5-10-17

$f : \Omega \to \mathbb{R}$. $\quad \Omega \subset \mathbb{R}^n$

Derivative of f at $x = x_0$ with respect to a vector $y \in \mathbb{R}^n$ is defined as

$$f'(x_0; y) := \lim_{h \to 0} \frac{f(x_0 + hy) - f(x_0)}{h}$$ if the limit exists.

If y = unit vector, say e, then this derivative is called the directional derivative of f in the direction of e

If $e = e_i = (0, \ldots 0, 1, 0 \ldots 0)$ then this derivative is called the partial derivative of f in the direction of e_i, or i-th partial derivative.

$$\frac{\partial f}{\partial x^i}(x_0) \quad \text{or} \quad D_i f(x_0), \quad f_i(x_0), \quad f_i'(x_0).$$

$$\frac{\partial f}{\partial x^i} := \lim_{h \to 0} \frac{f(x_0^1, x_0^2, \dots, x_0^i + h, x_0^{i+1}, \dots x_0^n) - f(x_0^1, x_0^2, \dots, x_0^n)}{h}.$$

- Directional derivative for the direction of $\bar{e}$ can also be realized

$$\frac{d}{dt} f(x_0 + et)\Big|_{t=0}.$$

Rk. $f: \mathbb{R}^2 \to \mathbb{R}$.

Ex. $f(x,y,z) = 3x^2y + z^3\sin x$ $\qquad f = \mathbb{R}^3 \to \mathbb{R}$.

$$\frac{\partial f}{\partial x} = 6xy + z^3\cos x \qquad \frac{\partial f}{\partial y} = 3x^2 \qquad \frac{\partial f}{\partial z} = 3z^2\sin x$$

First order partial derivatives. They create new scalar fields

$$\frac{\partial f}{\partial x} = \mathbb{R}^3 \to \mathbb{R}$$
$$(x,y,z) \to 6xy + z^3\cos x$$

We can consider

$$\underbrace{\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)}_{\frac{\partial^2 f}{\partial y\,\partial x}} \qquad \frac{\partial}{\partial z}\left(\frac{\partial f}{\partial x}\right) \qquad \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right).$$

$$\underbrace{\frac{\partial^2 f}{\partial y\,\partial x} \qquad \frac{\partial^2 f}{\partial z\,\partial x} \qquad \frac{\partial^2 f}{\partial x^2}}_{\text{2nd order partial derivatives.}}$$

Ex. $$f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0). \end{cases}$$

Except for $(x,y) = (0,0)$, we can evaluate the partial derivatives of f using product/quotient rule for derivatives of functions of one variable.

For $(x,y) = (0,0)$, we need to use the definition of partial der. as limits.

$$\frac{\partial f}{\partial x}(0,0) := \lim_{h \to 0} \frac{f\big((0,0) + h(1,0)\big) - f(0,0)}{h}$$

$$= \lim_{\substack{h \to 0 \\ h \neq 0}} \frac{f(h,0) - f(0,0)}{h} = \lim_{h \to 0} \frac{\frac{h \cdot 0}{h^2 + 0^2} - 0}{h} = 0$$

Similarly

$$\frac{\partial f}{\partial y}(0,0) = 0.$$

So at $(0,0)$ both partial derivatives exist. But we've already seen that this function is not continuous at $(0,0)$

In fact it can happen that the directional derivative of a function exists along any vector $e=(a,b)$ yet f is not continous.

eg: $f(x,y) = \begin{cases} \dfrac{xy^2}{x^2+y^4} & (x,y)\neq(0,0) \\ 0 & (x,y)=(0,0) \end{cases}$ has directional derivatives along any direction

but it is not continous at $(0,0)$. (Exercise).

<u>Moral of the story</u> For a function $f:\mathbb{R}\to\mathbb{R}$

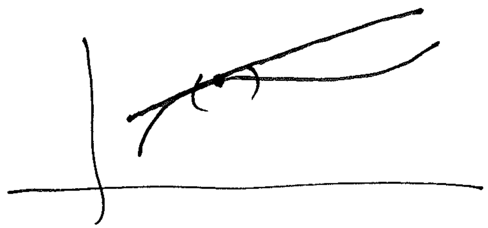
we have that f is diff. at $x_0 \Rightarrow f$ is cont. at x_0

For functions $f:\mathbb{R}^n\to\mathbb{R}$

f has derivatives at x_0 in <u>any</u> direction e $\not\Rightarrow$ f is continous at x_0.

We need a stronger differentiability criteria.
This is given by the concept of "total differential".

