Lec 1/31

Tuesday, January 31, 2017 09:07

Theorem (Chain rule) Suppose $\vec{f}: \underset{\text{open}}{U} \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$ is differentiable at $\vec{a} \in U$ and $\vec{g}: V \to \mathbb{R}^p$ is differentiable at $\vec{f}(\vec{a}) \in V$. Then $\vec{g} \circ \vec{f}: U \to \mathbb{R}^p$ is differentiable at $\vec{a}$ and

$$\underbrace{D(\vec{g} \circ \vec{f})(\vec{a})}_{p \times n} = \underbrace{D\vec{g}(\vec{f}(\vec{a}))}_{p \times m} \cdot \underbrace{D\vec{f}(\vec{a})}_{m \times n}$$

$$\left(A = D\vec{f}(\vec{a}),\ B = D\vec{g}(\vec{f}(\vec{a})) \right)$$

Proof: Given:

$$\vec{f}(\vec{a}+\vec{h}) = \vec{f}(\vec{a}) + A\vec{h} + |\vec{h}|\vec{\varepsilon}(\vec{h}) \qquad \vec{h} \in B(r, \vec{0}) \qquad \vec{\varepsilon}(\vec{h}) \to \vec{0} \text{ as } \vec{h} \to \vec{0}.$$

$$\vec{g}(\vec{f}(\vec{a})+\vec{k}) = \vec{g}(\vec{f}(\vec{a})) + B\vec{k} + |\vec{k}|\vec{\gamma}(\vec{k}) \qquad \vec{k} \in B(s, \vec{0}) \qquad \vec{\gamma}(\vec{k}) \to \vec{0} \text{ as } \vec{k} \to \vec{0}.$$

wts

$$(\vec{g} \circ \vec{f})(\vec{a}+\vec{h}) = (\vec{g} \circ \vec{f})(\vec{a}) + BA\vec{h} + |\vec{h}|\vec{\delta}(\vec{h}) \qquad \vec{h} \in B(r', \vec{0}) \text{ where } r' \leq r,\ \vec{\delta}(\vec{h}) \to \vec{0} \text{ as } \vec{h} \to \vec{0}.$$

Plug 1 into 2:

$$\vec{g}(\vec{f}(\vec{a}+\vec{h})) = \vec{g}(\vec{f}(\vec{a}) + \overbrace{A\vec{h} + |\vec{h}|\vec{\varepsilon}(\vec{h})}^{k}) \qquad \text{choose } r' \leq r \text{ s.t. } \vec{h} \in B(r', \vec{0}) \Rightarrow k \in B(s, \vec{0}).$$

$$= \vec{g}(\vec{f}(\vec{a})) + B(A\vec{h} + |\vec{h}|\vec{\varepsilon}(\vec{h})) + |\vec{k}|\vec{\gamma}(\vec{k})$$

$$= \vec{g}(\vec{f}(\vec{a})) + BA\vec{h} + \underbrace{|\vec{h}|B\vec{\varepsilon}(\vec{h}) + |\vec{k}|\vec{\gamma}(\vec{k})}$$

$$|\vec{h}|\vec{\delta}(\vec{h}) \Rightarrow \vec{\delta}(\vec{h}) = \frac{|\vec{h}|B\vec{\varepsilon}(\vec{h}) + |\vec{k}|\vec{\gamma}(\vec{k})}{|\vec{h}|}$$

Now want to show $\vec{\delta}(\vec{h}) \to \vec{0}$ as $\vec{h} \to \vec{0}$:

$$\lim_{\vec{h}\to\vec{0}} B\vec{\varepsilon}(\vec{h}) + \frac{|\vec{k}|}{|\vec{h}|}\vec{\gamma}(\vec{k}) = \overset{\substack{\text{linear maps are continuous.} \\ \downarrow}}{0} + \lim_{\vec{h}\to\vec{0}} \frac{|A\vec{h} + |\vec{h}|\vec{\varepsilon}(\vec{h})|}{|\vec{h}|}\vec{\gamma}(\vec{k})$$

