Differential calculus for vector functions

1 Vector functions of real variable

We start by considering derivatives of vector functions, of real variable. The general context is the following:

$$\begin{cases} A \subseteq \mathbb{R} \\ a \in A \cap A' \\ f : A \to \mathbb{R}^m \end{cases}$$

If

$$\exists \quad \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \in \mathbb{R}^m,$$

then this limit is called **the derivative of f at a** and is usually denoted by:

$$f'(a) = \frac{df}{dx}(a).$$

Since f is a vector function, of real variable, it actually means that it can be represented through m **real** functions **of real variable**

$$\forall i \in \{1, ..., m\}, \quad \exists f_i : A \to \mathbb{R}, \quad \text{and} \quad f = (f_1, ..., f_m).$$

The function f has a derivative at a if and only if, for all $i \in \{1, ..., m\}$, each function f_i has a derivative at a. Moreover, the following equality holds:

$$f'(a) = (f_1'(a), ..., f_m'(a)).$$

2 Vector Functions of Vector Variable

The general context is the following:

$$\begin{cases} A \subseteq \mathbb{R}^n \\ a \in intA \\ f : A \to \mathbb{R}^m \end{cases}$$

If there exists a linear function $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{x \to a} \frac{1}{\|x-a\|} \cdot \Big\{ f(x) - f(a) - \varphi(x-a) \Big\} = 0_m, \tag{1}$$

then the function f is said to be **differentiable at a**, and the linear function φ is said to be **the differential of f at a, and is denoted by**

$$df(a).$$

Notice that the differential is a **function**, which has the following properties:

- $: \mathbb{R}^n \to \mathbb{R}^m$**;**
- **is linear;**
- **it satisfeis** (1)
- **the differential is a notion attached to a function at a single point. So if we change the point, we may expect, if f is still differentiable at the new point, to even get a different differential function.**

Theorem: If the function f is differentiable at a, then f is continuous at a.

In practice, when computing effectively the differential of a function f at a point a, instead of computing a limit such as (1) we may study instead

$$\lim_{h \to 0_n} \frac{1}{\|h\|} \cdot \Big\{ f(a+h) - f(a) - \varphi(h) \Big\} = 0_m, \tag{2}$$

this is simply done because we replace $x-a$ by h.

Remark: The reason behind searching for something different from the partial derivatives lies in the fact that a function may have partial derivatives with respect to all variables at a point, and in the same time, it may not be continuous at that point. We had at least two such examples in Seminar 5. In contrast, if a functionis differentiable at a point, then it is continuous at that point which is a property inherited from the derivatives of real functions of real variable.

3 The Determination of the Differential Function - mostly theoretic

3.1 Once Again The Case of Vector Functions of real variable

To begin with we go back to the case when $n = 1$, thus

$$A \subseteq \mathbb{R}\,.$$

Then

$$f \textbf{ has a derivative at } a \iff f \textbf{ is differentiable at } a$$

and the following equality holds

$$df(a)(x) = xf'(a), \quad \forall x \in \mathbb{R}\,.$$

3.2 Once Again The Case of Vector Functions of Vector Variable

Let now $n > 1$. Because the codomain of the function is $\mathbb{R}^m$, f is represented as

$$f = (f_1, ...f_m), \quad \text{with } f_i : A \to \mathbb{R} \quad \forall i \in \{1, ..., m\}.$$

Since f is a vector function, of real variable, it actually means that

$$\forall i \in \{1, ..., m\}, \quad \exists f_i : A \to \mathbb{R}, \quad \text{and} \quad f = (f_1, ..., f_m).$$

Then

$$f \textbf{ is differentiable } a \iff \forall i \in \{1, ..., m\} f_i \textbf{ is differentiable at } a$$

and the following equality holds

$$df(a)(x) = (df_1(a), ..., df_m(a))\,. \quad \forall x \in \mathbb{R}^n\,.$$

4 Actual Algorithm for Determining the Differential

4.1 Real Functions of Vector Variables

$$\begin{cases} A \subseteq \mathbb{R}^n \\ a \in intA \\ f : A \to \mathbb{R} \end{cases}$$

Step 1: Study if f has partial derivatives at a with respect to all variables.

$$\begin{cases} YES & \to \text{ go to Step 2} \\ NO & \to \text{STOP the function does not have a differential at a} \end{cases}$$

In fact, we have to determine

$$\frac{\partial f}{\partial x_j}(a) \quad \forall j \in \{1, ..., n\}.$$

Step 2: Study the limit

$$l = \lim_{h \to 0_n} \frac{1}{\|h\|} \Big\{ f(a+h) - f(a) - \langle h, \nabla f(a) \rangle \Big\}. \tag{3}$$

