Lectures notes on:
Multivariable Calculus

Jean-Marc Bonnisseau[1]

March 22, 2021

[1]Paris School of Economics, Université Paris 1 Panthéon Sorbonne, 106-112 Boulevard de l'Hôpital, 75647 Paris Cedex 13, France, Jean-marc.Bonnisseau@univ-paris1.fr

Contents

Chapter 1

Multivariable Calculus

1.1 Derivative of $f : \mathbb{R} \to \mathbb{R}^p$

Definition 1 Let f be a mapping of an interval J into $\mathbb{R}^p$. We assume that the interval has more than one point, but the interval may contain its end points. We say that f is differentiable at a number t in its interval of definition if $\lim_{h\to 0} \frac{f(t+h)-f(t)}{h}$ exists, in which case this limit is called the derivative of f at t and is denoted by $f'(t)$.
We say that f is differentiable (on J) if it is differentiable at every $t \in J$, and in that case, f' is a mapping of J into $\mathbb{R}^p$.
If f has p continuous derivatives, we say f is of class $\mathcal{C}^p$.
If f is infinitely differentiable, we say that f is $\mathcal{C}^\infty$.

Remark 1 $f : J \to \mathbb{R}^p$ can be represented by coordinate functions,
$f(t) = (f_1(t), \ldots, f_p(t))$ and $\frac{f(t+h)-f(t)}{h} = \left(\frac{f_1(t+h)-f_1(t)}{h}, \ldots, \frac{f_p(t+h)-f_p(t)}{h}\right)$.
The limit can be taken componentwise, and consequently f is differentiable if and only if each coordinate function is differentiable, and then
$f'(t) = (f_1'(t), \ldots, f_p'(t))$.
One usually views a map f such as above as a parametrized curve in $\mathbb{R}^p$.

Examples: Let $f(t) = (cos(t),\ sin(t))$ parametrizes the circle. We have $f'(t) = (-sin(t),\ cos(t))$.
Let $f(t) = (cos(t),\ sin(t),\ t)$. Then $f(t)$ describes a spiral. Its projection in the plane of the first two coordinates is of course the circle.

The examples give a curve in $\mathbb{R}^2$ and $\mathbb{R}^3$ respectively.

To distinguish such curves from those given by an equation like $x^2 + y^2 = 1$ we also call them parametrized curves. If f is a differentiable curve, then the derivative f' is called the velocity of the curve. The second derivative f'', if it exists, is called the acceleration of the curve.

Proposition 1 *Let f and g from $J \to \mathbb{R}^p$, if f and g are differentiable at t, then*

so is $f+g$ and $(f+g)'(t) = f'(t) + g'(t)$.
If $f : J \to \mathbb{R}^p$ and $g : J \to \mathbb{R}^p$ are differentiable at t, let $f \cdot g$ be defined by $(f \cdot g)(t) = f(t) \cdot g(t)$. Then:
$(f \cdot g)'(t) = f(t) \cdot g'(t) + f'(t) \cdot g(t)$.

Proposition 2 (Chain rule) *Let J_1, J_2 be intervals. Let $f : J_1 \to J_2$ and $g : J_2 \to \mathbb{R}^p$ be maps. Let $t \in J_1$. If f is differentiable at t and g is differentiable at $f(t)$, then $g \circ f$ is differentiable at t and*
$(g \circ f)'(t) = g'(f(t))f'(t)$.

1.2 Derivative of $f : \mathbb{R}^n \to \mathbb{R}$

1.2.1 Partial Derivatives

Definition 2 Let U be an open set of $\mathbb{R}^n$, and let $f : U \to \mathbb{R}$ be a function. We define its partial derivative at a point $x = (x_1, \ldots, x_n) \in U$ by
$\frac{\partial f}{\partial x_i}(x) = \lim_{h \to 0, h \neq 0} \frac{f(x+he^i)-f(x)}{h} = \lim_{h \to 0, h \neq 0} \frac{f(x_1,\ldots,x_i+h,\ldots,x_n)-f(x_1,\ldots,x_n)}{h}$ if the limit exists.
$ei = (0, \ldots, 1, \ldots, 0)$ is the unit vector with the i-th component being equal to 1 and all others equal to 0. Note that $f(x + he^i)$ is well defined if h is small enough since U, the domain of f, is open and x belongs to U.

