LECTURE 3

DATE : 11 OCTOBER 2021

3. Differential Calculus for functions of several variables II : Differentiability and Properties

$f : \mathbb{R}^d \to \mathbb{R}$, $f = f(x_1, \dots, x_d)$

$x = (x_1, \dots, x_d)$, $y = (y_1, \dots, y_d) \in \mathbb{R}^d$

$x \cdot y = x_1 y_1 + x_2 y_2 + \dots + x_d y_d$

$\|x\| = \sqrt{x \cdot x}$

Def (continuity): $f : \mathbb{R}^d \to \mathbb{R}$ is cont at $x \in \mathbb{R}^d$
if $\forall \varepsilon > 0 \ \exists \delta > 0$ s.t. $|f(x) - f(y)| < \varepsilon \ \forall y \in \mathbb{R}^d$ with $\|x - y\| < \delta$

Partial derivatives = normal / standard derivatives but w.r.t. one of the $x_1, \dots, x_d$ (the rest being fixed)

ex: $f(x_1, x_2) = x_1 x_2^3$ $\qquad$ $g(x_1, x_2) = x_1 x_2 + x_2$

$\dfrac{\partial f}{\partial x_1}(x_1, x_2) = 1\, x_2^3$ $\qquad$ $\dfrac{\partial f}{\partial x_2}(x_1, x_2) = x_1 \cdot 3 \cdot x_2^2$

$\dfrac{\partial g}{\partial x_1}(x_1, x_2) = 1 \cdot x_2 + 0$ $\qquad$ $\dfrac{\partial g}{\partial x_2}(x_1, x_2) = x_1 \cdot 1 + 1$

Partial derivative = partial info // Full info : The Gradient

$$\nabla f(x_1, \dots, x_d) = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_d} \right)$$

Theorem 1 (CHAIN rule):

$f : \mathbb{R}^d \to \mathbb{R}$, $\qquad$ $x_1, \dots, x_d : [a, b] \subset \mathbb{R} \to \mathbb{R}$

has cont $\dfrac{\partial}{\partial x_i}$ $\qquad$ all are differentiable

Then $F : [a,b] \to \mathbb{R}$, $F(t) = f(x_1(t), \dots, x_d(t))$
"$F = f \circ (x_1, \dots, x_d)$" is differentiable

and $\dfrac{d}{dt} F(t) = \nabla f(x_1(t), \dots, x_d(t)) = \dfrac{d}{dt} x(t)$

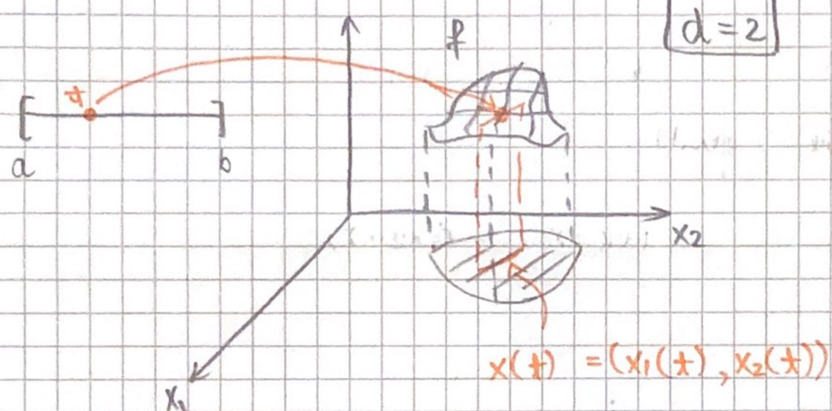
$\dfrac{d}{dt}$ = standard derivative w. r. t. one variable (t)

$\dfrac{d}{dt} x(t) = \left(\dfrac{d}{dt} x_1(t), \dots, \dfrac{d}{dt} x_d(t)\right) = (x_1'(t), \dots x_d'(t))$

$x(t) = (x_1(t), \dots, x_d(t))$

Leibniz notation for a derivative

$\boxed{d=2}$



Theorem $(d=1)$ $f : [a,b] \to \mathbb{R}$ **LAGRANGE**

If : • f cont on $[a,b]$ • f diffable (a,b) | Then $\exists\, \xi \in (a,b)$ s.t

$$f'(\xi) = \frac{f(b) - f(a)}{b-a}$$

(see Lecture 1)

$$\Leftrightarrow f(b) - f(a) = f'(\xi)(b-a)$$

Theorem 2 (Lagrange for $d>1$)

If $D \subseteq \mathbb{R}^d$ convex , $a, b \in D$ $(a \neq b)$

$f : D \to \mathbb{R}$ (or $\mathbb{R}^d$) has cont $\dfrac{\partial}{\partial x_i}$

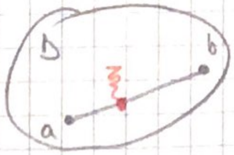
open segment

Then $\exists\, \xi \in (a,b)$ s.t.

