Why equivalent to the usual?

$$\lim_{h\to 0}\frac{1}{h}\left(f(a+h)-f(a)-mh\right) = 0 \quad \text{(and exists, in particular)}$$

$\Updownarrow$ since $\lim_{h\to 0}\frac{1}{h}(mh) = m$ <u>exists</u>, can add it to both sides

$$\lim_{h\to 0}\frac{1}{h}\left(f(a+h)-f(a)-mh\right) + \lim_{h\to 0}\frac{1}{h}(mh) = m$$

$\Updownarrow$ limit laws

$$\lim_{h\to 0}\frac{1}{h}\left(f(a+h)-f(a)\right) = m \quad \leftarrow \boxed{\text{usual definition}}$$

Rather than doing something naive (and wrong) for $\bar{f}: \mathbb{R}^n \to \mathbb{R}^m$, like defining $\bar{f}'(\bar{a}) = \lim_{\bar{h}\to\bar{0}} \frac{\bar{f}(\bar{a}+\bar{h})-\bar{f}(\bar{a})}{|\bar{h}|}$ (wrong even for $n=m=1$, since $|h|$ is always <u>positive</u>)

we ask for a <u>linear</u> function $L: \mathbb{R}^n \to \mathbb{R}^m$ that plays the above role of $h \mapsto mh$ ...

<u>DEF'N</u> (1.7.10, essentially) For $\bar{f}: \underset{\mathbb{R}^n}{\overset{\text{open}}{U}} \to \mathbb{R}^m$ and $\bar{a} \in U$, ($U \subset \mathbb{R}^n$)

say that $\bar{f}$ is <u>differentiable at $\bar{a}$</u>

if $\exists$ some <u>linear</u> transformation $L: \mathbb{R}^n \to \mathbb{R}^m$, $\bar{x} \mapsto L(\bar{x})$

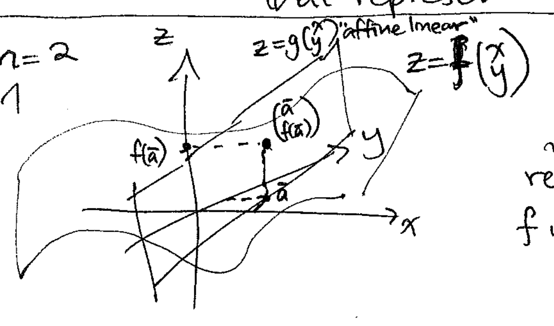
with $\lim_{\bar{h}\to\bar{0}} \frac{1}{|\bar{h}|}\left(\left(\bar{f}(\bar{a}+\bar{h})-\bar{f}(\bar{a})\right)-L(\bar{h})\right) = 0.$

In this case we write $D\bar{f}(\bar{a}) = L$

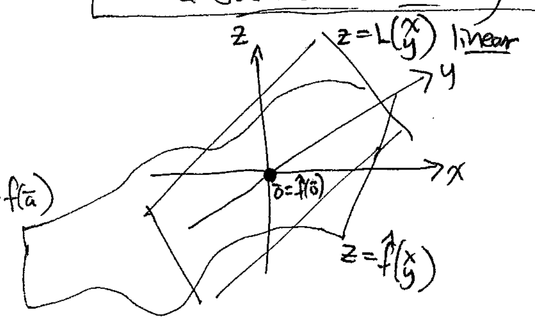
(and we'll see shortly how to compute the <u>matrix</u> $[D\bar{f}(\bar{a})] = [L]$ that represents $L: \mathbb{R}^n \to \mathbb{R}^m$, via partial derivatives & Jacobian matrix)

10/12/2016

e.g. $n=2$, $m=1$



replace f with $\hat{f}\binom{x}{y} = f(\bar{x}+\bar{a}) - f(\bar{a})$



So how should we get the matrix $\overset{[D\bar{f}(\bar{a})]}{\overset{\|}{[L]}}$ for $L: \mathbb{R}^n \to \mathbb{R}^m$

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

given by $\bar{x} \overset{\bar{f}}{\longmapsto} \begin{bmatrix} f_1(\bar{x}) \\ \vdots \\ f_m(\bar{x}) \end{bmatrix}$?

