Recollection of linear algebra

U with a basis $\mathbf{u}_1, \ldots, \mathbf{u}_n$, $\mathbf{x} \in U$

$\mathbf{x} = x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n$,

$\mathbf{x} \mapsto (x_1, \ldots, x_n)$ (coordinates)

A linear mapping $L : U \to \mathbb{R}$ can be written as

$$L(\mathbf{x}) = \sum_i x_i L(\mathbf{u}_i) = \sum_i A_i x_i$$

where $L(\mathbf{u}_i) = A_i$ (compare with tot. diff.).

V with a basis $\mathbf{v}_1, \ldots, \mathbf{v}_m$, α a linear map $U \to V$. Then $\alpha(\mathbf{u}_i) = \sum_j A_{ij}\mathbf{v}_j$ and $\mathbf{A} = (a_{ij})_{ij}$

$$\alpha(\mathbf{x}) = \sum_i x_i L(\mathbf{u}_i) = \sum_i x_i \sum_j A_{ij}\mathbf{v}_j =$$

$$= \sum_j (\sum_i x_i A_{ij})\mathbf{v}_j = \mathbf{x}\mathbf{A}$$

We multiply $\mathbf{x}$ represented as $(x_1, \ldots, x_n)$ and $\mathbf{A}$ as matrices.

If we similarly have a $\beta : V \to W$ with a matrix $\mathbf{B}$ we obtain for the composed map

$$\mathbf{x} \mapsto (\mathbf{x}\mathbf{A})\mathbf{B} = \mathbf{x}(\mathbf{A}\mathbf{B}).$$

Hence:

if we have linear maps represented by matrices $\mathbf{A}$ *and* $\mathbf{B}$, *then their composition is represented by the matrix product*

$$\mathbf{AB}.$$

Repetition. *Product*, in particular recall in $(\prod_{i=1}^{n} X_i, d) = \prod_{i=1}^{n}(X_i, d_i)$ the distance

$$d((x_1, \dots, x_n), (y_1, \dots, y_n)) = \max_i d_i(x_i, y_i).$$

Note that

$$(\mathbb{E}_n, \sigma) = \overbrace{\mathbb{R} \times \cdots \times \mathbb{R}}^{n \text{ times}} = \mathbb{R}^n.$$

Partial derivatives are standard derivatives, that is, limits

$$\lim_{h \to 0} \frac{f(\dots x_{k-1}, x_k + h, x_{k+1} \dots) - f(x_1, \dots)}{h}.$$

Notation

$$\frac{\partial f(x_1, \dots, x_n)}{\partial x_k}$$

P.d. are not very satisfactory. They even do not imply continuity (no surprise after what we already know).

In particular we miss a counterpart of the formula

$$(*)\ f(x+h) - f(x) = Ah + |h| \cdot \mu(h)$$

with the geometric (tangent) and aproximative connotations.

The concept of total differential introduced:

μ continuous in a neighborhood U of $\mathbf{o}$ such that $\mu(\mathbf{o}) = 0$,

and numbers $A_1, \ldots, A_n$for which

$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = \sum_{k=1}^{n} A_k h_k + \|\mathbf{h}\| \mu(\mathbf{h})$$

This extends $(*)$, explain.

Partial derivatives are implied, not the other way round.

BUT: TD follows from *continuous* partial derivatives. Comment.

Computing with partial derivatives: arithmetic rules are the same as for standard partial derivatives, composition is not quite so simple.