Remark 2 We see that $\frac{\partial f}{\partial x_i}$ is an ordinary derivative which keeps all variables fixed but not the i-th variable. In particular, we know that the derivative of a sum, and the derivative of a constant times a function follow the usual rules, that is $\frac{\partial f+g}{\partial x_i} = \frac{\partial f}{\partial x_i} + \frac{\partial g}{\partial x_i}$ and $\frac{\partial cf}{\partial x_i} = c\frac{\partial f}{\partial x_i}$ for any constant c.

Example: If $f(x, y) = 3x^3y^2$ then $\frac{\partial f}{\partial x}(x, y) = 9x^2y^2$ and $\frac{\partial f}{\partial y}(x, y) = 6x^3y$. Of course we may iterate partial derivatives. In this example, we have $\frac{\partial^2 f}{\partial x^2}(x, y) = 18xy^2$, $\frac{\partial^2 f}{\partial y^2}(x, y) = 6x^3$ and $\frac{\partial^2 f}{\partial x \partial y}(x, y) = 18x^2y$,
$\frac{\partial^2 f}{\partial y \partial x}(x, y) = 18x^2y$.
Observe that the two last iterated partials are equal. This is not an accident, and is a special case of the following general theorem.

Theorem 1 (Schwarz) *Let f be a function on an open set $U \in \mathbb{R}^2$. Assume that the partial derivatives $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial^2 f}{\partial x \partial x}, \frac{\partial^2 f}{\partial y \partial y}, \frac{\partial^2 f}{\partial x \partial y}, \frac{\partial^2 f}{\partial y \partial x}$ exist and are continuous. Then $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$.*

Definition 3 We define the gradient of f at any point x at which all partial derivatives exist to be the vector $\nabla f(x) = \left(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x)\right)$.

Definition 4 We define the Hessian matrix of f at any point

$x = (x_1, \dots, x_n)$ by

$$H(f,x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_2^2}(x) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \frac{\partial^2 f}{\partial x_n \partial x_2}(x) & \dots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix}$$

Remark on higher order partial derivatives

Let f be a function on an open set U of $\mathbb{R}^n$. We may take iterated partial derivatives (if they exist) of the form $\left(\frac{\partial}{\partial x_1}\right)^{i_1} \dots \left(\frac{\partial}{\partial x_n}\right)^{i_n} f$ where $i_1, \dots, i_n$ are integers ≥ 0. It does not matter in which order we take the partials (provided they exist and are continuous) according to the Schwarz's theorem. If $c_{i_0 \dots i_n}$ are numbers, we may form finite sums $\sum c_{i_1 \dots i_n} \left(\frac{\partial}{\partial x_1}\right)^{i_1} \dots \left(\frac{\partial}{\partial x_n}\right)^{i_n}$ which we view as applicable to functions which have enough partial derivatives. More precisely, we say that a function f on U is of class $\mathcal{C}^p$, for some integer $p \geq 0$, if all partial derivatives $\left(\frac{\partial}{\partial x_1}\right)^{i_1} \dots \left(\frac{\partial}{\partial x_n}\right)^{i_n} f$ exist for $i_1, \dots, i_n \leq p$ and are continuous. It is clear that the functions of class $\mathcal{C}^p$ form a vector space, that is the sum of two functions of class $\mathcal{C}^p$ is of class $\mathcal{C}^p$ and the product of a function of class $\mathcal{C}^p$ by a real number is a function of class $\mathcal{C}^p$.. Let $i_1, \dots, i_n$ be integers ≥ 0 such that $i_1 + \dots + i_n = r \leq p$.

Let F_p be the vector space of functions of class $\mathcal{C}^p$. (For $p = 0$, this is the vector space of continuous functions on U.) Then any monomial $\left(\frac{\partial}{\partial x_1}\right)^{i_1} \dots \left(\frac{\partial}{\partial x_n}\right)^{i_n}$ may be viewed as a linear map $F_p \to F_{p-r}$ given by $f \mapsto \left(\frac{\partial}{\partial x_1}\right)^{i_1} \dots \left(\frac{\partial}{\partial x_n}\right)^{i_n} f$.

We say that f is of class $\mathcal{C}^\infty$ if it is of class $\mathcal{C}^p$ for every positive integer p.

If f is of class $\mathcal{C}^\infty$, then $\left(\frac{\partial}{\partial x_1}\right)^{i_1} \dots \left(\frac{\partial}{\partial x_n}\right)^{i_n} f$ is also of class $\mathcal{C}^\infty$.