Let's start with recalling what we know for $f: \mathbb{R} \to \mathbb{R}$ that is differentiable.
If f is differentiable at x_0, then f can be well approximated by the affine linear map.



$$\underbrace{f(x_0) + f'(x_0)(x - x_0)}_{\text{Tangent line}}.$$

Well approximated means that

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + R(x, x_0)$$

with $\displaystyle\lim_{x \to x_0} \frac{R(x, x_0)}{|x - x_0|} = 0.$ i.e. The error of approxim. goes to 0 faster than $|x - x_0|$ goes to zero.

The number $a := f'(x_0)$ can be thought as the 1×1 matrix that represents the linear map

$$A : \mathbb{R} \to \mathbb{R}, \quad x_0 \mapsto a x_0. \qquad a = f'(x_0).$$

<u>Defn</u> $f : \mathbb{R} \to \mathbb{R}$ is in x_0 <u>differentiable</u> if there exists a <u>linear map</u> $A_{x_0} : \mathbb{R} \to \mathbb{R}$ such that

$$\lim_{\substack{x \to x_0 \\ x \neq x_0}} \frac{f(x) - f(x_0) - A(x - x_0)}{|x - x_0|} = 0.$$

or $\displaystyle\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - A(h)}{h} = 0.$

Defn. (7.1.2 Struwe). (**Differential**).

The function $f: \mathbb{R}^n \to \mathbb{R}$ is in $x_0 \in \mathbb{R}^n$ (total) **differentiable** if there exists a **linear map** $A = A_{x_0} : \mathbb{R}^n \to \mathbb{R}$ so that

$$\lim_{x \to x_0} \frac{|f(x) - f(x_0) - A_{x_0}(x - x_0)|}{\| x - x_0 \|} = 0. \quad \Big\} (*)$$

i.e. $\exists$ a real valued function $R(x, x_0)$ so that

$$f(x) = f(x_0) + A(x - x_0) + R(x, x_0)$$

with

$$\lim_{x \to x_0} \frac{|R(x, x_0)|}{\| x - x_0 \|} = 0.$$

If that is the case, the linear transformation $A_{x_0} : \mathbb{R}^n \to \mathbb{R}$ is called the <u>(total) derivative</u> of f at x_0

It is denoted by $df(x_0)$ or $d_{x_0} f$

Note!!! The total derivative $d_{x_0} f : \mathbb{R}^n \to \mathbb{R}$ is a linear map not a number.

Even in the case of $f : \mathbb{R} \to \mathbb{R}$, df is a linear map $df : \mathbb{R} \to \mathbb{R}$ which is represented by a 1×1 matrix which is a number.

Rk. 1) In the defn in (*) the norm in the numerator is the norm in $\mathbb{R}$, the norm in the denominator is the norm in $\mathbb{R}^n$.

2) If such a linear map exists then it is unique!

Pf. Exercise.

3) Sometimes one can guess what the function $A: \mathbb{R}^n \to \mathbb{R}$ should be in (*) $\lim \frac{\|f(x) - f(x_0) - A(x - x_0)\|}{\|x - x_0\|} = 0$.

And if you can verify that (*) holds for this choice of A then it must be the derivative because of Rmk 2)

Example : $f : \mathbb{R}^n \longrightarrow \mathbb{R}$.
$\qquad x \longrightarrow c$

We're looking for a linear map $A_{x_0} : \mathbb{R}^n \to \mathbb{R}$.

$$\frac{f(x) - f(x_0) - A(x - x_0)}{|x - x_0|} \to 0.$$

Let $A : \mathbb{R}^n \to \mathbb{R}$
$\qquad x \to 0.$

then $f(x) - f(x_0) - A_{x_0}(x - x_0) = c - c - 0 = 0.$

Hence $\lim_{x \to x_0} \dfrac{f(x) - f(x_0) - A(x - x_0)}{|x - x_0|} = 0.$

Hence $df(x_0) = A$ where $A : \mathbb{R}^n \to \mathbb{R}$
$\qquad x \to 0.$

$f : \mathbb{R} \longrightarrow \mathbb{R}$
$x \longrightarrow x^3$

$A_{x_0} : \mathbb{R} \longrightarrow \mathbb{R}$
$x \longrightarrow a x$
where $a = f'(x_0)$

$f'(x) = 3x^2$

$df : \mathbb{R} \longrightarrow \mathcal{L}(\mathbb{R}, \mathbb{R})$.
$x_0 \longrightarrow A_{x_0} : \mathbb{R} \to \mathbb{R}$.
$x \to a x$
where $a = f'(x_0)$.

$x_0 \longrightarrow A_{x_0} : \mathbb{R} \to \mathbb{R}$.
$x \to \underbrace{(3x_0^2)}_{a.} x$.

