Chain Rule

Professor David Barrett
Transcribed by Thomas Cohn

9/24/18

Goal: (*) $\frac{g(f(\vec{a}+\vec{h}))-g(f(\vec{a}))-D_g(\vec{b})(D_f(\vec{a})(\vec{h}))}{||\vec{h}||} \to \vec{0}$ as $\vec{h} \to \vec{0}$.

Proof: $F(\vec{h}) = \begin{cases} \frac{f(\vec{a}+\vec{h})-f(\vec{a})-D_f(\vec{a})(\vec{h})}{||\vec{h}||} & \vec{h} \neq \vec{0} \\ \vec{0} & \vec{h} = \vec{0} \end{cases}$.

Note that $D_f(\vec{a})$ is a bounded, linear map from V to W (according to the handout).

F is continuous at $\vec{0}$.
$\vec{k} = f(\vec{a}+\vec{h}) - f(\vec{a}) = D_f(\vec{a})(\vec{h}) + \left|\left|\vec{h}\right|\right| F(\vec{h})$.
Similarly, $g(\vec{b}+\vec{k}) - g(\vec{b}) = D_g(\vec{b})(\vec{k}) + ||k||\, G(\vec{k})$ with $G(\vec{0}) = \vec{0}$, so G is continuous at $\vec{0}$.

$$
\begin{aligned}
\text{So } (*) &= \frac{g(\vec{b}+\vec{k})-g(\vec{b})-D_g(\vec{b})(D_f(\vec{a})(\vec{h}))}{||h||} \\
&= \frac{D_g(\vec{b})(\vec{k})+\left|\left|\vec{k}\right|\right|G(\vec{k})-D_g(\vec{b})(D_f(\vec{a})(\vec{h}))}{\left|\left|\vec{h}\right|\right|} \\
&= \frac{\cancel{D_g(\vec{b})(D_f(\vec{a})(\vec{h}))}+\left|\left|\vec{h}\right|\right|D_g(\vec{b})(F(\vec{h}))+\left|\left|\vec{k}\right|\right|G(\vec{k})-\cancel{D_g(\vec{b})(D_f(\vec{a})(\vec{h}))}}{\left|\left|\vec{h}\right|\right|} \\
&= \frac{\left|\left|\vec{h}\right|\right|D_g(\vec{b})(F(\vec{h}))+\left|\left|\vec{k}\right|\right|G(\vec{k})}{\left|\left|\vec{h}\right|\right|} \\
&= D_g(\vec{b})(F(\vec{h})) + \frac{\left|\left|\vec{k}\right|\right|}{\left|\left|\vec{h}\right|\right|}G(\vec{k}).
\end{aligned}
$$

If $\vec{h} \to \vec{0}$, then $\vec{k} \to \vec{0}$, so $G(\vec{k}) \to \vec{0}$. And $F(\vec{h}) \to \vec{0}$.
$\frac{\left|\left|\vec{k}\right|\right|}{\left|\left|\vec{h}\right|\right|} = \left|\left|D_f(\vec{a})\left(\frac{\vec{h}}{\left|\left|\vec{h}\right|\right|}\right) + F(\vec{h})\right|\right|$. $F(\vec{h}) \to \vec{0}$, and $\frac{\vec{h}}{\left|\left|\vec{h}\right|\right|} = 1$, so $D_f(\vec{a})\left(\frac{\vec{h}}{\left|\left|\vec{h}\right|\right|}\right)$ is bounded; thus $\frac{\left|\left|\vec{k}\right|\right|}{\left|\left|\vec{h}\right|\right|} \to \vec{0}$.
So $D_g(\vec{b})(F(\vec{h})) + \frac{\left|\left|\vec{k}\right|\right|}{\left|\left|\vec{h}\right|\right|}G(\vec{k}) \to \vec{0}$ as $\vec{h} \to \vec{0}$. $\square$

Ex: $\left.\begin{array}{l} \alpha : W \times W \to W \\ (\vec{w_1}, \vec{w_2}) \mapsto \vec{w_1} + \vec{w_2} \end{array}\right]$-bounded, linear. $D_\alpha(\vec{w_1}, \vec{w_2}) = \alpha$

Ex: Suppose $A \underset{f_2}{\overset{f_1}{\rightrightarrows}} W$ is differentiable at $\vec{a} \in A$ (where A is an open subset of V).