1.2.2 Reminder on Euclidean algebra

Orthogonal spaces

Definition 5 $(u_1, \dots, u_n)$ is an orthogonal basis of $\mathbb{R}^n$ if $(u_1, \dots, u_n)$ is a basis of $\mathbb{R}^n$ and $u_i \cdot u_j = 0$ for all (i,j), $i \neq j$.

$(u_1, \dots, u_n)$ os an orthonormal basis of $\mathbb{R}^n$ if $(u_1, \dots, u_n)$ is an orthogonal basis of $\mathbb{R}^n$ and $\|u_i\| = 1$ for all i.

The canonical basis of $\mathbb{R}^n$ is orthonormal.

Proposition 3 *Let $\mathcal{B} = (u_1, \dots, u_n)$ be an orthonormal basis of $\mathbb{R}^n$ and let x and y be two vectors of $\mathbb{R}^n$. Let $(\xi_1, \dots, \xi_n)$ be the coordinates of x is the basis $\mathcal{B}$ and $(\zeta_1, \dots, \zeta_n)$ be the coordinates of y in the basis $\mathcal{B}$. Then, for all i, $\xi_i = x \cdot u_i$,*

$\zeta_i = y \cdot u_i$ *and*

$$x \cdot y = \sum_{i=1}^{n} \xi_i \zeta_i \text{ and } \|x\| = \sqrt{\sum_{i=1}^{n} \xi_i^2}$$

Let E be a linear subspace of $\mathbb{R}^n$ and $(u_1, \dots, u_p)$ a basis of E. One can built an orthogonal basis of E, $(v_1, \dots, v_p)$, starting from $(u_1, \dots, u_p)$ using the Gram-Schmidt orthogonalisation method as follows:

$v_1 = u_1;$

$v_2 = u_2 - \frac{v_1 \cdot u_2}{\|v_1\|^2} v_1$

$\vdots$

$v_k = u_k - \frac{v_1 \cdot u_k}{\|v_1\|^2} v_1 - \frac{v_2 \cdot u_k}{\|v_2\|^2} v_2 - \dots - \frac{v_{k-1} \cdot u_k}{\|v_{k-1}\|^2} v_{k-1}$

$\vdots$

$v_p = u_p - \frac{v_1 \cdot u_p}{\|v_1\|^2} v_1 - \frac{v_2 \cdot u_p}{\|v_2\|^2} v_2 - \dots - \frac{v_{p-1} \cdot u_p}{\|v_{p-1}\|^2} v_{p-1}$

From which, we deduces that all linear subspace of $\mathbb{R}^n$ has an orthogonal basis.

Exercise 1 Let $(u = (1, 0, 1), v = (2, -1, 1), w = (-1, -1, 2))$ be three vectors of $\mathbb{R}^3$. Show that this is a basis of $\mathbb{R}^3$. Apply the Gram-Schmidt orthogonalisation method to find an orthogonal basis of $\mathbb{R}^3$.

Same question with $((1, 0, 0, 0), (1, 1, 0, 0), (1, 1, 1, 0), (1, 1, 1, 1))$ in $\mathbb{R}^4$.

Let E be a linear subspace of $\mathbb{R}^n$. The orthogonal complement of E denoted $E^\perp$is the set defined by:

$$E^\perp = \{v \in \mathbb{R}^n \mid \forall u \in E, u \cdot v = 0\}$$

Proposition 4 *$E^\perp$ is a linear subspace of $\mathbb{R}^n$. $E \cap E^\perp = \{0\}$.*

Let $(u_1, \dots, u_p)$ be a basis of E, then

$$E^\perp = \{v \in \mathbb{R}^n \mid \forall i = 1, \dots, p \in E, u_i \cdot v = 0\}$$

In other words, $E^\perp$ is the kernel of the linear mapping f from $\mathbb{R}^n$ to $\mathbb{R}^p$ defined by $f(v) = (u_1 \cdot v, \dots, u_p \cdot v)$.

Let E be a linear subspace of $\mathbb{R}^n$ and $(u_1, \dots, u_p)$ be an orthogonal basis of E. we know that there exists $(u_{p+1}, \dots, u_n) \in (\mathbb{R}^n)_{n-p}$ such that $(u_1, \dots, u_p, u_{p+1}, \dots, u_n)$ is a basis of $\mathbb{R}^n$. Using the Gram-Schmidt orthogonalisation method, we build orthogonal basis of $\mathbb{R}^n$ $(v_1, \dots, v_p, v_{p+1}, \dots, v_n)$ from $(u_1, \dots, u_p, u_{p+1}, \dots, u_n)$. Since $(u_1, \dots, u_p)$ is an orthogonal basis of E, we remark that$v_1 = u_1$, $v_2 = u_2$, $\dots$, $v_p = u_p$. So $(v_{p+1}, \dots, v_n)$ are linearly independent vectors of $E^\perp$ and they are a basis of $E^\perp$. So we conclude that