Since the limits converge componentwise, we ought to be able to deal with each $f_i(\bar{x})$ separately for $i=1,2,\dots,m$, i.e. reduce to $m=1$

Also, since L is linear its matrix $[L] = \begin{bmatrix} | & & | \\ L(\bar{e}_1) & \cdots & L(\bar{e}_n) \\ | & & | \end{bmatrix}$,

so we should deal with $\bar{h} \to \bar{0}$ along the coordinate axes first,

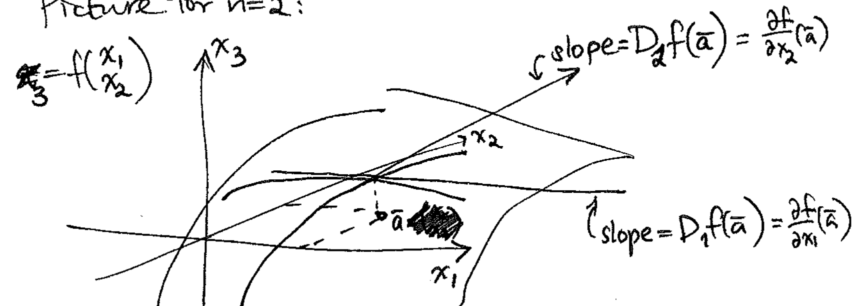
i.e. $\bar{h} = h\bar{e}_j = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ h \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ as $h \to 0$

DEFIN 1.7.3 The j^{th} partial derivative for $j=1,2,\dots,n$ of $f: \underset{\mathbb{R}^n}{\overset{\text{open}}{U}} \subset \to \mathbb{R}$ at $\bar{a} \in U$, $\bar{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$

is
$$D_j f(\bar{a}) := \lim_{h\to 0} \frac{1}{h}\left(f\begin{pmatrix} a_1 \\ \vdots \\ a_j+h \\ \vdots \\ a_n \end{pmatrix} - f\begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} \right)$$
$$= \lim_{h\to 0} \frac{1}{h}\left(f(\bar{a}+h\bar{e}_j) - f(\bar{a}) \right)$$

Also written $\left.\frac{\partial f}{\partial x_j}\right|_{\bar{x}=\bar{a}}$ or $\frac{\partial f}{\partial x_j}(\bar{a})$ or $f_{x_j}(\bar{a})$

Picture for $n=2$:



The matrix representing $L: \mathbb{R}^n \to \mathbb{R}$ linearly approximating $f: \mathbb{R}^n \to \mathbb{R}$

then ought to be (if L exists) given by $[D_1 f(\bar{a}) \; D_2 f(\bar{a}) \cdots D_n f(\bar{a})] = \nabla f(\bar{a})^T$

where $\nabla f(\bar{a})$ = gradient of f at $\bar{a}$ $= \begin{bmatrix} D_1 f(\bar{a}) \\ \vdots \\ D_n f(\bar{a}) \end{bmatrix}$

EXAMPLE: $f: \mathbb{R}^2 \to \mathbb{R}$, $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \sin(x^2 y)$ has gradient at $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \pi/2 \\ 0 \end{pmatrix} = \bar{a}$ given by

$$\nabla f\begin{pmatrix} \pi/2 \\ 0 \end{pmatrix} = \begin{bmatrix} D_1 f(\bar{a}) \\ D_2 f(\bar{a}) \end{bmatrix} = \begin{bmatrix} \left.\frac{\partial \sin(x^2y)}{\partial x}\right|_{\binom{x}{y}=\binom{\pi/2}{0}} \\ \left.\frac{\partial \sin(x^2y)}{\partial y}\right|_{\binom{x}{y}=\binom{\pi/2}{0}} \end{bmatrix}$$