$A_{x_0} = (a)_{1\times1}$ $\quad a = 3x_0^2$.

<u>Recall</u>. A map $A: \mathbb{R}^n \to \mathbb{R}$ is linear

if $\forall x, y \in \mathbb{R}^n$, and $\alpha, \beta \in \mathbb{R}$

$$A(\alpha x + \beta y) = \alpha A(x) + \beta A(y).$$

Every linear map is uniquely determined by its values at a given Basis.

In particular if we choose as a basis $\{e_1, e_2 .. , e_n\}$

A is uniquely determined by $A(e_i)$ $1 \leq i \leq n$.

If we write $A(e_1) = A_1$ $A(e_2) = A_2$, .. $A(e_n) = A_n$.

then the $1 \times n$ Matrix $(A_1 \; A_2 \; .. \; A_n)$ is the Matrix repn of $A: \mathbb{R}^n \to \mathbb{R}$ wrt the standard basis.

$$A(x) = (A_1 \; .. \; A_n)\begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix} = \sum_{i=1}^{n} A_i x^i$$

Our goal: To determine the Matrix repn of the total differential $d_{x_0}f$

Ex. 1) Every Affine Linear function defined by
$$f: \mathbb{R}^n \to \mathbb{R}$$
$$x \to Ax + b$$
where A is a $1 \times n$ matrix, $b \in \mathbb{R}$ is differentiable at every point $x = x_0$ with total differential
$$df: \mathbb{R}^n \to \mathbb{R}$$
$$x \to Ax$$
i.e. For any point x_0 the Matrix rep. of $d_{x_0}f = A$.

Why?
$$\begin{aligned} R(x, x_0) := &= f(x) - f(x_0) - A(x - x_0) \\ &= (Ax + b) - (Ax_0 + b) - A(x - x_0) \\ &= Ax + \cancel{b} - Ax_0 - \cancel{b} - A\cancel{x} + Ax_0 = 0. \end{aligned}$$

Hence $\lim \frac{R(x, x_0)}{|x - x_0|} = 0$.

2) A special case of an affine linear map.

is

$$x^i : \mathbb{R}^n \to \mathbb{R}$$
$$x = (x^1, x^2 .. x^n) \to x^i$$

the i-th coordinate function.

~~This~~ is the affine map with $A = (0, .. \underset{\substack{\uparrow \\ i\text{th component}}}{1}, .. 0)$

$$x \to Ax = (0, 1 .. 0)\begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix} = x^i$$

$$dx^i(x_0) = A(x_0).$$

How can we find the matrix repn. of the linear map $d_{x_0} f$?

Thm (Bmk 7.1.1) Assume f is differentiable at x_0 with total derivative $d_{x_0} f$. Then the the derivative of f at x_0 with respect to any $y \in \mathbb{R}^n$, $f'(x_0, y)$, exists and ~~~~ ~~~~ we have

$$f'(x_0, y) = (d_{x_0} f)(y)$$

In particular $y = e_i$ we have

$$f'(x_0, e_i) = \frac{\partial f}{\partial x^i}(x_0) = (d_{x_0} f)(e_i)$$

i.e. The Matrix representation of the linear map $d_{x_0} f : \mathbb{R}^n \to \mathbb{R}$ with the respect to the standard basis is the Matrix $\left(\frac{\partial f}{\partial x^1}(x_0) \quad \frac{\partial f}{\partial x^2}(x_0), \ldots \frac{\partial f}{\partial x^n}(x_0) \right)$

Moreover if $y = (y^1, y^2, \ldots, y^n)$ then we have

$$f'(x_0, y) = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(x_0)\, y^i = \left(\frac{\partial f}{\partial x^1}(x_0) \cdots \frac{\partial f}{\partial x^n}(x_0)\right)\begin{pmatrix} y^1 \\ y^2 \\ \vdots \\ y^n \end{pmatrix}.$$

Defn. The vector of partial derivatives

$$\left(\frac{\partial f}{\partial x^1}(x_0), \ldots, \frac{\partial f}{\partial x^n}(x_0)\right)$$

is called the **gradient of f** denoted by $\nabla f(x_0)$ (Nabla of f) or grad f

Note $\nabla f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$. is a vector field.

$$x_0 \longrightarrow \nabla f(x_0)$$

which is defined at each point x_0 where the partial derivatives $\frac{\partial f}{\partial x^i}(x_0)$ $i = 1, \ldots, n$ exist.

Rmk ① The derivative of f at a pt $x = x_0$ wrt y is then given by

$$f'(x_0, y) = \langle \nabla f(x_0), y \rangle.$$

It is the scalar product of the gradient vector $\nabla f(x_0)$ with the vector y.