Recall that $a = (a_1, ..., a_n)$ and $h = (h_1, ..., h_n) \in \mathbb{R}^n$

$$a + h = (a_1 + h_1, ..., a_n + h_n) \quad \text{and} \quad \|h\| = \sqrt{h_1^2 + ... + h_n^2}.$$

Moreover, the gradient of a function at a point is the vector composed of all of its partial derivatives at that point. Thus

$$\langle h, \nabla f(a) \rangle = h_1 \cdot \frac{\partial f}{\partial x_1}(a) + ... + h_n \cdot \frac{\partial f}{\partial x_n}(a) = \sum_{j=1}^{n} h_j \cdot \frac{\partial f}{\partial x_j}(a).$$

This means that the limit (3) is actually

$$l = \lim_{h \to 0_n} \frac{f(a_1 + h_1, ..., a_2 + h_2) - f(a_1, ..., a_n) - \sum_{j=1}^n h_j \cdot \frac{\partial f}{\partial x_j}(a)}{\sqrt{h_1^2 + ... + h_n^2}}. \tag{4}$$

We encounter the following cases

$$\begin{cases} l = 0 & \to \text{ go to Step 3} \\ otherwise & \to \text{STOP the function does not have a differential at a} \end{cases}$$

Step 3: We are in the case when $l = 0$. This means that the function f is differentiable at the point a. Its differential is the linear function

$$df(a) : \mathbb{R}^n \to \mathbb{R}$$

defined by

$$df(a)(h) = \langle h, \nabla f(a) \rangle = \sum_{j=1}^{n} h_j \cdot \frac{\partial f}{\partial x_j}(a), \quad \forall h \in \mathbb{R}^n .$$

4.2 Vector Function of Vector Variable

$$\begin{cases} A \subseteq \mathbb{R}^n \\ a \in intA \\ f : A \to \mathbb{R}^m \end{cases}$$

Recall that f is represented as

$$f = (f_1, ...f_m), \qquad \text{with } f_i : A \to \mathbb{R} \quad \forall i \in \{1, ..., m\}.$$

For each $i \in \{1, ..., m\}$, the function f_i must be analysed as described in section 4.1, since it is a real function of vector variable. At the lecture, there is a theorem that prove that

f is differentiable at a $\iff$ $\forall i \in \{1, ..., m\}$, f_i is differentiable at a. Moreover $df(a) : \mathbb{R}^n \to \mathbb{R}^m$ is equal to

$$df(a) = (df_1(a), df_2(a), ..., df_m(a)).$$

5 The Jacobi Matrix

If the function $f : A \to \mathbb{R}^m$, with $A \subseteq \mathbb{R}^n$ is partially derivable at a, then one can construct the so-called **Jacobi matrix of the function f at a as**

$$J(f)(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & ... & \frac{\partial f_1}{\partial x_n}(a) \\ ... & ... & ... & ... \\ ... & ... & ... & ... \\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & ... & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix}$$

This matrix is used to simplify the writing of the differential function of a vector function of a vector variable.

So, for a differentiable function, with $A \subseteq \mathbb{R}^n$, with $f : A \to \mathbb{R}^m$, the differential function at the point $a \in intA$ is $df(a) : \mathbb{R}^n \to \mathbb{R}^m$, such that

$$\forall h = (h_1, ..., h_n) \in \mathbb{R}^n, \quad df(a)(h) = J(f)(a) \cdot h.$$

This means that for a random $h = (h_1, ..., h_n) \in \mathbb{R}^n$, $df(a)(h)$ is a vector in $\mathbb{R}^m$, and, as we know from algebra, a vector may be represented as a column matrix. Therefore,

$$df(a)(h) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & ... & \frac{\partial f_1}{\partial x_n}(a) \\ ... & ... & ... & ... \\ ... & ... & ... & ... \\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & ... & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix} \cdot \begin{pmatrix} h_1 \\ ... \\ ... \\ h_n \end{pmatrix}.$$

Once me compute the matrix multiplication we reach the result

$$df(a)(h) = \begin{pmatrix} \sum_{i=1}^n \frac{\partial f_1}{\partial x_i}(a) \cdot h_i \\ ... \\ ... \\ \sum_{i=1}^n \frac{\partial f_1}{\partial x_i}(a) \cdot h_i \end{pmatrix}$$

which is the matrix form of the vector $df(a)(h)$.