$$= \begin{bmatrix} \cos(x^2y)\cdot 2x\big|_{\binom{x}{y}=\binom{\pi/2}{0}} \\ \cos(x^2y)\cdot x^2\big|_{\binom{x}{y}=\binom{\pi/2}{0}} \end{bmatrix} = \begin{bmatrix} \pi \\ \pi^2/4 \end{bmatrix}$$

and if $D\bar{f}(\bar{a})$ exists, it should have matrix $[\pi \quad \pi^2/4] : \mathbb{R}^2 \to \mathbb{R}$, with $\pi = D_1 f(\bar{a})$, $\pi^2/4 = D_2 f(\bar{a})$;

$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \underbrace{[\pi \quad \pi^2/4]}_{\nabla f\binom{\pi/2}{0}^t} \begin{bmatrix} x \\ y \end{bmatrix}$$

Putting it together componentwise for $i = 1, 2, \ldots, m$ ...

DEF'N 1.7.7: If $\bar{f}: U \to \mathbb{R}^m$ (open, $U \subset \mathbb{R}^n$), $\bar{x} \mapsto \begin{pmatrix} f_1(\bar{x}) \\ \vdots \\ f_m(\bar{x}) \end{pmatrix}$ has a derivative $L: \mathbb{R}^n \to \mathbb{R}^m$ at $\bar{x} = \bar{a}$

then its Jacobian matrix is

$$J\bar{f}(\bar{a}) = \begin{bmatrix} D_1 f_1(\bar{a}) & \cdots & D_n f_1(\bar{a}) \\ D_1 f_2(\bar{a}) & \cdots & D_n f_2(\bar{a}) \\ \vdots & & \\ D_1 f_m(\bar{a}) & \cdots & D_n f_m(\bar{a}) \end{bmatrix}$$

$$= \begin{bmatrix} \nabla f_1(\bar{a})^T \\ \vdots \\ \nabla f_m(\bar{a})^T \end{bmatrix}$$

(and ought to be the matrix $[L] = J\bar{f}(\bar{a})$!)

EXAMPLE:

$$\bar{f}: U = \left\{\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : x \neq 0\right\} \to \mathbb{R}^3, \quad U \subset \mathbb{R}^2$$

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \bar{f}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \sin(x^2y) \\ e^{x+y} \\ \frac{1}{x} \end{pmatrix}$$

at $\bar{x} = \bar{a} = \begin{pmatrix} \pi/2 \\ 0 \end{pmatrix}$ has Jacobian matrix

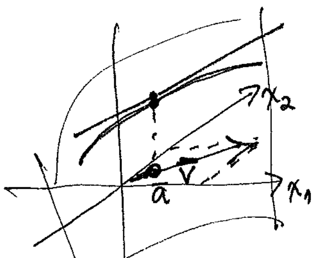
$$J\bar{f}\begin{pmatrix} \pi/2 \\ 0 \end{pmatrix} = \left.\begin{bmatrix} D_1 \sin(x^2y) & D_2 \sin(x^2y) \\ D_1 e^{x+y} & D_2 e^{x+y} \\ D_1\left(\frac{1}{x}\right) & D_2\left(\frac{1}{x}\right) \end{bmatrix}\right|_{\binom{x}{y}=\binom{\pi/2}{0}} = \begin{bmatrix} \pi & \pi^2/4 \\ e^{\pi/2} & e^{\pi/2} \\ \frac{-1}{(\pi/2)^2} & 0 \end{bmatrix} : \mathbb{R}^2 \to \mathbb{R}^3$$

So now we can compute a multivariate derivative...