②. We can write the first order approximation to f at x_0 as

$$f(x) = f(x_0) + (d_{x_0} f)(x - x_0) + R(x, x_0)$$

$$\boxed{f(x) = f(x_0) + (\nabla f(x_0)) \cdot (x - x_0) + R(x, x_0).}$$

with $\lim\limits_{x \to x_0} \frac{R(x, x_0)}{|x - x_0|} = 0$. It resembles the first order Taylor formula.

$$\boxed{f(x) = f(x_0) + f'(x_0)(x - x_0) + R(x, x_0) \quad \text{for } f : \mathbb{R} \to \mathbb{R}.}$$

$$f(x) = f(x_0) + \left(\frac{\partial f}{\partial x^1}(x_0), \dots \frac{\partial f}{\partial x^n}(x_0)\right) \cdot \left(x^1 - x_0^1, x^2 - x_0^2, \dots x^n - x_0^n\right) + R(x, x_0).$$

$$= f(x_0) + \frac{\partial f}{\partial x^1}(x_0)\,(x^1 - x_0^1) + \frac{\partial f}{\partial x^2}(x_0)\,(x^2 - x_0^2) + \dots + \frac{\partial f}{\partial x^n}(x_0)(x^n - x_0^n) + R(x, x_0).$$

eg. $f = \mathbb{R}^2 \to \mathbb{R}$, $(x,y) \to f(x,y)$

$$f(x,y) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)\,(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) + R((x,y), (x_0, y_0)).$$

$$f(x,y) = \underbrace{f(x_0, y_0) + A_1 (x - x_0) + A_2 (y - y_0)} + \text{Rest}$$

were $A_1 = \frac{\partial f}{\partial x}(x_0, y_0)$ $\quad A_2 = \frac{\partial f}{\partial y}(x_0, y_0)$.

$$\boxed{z = f(x_0, y_0) + A_1 (x - x_0) + A_2 (y - y_0).}$$

$z = z_0 + A_1(x - x_0) + A_2(y - y_0)$. $\leftarrow$ This is the eqn of a plane.

From the Taylor formula it is easy to prove

<u>Thm 2.</u> Let $f : \mathbb{R}^n \to \mathbb{R}$ a scalar field. If f is differentiable at x_0 then f is continuous.

<u>Pf.</u> If f is total diff at $x = x_0$ then.

$$f(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0) + R(x, x_0)$$

with $\frac{R(x, x_0)}{|x - x_0|} \to 0$.

Hence $0 < |f(x) - f(x_0)| \leq \|\nabla f(x_0)\| \underbrace{\|x - x_0\|}_{\downarrow 0} + \underbrace{R(x, x_0)}_{\downarrow 0}$.

Taking the limit as $x \to x_0$. we get $\lim_{x \to x_0} f(x) = f(x_0)$.

Let's look at proof of Thm 1.

Proof: f is total diff. at $x = x_0$.

i.e. $\exists$ a lin. map $d_{x_0}f : \mathbb{R}^n \to \mathbb{R}$ so that

$\|y\| = 1$.

$$\lim_{x \to x_0} \left| \frac{f(x) - f(x_0) - (d_{x_0}f)(x - x_0)}{|x - x_0|} \right| = 0.$$

$$\Rightarrow \lim_{h \to 0} \frac{f(x_0 + hy) - f(x_0) - (d_{x_0}f)(hy)}{h} = 0.$$

$\downarrow$ $d_{x_0}f$ is linear

$$\Rightarrow \lim_{h \to 0} \frac{f(x_0 + hy) - f(x_0)}{h} - \frac{h\,(d_{x_0}f)(y)}{h}.$$

$$\Rightarrow \underbrace{\lim_{h \to 0} \frac{f(x_0 + hy) - f(x_0)}{h}}_{\text{is exactly}} = (d_{x_0}f)(y).$$

is exactly ~~thats~~ the directional derivative of f in x_0 in direction of y.

In particular for $y = e_i$ we get on the RHS $(d_{x_0} f)(e_i)$ which is the i-th column in the matrix repn of $d_{x_0} f : \mathbb{R}^n \to \mathbb{R}$.

On the LHS $\lim_{h \to 0} \frac{f(x_0 + he_i) - f(x_0)}{h} = \frac{\partial f}{\partial x^i}(x_0)$.

Hence the matrix repn of $d_{x_0} f$ with respect to the standard basis is $\left(\frac{\partial f}{\partial x^1}(x_0), \ldots \frac{\partial f}{\partial x^n}(x_0)\right) = \nabla f(x_0)$.