Proposition 5 *1) E are $E^\perp$ are complements in $\mathbb{R}^n$, $\mathbb{R}^n = E \oplus E^\perp$ and $\dim E^\perp = n - \dim E$.*
2) THe orthogonal complement of $E^\perp$ is E: $(E^\perp)^\perp = E$

For all $x \in \mathbb{R}^n$, there exists a unique pair $(y, z) \in E \times E^\perp$ such that $x = y + z$. y is the orthogonal projection of x on E, z is the orthogonal projection of x on $E^\perp$. They are denoted $\mathrm{proj}_E^\perp(x)$ and $\mathrm{proj}_{E^\perp}^\perp(x)$.

Remark that $\mathrm{proj}_E^\perp(x) \cdot \mathrm{proj}_{E^\perp}^\perp(x) = 0$.

Proposition 6 *1) The mappings* $\mathrm{proj}_E^\perp$ *are* $\mathrm{proj}_{E^\perp}^\perp$ *are linear;*

2) The kernel of $\mathrm{proj}_E^\perp$ *(resp. the range* $\mathrm{proj}_{E^\perp}^\perp$*) is* $E^\perp$*, the range of* $\mathrm{proj}_E^\perp$ *(resp. the kernel of* $\mathrm{proj}_{E^\perp}^\perp$*) is* E*.*

3) $\mathrm{proj}_E^\perp \circ \mathrm{proj}_E^\perp = \mathrm{proj}_E^\perp$.

4) $\mathrm{proj}_E^\perp + \mathrm{proj}_{E^\perp}^\perp = \mathrm{Id}$.

We remark that all linear subspaces of $\mathbb{R}^n$ is the kernel of a linear mapping. We can also represent a linear subspace E of $\mathbb{R}^n$ of dimension p by $n-p$ independent linear equations. Indeed, if $(v_1, \ldots, v_{n-p})$ is a basis of $E^\perp$, then

$$E = \{x \in \mathbb{R}^n \mid \forall j = 1, \ldots, n-p, v_j \cdot x = 0\}$$

Proposition 7 *Let* E *and* F *be two linear subspaces of* $\mathbb{R}^n$*. Then* $(E \cap F)^\perp = E^\perp + F^\perp$ *and* $(E+F)^\perp = E^\perp \cap F^\perp$*. If* $E \subset F$*, then* $F^\perp \subset E^\perp$*.*

Let u be a non zero vector in $\mathbb{R}^n$; we denote by $u^\perp$ the orthogonal complement of the line D generated by u: $D = \{tu \mid t \in \mathbb{R}\}$. We remark that $u^\perp$ is an hyperplan, that is, a linear subspace of dimension $n-1$ and the projections on $u^\perp$ and on D are defined as follows:

$$\mathrm{proj}_{u^\perp}^\perp(x) = x - \frac{x \cdot u}{\|u\|^2} u \text{ et } \mathrm{proj}_D^\perp(x) = \frac{x \cdot u}{\|u\|^2} u$$

Linear mappings and inner product

Let f be a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^p$. Let M its $p \times n$ matrix. in the canonical basis of $\mathbb{R}^n$ and $\mathbb{R}^p$. we denote by $(\ell_j)_{j=1}^p$ the rows of the matrix M which are vectors in $\mathbb{R}^n$ and by $(c_i)_{i=1}^n$ the columns of M which are vectors in $\mathbb{R}^p$. The transpose of M denoted M^t is the $n \times p$ matrix whose column vectors are the row vectors of M.

For all $x \in \mathbb{R}^n$, we have two ways to compute the image of x by f.

$$f(x) = \sum_{i=1}^n x_i c_i = (\ell_j \cdot x)_{j=1}^p$$

The transpose of f is the unique linear mapping f^t from $\mathbb{R}^p$ to $\mathbb{R}^n$ satisfying for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^p$, $y \cdot f(x) = f^t(y) \cdot x$.

The matrix of f^t in the canonical basis is M^t. We remark that the transpose of the transpose of f is equal to f.