$$\leq \lim_{\vec{h}\to\vec{0}} \left|A\frac{\vec{h}}{|\vec{h}|}\vec{\gamma}(\vec{k})\right| + \left|\vec{\varepsilon}(\vec{h})\ \vec{\gamma}(\vec{k})\right|$$

$$= \lim_{\vec{h}\to\vec{0}} \left|A\frac{\vec{h}}{|\vec{h}|}\vec{\gamma}(\vec{k})\right| + 0 \quad \leftarrow \text{both approach 0}$$

$$= \lim_{\vec{h}\to\vec{0}} \left|A\frac{\vec{h}}{|\vec{h}|}\right| \cdot \lim_{\vec{h}\to\vec{0}} \vec{\gamma}(\vec{k})$$

$$= \lim_{\vec{h}\to\vec{0}} \left|A\frac{\vec{h}}{|\vec{h}|}\right| \cdot 0$$

↳ this is some number because $\lim_{\vec{h}\to\vec{0}} \frac{\vec{h}}{|\vec{h}|}$ is the unit-ized $\vec{h}$.

$A: \underset{\{\vec{u} \in \mathbb{R}^n : |\vec{u}| = 1\}}{S^{n-1}} \to \mathbb{R}$ is continuous on a compact set, so by EVT it has a max/min and does not blow up

$\{\vec{u} \in \mathbb{R}^n : |\vec{u}| = 1\}$ and does not blow up

So $\lim_{\vec{h} \to 0} \vec{\delta}(\vec{h}) = \vec{0}$.

Applications:

(1) Directional Derivatives

Definition a direction in $\mathbb{R}^n$ is a unit vector in $\mathbb{R}^n$.

The directional derivative of $f: \mathcal{U} \to \mathbb{R}$ ($\mathcal{U} \subseteq \mathbb{R}^n$) at $\vec{a} \in \mathcal{U}$ is the derivative of the composite

$$(-\varepsilon, \varepsilon) \xrightarrow{\vec{g}} \mathcal{U} \xrightarrow{f} \mathbb{R}$$

small enough interval

straight line function passing through $\vec{a}$ at $\vec{g}(0)$, in the direction $\vec{u}$, $\vec{g}(t) = \vec{a} + t\vec{u}$. $\frac{\partial}{\partial t}\vec{g}(t) = \vec{u}$

Apply chain rule: $\frac{\partial}{\partial t}(f \circ \vec{g})(t)\Big|_{t=0} = Df(\vec{a})\, D\vec{g}(0) = \nabla f(\vec{a}) \cdot \vec{u}$

$= |\nabla f(\vec{a})| \cdot \cos(\theta)$ where θ is angle between $\vec{u}$ & $\nabla f(\vec{a})$

$\Rightarrow$ gradient $\nabla f(\vec{a})$ points in direction of maximal change.

(2) Suppose that $\{\vec{x} \in \mathbb{R}^n : f(\vec{x}) = 0\} =: S$ specifies a smooth hypersurface.

Ex: $\{\vec{x} \in \mathbb{R}^n : x_1^2 + x_2^2 + \dots + x_n^2 - 1 = 0\} = S^{n-1}$ (the $n-1$ dimensional sphere in $\mathbb{R}^n$).

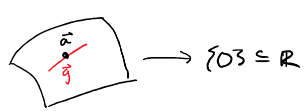
Then the tangent hyperplane to S at $\vec{a} \in S$ is specified by

$\{\vec{x} \in \mathbb{R}^n : 0 = \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a})\}$

Proof: suppose that $\vec{g}: (-\varepsilon, \varepsilon) \to S \subseteq \mathbb{R}^n$ is a smooth curve in S with $\vec{g}(0) = \vec{a}$.

Then the composite $\underbrace{(-\varepsilon, \varepsilon) \xrightarrow{g} S \xrightarrow{f} \mathbb{R}}_{0}$ is the 0 function.

So $0 = \frac{\partial}{\partial t}(f \circ g)(0) = \nabla f(\vec{a}) \cdot g'(0) \Rightarrow \nabla f(\vec{a})$ is a normal vector to S at $\vec{a}$.



$g'(0)$ is tangent to a curve in S, so it must be in the tangent plane.

Compare this w/ previous eqn of hyperplane.