Chapter 1

Differential Calculus

1.1 Topology Concepts - Reminders

In this section, we review the basic notions of topology, including:
- Open sets. - Neighborhoods. - Open ball, closed ball....

Normed vectorial spaces and Banach spaces

Consider the vectorial space E over $\mathbb{R}$ or $\mathbb{C}$.
Define the norm, the application $\|.\| : E \to \mathbb{R}_+$, verifying:

1. $\|x\| = 0$ iff $x = 0$.
2. $\|\alpha x\| = |\alpha|\|x\|$ for all $\alpha \in \mathbb{R}$ or $\mathbb{C}$ and for all $x \in E$.
3. $\|x + y\| \leq \|x\| + \|x\|$ for all $x, y \in E$.

We can then associate an application called distance, denoted by d and defined by $d(x, y) = \|x - y\|$ for all x and $y \in E$.

1.2 Continuity Concepts

In this section, we review the basic topological notions of continuity of applications and linear operators.

Definition 1.1 *(Cauchy sequence) A sequence of elements* $(x_n) \in E$ *is called Cauchy if* $\forall \varepsilon > 0, \exists N \in \mathbb{N}$ *such that* $\forall n \geq N$ *and* $p \geq 0, d(x_n, x_{n+p}) < \varepsilon$.

We say that E is a Banach space (or a complete normed vector space) if every Cauchy sequence in E converges in E.

Theorem 1.2 *(Continuous linear applications) Let* E *and* F *be two normed vector spaces and* $f : E \to F$ *a given mapping. The following three statements are equivalent:*

1. f *is continuous*
2. f *is continuous at* 0.
3. *There exists a strictly positive constant* M *such that* $\|f(x)\| \leq M$ *for all* $x \in E$ *with* $\|x\| \leq 1$.

In other words, f is bounded on the unit ball. This property is equivalent to $\exists M > 0$ such that

$$\forall x \in E; \|f(x)\| \leq M\|x\|$$

.

1.3 Differentiability Concepts - Derivability

Let E, F be two vectorial spaces and $\mathcal{O}$ an open set in E (in general $E = \mathbb{R}^n$ and $F = \mathbb{R}^p$).

Definition 1.3 *: Let $f : \mathcal{O} \subset E \to F$ be an application, and $a \in \mathcal{O}$ a point. We say that f is differentiable at a if there exists a linear application $L : E \to F$ such that, for all $h \in E$, $a + h \in \mathcal{O}$,*

$$f(a+h) = f(a) + L(h) + \|h\|\varepsilon(h)$$

and

$$\lim_{\|h\| \to 0} \|\varepsilon(h)\| = 0$$

Proposition 1.4 *: The application L is unique.*

Definition 1.5 *: The unique linear application L is called the differential of f at the point a or the tangent application of f at a, and is denoted by df_a, $T_a f$, or $f'(a)$.*

Remark 1.6 *: If E and F are finite-dimensional (e.g., $\mathbb{R}^n, \mathbb{R}^p, \mathbb{R}^q, \ldots$ etc.), any linear application $L : E \to F$ is continuous.*

Example 1.7 *: Consider the application*

$$f : \mathbb{R}^2 \to \mathbb{R}$$
$$(x, y) \mapsto f(x, y) = x^2 + y^2 - 1$$

For all $a \in \mathcal{O} \subset \mathbb{R}^2$, $df_a = L$, $a = (a_1, a_2) \in \mathbb{R}^2$, $h \in E$, $h = (h_1, h_2)$,

$$\begin{aligned} f(a+h) - f(a) &= f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2) \\ &= (a_1 + h_1)^2 + (a_2 + h_2)^2 - 1 - (a_1^2 + a_2^2 - 1) \\ &= 2a_1 h_1 + 2a_2 h_2 + h_1^2 + h_2^2 \\ &= L(h) + \|h\|\varepsilon(h) \end{aligned}$$

with $\varepsilon(h) = \frac{h_1^2 + h_2^2}{\|h\|}$, $df_a = f'(a) = L$, thus

$$L : \mathbb{R}^2 \to \mathbb{R}$$
$$\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \mapsto 2a_1 h_1 + 2a_2 h_2 = (2a_1 - 2a_1) \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$$

Remarks 1.8 *:*

1. *If f is differentiable at a, then f is continuous at a.*

2. *If $f \ : \ \mathbb{R}^n \to \mathbb{R}^p$ with $n = p = 1$, then f is differentiable at a if and only if f is derivable at a.*

3. *The previous definition remains valid for arbitrary Banach spaces, but the continuity of the application df_a must be added.*