Proposition 8 *Let* f *be a linear mapping from* $\mathbb{R}^n$ *to* $\mathbb{R}^p$*.*

1) $\mathrm{Ker} f = (\mathrm{Im} f^t)^\perp$, $\mathrm{Im} f = (\mathrm{Ker} f^t)^\perp$*;*

2) $\mathrm{Ker} f^t = (\mathrm{Im} f)^\perp$, $\mathrm{Im} f^t = (\mathrm{Ker} f)^\perp$*;*

3) f and f^t have the same rank, the dimension of their ranges are equal.

Properties of the symmetric matrices Let M be a $n \times n$ symmetric matrix, that is, M is equal to its transpose, $M^t = M$. M is the matrix of a linear mapping f from $\mathbb{R}^n$ to $\mathbb{R}^n$ defined by the matrix-vector product $f(x) = Mx$. This linear mapping satisfies $y \cdot f(x) = y \cdot Mx = M^t y \cdot x = My \cdot x = f(y) \cdot x$ and it is called a symmetric linear mapping.

We recall the fundamental spectral theorem on the symmetric matrices. An orthonormal basis of $\mathbb{R}^n$ is a basis $\mathcal{B} = (u_1, \ldots, u_n)$ such that $u_i \cdot u_j = 0$ for all (i, j) with $i \neq j$ and $\|u_i\| = 1$ for all i.

Theorem 2 *Let f be a symmetric linear mapping on $\mathbb{R}^n$ and M its symmetric matrix in the canonical basis. Then it exists an orthonormal basis $\mathcal{B} = (u_1, \ldots, u_n)$ such that for all i, there exists a real number λ_i such that $f(u_i) = \lambda_i u_i$. In other words, the matrix of f in the basis $\mathcal{B}$ is diagonal and λ_i is the term on the diagonal and on the ith row. Equivalently, we can say that there exists a $n \times n$ matrix P such that $P^{-1} = P^t$ and $P^{-1}MP$ is a diagonal matrix.*

Definition 6 Let M be a $n \times n$ symmetric matrix. Then M is

positive definite if all its eigenvalues are positive;

positive semi-definite if all its eigenvalues are non negative;

negative definite if all its eigenvalues are negative;

negative semi-definite if all its eigenvalues are non positive;

Proposition 9 *Let M be a $n \times n$ symmetric matrix. If M is positive definite (resp. negative definite), then M is invertible and its inverse is positive definite (resp. negative definite).*

From a symmetric $n \times n$ matrix M, we define a quadratic form q from $\mathbb{R}^n$ to $\mathbb{R}$ and a bilinear symmetric form φ from $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}$ as follows:
$q(x) = x \cdot Mx$;
$\varphi(x, y) = x \cdot My$.

We note that $q(x) = \varphi(x, x)$ and $\forall (x, y, z) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$, $\forall t \in \mathbb{R}$,
- $\varphi(x, y) = \varphi(y, x)$
- $\varphi(x + z, y) = \varphi(x, y) + \varphi(z, y)$
- $\varphi(x, y + z) = \varphi(x, y) + \varphi(x, z)$
- $\varphi(tx, y) = \varphi(x, ty) = t\varphi(x, y)$
- $q(tx) = t^2 q(x)$
- $q(x + y) = q(x) + q(y) + 2\varphi(x, y)$
- $\varphi(x, y) = \frac{1}{4}(q(x + y) - q(x - y))$

Remark 3 M is positive definite (resp. positive semi-definite, negative semi-definite, negativedefinite) if and only if $q(x) > 0$ (resp. ≥ 0, ≤ 0, < 0) for all $x \neq 0$. More precisely, if $\underline{\lambda}$ is the smallest eigenvalue of M and $\overline{\lambda}$ the largest eigenvalue of M, then

$$\underline{\lambda}\|x\|^2 \leq q(x) \leq \overline{\lambda}\|x\|^2$$

Exercise 2 Let M be a $p \times n$ matrix. Let P be the $p \times p$ matrix defined by $P = MM^t$.
1) Show that P is a symmetric positive semi-definite matrix.
2) Show that if the rank of M is equal to p, then P is positive definite.

Let N be a $n \times n$ symmetric positive definite matrix. Same questions with $Q = MNM^t$.

Criterion for a positive definite symmetric matrix

If M is a 2×2 symmetric matrix. M is positive definite if both trace and the determinant are positive.