THMS 1.7.9, 1.7.14: If $\bar{f}: \underset{\mathbb{R}^n}{\overset{\text{open}}{U}} \to \mathbb{R}^m$, $\bar{x} \mapsto \begin{pmatrix} f_1(\bar{x}) \\ \vdots \\ f_m(\bar{x}) \end{pmatrix}$ is differentiable at $\bar{a}$

with derivative $L: \mathbb{R}^n \to \mathbb{R}^m$, then

(i) L has matrix $[L] = J\bar{f}(\bar{a})$ ← Jacobian matrix



(ii) for any $\bar{v} \in \mathbb{R}^n - \{\bar{0}\}$, the directional derivative of $\bar{f}$ in direction $\bar{v}$ exists, and equals $L(\bar{v}) = J\bar{f}(\bar{a})\bar{v}$

i.e. $\lim_{h \to 0} \frac{1}{h}\left(\bar{f}(\bar{a}+h\cdot\bar{v}) - f(\bar{a})\right) = J\bar{f}(\bar{a})\bar{v}$

proof: Let's check (ii) first. Since $L: \mathbb{R}^n \to \mathbb{R}^m$ is linear and has

$$\lim_{\bar{h} \to \bar{0}} \frac{1}{|\bar{h}|}\left(\left(\bar{f}(\bar{a}+\bar{h}) - f(\bar{a})\right) - L(\bar{h})\right) = \bar{0}$$

taking $\bar{h} = h\cdot\bar{v}$ with $h \to 0$ gives

$$\lim_{h \to 0} \frac{1}{|h\cdot\bar{v}|}\left(\left(\bar{f}(\bar{a}+h\cdot\bar{v}) - f(\bar{a})\right) - L(h\cdot\bar{v})\right) = \bar{0}$$

$$\frac{1}{|\bar{v}|}\lim_{h \to 0} \frac{1}{|h|}\left(\left(\bar{f}(\bar{a}+h\cdot\bar{v}) - f(\bar{a})\right) - hL(\bar{v})\right) = \bar{0}$$

$$\text{i.e. } \bar{0} = \lim_{h \to 0} \frac{h}{|h|}\left(\left(\frac{\bar{f}(\bar{a}+h\cdot\bar{v}) - f(\bar{a})}{h}\right) - L(\bar{v})\right)$$

$\frac{h}{|h|} = \begin{cases} +1 & \text{if } h>0 \\ -1 & \text{if } h<0 \end{cases}$

$$\Leftarrow \begin{cases} \lim_{h \to 0^+} \left(\frac{\bar{f}(\bar{a}+h\cdot\bar{v}) - f(\bar{a})}{h}\right) - L(\bar{v}) \\ \lim_{h \to 0^-} L(\bar{v}) - \left(\frac{\bar{f}(\bar{a}+h\bar{v}) - f(\bar{a})}{h}\right) \end{cases}$$

i.e. $L(\bar{v}) = \lim_{h \to 0} \frac{\bar{f}(\bar{a}+h\bar{v}) - f(\bar{a})}{h}$ = dir. deriv. of $\bar{f}$ in dir. $\bar{v}$ exists

To deduce (i) from this, note L linear implies its matrix is

$[L] = \begin{bmatrix} | & & | \\ L(\bar{e}_1) & \cdots & L(\bar{e}_n) \\ | & & | \end{bmatrix}$, but $L(\bar{e}_j) = \lim_{h \to 0} \frac{1}{h}\left(\bar{f}(\bar{a}+h\bar{e}_j) - f(\bar{a})\right) = \begin{bmatrix} D_j f_1(\bar{a}) \\ \vdots \\ D_j f_m(\bar{a}) \end{bmatrix}$

so $[L] = \begin{bmatrix} D_1 f_1(\bar{a}) & \cdots & D_n f_1(\bar{a}) \\ D_1 f_m(\bar{a}) & \cdots & D_n f_m(\bar{a}) \end{bmatrix} = J\bar{f}(\bar{a})$ ∎

10/14/2016 >