Images and preimages

$f : X \to Y, \quad A \subseteq X, \quad B \subseteq Y$

The image of a subset $A \subseteq X$ in Y:

$$f[A] = \{f(x) \mid x \in A\}$$

The preimage of a subset $B \subseteq Y$ in X:

$$f^{-1}[B] = \{x \mid f(x) \in B\}$$

Thus we have maps

$$\mathfrak{P}(X) \underset{f^{-1}[-]}{\overset{f[-]}{\rightleftarrows}} \mathfrak{P}(Y).$$

One has:

$$f[A] \subseteq B \;\equiv\; A \subseteq f^{-1}[B],$$
$$f[f^{-1}[B]] \subseteq B \quad f^{-1}[f[A]] \supseteq A$$

Caution: f^{-1} appears in two roles: inverse $f^{-1} : Y \to X$, may not exist, part of $f^{-1}[-]$, always making sense

Exercise. 1. Is there some relation between these two f^{-1} a $f^{-1}[-]$?

2. When one has $f^{-1}[f[A]] = A$?

3. When one has $f[f^{-1}[B]] = B$?

Repetition. Trouble with functions of several variables, example

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{pro} \quad (x,y) \neq (0,0), \\ 0 \quad \text{for} & (x,y) = (0,0). \end{cases}$$

We will need it again.

Metric space, metric (distance), triangle inequality. Subspaces. $\mathbb{R}$, $\mathbb{C}$, $\mathbb{E}_n$

Neighborhood, open and closed sets, closure.

Continuity, convergence, preimages of open resp. closed sets.

Equivalent and strongly equivalent metrics

$$\alpha \cdot d_1(x,y) \leq d_2(x,y) \leq \beta \cdot d_1(x,y).$$

In $\mathbb{E}_n$ we can replace $\sqrt{\sum_i (x_i - y_i)^2}$ by

$$\max_i |x_i - y_i|.$$

A real function in n variables will be here

$$f : D \to \mathbb{R}, \quad D \subseteq \mathbb{E}_n$$

Similarly like in the case of one variable we cannot restrict ourselves to the case of domains being the whole of $\mathbb{E}_n$. In case of one variable the domains were usually intervals or simple unions of intervals. Here the domains D will be more complicated, often (but not always) open sets in $\mathbb{E}_n$.

Products. For (X_1, d_i), $i = 1, \dots, n$ we endow the cartesian product $\prod_{i=1}^n X_i$ with metric

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

The obtained

$$(\prod_{i=1}^n X_i, d) = \prod_{i=1}^n (X_i, d_i)$$

is called the product of spaces (X_i, d_i).

We also write

$$(X_1, d_1) \times \cdots \times (X_n, d_n).$$

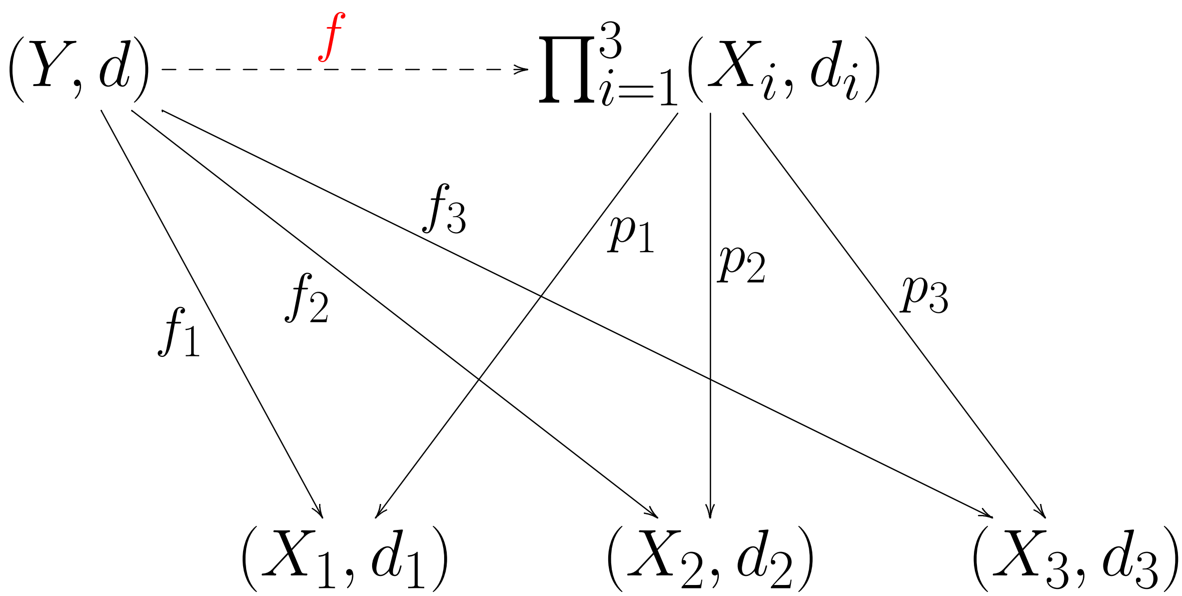
Thus,

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

Theorem. 1. *The projections* $p_j = ((x_i)_i \mapsto x_j) : \prod_{i=1}^n (X_i, d_i) \to (X_j, d_j)$ *are continuous maps.*