Theorem. *Let* $f(\mathbf{x})$ *have a* **total differential** *in* $\mathbf{a}$*. Let* $g_k(t)$ *have derivatives in* b *and let* $g_k(b) = a_k$ *for* $k = 1, \dots, n$*. Set*

$$F(t) = f(\mathbf{g}(t)) = f(g_1(t), \dots, g_n(t)).$$

Then F *has a derivative in* b*, namely*

$$F'(b) = \sum_{k=1}^{n} \frac{\partial f(\mathbf{a})}{\partial x_k} \cdot g_k'(b).$$

Corollary. (Chain Rule) *Let* $f(\mathbf{x})$ *have a total differential in* $\mathbf{a}$. *Let* $g_k(t_1, \ldots, t_r)$ *have partial derivatives in* $\mathbf{b} = (b_1, \ldots, b_r)$ *and let* $g_k(\mathbf{b}) = a_k$ *for* $k = 1, \ldots, n$. *Then the function*

$$(f \circ \mathbf{g})(t_1, \ldots, t_r) = f(\mathbf{g}(t)) = f(g_1(t), \ldots, g_n(t))$$

has all the partial derivatives in $\mathbf{b}$ *and*

$$\frac{\partial (f \circ \mathbf{g})(\mathbf{b})}{\partial t_j} = \sum_{k=1}^{n} \frac{\partial f(\mathbf{a})}{\partial x_k} \cdot \frac{\partial g_k(\mathbf{b})}{\partial t_j}.$$

We have composed

$$\mathbb{E}_k \xrightarrow{\mathbf{g}} \mathbb{E}_n \xrightarrow{f} \mathbb{R}$$

Now consider an m-tuple of functions

$\mathbf{f} = (f_1 \ldots, f_m)$, that is, $\mathbf{f} : \mathbb{E}_n \to \mathbb{E}_m$

$$\mathbb{E}_k \xrightarrow{\mathbf{g}} \mathbb{E}_n \xrightarrow{\mathbf{f}} \mathbb{E}_m$$

The rule from the previous theorem yields

$$\frac{\partial(f_i \circ \mathbf{g})(\mathbf{b})}{\partial t_j} = \sum_{k=1}^{n} \frac{\partial f_i(\mathbf{a})}{\partial x_k} \cdot \frac{\partial g_k(\mathbf{b})}{\partial t_j}.$$

Introducing $D\mathbf{f} = \left(\frac{\partial h_i(\mathbf{a})}{\partial x_k}\right)_{ik}$ we get

$D(\mathbf{f} \circ \mathbf{g}) = D\mathbf{f} \cdot D\mathbf{g}$ (matrix multiplication)

and this is how it should be.

$D\mathbf{h}$ is the matrix of the linear approximation of $\mathbf{h}$ and:

linear aproximations compose in parallel with the mappings approximated.

Arithmetic rules obtained from the chain one.

Multiplication.

$f(u,v) = u \cdot v$. Then $\frac{\partial f}{\partial u} = v$, $\frac{\partial f}{\partial v} = u$ and for $u = \phi(x)$ and $v = \psi(x)$

$$(\phi(x).\psi(x))' = \frac{\partial f}{\partial u}\phi'(x) + \frac{\partial f}{\partial v}\psi'(x) =$$
$$= \psi(x)\phi'(x) + \phi(x)\psi'(x)$$

Division.

$f(u,v) = \frac{u}{v}$. Then $\frac{\partial f}{\partial u} = \frac{1}{v}$, $\frac{\partial f}{\partial v} = -\frac{u}{v^2}$ and for $u = \phi(x)$ and $v = \psi(x)$

$$\left(\frac{\phi(x)}{\psi(x)}\right)' = \frac{\partial f}{\partial u}\phi'(x) - \frac{\partial f}{\partial v}\psi'(x) =$$
$$= \frac{1}{\psi(x)}\phi'(x) + \frac{\phi(x)}{\psi(x)^2}\psi'(x) =$$
$$= \frac{\psi(x)\phi'(x) - \phi(x)\psi'(x)}{\psi(x)^2}$$

$U \subseteq \mathbb{E}_n$ is *convex* if

$$\mathbf{x}, \mathbf{y} \in U \quad \Rightarrow \quad \forall t,\ 0 \le t \le 1,\ (1-t)\mathbf{x}+t\mathbf{y} = \mathbf{x}+t(\mathbf{y}-\mathbf{x}) \in U.$$

Lagrange in several variables.