If M is a $n \times n$ symmetric matrix. We denote M^p the $p \times p$ submatrix containing the first p columns and the first p rows of the matrix M. M is postive definite if the determinant of the matrices M^p with $p = 1, \ldots, n$ are positive.

$$M = \begin{pmatrix} m_{11} & m_{12} & \ldots & m_{1n} \\ m_{21} & m_{22} & \ldots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \ldots & m_{nn} \end{pmatrix} M^p = \begin{pmatrix} m_{11} & m_{12} & \ldots & m_{1p} \\ m_{21} & m_{22} & \ldots & m_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ m_{p1} & m_{p2} & \ldots & m_{pp} \end{pmatrix}$$

Exercise 3 Let $a \in \mathbb{R}$ and q be a quadratic function define on $\mathbb{R}^3$ as:

$$q(x, y, z) = x^2 + (1 + a)y^2 + (1 + a + a^2)z^2 + 2xy - 2ayz$$

1) Compute the bilinear form φ associated to q.
2) Give the matrix of q in the canonical basis of $\mathbb{R}^3$.
3) For which values of a, φ is positive definite?

Exercise 4 Let q be the quadratic form defined by its matrix in the canonical basis:

$$M = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

1) Compute the bilinear form φ associated to q.
2) Show that φ is positive definite?

1.2.3 Differentiable Functions

Definition 7 A function $f : U \to \mathbb{R}$, where U is an open set of $\mathbb{R}^n$, is differentiable at a point x if there exists a vector $g \in \mathbb{R}^n$ and a mapping ϵ defined on an open set containing 0 such that $f(x + h) = f(x) + g \cdot h + \|h\|\epsilon(h)$ with $\lim_{h\to 0} \epsilon(h) = 0$.

Proposition 10 *Let f be a function $U \to \mathbb{R}$, where U is an open set of $\mathbb{R}^n$. If f is differentiable at a point x, then it is continuous at x.*

Theorem 3 *Let f be differentiable at a point x and let g be a vector such that $f(x+h) = f(x) + g \cdot h + \|h\|\epsilon(h)$ with $\lim_{h \to 0}\ \epsilon(h) = 0$. Then all partial derivatives of f at x exist, and $g = \nabla f(x)$.*
Conversely, assume that all partial derivatives of f exist in some open set containing x and are continuous functions. Then f is differentiable at x.

Remark 4 Note that a function may have partial derivatives everywhere and not being differentiable. For example:

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

You can check that $\frac{\partial f}{\partial x_1}(x,y)$ and $\frac{\partial f}{\partial x_2}(x,y)$ are defined for all (x,y), included at $(0,0)$, but f is not differentiable at $(0,0)$. It is not even continuous at the origin.

Definition 8 A function f from U an open set of $\mathbb{R}^n$ to $\mathbb{R}$ is differentiable on U if it is differentiable at every point of U. It is continuously differentiable on U if all partial derivatives are continuous on U.

Proposition 11 *Let f and g be two differentiable functions from U an open set of $\mathbb{R}^n$ to $\mathbb{R}$. Then, for all $x \in U$, $\nabla(f+g)(x) = \nabla f(x) + \nabla g(x)$, $\nabla(fg)(x) = f(x)\nabla g(x) + g(x)\nabla f(x)$ and for all $c \in \mathbb{R}$, $\nabla(cf)(x) = c\nabla f(x)$.*

Remark 5 Suppose f is defined on an open set U, and let $\varphi : [a,b] \to U$ be a differentiable curve. Then we may form the composite function $f \circ \varphi$ given by $(f \circ \varphi)(t) = f(\varphi(t))$. We may think of φ as parametrization of a curve, or we may think of $\varphi(t)$ as representing the position on a curve at time t. If $f(x)$ represents, say, the value of the basket of commodities x, then $f(\varphi(t))$ is the value of the commodities at time t of $x = \varphi(t)$. The rate of change of the value along the curve is then given by the derivative $\frac{\partial f(\varphi(t))}{\partial t}$. The chain rule which follows gives an expression for this derivative in terms of the gradient, and generalises the usual chain rule to n variables.

Theorem 4 *Let $\varphi : J \to U$ be a differentiable function defined on some interval, and with values in an open set U of $\mathbb{R}^n$. Let $f : U \to \mathbb{R}$ be a differentiable function. Then $f \circ \varphi : J \to \mathbb{R}$ is differentiable, and $(f \circ \varphi)'(t) = \nabla f(\varphi(t)) \cdot \varphi'(t)$.*

1.2.4 Geometric properties of the gradient

From the chain rule, we deduce a geometric interpretation for the gradient.

Definition 9 Let x be a point of U and let v be a fixed vector. We define the directional derivative of f at x in the direction of v to be $f'(x,v) = \lim_{t \to 0, t \neq 0}\ \frac{1}{t}(f(x+tv) - f(x))$.

