Math 600 Lecture 35

Functions of several variables

We now consider functions of the form

$$f:\mathbb{R}^n\to\mathbb{R}^m$$

(or $f:E\to\mathbb{R}^m$, where $E\subset\mathbb{R}^n$). Four special cases deserve mention:

- $f:\mathbb{R}\to\mathbb{R}$ $(m=n=1)$
- $f:\mathbb{R}^n\to\mathbb{R}$ $(m=1, n>1)$
- $f:\mathbb{R}\to\mathbb{R}^m$ $(n=1, m>1)$
- $f:\mathbb{R}^n\to\mathbb{R}^m$ $(m>1, n>1)$

Mostly, we discuss the general case ($f:\mathbb{R}^n\to\mathbb{R}^m$), but sometimes we focus on one of the special cases.

__Definition__: $L:\mathbb{R}^n\to\mathbb{R}^m$ is called __linear__ iff

$$L(\alpha x+\beta y)=\alpha L(x)+\beta L(y)\ \ \forall x,y\in\mathbb{R}^n\ \forall \alpha,\beta\in\mathbb{R}.$$

__Definition__: Let $E\subset\mathbb{R}^n$ be open and let $f:E\to\mathbb{R}^m$. We say that f is __differentiable__ at $x\in E$ iff there exists a linear map $L:\mathbb{R}^n\to\mathbb{R}^m$ such that

$$\lim_{h\to 0}\frac{\|f(x+h)-f(x)-L(h)\|}{\|h\|}=0$$

(Here we use the Euclidean norm on $\mathbb{R}^n$ and $\mathbb{R}^m$, which reduces to the absolute value if $m=1$ or $n=1$.) If f is differentiable at x, L is called the __derivative__ of f at x: $L=Df(x)$.

Note: $(m=n=1)$ Every linear map $L:\mathbb{R}\to\mathbb{R}$ is of the form

$$L(x)=mx \quad \forall x\in\mathbb{R},$$

where m is a real constant. Note that

$$\lim_{h\to 0}\frac{|f(x+h)-f(x)-mh|}{|h|}=0$$

$$\iff \lim_{h\to 0}\left|\frac{f(x+h)-f(x)-mh}{h}\right|=0$$

$$\iff \lim_{h\to 0}\left|\frac{f(x+h)-f(x)}{h}-m\right|=0$$

$$\iff \lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=m.$$

Thus the above definition is basically the same as our previous definition, except now

$$Df(x)=L, \text{ where } L:\mathbb{R}\to\mathbb{R} \text{ is defined by } L(x)=mx\ \forall x\in\mathbb{R},$$

whereas

$$f'(x)=m.$$

The way to think about this is as follows: $f'(x)$ is the representative of $Df(x)$ (every linear map $L:\mathbb{R}\to\mathbb{R}$ is represented by a real number).

Theorem: Let $L:\mathbb{R}^n\to\mathbb{R}^m$ be linear. Then there exists a unique $M\in\mathbb{R}^{m\times n}$ such that

$$L(x)=Mx \quad \forall x\in\mathbb{R}^n.$$

$(m=n=1 \Rightarrow M \in \mathbb{R}^{1\times 1} \cong \mathbb{R})$.

The general case: $m>1, n>1$

$$f(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_m(x) \end{bmatrix}, \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$Df(x)$ is represented by an $m\times n$ matrix $f'(x)$, called the Jacobian (matrix) of f at x:

$$f'(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \frac{\partial f_m}{\partial x_2}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{bmatrix}$$

How do we derive this formula? Let $e_j \in \mathbb{R}^n$ be the jth standard basis vector. Assuming f is differentiable at x, we have

$$\lim_{h\to 0^+} \frac{\|f(x+he_j) - f(x) - L(he_j)\|}{\|he_j\|} = 0 \qquad (h \in \mathbb{R}^+)$$

$$\Rightarrow \lim_{h\to 0^+} \frac{\|f(x+he_j) - f(x) - hL(e_j)\|}{h} = 0$$

$$\Rightarrow \lim_{h\to 0^+} \left\| \frac{f(x+he_j) - f(x)}{h} - L(e_j) \right\| = 0$$

$$\Rightarrow \lim_{h\to 0} \frac{f(x+he_j)-f(x)}{h} = L(e_j)$$

$$\Rightarrow L(e_j) = \begin{bmatrix} \frac{\partial f_1}{\partial x_j}(x) \\ \frac{\partial f_2}{\partial x_j}(x) \\ \vdots \\ \frac{\partial f_m}{\partial x_j}(x) \end{bmatrix} \quad \text{(the } j\text{th column of } f'(x)\text{).}$$

Special case: $m=n=1$

$$Df(x)h = f'(x)h \quad \forall h \in \mathbb{R}, \text{ where } f'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} \in \mathbb{R}, \text{ as usual}$$

Special case: $m>1, n=1$

Now f is a vector-valued function of a real variable. Let's use x instead of f (i.e. we consider $x(t)$ instead of $f(x)$). From above, $x'(t)$ is an $m\times 1$ matrix, i.e. an m-vector, often write $\dot{x}$:

$$x'(t) = \dot{x}(t) = \begin{bmatrix} x_1'(t) \\ x_2'(t) \\ \vdots \\ x_m'(t) \end{bmatrix}$$

We have

$$Dx(t)h = h\dot{x}(t).$$

Special case: $m=1, n>1$

Now the representative $f'(x)$ of $Df(x)$ is a $1 \times n$ matrix:

$$f'(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) & \frac{\partial f}{\partial x_2}(x) & \cdots & \frac{\partial f}{\partial x_n}(x) \end{bmatrix}$$

and, for $h \in \mathbb{R}^n$,

$$Df(x)h = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) & \frac{\partial f}{\partial x_2}(x) & \cdots & \frac{\partial f}{\partial x_n}(x) \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix}$$

$$= \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) h_j \quad \text{(a dot product)}$$

$$= \nabla f(x) \cdot h,$$

where

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) \\ \frac{\partial f}{\partial x_2}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{bmatrix} \in \mathbb{R}^n.$$

We usually use $\nabla f(x)$, rather than $f'(x)$, as the representative of $Df(x)$.

(This is an example of the Riesz representation theorem.)