Lec 1/30

Monday, January 30, 2017 09:12

Differentiability of $\mathbb{R}^n \to \mathbb{R}^m$

Proposition: Let $U \subseteq \mathbb{R}^n$ open. Then f is differentiable at x_0 iff there is a function $\varepsilon : B(r,\vec{0}) \to \mathbb{R}$ defined on some ball of small enough radius r such that $\forall \vec{h} \in B(r,\vec{0})$ we have $f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + \nabla f(\vec{x}_0)\cdot\vec{h} + |\vec{h}|\,\varepsilon(\vec{h})$ with $\lim_{\vec{h}\to\vec{0}} \varepsilon(\vec{h}) = 0$

Proof:
$$\varepsilon(\vec{h}) = \begin{cases} \dfrac{f(\vec{x}_0 - \vec{h}) - f(\vec{x}_0) - \nabla f(\vec{x}_0)\cdot\vec{h}}{|\vec{h}|} & \vec{h} \neq \vec{0} \\ 0 & \vec{h} = \vec{0} \end{cases}$$

Want to extend the notion of differentiability to vector-valued functions, i.e. taking values in $\mathbb{R}^m$.

dot product can be reformulated as a matrix product.

$$\nabla f(\vec{x}_0)\cdot\vec{h} = \underset{1\times n}{(\partial_1 f(\vec{x}_0), \dots, \partial_n f(\vec{x}_0))} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}_{n\times 1}$$

Definition: Let $\vec{f} : U \to \mathbb{R}^m$ be defined on an open set $U \subseteq \mathbb{R}^n$. Let $\vec{x}_0 \in U$. We say that $\vec{f}$ is differentiable at $\vec{x}_0$ if there is a linear map $A : \mathbb{R}^n \to \mathbb{R}^m$ (given by $m\times n$ matrix) and a function $\vec{\varepsilon} : B(r,\vec{0}) \to \mathbb{R}^m$ defined on some ball with small enough radius $r > 0$ s.t.

$m\times 1$ column vectors $\hookrightarrow$
$$\vec{f}(\vec{x}_0 + \vec{h}) = \vec{f}(x_0) + \underset{m\times n}{A}\,\underset{n\times 1}{\vec{h}} + |\vec{h}|\,\vec{\varepsilon}(\vec{h}) \qquad (*)$$

for all $\vec{h} \in B(r,\vec{0})$ and $\lim_{\vec{h}\to\vec{0}} \vec{\varepsilon}(\vec{h}) = 0$.

Proposition: $\vec{f}$ as in the definition is differentiable at $\vec{x}_0$ iff each component function $(f_i : U \to \mathbb{R})$ is diffable for $i = 1, \dots, m$. Moreover, A is the Jacobian matrix:

$$A = (\partial_j f_i(\vec{x}_0))_{\substack{1\le i\le m \\ 1\le j\le n}}$$

Proof: $(*)$ is equivalent to a system of m equations

$$f_i(\vec{x}_0+\vec{h}) = f_i(x_0) + A_i\vec{h} + |\vec{h}|\,\varepsilon_i(\vec{h}) \qquad i=1,2,\dots,m$$

We need to show that $\lim\limits_{\vec{h}\to\vec{0}} \vec{\varepsilon}(\vec{h}) = \vec{0} \iff \lim\limits_{\vec{h}\to\vec{0}} \varepsilon_i(\vec{h}) = 0$ for $i=1,\dots,m$

We have $\Updownarrow$

$$\lim_{\vec{h}\to\vec{0}} |\vec{\varepsilon}(\vec{h})| = 0$$

triangle ineq.

$$\underset{\substack{\downarrow\\0}}{|\varepsilon_i(\vec{h})|} \le \underset{\substack{\downarrow\\0}}{|\vec{\varepsilon}(\vec{h})|} \le \underset{\substack{\downarrow\\0}}{\sum_{i=1}^{n} |\varepsilon_i(\vec{h})|}$$

$$\Leftarrow \qquad \Leftarrow$$

We have already shown that if f_i is diffable at $\vec{x}_0$ then $A_i = \nabla f_i(\vec{x}_0)$. ■

Definition: If $\vec{f}: U \to \mathbb{R}^m$, $U \subseteq \mathbb{R}^n$, is differentiable at $\vec{x}_0$ we call

$$\left(\partial_j f_i(\vec{x}_0)\right)_{\substack{1\le i\le m\\1\le j\le n}} = D\vec{f}(x_0)$$

The derivative of $\vec{f}$ at $\vec{x}_0$

(if $m=n=1$, $f: U \to \mathbb{R}$, $U \subseteq \mathbb{R}$, $Df(x_0) = (f'(x_0))$ 1×1 matrix.)