For the particular case $m = 1$, the Jacobi matrix is the Gradient

6 Exercise/applications

In practice, we often use the following theorem:

Let $A \subseteq \mathbb{R}^n$ is an open set, and $f : A \to \mathbb{R}^m$ be a function which has all partial derivatives. If all partial derivatives are continuous on A, then f is differentiable on A.

This means that is all the partial derivatives of a function are continuous, we may skip Step 2 in the Algorithm from Section 4, and jump directly to Step 3, which is actually the construction of the differential function, by means of :

- the Gradient of f, if f is a real function of vector variable
- the Jacobi matrix of f, if f is a vector function of vector variable.

Recall Exercise 4 from Seminar 5:

Determine all first-order and all second-order partial derivatives of the function

$$f\colon \mathbb{R}^3 \to \mathbb{R}, f(x, y, z) = z^2 \cos(x - y).$$

Determine now its differential at the point $(1, 1, 1)$.

Solution: Let $(x, y, z) \in \mathbb{R}^3$ be randomly chosen. Then

$$\frac{\partial f}{\partial x}(x, y, z) = -z^2 \sin(x - y),\ \frac{\partial f}{\partial y}(x, y, z) = -z^2 \sin(x - y)(x - y)'_y = z^2 \sin(x - y) \text{ and}$$

$$\frac{\partial f}{\partial z}(x, y, z) = 2z \cos(x - y).$$

Hence, the Gradient of the function f at (x, y, z) is

$$\nabla f(x, y, z) = \left(\frac{\partial f}{\partial x}(x, y, z), \frac{\partial f}{\partial y}(x, y, z), \frac{\partial f}{\partial z}(x, y, z)\right) =$$

$$= \left(-z^2 \sin(x - y), z^2 \sin(x - y), 2z \cos(x - y)\right).$$

Since all of these partial derivatives are continuous functions on $\mathbb{R}^3$, we conclude that the function f is differentiable at each $a \in \mathbb{R}^3$.

We will compute this differential for two different points.

Case 1: Consider now $a = (1, 1, 1) \in \mathbb{R}^3$. The differential function associated to f at the point $(1, 1, 1)$ is

$$df(1, 1, 1) : \mathbb{R}^3 \to \mathbb{R}$$

defined for all $h=(h_1,h_2,h_3)\in\mathbb{R}^n$, by

$$df(a)(h)=df(1,1,1)(h_1,h_2,h_3)=\langle h,\nabla f(a)\rangle=$$

$$=h_1\frac{\partial f}{\partial x}(1,1,1)+h_2\frac{\partial f}{\partial y}(1,1,1)+h_1\frac{\partial f}{\partial z}(1,1,1)=$$

$$=h_1\cdot(-1\sin 0)+h_2\cdot(1\sin 0)+h_3\cdot(2\cdot 1\cos 0)=h_1\cdot 0+h_2\cdot 0+h_3\cdot 2=$$

$$=2h_3.$$

Case 2: Consider now $a=\left(\frac{\pi}{2},0,1\right)\in\mathbb{R}^3$. The differential function associated to f at the point $\left(\frac{\pi}{2},0,1\right)$ is

$$df\left(\frac{\pi}{2},0,1\right):\mathbb{R}^3\to\mathbb{R}$$

defined for all $h=(h_1,h_2,h_3)\in\mathbb{R}^n$, by

$$df(a)(h)=df\left(\frac{\pi}{2},0,1\right)(h_1,h_2,h_3)=\langle h,\nabla f(a)\rangle=$$

$$=h_1\frac{\partial f}{\partial x}\left(\frac{\pi}{2},0,1\right)+h_2\frac{\partial f}{\partial y}\left(\frac{\pi}{2},0,1\right)+h_1\frac{\partial f}{\partial z}\left(\frac{\pi}{2},0,1\right)=$$

$$=h_1\cdot(-1\sin\frac{\pi}{2})+h_2\cdot(1\sin\frac{\pi}{2})+h_3\cdot(2\cdot 1\cos\frac{\pi}{2})=h_1\cdot 1+h_2\cdot 1+h_3\cdot 2\cdot 0=$$

$$=-h_1+h_2.$$

Conlusion: for the two distinct particular cases considered, we have determined two **different differential functions**, namely

$$df(1,1,1):\mathbb{R}^{\to}\mathbb{R}\qquad df(1,1,1)(h_1,h_2,h_3)=2h_3,\quad\forall(h_1,h_2,h_3)\in\mathbb{R}^3.$$

and

$$df\left(\frac{\pi}{2},0,1\right):\mathbb{R}^3\to\mathbb{R},\quad df\left(\frac{\pi}{2},0,1\right)=-h_1+h_2,\quad\forall(h_1,h_2,h_3)\in\mathbb{R}^3.$$

Recall Exercise 5 from Seminar 5:

Determine the gradient of the function f at the point a in the following cases:

a) $f:\mathbb{R}^2\to\mathbb{R}$, $f(x,y)=e^{-x}\sin(x+2y)$ **and** $a=\left(0,\frac{\pi}{2}\right)$;

b) $f:\mathbb{R}^3\to\mathbb{R}$, $f(x,y,z)=(x-y)\cos(\pi z)$ **and** $a=\left(1,0,\frac{1}{2}\right)$.