Proposition 1.9 *: (Composition) Let $\mathcal{O}$ be an open set in $\mathbb{R}^n$ and W an open set in $\mathbb{R}^p$. If $f : \ \mathbb{R}^n \to \mathbb{R}^p$ is differentiable on $\mathcal{O}$ and $g \ : \mathbb{R}^p \to \mathbb{R}^q$ is differentiable on W, such that $f(\mathcal{O}) \subset W$, then $g \circ f$ is differentiable on $\mathcal{O}$, and for all $a \in \mathcal{O}$,*

$$d(g \circ f)_a = (dg)_{f(a)} \circ df_a$$

Definition 1.10 *: We say that f is differentiable on $\mathcal{O}$ if f is differentiable at every point of $\mathcal{O}$.*

1.4 First-Order Partial Derivatives and Gradient

The partial derivative of a function is the derivative with respect to one of its variables, keeping the others constant. The partial derivative with respect to the variable x is denoted by $\frac{\partial f}{\partial x}$ or $\partial_x f$.

$$f : \mathcal{O} \subset \mathbb{R}^n \longrightarrow \mathbb{R}^p$$
$$(x_1, \ldots, x_n) \mapsto \begin{pmatrix} f_1(x_1, \ldots, x_n) \\ \vdots \\ f_p(x_1, \ldots, x_n) \end{pmatrix}$$

The first derivative at $a \in O$ with respect to x_j is the following limit, if it exists:

$$\lim_{h \to 0} \frac{f(a_1, \ldots, x_j + h, \ldots, a_n) - f(a_1, \ldots, a_n)}{h}$$

This limit is denoted by $\frac{\partial f}{\partial x_j}(a)$.

The gradient of f at (x_0, y_0), denoted by ∇f or $\mathrm{grad} f(x_0, y_0)$, is the vector whose components are the first partial derivatives.

1.4.1 Function of Class C^k

A function $f : \mathcal{O} \subset \mathbb{R}^n \longrightarrow \mathbb{R}^p$ is of class C^1 if it admits first partial derivatives with respect to all variables at each point $a \in \mathcal{O}$, and these derivatives are continuous. Similarly, f is of class C^2 if it admits partial derivatives with respect to all variables, and the applications $a \mapsto \frac{\partial f}{\partial x_j}(a)$, $j = 1, n$, are all of class C^1 on $\mathcal{O}$.

Remark 1.11 *: A function is said to be of class $C^0 = C$ if it is continuous, of class C^k (for $k \geq 1$) if its partial derivatives up to order k exist and are continuous, and of class C^∞ (infinitely differentiable) if it is of class C^k for all $k \in \mathbb{N} \cup \{+\infty\}$.*

Conclusion: $f \in C^1 \implies \frac{\partial f_i}{\partial x_i} \in C^0$.

1.5 Jacobian Matrix

$$f : \mathcal{O} \subset \mathbb{R}^n \to \mathbb{R}^p$$
$$(x_1, \ldots, x_n) \mapsto \begin{pmatrix} f_1(x_1, \ldots, x_n) \\ \vdots \\ f_p(x_1, \ldots, x_n) \end{pmatrix}$$

The partial derivatives of these functions at a point a, if they exist, can be arranged in a matrix $df_a = f'(a) = J_f(a) = Jf_a$, such that

$$J_f(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1} & \cdots & \frac{\partial f_p}{\partial x_n} \end{pmatrix}$$

Remarks 1.12 *:*

1. *The Jacobian matrix is the matrix of first-order partial derivatives of a differentiable function.*
2. *The Jacobian matrix of f at a is the matrix of the linear application df_a in the canonical bases of $\mathbb{R}^n$ and $\mathbb{R}^p$.*
3. *If $n = p$, the Jacobian matrix is square, and its determinant $\det(Jf_a)$ is called the Jacobian determinant or Jacobian.*

Examples 1.13 *:*

1. *Let f be defined by*

$$\begin{array}{ccc} f : \mathbb{R}^3 & \longrightarrow & \mathbb{R}^4 \\ (x_1 x_2, x_3) & \longmapsto & (x_1, 5x_3, 4{x_2}^2 - 2x_3, x_3 \sin x_1\ ) \end{array}$$

2. *Let g be defined by*

$$\begin{array}{ccc} g : \mathbb{R}_+ \times [0, 2\pi] & \longrightarrow & \mathbb{R}^2 \\ (r, \theta) & \longmapsto & (r\cos\theta\ , r\ \sin\theta\ ) \end{array}$$

Question: Determine $Jf(x_1, x_2, x_3)$, $\det(J_g(r, \theta))$.

$$J_f = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 5 \\ 0 & 8x_2 & -2 \\ x_3 \cos x_1 & 0 & \sin x_1 \end{pmatrix}$$

$$J_g = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

Thus, $\det(J_g) = r\cos^2\theta + r\sin^2\theta = r$.