Chain Rule (multivariable). Suppose $\vec{f}: U \to V$, $U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^m$ open, is differentiable at $x_0 \in U$, $\vec{g}: V \to \mathbb{R}^p$ is differentiable at $\vec{f}(\vec{x}_0) \in V$. Then the composite function $\vec{g}\circ\vec{f}: U \to \mathbb{R}^p$ is differentiable at $\vec{x}_0 \in U$, and

$$D(\vec{g}\circ\vec{f})(\vec{x}_0) = \underbrace{D\vec{g}(\vec{f}(x_0))}_{p\times m} \cdot \underbrace{D\vec{f}(x_0)}_{m\times n}$$

matrix product.

If we denote the coordinates of $\mathbb{R}^n$ by $(x_1,\dots,x_n)$

" " $\mathbb{R}^m$ by $(y_1,\dots,y_m)$

" " $\mathbb{R}^p$ by $(z_1,\dots,z_p)$

then $\vec{y} = \vec{f}(\vec{x})$, $\vec{z} = \vec{g}(\vec{y})$ and

Then $\vec{y} = \vec{f}(\vec{x})$, $\vec{z} = \vec{g}(\vec{y})$ and

$i = 1, \dots, p$
$j = 1, \dots, n$

$$\frac{\partial z_i}{\partial x_j} = \sum_{k=1}^{m} \frac{\partial z_i}{\partial y_k} \frac{\partial y_k}{\partial x_j}$$

$D(\vec{g} \circ \vec{f})_{i,j}$

$$\frac{\partial z_i}{\partial x_j} = \underbrace{\left(\frac{\partial z_i}{\partial y_1}, \dots, \frac{\partial z_i}{\partial y_m}\right)}_{\text{i-th row of } D\vec{g}(\vec{f}(x_0))} \begin{pmatrix} \frac{\partial y_1}{\partial x_j} \\ \vdots \\ \frac{y_m}{\partial x_j} \end{pmatrix} \leftarrow \text{jth column of } D\vec{f}(x_0)$$

Proof: We have

$$(1) \quad \vec{f}(\vec{x}_0 + \vec{h}) = \vec{f}(\vec{x}_0) + \overbrace{D\vec{f}(x_0)\cdot\vec{h} + |\vec{h}|\,\vec{\varepsilon}(\vec{h})}^{= \vec{k}} \quad \text{for } \vec{h} \in B(r, \vec{0}) \subseteq \mathbb{R}^n$$

$$(2) \quad \vec{g}(\vec{f}(\vec{x}_0) + \vec{k}) = \vec{g}(\vec{f}(\vec{x}_0)) + D\vec{g}(\vec{f}(x_0))\cdot\vec{k} + |\vec{k}|\,\vec{\gamma}(\vec{k}) \quad \text{for } \vec{k} \in B(s, \vec{0}) \subseteq \mathbb{R}^m$$

Since $\lim_{\vec{h} \to \vec{0}} \vec{k} = \vec{0}$ we can find $r' \le r$ such that $\vec{h} \in B(r', \vec{0}) \Rightarrow \vec{k} \in B(s, \vec{0})$

Combining (1) and (2), we get

$$\vec{g}(\vec{f}(\vec{x}_0) + \vec{k}) = \vec{g}(\vec{f}(\vec{x}_0)) + D\vec{g}(\vec{f}(\vec{x}_0))\left(D\vec{f}(\vec{x}_0)\cdot h + |\vec{h}|\,\vec{\varepsilon}(\vec{h})\right) + |\vec{k}|\,\vec{\gamma}(\vec{k})$$

$$= \vec{g}(\vec{f}(\vec{x}_0)) + \left(D\vec{g}(\vec{f}(\vec{x}_0))\, D\vec{f}(\vec{x}_0)\right)\cdot\vec{h} + \underbrace{|\vec{h}|\, D\vec{g}(\vec{f}(\vec{x}_0))\cdot\vec{\varepsilon}(\vec{h}) + |\vec{k}|\,\vec{\gamma}(\vec{k})}_{\overset{?}{=}\ |\vec{k}|\,\vec{\zeta}(\vec{h}) \text{ where } \vec{\zeta}(\vec{h}) \to \vec{0} \text{ as } \vec{h} \to \vec{0}}$$

It remains to show:

$$\frac{|\vec{h}|\, D\vec{g}(\vec{f}(x_0))\cdot\vec{\varepsilon}(\vec{h}) + |\vec{k}|\,\vec{\gamma}(\vec{k})}{|\vec{k}|} \to \vec{0} \quad \text{as} \quad \vec{h} \to \vec{0}.$$

$$\parallel$$

$$\underbrace{\frac{|\vec{h}|\, D\vec{g}(\vec{f}(\vec{x}_0))\cdot\vec{\varepsilon}(\vec{h})}{|\vec{k}|}}_{\text{need to analyze this } \frac{0}{0} \text{ limit}} + \underbrace{\vec{\gamma}(\vec{k})}_{\downarrow\ \vec{0} \text{ as } \vec{k}\to\vec{0} \text{ as } \vec{h}\to\vec{0}}$$