Determine now the differential function of f at the given a.

Solution:

a) Let $(x,y)\in\mathbb{R}^2$ be randomly chosen. Then $\frac{\partial f}{\partial x}(x,y)=-e^{-x}\sin(x+2y)+e^{-x}\cos(x+2y)$ and $\frac{\partial f}{\partial y}(x,y)=2e^{-x}\cos(x+2y)$.

We notice that the partial derivatives with respect to both variables are continuous functions. This means that f has a differential function at all the points in $\mathbb{R}^2$.

For the given $a = \left(0, \frac{\pi}{4}\right)$, the gradient is

$$\nabla f\left(0, \frac{\pi}{4}\right) = \left(\frac{\partial f}{\partial x}\left(0, \frac{\pi}{4}\right), \frac{\partial f}{\partial y}\left(0, \frac{\pi}{4}\right)\right)$$
$$= \left(-e^{-0}\sin\left(0+\frac{\pi}{2}\right) + e^{-0}\cos\left(0+\frac{\pi}{2}\right), 2e^{-0}\cos\left(0+\frac{\pi}{2}\right)\right) = (-1, 0).$$

The differential is the function $df\left(0, \frac{\pi}{4}\right) : \mathbb{R}^2 \to \mathbb{R}$, **defined by**

$$df\left(0, \frac{\pi}{4}\right)(h_1, h_2) = \langle h, \nabla f\left(0, \frac{\pi}{4}\right)\rangle = h_1 \cdot (-1) + h_2 \cdot 0 = -h_1, \quad \forall h = (h_1, h_2) \in \mathbb{R}^2 .$$

b) Let $(x, y, z) \in \mathbb{R}^3$ be randomly chosen. Then $\frac{\partial f}{\partial x}(x, y, z) = \cos \pi z$, $\frac{\partial f}{\partial x}(x, y, z) = -\cos \pi z$ and $\frac{\partial f}{\partial x}(x, y, z) = -(x-y)\pi \sin \pi z$.

We notice that the partial derivatives with respect to both variables are continuous functions. This means that f has a differential function at all the points in $\mathbb{R}^3$.

For the given $\left(1, 0, \frac{1}{2}\right)$, the Gradient of f is

$$\nabla f\left(1, 0, \frac{1}{2}\right) = \left(\cos\frac{\pi}{2}, -\cos\frac{\pi}{2}, -(1-0)\pi \sin\frac{\pi}{2}\right) = (0, 0, -\pi).$$

The differential function associated to f at the point $\left(1, 0, \frac{1}{2}\right)$ is $df\left(1, 0, \frac{1}{2}\right) : \mathbb{R}^3 \to \mathbb{R}$, **given by**

$$df\left(1, 0, \frac{1}{2}\right)(h_1, h_2, h_3) = \langle h, \nabla f\left(1, 0, \frac{1}{2}\right)\rangle = h_1 \cdot 0 + h_2 \cot 0 + h_3 \cdot (-\pi).$$

Thus

$$df\left(1, 0, \frac{1}{2}\right)(h_1, h_2, h_3) = -\pi \cdot h_3 \quad \forall h = (h_1, h_2, h_3) \in \mathbb{R}^3 .$$

Exercise: Determine the differential of the following function, $f : \mathbb{R}^2 \to \mathbb{R}^3$, defined by

$$f(x, y) = (\cos x + \sin y, \sin x + \cos y, e^{x-y}), \quad \forall (x, y) \in \mathbb{R}^2,$$

at a random point $(x, y) \in \mathbb{R}^2$.