Corollary 1.14 *: The Jacobian matrix of the composition $g \circ f$ is:*

$$J(g \circ f)_a = J(g)_{f(a)} J f_a$$

i.e.,

$$J_{(g\circ f)(a)} = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(f_1(a)) & \dots & \frac{\partial g_n}{\partial x_n}(f_1(a)) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_p}{\partial x_1}(f_p(a)) & \dots & \frac{\partial g_p}{\partial x_n}(f_p(a)) \end{pmatrix} \times \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1} & \dots & \frac{\partial f_p}{\partial x_n} \end{pmatrix}$$

Definition 1.15 *: (Directional Derivative - Gâteaux Sense)*
Let f be an application from an open set U of E to F, v an element of E, and $a \in U$. We say that f admits a derivative at a in the direction of v if the following limit exists:

$$\frac{\partial f}{\partial v}(a) = \lim_{t\to 0} \frac{f(a+tv) - f(a)}{t}$$

We say that f is Gâteaux differentiable at a if the directional derivatives $\frac{\partial f}{\partial v}(a)$ are defined for all $v \in E$, and the application $v \mapsto \frac{\partial f}{\partial v}(a)$ is linear.

Proposition 1.16 *Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$ be a function differentiable at a. Then we have the formula:*

$$df(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) dx_i$$

Definition 1.17 *(Fréchet Derivative) Let E, F be two normed spaces. Let f be a mapping from E to F. We say that f is Fréchet differentiable at a point $a \in E$ if there exists a continuous linear mapping $L : E \to F$ such that:*

$$f(x+h) - f(x) = L \cdot h + \epsilon(x,h)$$

where:

$$\lim_{h\to 0} \frac{\|\epsilon(x,h)\|}{\|h\|} = 0 \text{ as } \|h\| \to 0.$$

Remarks 1.18

1. *If a function f is differentiable in the Fréchet sense, then it is differentiable in the Gâteaux sense in every direction.*

2. *The converse is not true.*

Definition 1.19 *Let E, F be two normed spaces, $\mathcal{O}$ an open subset of E and V an open subset of F. A mapping $f : \mathcal{O} \to V$ is called a :*

1. *Diffeomorphism if:*

   (a) f is bijective.

   (b) f is differentiable on $\mathcal{O}$.

   (c) $f^{-1} : V \to \mathcal{O}$ is differentiable on V.

2. *C^k-diffeomorphism if:*

(a) f is bijective.

(b) f is C^k-differentiable.

(c) $f^{-1} : V \to \mathcal{O}$ is C^k-differentiable.

3. If $k = 0$, the mapping f is called a homeomorphism, i.e.,

(a) f is bijective.

(b) f is continuous.

(c) $f^{-1} : V \to \mathcal{O}$ is continuous.

Remarks 1.20

1. $Isom(E, F) = \{f : E \to F$ linear and bijective$\}$
2. Every C^k-diffeomorphism is a homeomorphism, but the converse is false.
3. A homeomorphism preserves topological properties (open, closed, compact, convex, etc.).
4. A diffeomorphism preserves both topological and geometric properties.

1.6 Rank Theorem

Definition 1.21 *Let E, F be two vector spaces and $\varphi : E \to F$ a linear mapping. The rank of the mapping φ, denoted rang(φ) or rg(φ), is $r = \dim(Im(\varphi))$. We have:*

1. $r \leq \dim(E)$, $r \leq \dim(F)$.
2. $r = \dim(E) \Longleftrightarrow \varphi$ is injective.
3. $r = \dim(F) \Longleftrightarrow \varphi$ is surjective.

Definition 1.22

1. Let $\varphi : \mathcal{O} \subset E \to F$ be of class C^k and $x \in \mathcal{O}$. We say that φ is an immersion at x if $d\varphi_x$ is injective.
2. We say that φ is a submersion at x if $d\varphi_x$ is surjective.
3. If $\varphi : \mathcal{O} \to F$ is both a submersion and an immersion at x, we say that φ is étale.

Definition 1.23 *A mapping f of class C^k $(k \geq 1)$ from $\mathcal{O}$ to V is said to be étale at x if:*

$$df_x \in Isom(E, F)$$

Proposition 1.24 *Let $\varphi : \mathcal{O} \subset \mathbb{R}^n \to \mathbb{R}^p$ be a mapping of class C^k. Then:*

1. φ is an immersion on $\mathcal{O} \Longleftrightarrow rank(d\varphi_x) = n$, $\forall x \in \mathcal{O}$.
2. φ is a submersion on $\mathcal{O} \Longleftrightarrow rank(d\varphi_x) = p$, $\forall x \in \mathcal{O}$.