Then $D_{(f_1,f_2)}(\vec{a}) = (D_{f_1}(\vec{a}), D_{f_2}(\vec{a}) \in B(V, W \times W)$ where $B(V, W \times W)$ is the set of bounded linear maps.

$f_1 + f_2 = \alpha \circ (f_1, f_2) : A \to W$.
Chain Rule: $f_1 + f_2$ differentiable at $\vec{a}$, so
$D_{f_1+f_2}(\vec{a}) = D_\alpha(f_1(\vec{a}), f_2(\vec{a}))(D_{f_1}(\vec{a}), D_{f_2}(\vec{a})) = D_{f_1}(\vec{a}) + D_{f_2}(\vec{a})$

Ex: $\mu : \mathbb{R}^2 \to \mathbb{R}$, $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1 x_2$.

From last week: $D_\mu \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{bmatrix} D_{1\mu}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & D_{2\mu}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \end{bmatrix} = \begin{bmatrix} x_2 & x_1 \end{bmatrix}$.

This is, of course, assuming the derivative exists. Prove that the derivative exists.

Ex: Suppose $A \underset{f_2}{\overset{f_1}{\rightrightarrows}} \mathbb{R}$ is differentiable at $\vec{a} \in A$.

$\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} : A \to \mathbb{R}^2$

$f_1 f_2 = \mu \circ \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} : A \to \mathbb{R}$

Chain Rule: $f_1 f_2$ differentiable at $\vec{a}$.

$$D_{f_1 f_2}(\vec{a})(\vec{u}) = \left(D_\mu \begin{pmatrix} f_1(\vec{a}) \\ f_2(\vec{a}) \end{pmatrix} \circ \begin{pmatrix} D_{f_1}(\vec{a}) \\ D_{f_2}(\vec{a}) \end{pmatrix} \right)(\vec{u})$$
$$= \begin{bmatrix} f_2(\vec{a}) & f_1(\vec{a}) \end{bmatrix} \begin{bmatrix} D_{f_1}(\vec{a})(\vec{u}) \\ D_{f_2}(\vec{a})(\vec{u}) \end{bmatrix}$$

Rewrite: $D_{(f_1,f_2)}(\vec{a}) = f_2(\vec{a}) \cdot D_{f_1}(\vec{a}) + f_1(\vec{a}) \cdot D_{f_2}(\vec{a})$.

What about multiplication of vector-valued f_1, f_2 (that is, f_1 and f_2 output vectors)?

- Not defined in general.
- Works if f_1 is $\mathbb{R}$-valued and f_2 is vector-valued.
- Works if f_1, f_2 map into some inner product space W.
- Works if $f_1 \to W_1$, $f_2 \to W_2$, with $\vec{a} \mapsto (f_1(\vec{a}), f_2(\vec{a})) \in W_1 \times W_2$.
- Works if $f_1 \to \text{Mat}\,(n, m, \mathbb{R})$, $f_2 \to \text{Mat}\,(m, p, \mathbb{R})$, $\vec{a} \mapsto f_1(\vec{a}) f_2(\vec{a}) \in \text{Mat}\,(n, p, \mathbb{R})$.

Guess: f_1, f_2 matrix-valued functions.
$D_{(f_1,f_2)}(\vec{a})(\vec{h}) = D_{f_1}(\vec{a})(\vec{h}) f_2(\vec{a}) + f_1(\vec{a}) D_{f_2}(\vec{a})(\vec{h})$.
Exercise: check this. It will be on an upcoming homework!

Other examples:

f_1, f_2 are $\mathbb{C}$-valued. Identify $a + ib$ with $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$.

The multiplication of "tensors"

Now suppose $A \subset V$ open, with $f : A \to W$ differentiable at each $\vec{a} \in A$. Then f is continuous on A.
Get $D_f : A \to B(V, W)$ (on a normed vector space – see handout).

Defn: f is <u>continuously differentiable</u> on $A \leftrightarrow D_f$ is continuous on A. If this is the case, we say that f is C^1, i.e., $f \in C^1$.

Note that continuity of the derivative is *not* automatic.