Solution:

Let $(x, y) \in \mathbb{R}^2$ be randomly chosen. To begin with, we notice that our function is a vector one, of vector variables, so it takes vectors of dimension two and maps them into vectors of dimension three. In order to determine the differential function, we need (if they exist and are continuous functions) all the partial derivatives. We will consider

$$f = (f_1, f_2, f_3), \quad \text{with } f_1, f_2, f_3 : \mathbb{R}^2 \to \mathbb{R}$$

$$f_1(x,y)=\cos x+\sin y,\quad f_2(x,y)=\sin x+\cos y,\quad f_3(x,y)=e^{x-y},\quad \forall (x,y)\in\mathbb{R}^2.$$

We start computing the partial derivatives of f_1.

$$\frac{\partial f_1}{\partial x}(x,y)=(\cos x+\sin y)'_x=-\sin x+0=-\sin x,$$

and

$$\frac{\partial f_1}{\partial y}(x,y)=(\cos x+\sin y)'_y=0+\cos y=\cos y.$$

We continue with

$$\frac{\partial f_2}{\partial x}(x,y)=(\sin x+\cos y)'_x=\cos x+0=\cos x,$$

and

$$\frac{\partial f_2}{\partial y}(x,y)=(\sin x+\cos y)'_y=0-\sin y=-\sin y.$$

And conclude with

$$\frac{\partial f_3}{\partial x}(x,y)=(e^{x-y})'_x=e^{x-y}\cdot(x-y)'_x=e^{x-y}\cdot(1-0)=e^{x-y}$$

and

$$\frac{\partial f_3}{\partial y}(x,y)=(e^{x-y})'_y=e^{x-y}\cdot(x-y)'_y=e^{x-y}\cdot(0-1)=-e^{x-y}.$$

We construct the Jacobi matrix associated to the function f at the point (x,y),

$$J(f)(x,y)=\begin{pmatrix}\frac{\partial f_1}{\partial x}(x,y) & \frac{\partial f_1}{\partial y}(x,y)\\ \frac{\partial f_2}{\partial x}(x,y) & \frac{\partial f_2}{\partial y}(x,y)\\ \frac{\partial f_3}{\partial x}(x,y) & \frac{\partial f_3}{\partial y}(x,y)\end{pmatrix}=\begin{pmatrix}-\sin x & \cos y\\ \cos x & -\sin y\\ e^{x-y} & -e^{x-y}\end{pmatrix}$$

We notice that all the partial derivatives are continuous functions, thus f is differentiable at each random point $(x,y)\in\mathbb{R}^2$. The differentiable function is $df(x,y):\mathbb{R}^2\to\mathbb{R}^3$ given by

$$df(x,y)(h_1,h_2)=J(f)(x,y)\cdot\begin{pmatrix}h_1\\ h_2\end{pmatrix}=$$

$$=\begin{pmatrix}-\sin x & \cos y\\ \cos x & -\sin y\\ e^{x-y} & -e^{x-y}\end{pmatrix}\cdot\begin{pmatrix}h_1\\ h_2\end{pmatrix}=\begin{pmatrix}-\sin x\cdot h_1+\cos y\cdot h_2\\ \cos x\cdot h_1-\sin y\cdot h_2\\ e^{x-y}\cdot h_1-e^{x-y}\cdot h_2\end{pmatrix}$$

This is the matrix form of the linear map $df(x,y)$. It vector form is

$$d(f)(x,y)(h_1,h_2) = \left(-\sin x \cdot h_1 + \cos y \cdot h_2, \quad \cos x \cdot h_1 - \sin y \cdot h_2, \quad e^{x-y} \cdot h_1 - e^{x-y} \cdot h_2\right),$$

for all $(h_1,h_2) \in \mathbb{R}^2$.

For the particular case when $(x,y) = \left(\frac{\pi}{2}, \frac{\pi}{2}\right)$, the Jacobi matrix is

$$J(f)\left(\frac{\pi}{2}, \frac{\pi}{2}\right) = \begin{pmatrix} -1 & 0 \\ \\ 0 & -1 \\ \\ 1 & -1 \end{pmatrix}$$

And the differential at this point is $df\left(\frac{\pi}{2}, \frac{\pi}{2}\right) : \mathbb{R}^2 \to \mathbb{R}^3$, defined by

$$df\left(\frac{\pi}{2}, \frac{\pi}{2}\right)(h_1,h_2) = (-h_1, -h_2, h_1 - h_2), \quad \forall (h_1,h_2) \in \mathbb{R}^2.$$

Let us recall the fact that a vector in $\mathbb{R}^n$ may be written in its matrix form, as a matrix with n rows and 1 column, thus

$$x = (x_1, ..., x_n) = \begin{pmatrix} x_1 \\ x_2 \\ ... \\ x_3 \end{pmatrix}$$

As **Homework**, please solve completely exercises 1, 2, 3 and 4 from Seminar5-ttrif