Remark 6 This means that if we let $g(t) = f(x + tv)$ then $f'(x, v) = g'(0)$. By the chain rule, $g'(t) = \nabla f(x + tv) \cdot v$ whence $f'(x, v) = \nabla f(x) \cdot v$.

From this formula we obtain an interpretation for the gradient. We use the standard expression for the dot product, namely $f'(x, v) = \|f(x)\|\|v\|cos(\theta)$ where θ is the angle between v and $\nabla f(x)$. Depending on the direction of the vector v, the number $cos(\theta)$ ranges from -1 to 1. The maximal value occurs when v has the same direction as $\nabla f(x)$, in which case for such unit vector v we obtain $f'(x, v) = \|\nabla f(x)\|\|v\|$. Therefore we get an interpretation for the direction and norm of the gradient:
The direction of $\nabla f(x)$ is the direction of maximal increase of the function f at x. The norm $\|\nabla f(x)\|$ is equal to the rate of change of f in its normalized direction of maximal increase.

Example: Find the directional derivative of the function $f(x, y) = x^2y^3$ at $(1, -2)$ for $v = \frac{1}{\sqrt{10}}(3, 1)$.
We have $\nabla f(x, y) = (2xy^3, 3x^2y^2)$ and $\nabla f(1, -2) = (-16, 12)$. Hence the desired directional derivative is $f'((1, -2), v) = (-16, 12).\frac{1}{\sqrt{10}}(3, 1) = \frac{1}{\sqrt{10}}(-36)$.

1.2.5 Tangent plane to a surface

Consider the set of all $x \in U$ such that $f(x) = 0$; or given a number c, the set of all $x \in U$ such that $f(x) = c$. This set, denoted by S_c, is called the level hypersurface at c. Let $x \in S_c$ and assume again that $\nabla f(x) \neq 0$.

It will be shown later as a consequence of the implicit function theorem that given any direction v perpendicular to the gradient, there exists a differentiable curve $\alpha : J \to U$ defined on some interval J containing 0 such that $\alpha(0) = x$ and $\alpha'(0) = v$ and $f(\alpha(t)) = c$ for all $t \in J$. In other words, the curve is contained in the level hypersurface. Conversely, we see from the chain rule that if we have a curve α lying in the hypersurface such that $\alpha(0) = x$, then $0 = \frac{\partial f}{\partial t}(\alpha(t)) = \nabla f(\alpha(t)).\alpha'(t)$.
In particular, for $t = 0$, $0 = \nabla f(\alpha(0)).\alpha'(0) = \nabla f(x).\alpha'(0)$. Hence the velocity vector $\alpha'(0)$ of the curve at $t = 0$ is perpendicular to $\nabla f(x)$. From this result, we make the geometric conclusion that $\nabla f(x)$ is perpendicular to the level hypersurface at x.

So we get the formal definition of the tangent plane to the level surface as follows:

Definition 10 Let f be a differentiable mapping for U, an open subset of $\mathbb{R}^n$, to $\mathbb{R}$. Let $x \in U$ such that $\nabla f(x) \neq 0$. Let $c = f(x)$. The set $S_c = \{x' \in U \mid f(x') = c\}$ is the level surface of f at the level c. The tangent hyperplane of S_c at x denoted $T_{S_c}(x)$ is defined by:

$$T_{S_c}(x) = \{u \in \mathbb{R}^n \mid u \cdot \nabla f(x) = 0\}$$

or, in other words, $T_{S_c}(x)$ is the orthogonal space to $\nabla f(x)$.

Note that we often consider a translation of the tangent plan which contains the point x and which is defined as $\{u \in \mathbb{R}^n \mid u \cdot \nabla f(x) = x \cdot \nabla f(x)\}$. Sometimes, there is a confusion between the two plans.

Example: Let $f(x, y, z) = x^2 + y^2 + z^2$. The surface S of points $X = (x, y, z)$ such that $f(X) = 4$ is the sphere of radius 2 centered at the origin. Let $P = (1, 1, \sqrt{2})$. We have $\nabla f(x, y, z) = (2x, 2y, 2z)$ and so $\nabla f(P) = (2, 2, 2\sqrt{2})$. Hence the tangent plane at P is given by the equation $2x + 2y + 2\sqrt{2}z = 0$.

