}{\partial y_k}(G(\vec a))\right]\left[\frac{\partial g_k}{\partial x_j}(\vec a)\right],$$

i.e. for each fixed i, j $\quad \partial h_i/\partial x_j = \sum_{k=1}^{n} \frac{\partial f_i}{\partial y_k}\frac{\partial g_k}{\partial x_j}$.

Ex/ Consider the composition

$$\begin{tikzcd} \mathbb{R}^2 \arrow[r, "G"] \arrow[rr, bend right, "\tilde f"'] & \mathbb{R}^2 \arrow[r, "f"] & \mathbb{R} \end{tikzcd}$$

(r,θ) $\quad (x,y)$ $\quad z$

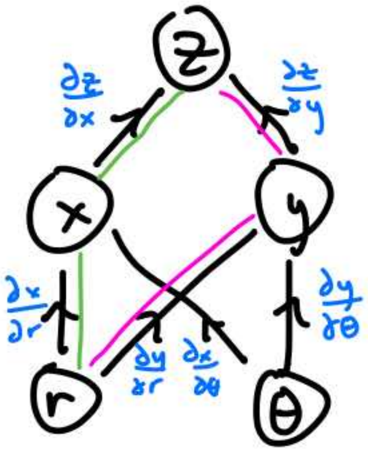
where $G(r,\theta) := (\underset{g_1}{r\cos\theta}, \underset{g_2}{r\sin\theta})$.

Then f' means $\vec\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$, $\tilde f'$ means $\vec\nabla \tilde f = \left(\frac{\partial \tilde f}{\partial r}, \frac{\partial \tilde f}{\partial \theta}\right)$,

and $G' = \begin{pmatrix} \partial g_1/\partial r & \partial g_1/\partial\theta \\ \partial g_2/\partial r & \partial g_2/\partial\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$. The chain rule

says $\vec\nabla \tilde f = \vec\nabla f \cdot G' = \left(\frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta,\ -\frac{\partial f}{\partial x}r\sin\theta + \frac{\partial f}{\partial y}r\cos\theta\right)$.

A more efficient way to work this out / memory device is given by the picture:



which you are supposed to read as

$$\frac{\partial z}{\partial r} = \underbrace{\frac{\partial z}{\partial x}\frac{\partial x}{\partial r}} + \underbrace{\frac{\partial z}{\partial y}\frac{\partial y}{\partial r}}$$

sum over paths from r to z
the products of labels along the path.

Clairaut's Theorem

Let $f(x,y)$ be a function whose second partials exist in a neighborhood of $(0,0)$. Is there a difference between

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = f_{yx} \quad \text{and} \quad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = f_{xy} \ ?$$

The definition says that $f_x(0,y) = \lim_{\Delta x \to 0} \dfrac{f(\Delta x, y) - f(0,y)}{\Delta x}$, so

$$f_{xy}(0,0) = \lim_{\Delta y \to 0} \frac{\left(\lim_{\Delta x \to 0} \frac{f(\Delta x, \Delta y) - f(0,\Delta y)}{\Delta x}\right) - \left(\lim_{\Delta x \to 0} \frac{f(\Delta x, 0) - f(0,0)}{\Delta x}\right)}{\Delta y}$$

$$= \lim_{\Delta y \to 0} \lim_{\Delta x \to 0} \frac{f(\Delta x, \Delta y) - f(0,\Delta y) - f(\Delta x, 0) + f(0,0)}{\Delta y \, \Delta x}.$$

This is symmetric in x & y except for the order of the limits.

The question is whether we can switch those.

Ex/ $f(x,y) = \begin{cases} \dfrac{xy(x^2-y^2)}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0). \end{cases}$

$$f_x(0,y) = \frac{x^4y + 4x^2y^3 - y^5}{(x^2+y^2)^2}\Big|_{x=0} = -\frac{y^5}{y^4} = -y$$

$$\Rightarrow f_{xy}(0,0) = -1$$

$$f_y(x,0) = \frac{x^5 - 4x^3y^2 - xy^4}{(x^2+y^2)^2}\Big|_{y=0} = \frac{x^5}{x^4} = x$$

$\Rightarrow f_{yx}(0,0) = 1$. Notice also that f, f_x, f_y are continuous about $(0,0)$ (use the polar form + squeeze thm.) hence f is actually <u>differentiable</u> there! What is going on?

Well, the full $f_{xy}(x,y) = \dfrac{x^6 + 9x^4y^2 - 9x^2y^4 + 15y^6}{(x^2+y^2)^3}$ is evidently <u>not</u> continuous at $(0,0)$ (consider the limits along the vertical & horizontal axes)... this suggests a way out:

<u>Theorem</u>: Assume f_{xy} & f_{yx} are continuous in a neighborhood of $(0,0)$. Then $f_{xy}(0,0) = f_{yx}(0,0)$.

<u>Proof</u>: Write $\Delta(h) := f(h,h) - f(h,0) - f(0,h) + f(0,0)$.

Setting $g(x) := f(x,h) - f(x,0)$, for some $a \in [0,h]$

$$(*) \quad \frac{\Delta(h)}{h} = \frac{g(h) - g(0)}{h} \underset{MVT}{=\!=} g'(a) = f_x(a,h) - f_x(a,0).$$

Setting $G(y) := f_x(a,y)$, for some $b \in [0,h]$

$$\frac{\Delta(h)}{h^2} \underset{(*)}{=} \frac{G(h) - G(0)}{h} \underset{MVT}{=\!=} G'(b) = f_{xy}(a,b)$$

and $\lim_{h \to 0} \frac{\Delta(h)}{h^2} = \lim_{(a,b)\to(0,0)} f_{xy}(a,b) = f_{xy}(0,0)$ by continuity of f_{xy}.

An exactly symmetric argument (in x & y) gives

$$\lim_{h\to 0} \frac{\Delta(h)}{h^2} = f_{yx}(0,0), \qquad \text{proving the Theorem.} \qquad \square$$