1.7 Local and Global Inverse Function Theorem

Let I be an interval of $\mathbb{R}$ and $f : I \to \mathbb{R}$ a function of class C^1. A classical result from L1 assures us that if f' does not vanish on I, then f is a bijection from I to $f(I)$ (which is still an interval) and $f^{-1} : f(I) \to I$ is also of class C^1.

Theorem 1.25 *(Local Inverse Function Theorem) Let U, W be two open subsets of E and F respectively, and let $f \in C^k(U, W)$ be a mapping that is étale at $x \in U$. Then there exists an open subset V of U containing x such that $f|_V : V \to f(V) \subset F$ is a C^k-diffeomorphism from V to $f(V)$. Moreover, denoting f^{-1} as the inverse diffeomorphism, we have:*

$$\forall y \in f(V), \quad df^{-1}(y) = [d(f^{-1}(y))]^{-1}.$$

Theorem 1.26 *(Global Inverse Function Theorem) Let U be an open subset of E and φ a mapping from U to E. Suppose that:*

1. *φ is of class C^k on U.*
2. *φ is bijective from U to $\varphi(U)$.*
3. *For all $x \in U$, $d\varphi_x$ is invertible.*

Then $\varphi(U)$ is an open subset and φ is a C^k-diffeomorphism from U to $\varphi(U)$.

Corollary 1.27 *Let I be an open interval of $\mathbb{R}$ and $f : I \to \mathbb{R}$ a function of class C^k such that f' does not vanish on I. Then $f(I)$ is an open interval and f is a C^k-diffeomorphism from I to $f(I)$. 0*

Remarks 1.28

1. *φ is a C^k-diffeomorphism $\Longrightarrow$ φ is étale.*
2. *φ is étale $\Longrightarrow$ φ is a C^k-diffeomorphism (locally).*
3. *φ is étale and bijective $\Longrightarrow$ φ is a C^k-diffeomorphism (globally).*

Example 1.29 *Consider the function:*

$$f : \mathbb{R}^* \times \mathbb{R} \to \mathbb{R}^2, \quad (r, \theta) \mapsto (r\cos\theta, r\sin\theta)$$

Is f étale?

$$df_{(r,\theta)} = Jf_{(r,\theta)} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

$$|Jf_{(r,\theta)}| = r \neq 0$$

Thus, $df_{(r,\theta)}$ is linear and bijective, i.e., $df_{(r,\theta)} \in Isom(\mathbb{R}^2, \mathbb{R}^2)$, so f is étale.
Is f a diffeomorphism?

$$f(r, \theta) = f(r, \theta + 2\pi) \Longrightarrow f \text{ is not injective}(\theta \neq \theta + 2\pi)$$

Thus, f is not a global diffeomorphism.

1.8 Implicit Function Theorem

The implicit function theorem allows us to determine when a curve defined by an equation of the form $f(x, y) = 0$ is locally the graph of a function $y = \varphi(x)$, where the regularity of φ is related to that of f.

Theorem 1.30 *(Implicit Function Theorem) Let f be a mapping of class C^1 on the open subset U of $\mathbb{R}^n \times \mathbb{R}^p$ with values in $\mathbb{R}^p$, and let (a, b) be a point in U such that:*

1. $f(a, b) = 0$.
2. $J_f^y(a, b)$ *is invertible.*

Then there exists an open subset V of $\mathbb{R}^n$, an open subset W of $\mathbb{R}^p$, and a mapping g of class C^1 defined on V with values in W such that:

1. $(a, b) \in V \times W \subset U$.
2. *For every pair $(x, y) \in V \times W$, we have $f(x, y) = 0 \Longleftrightarrow y = g(x)$.*
3. *For every pair $(x, y) \in V \times W$, the matrix $J_f^y(x, y)$ is invertible, and the Jacobian matrix of g is given by:*

$$J_g(u) = -[J_f^y(u, g(u))]^{-1} \cdot [J_f^x(u, g(u))].$$

Example 1.31 *Consider the function:*

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad (x, y) \mapsto x^2 + y^2 - 1$$

We want to solve the equation $f(a, b) = 0$ for a value near a, i.e., if we can find a function g defined near a such that for x near a, $f(x, g(x)) = 0$.

$$\exists g = ?, \quad f(a, b) = 0$$
$$(x, y) \in V \times W \quad f(x, y) = 0 \Longleftrightarrow y = g(x)$$
$$V = g(x) = \sqrt{1 - x^2}$$