Proposition. *Let f have continuous partial derivatives in a convex open $U \subseteq \mathbb{E}_n$. Then for any two $x, y \in D$ there is a θ, $0 \le \theta \le 1$, such that*

$$f(\mathbf{y}) - f(\mathbf{x}) = \sum_{j=1}^{n} \frac{\partial f(\mathbf{x}+\theta(\mathbf{y}-\mathbf{x}))}{\partial x_j}(y_j - x_j).$$

Proof. $F(t) = f(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))$ is $F = f \circ \mathbf{g}$ with $\mathbf{g}$ where $g_j(t) = x_j + t(y_j - x_j)$, and

$$F'(t) = \sum_{j=1}^{n} \frac{\partial f(\mathbf{g}(t))}{\partial x_j} g_j'(t) = \sum_{j=1}^{n} \frac{\partial f(\mathbf{g}(t))}{\partial x_j}(y_j - x_j).$$

By Lagrange's theorem

$$f(\mathbf{y}) - f(\mathbf{x}) = F(1) - F(0) = F'(\theta).$$

Note. This formula is often used in the form

$$f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) = \sum_{j=1}^{n} \frac{\partial f(\mathbf{x}+\theta\mathbf{h})}{\partial x_j} h_j.$$

Compare with the formula for total differential:

$$f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) = \sum_{j=1}^{n} \frac{\partial f(\mathbf{x})}{\partial x_j} h_j + \|\mathbf{h}\|\mu(\mathbf{h})$$

If partial derivatives $\frac{\partial f(x_1,\ldots,x_n)}{\partial x_k}$ exist for all $(x_1, \ldots, x_n)$ in a domain D' we have a function

$$\frac{\partial f}{\partial x_k} : D' \to \mathbb{R}.$$

If we have $g(\mathbf{x}) = \frac{\partial f(\mathbf{x})}{\partial x_k}$ then similarly like asking about the second derivative of a function in one variable we can consider second derivatives of $f(\mathbf{x})$, that is,

$$\frac{\partial g(\mathbf{x})}{\partial x_l}.$$

The result, if it exists, is then denoted by

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_k \partial x_l}.$$

Iterating this procedure we obtain

$$\frac{\partial^r f(\mathbf{x})}{\partial x_{k_1} \partial x_{k_2} \ldots \partial x_{k_r}},$$

partial derivatives of order r.

The order is determined by the number of taking derivatives, not by repetitions in individual variables.

$$\frac{\partial^3 f(x,y,x)}{\partial x \partial y \partial z} \quad \text{and} \quad \frac{\partial^3 f(x,y,x)}{\partial x \partial x \partial x}$$

are third order derivatives.

Consecutive taking derivatives by the same variable is written as an exponent, e.g.,

$$\frac{\partial^5 f(x,y)}{\partial x^2 \partial y^3} = \frac{\partial^5 f(x,y)}{\partial x \partial x \partial x \partial y \partial y},$$

$$\frac{\partial^5 f(x,y)}{\partial x^2 \partial y^2 \partial x} = \frac{\partial^5 f(x,y)}{\partial x \partial x \partial y \partial y \partial x}.$$

A suggestive example.

Compute "mixed" second order derivatives of

$$f(x, y) = x \sin(y^2 + x).$$

First we obtain

$$\frac{\partial f(x, y)}{\partial x} = \sin(y^2 + x) + x \cos(y^2 + x),$$

$$\frac{\partial f(x, y)}{\partial y} = 2xy \cos(y^2 + x).$$

and then the second order derivatives,

$$\frac{\partial^2 f}{\partial x \partial y} = 2y \cos(y^2 + x) - 2xy \sin(y^2 + x)$$

$$\frac{\partial^2 f}{\partial y \partial x} = 2y \cos(y^2 + x) - 2xy \sin(y^2 + x).$$

<u>We have got the same result!</u>

Proposition. *Let $f(x, y)$ have continuous partial derivatives* $\frac{\partial^2 f}{\partial x \partial y}$ *and* $\frac{\partial^2 f}{\partial y \partial x}$ *in a neighborhood of (x, y). Then*

$$\frac{\partial^2 f(x, y)}{\partial x \partial y} = \frac{\partial^2 f(x, y)}{\partial y \partial x}.$$

Note that we have assumed

continuous partial derivatives,

hence

more than just a total differential.