2. *Let* $f_j : (Y, d') \to (X_j, d_j)$ *be arbitrary continuous maps. Then the uniquely defined map* $f : (Y, d') \to \prod_{i=1}^n (X_i, d_i)$ *satisfying* $p_j \circ f = f_j$, *namely the map defined by* $f(y) = (f_1(y), \ldots, f_n(y))$, *is continuous.*

It looks as follows:



There is precisely one f with $p_i \circ f = f_i$ and it is continuous.

Hence, if we study continuous maps

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

unlike the situation with the domain the continity depends regarding range on the continuity in the individual coordinates only.

2. Partial derivatives.

This is in individual coordinates. For $f(x_1, \dots, x_n)$ consider

$$\phi_k(t) = f(x_1, \dots, x_{k-1}, t, x_{k+1}, \dots, x_n)$$

Partial derivative of f by x_k (in $(x_1, \dots, x_n)$) is

(the usual) derivative of the function ϕ_k,

that is,

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

Notation

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

For $f(x, y)$ we write

$$\frac{\partial f(x,y)}{\partial x} \quad \text{a} \quad \frac{\partial f(x,y)}{\partial y}, \quad \text{etc.}$$

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

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

When speaking of a partial derivative it will be always obvious whether we have in mind such a function or just a number (the value of the limit above).

Note that the discontinuous function f from the example of continuity by variables has both partial derivatives in every point. Thus,

the existence of partial derivatives does not imply continuity !

We will need something stronger.

Recall the statement equivalent with the existence of a derivative:

There is a μ converging to 0 for $h \to 0$ and A such that

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

Geometrically:

$f(x+h) - f(x) = Ah$ describes a tangent to the graph of f in $(x, f(x))$.

$|h| \cdot \mu(h)$ is a small error.

Let us view $f(x,y)$ similarly and consider the surface

$$S = \{(t,u,f(t,u)) \mid (t,u) \in D\}.$$

The two partial derivatives express the directions of two tangent lines S in $(x,y,f(x,y))$,

but not a tangent plane

and only that could be a suitable extension of the fact above.

For $\mathbf{x} \in \mathbb{E}_n$ set

$$\|\mathbf{x}\| = \max_i |x_i|$$

This will come instead of the absolute value, instead of the h there will be an n-tuple close to $\mathbf{o}$.

Total differential.

f has a *total differential* in $\mathbf{x}$ if there is a

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

and numbers $A_1, \ldots, A_n$ such that

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

Proposition. *Let f have a total differential in $\mathbf{a}$. Then we have that*

1. *f is continuous in $\mathbf{a}$ and*

2. *f has all partial derivatives in $\mathbf{a}$, with values*

$$\frac{\partial f(\mathbf{a})}{\partial x_k} = A_k.$$

1. We have

$$|f(\mathbf{x}-\mathbf{y})| \leq |\mathbf{A}(\mathbf{x}-\mathbf{y})|+|\mu(\mathbf{x}-\mathbf{y})\|\mathbf{x}-\mathbf{y}\|$$

and the limit on the right hand side for $\mathbf{y} \to \mathbf{x}$ is 0.

2. We have

$$\frac{1}{h}(f(\dots x_{k-1}, x_k + h, x_{k+1}, \dots) - f(x_1, \dots)) =$$

$$= A_k + \mu((\dots, 0, h, 0, \dots))\frac{\|(0, \dots, h, \dots, 0)\|}{h},$$

and the right hand side limit is obviously A_k.

This time we have continuity!

Note that in the case of one variable there is no difference between the existence of derivative and possessing total differential in a. In case of several variables the difference is fundamental.

It may come as a surprise that while the existence of partial derivatives is not of much help,

the existence of continuous partial derivatives

is quite another story.