$$f(b) - f(a) = \nabla f(\xi) \cdot (b-a)$$

IDEA of the proof : apply the $d=1$ Lagrange Theorem to

$F:[0,1]\to\mathbb{R}$, $F(t)=f(a+t(b-a))$



Key point : segment = one-dimension

You can use $d=1$ Lagrange on the segment

Def (directional derivative)

$V=(V_1,\dots,V_d)\in\mathbb{R}^d$ "the direction"

$f:\mathbb{R}^d\to\mathbb{R}$

$$\frac{\partial f}{\partial V}(x_1,\dots,x_d)=\nabla f(x_1,\dots,x_d)\cdot V$$

(geometrically = the projection of ∇f onto V)

Remark : Partial derivatives are directional derivatives w.r.t (special directions)

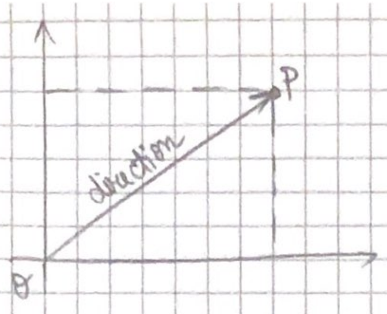
$$\frac{\partial f}{\partial x_1}=\frac{\partial f}{\partial V} \quad \text{for } V=(1,0,\dots,0)$$

$$\frac{\partial f}{\partial x_i}=\frac{\partial f}{\partial V} \quad \text{for } V=(0,\dots,0,\underset{i\text{-th position}}{1},0,\dots,0)$$

Remark : any $x=(x_1,\dots,x_d)\in\mathbb{R}^d$ can be decomposed w.r.t. these special (CHANONICAL) directions

$$x_i = x\cdot(0,\dots,0,\underset{i\text{-th}}{1},0,\dots,0)$$

$$\begin{array}{l}
x_1\,(1,0,0,\dots,0,0)\\
x_2\,(0,1,0,\dots,0,0)\\
\vdots\\
x_{d-1}\,(0,0,0,\dots,1,0)\\
x_d\,(0,0,0,\dots,0,1)\\
\hline
(x_1,x_2,\dots,x_{d-1},x_d)
\end{array}$$



Any Point (away of O) defines a "direction" = "position vector"

§ 3.2. Higher order partial derivatives

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)$$

Theorem 3 (H. A. Schwarz)

$f : \mathbb{R}^d \to \mathbb{R}$ admits continuous mixed second order partial derivatives (on a small ball around x) then

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x)$$

"order doesn't matter if continuous"

$\frac{\partial^2}{\partial x_i \partial x_j}$ only partial info Full info : Heise matrix

$$Hf(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right)_{i,j=\overline{1,d}}$$

§ 3.3. The Fréchet differential

- Linear functions of several variables

$T : \mathbb{R}^d \to \mathbb{R}$ is called linear if:

(i) $T(x+y) = T(x) + T(y) \quad \forall x, y \in \mathbb{R}^d$

(ii) $T(\alpha x) = \alpha T(x) \quad \forall x \in \mathbb{R}^d, \forall \alpha \in \mathbb{R}$

(i) and (ii) $\Leftrightarrow$ (iii) $T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$

Example :

$a = (a_1, \dots, a_d) \in \mathbb{R}^d$

$T_a(x) = a \cdot x$ is linear

Theorem 4 : $\forall T : \mathbb{R}^d \to \mathbb{R}$ linear

there exists unique $a_T \in \mathbb{R}^d$ such that $T(x) = a_T \cdot x$

"All linear functions are of the form $a \cdot x$"

- Quadratic functions

$Q : \mathbb{R}^d \to \mathbb{R}$ quadratic if

$$Q(x) = \sum_{i,j=1}^{d} a_{ij} x_i x_j \quad , \quad a_{ij} = a_{ji} \quad i,j = \overline{1,d}$$
$$a_{ij} \in \mathbb{R}$$

$x = (x_1, \dots, x_d)$

$A = (a_{ij})_{i,j=\overline{1,d}}$ the matrix of the quadratic function

Examples :

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad Q_1(x_1, x_2) = x_1^2 + x_2^2$$

$$A_2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad , \quad Q_2(x_1, x_2) = x_1^2 + x_1 x_2 + x_2 x_1 + x_2^2$$
$$= x_1^2 + 2x_1 x_2 + x_2^2$$

If the ∇f is the $\mathbb{R}^d$ version of f', how should we define higher order derivative equivalents in $\mathbb{R}^d$?

<u>Def</u> : $f : \mathbb{R}^d \to \mathbb{R}$ Fréchet differentiable at x if there exist a linear function $T : \mathbb{R}^n \to \mathbb{R}$ s.t.

$$\lim_{y \to x} \frac{|f(y) - f(x) - T(y-x)|}{\|y - x\|} = 0$$