Proof. Let us try to compute both the derivatives in one step, thus, let us compute the limit $\lim_{h\to 0} F(h)$ of the function

$$F(h) = \frac{f(x+h,y+h) - f(x,y+h) - f(x+h,y) + f(x,y)}{h^2}.$$

If we set

$$\varphi_h(y) = f(x+h,y) - f(x,y) \quad \text{and}$$
$$\psi_k(x) = f(x,y+k) - f(x,y),$$

we obtain for $F(h)$ two formulas:

$$F(h) = \frac{1}{h^2}(\varphi_h(y+h) - \varphi_h(y))$$
$$F(h) = \frac{1}{h^2}(\psi_h(x+h) - \psi_h(x)).$$

First: The function φ_h has a derivative (by y, it has no other variable)

$$\varphi_h'(y) = \frac{\partial f(x+h,y)}{\partial y} - \frac{\partial f(x,y)}{\partial y}$$

and hence by Lagrange's formula we obtain

$$F(h) = \frac{1}{h^2}(\varphi_h(y+h) - \varphi_h(y)) = \frac{1}{h}\varphi_h'(y+\theta_1 h) =$$
$$= \frac{\partial f(x+h, y+\theta_1 h)}{\partial y} - \frac{\partial f(x, y+\theta_1 h)}{\partial y}.$$

Then, again using the L. formula,

$$F(h) = \frac{\partial}{\partial x}\left(\frac{\partial f(x+\theta_2 h, y+\theta_1 h)}{\partial y}\right) \qquad (*)$$

for some θ_1, θ_2 between 0 and 1.

Second, $\frac{1}{h^2}(\psi_h(x+h) - \psi_h(x))$ yields similarly

$$F(h) = \frac{\partial}{\partial y}\left(\frac{\partial f(x+\theta_4 h, y+\theta_2 h)}{\partial x}\right). \qquad (**)$$

Both functions $\frac{\partial}{\partial y}(\frac{\partial f}{\partial x})$ and $\frac{\partial}{\partial x}(\frac{\partial f}{\partial y})$ are continuous in (x, y), and $\lim_{h\to 0} F(h)$ can be computed from any of the expressions $(*)$ or $(**)$:

$$\lim_{h\to 0} F(h) = \frac{\partial^2 f(x,y)}{\partial x \partial y} = \frac{\partial^2 f(x,y)}{\partial y \partial x}.$$

Iterating the exchanges as in the Proposition we obtain

Corollary. *Let a function f in n variables have continuous partial derivatives up to an order k. Then the values of these derivative depend only on the numbers of taking derivative in each of the individual variables* $x_1, \ldots, x_n$.

Under these assumptions, hence, we can write general partial derivatives of order $r \le k$ as

$$\frac{\partial^r f}{\partial x_1^{r_1} \partial x_2^{r_2} \ldots \partial x_n^{r_n}} \quad \text{kde} \quad r_1 + r_2 + \cdots + r_n = r$$

($r_j = 0$ indicates the absence of the symbol ∂x_j).

In the sequel we will need more about metric spaces, in particular a few facts about compactness and completenes. Recall the behavior of compact (closed bounded) intervals, in particular that

- in such intervals every sequence has a convergent subsequence, and this holds in no other type of interval,
- and that a continuous function on such interval attains a maximum and a minimum.

Also, refresh the concept of a cauchy sequence.

Informations and material to MA2

https://kam.mff.cuni.cz/ma2/

Details to the lectures: (In the text)

MA2.1: XIII,1,2,3,4

MA2.2: I; XIII,5; XIV,2,3,5

MA2.3: XIV,3,5,4