We have

Theorem. *Let f have continuous partial derivatives in a neighborhood of* $\mathbf{a}$. *Then it has in* $\mathbf{a}$ *a total differential.*

Set

$$\mathbf{h}^{(0)} = \mathbf{h},\ \mathbf{h}^{(1)} = (0, h_2, \ldots, h_n),\ \mathbf{h}^{(2)} = (0, 0, h_3, \ldots, h_n) \quad \text{etc.}$$

(so that $\mathbf{h}^{(n)} = \mathbf{o}$). Then we have

$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = \sum_{k=1}^{n}(f(\mathbf{a}+\mathbf{h}^{(k-1)}) - f(\mathbf{a}+\mathbf{h}^{(k)})) = M.$$

By Lagrange's Theorem there are $0 \le \theta_k \le 1$ such that

$$f(\mathbf{a}+\mathbf{h}^{(k-1)}) - f(\mathbf{a}+\mathbf{h}^{(k)}) = \frac{\partial f(a_1, \ldots, a_{k-1}, a_k + \theta_k h_k, a_{k+1}, \ldots, a_n)}{\partial x_k} h_k$$

nd we can proceed

$$\begin{aligned} M &= \sum \frac{\partial f(a_1, \ldots, a_k + \theta_k h_k, \ldots, a_n)}{\partial x_k} h_k = \\ &= \sum \frac{\partial f(\mathbf{a})}{\partial x_k} h_k + \sum \left(\frac{\partial f(a_1, \ldots, a_k + \theta_k h_k, \ldots, a_n)}{\partial x_k} - \frac{\partial f(\mathbf{a})}{\partial x_k} \right) h_k = \\ &= \sum \frac{\partial f(\mathbf{a})}{\partial x_k} h_k + \|\mathbf{h}\| \sum \left(\frac{\partial f(a_1, \ldots, a_k + \theta_k h_k, \ldots, a_n)}{\partial x_k} - \frac{\partial f(\mathbf{a})}{\partial x_k} \right) \frac{h_k}{\|\mathbf{h}\|}. \end{aligned}$$

Set

$$\mu(\mathbf{h}) = \sum \left(\frac{\partial f(a_1, \ldots, a_k + \theta_k h_k, \ldots, a_n)}{\partial x_k} - \frac{\partial f(\mathbf{a})}{\partial x_k} \right) \frac{h_k}{\|\mathbf{h}\|}.$$

Since $\left| \frac{h_k}{\|\mathbf{h}\|} \right| \le 1$ and since $\frac{\partial f}{\partial x_k}$ are continuous, $\lim_{\mathbf{h}\to\mathbf{o}} \mu(\mathbf{h}) = 0$.

We can conclude that

continuous PD $\Rightarrow$ TD $\Rightarrow$ PD

The rules for computing partial derivatives.

Arithmetic rules are the same as for standard derivatives (here, partial derivatives are just the standard ones).

The composition rule differs. Recall that even for the standard derivative it is proved from the formula

$$f(a+h) - f(a) = Ah + |h|\mu(h),$$

that is, using the differential (which, of course, is there the existence of derivative).

The composition rule in simplest form:

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

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

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

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

How one proves it :

$$\frac{1}{h}(F(b+h)-F(b)) = \frac{1}{h}(f(\mathbf{g}(b+h))-f(\mathbf{g}(b)) =$$
$$= \frac{1}{h}(f(\mathbf{g}(b)+(\mathbf{g}(b+h)-\mathbf{g}(b)))-f(\mathbf{g}(b)) =$$
$$= \sum_{k=1}^{n} A_k \frac{g_k(b+h)-g_k(b)}{h} + \mu(\mathbf{g}(b+h)-\mathbf{g}(b)) \max_k \frac{|g_k(b+h)-g_k(b)|}{h}.$$

We have $\lim_{h\to 0}\mu(\mathbf{g}(b+h)-\mathbf{g}(b)) = 0$ as g_k are continuous in b. Since g_k have derivatives, $\max_k \frac{|g_k(b+h)-g_k(b)|}{h}$ are bounded in a sufficiently small neighborhood of 0. Hence the limit of the last summand is 0 and

$$\lim_{h\to 0}\frac{1}{h}(F(b+h)-F(b)) = \lim_{h\to 0}\sum_{k=1}^{n} A_k \frac{g_k(b+h)-g_k(b)}{h} =$$
$$= \sum_{k=1}^{n} A_k \lim_{h\to 0}\frac{g_k(b+h)-g_k(b)}{h} = \sum_{k=1}^{n}\frac{\partial f(\mathbf{a})}{\partial x_k} g_k'(b).$$

What happens, geometrically: The tangent hyperplane expressed by the differential of f has no reason for preferring the main axes in which happen the derivatives in the functions g_k. Hence just partial derivatives would not suffice.

Corollary. (Chain rule) *Let* $f(\mathbf{x})$ *have total differential in* $\mathbf{a}$. *Let functions* $g_k(t_1, \ldots, t_r)$ *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 partial derivatives in b, *and one has*

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