Exercise 5 Let f be a differentiable function on $\mathbb{R}^n \setminus \{0\}$, depending only on the distance from the origin, that is, there exists a differentiable function g on $\mathbb{R}_{++}$ such that $f(x) = g(\|x\|)$ where $\|x\|$ is the Euclidean norm. Show that $\nabla f(x) = \dfrac{g'(\|x\|)}{\|x\|} x$.

1.2.6 Taylor Formula

By applying the result on the directional derivatives to the first order partial derivatives, we obtain the following result:

Proposition 12 *Let f be a $\mathcal{C}^2$ function from U, an open subset of $\mathbb{R}^n$, to $\mathbb{R}$. Let $\bar{x} \in U$ and $u \in \mathbb{R}^n$. Let φ be the function from the open interval I containing 0 in $\mathbb{R}$ defined by $\varphi(t) = f(\bar{x} + tu)$. then,*

$$\varphi''(t) = u^t H_f(\bar{x} + tu)u = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x} + tu) u_i u_j$$

Using the Taylor-Lagrange development of φ, we obtain the following result for f:

Proposition 13 *Let f be a $\mathcal{C}^2$ function from U, an open subset of $\mathbb{R}^n$, to $\mathbb{R}$. Let x and x' be two elements of U such that the segment $[x, x'] \subset U$. Then, it exists $\xi \in]x, x'[$ such that:*

$$f(x') = f(x) + Df(x)(x' - x) + \frac{1}{2}(x' - x)^t H_f(\xi)(x' - x)$$

Using the continuity of the second order partial derivatives, we obtain the following Taylor development:

Proposition 14 *Let f be a $\mathcal{C}^2$ function from U, an open subset of $\mathbb{R}^n$, to $\mathbb{R}$. Then, it exists a continuous function η from $U \times U$ to $\mathbb{R}$ such that:*

$$f(x') = f(x) + Df(x)(x' - x) + \frac{1}{2}(x' - x)^t H_f(x)(x' - x) + \|x' - x\|^2 \eta(x', x)$$

and $\eta(x, x) = 0$ for all $x \in U$.

Note that the continuity of η implies that for all $x \in U$, $\lim_{x' \to x} \eta(x', x) = 0$.

1.2.7 Euler's formula

Definition 11 For any real number k, a real-valued function f defined on a cone K[1] of $\mathbb{R}^n$ is homogeneous of degree k if $f(tx) = t^k f(x_1, \ldots, x_n)$ for all $x \in K$ and all $t > 0$.

Theorem 5 *Let $f(x)$ be a C^1 function on an open cone K of $\mathbb{R}^n$. If f is homogeneous of degree k, its first order partial derivatives are homogeneous of degree $k-1$.*

Theorem 6 (Euler's formula) *Let f be a C^1 homogeneous function of degree k on an open cone K of $\mathbb{R}^n$. Then for all x,*

$$x_1 \frac{\partial f}{\partial x_1}(x) + x_2 \frac{\partial f}{\partial x_2}(x) + \ldots + x_n \frac{\partial f}{\partial x_n}(x) = kf(x)$$

or using the gradient

$$x \cdot \nabla f(x) = kf(x)$$

1.3 Derivative of $f : \mathbb{R}^n \to \mathbb{R}^p$

1.3.1 The Frechet derivative as a linear map

Definition 12 Let U be an open subset of $\mathbb{R}^n$ and let $x \in U$. Let f be a mapping from U to $\mathbb{R}^p$. We shall say that f is (Frechet)-differentiable at x if there exists a continuous linear map $\varphi : \mathbb{R}^n \to \mathbb{R}^p$ and a map η defined for all sufficiently small $h \in \mathbb{R}^n$, with values in $\mathbb{R}^p$, such that $\lim_{h \to 0} \ \eta(h) = 0$ and $f(x+h) = f(x) + \varphi(h) + \|h\|\eta(h)$.

Remark 7 Setting $h = 0$ shows that we may assume that η is defined at 0 and that $\eta(0) = 0$. The preceding formula still holds.
We view the definition of the derivative as stating that near x, the values of f can be approximated by a affine map $f(x) + \varphi(x')$ with an error term described by the limit property of η at 0.

Theorem 7 *If f is (Frechet)-differentiable at x, then f is continuous at x.*

Definition 13 If f is (Frechet)-differentiable at every point x of U, then we say that f is (Frechet)-differentiable on U. In that case, the derivative Df is a mapping from U to the space of continuous linear mappings $\mathcal{L}(\mathbb{R}^n, \ \mathbb{R}^p)$, and thus to each $x \in U$, we have associated the linear map $Df(x) \in L(\mathbb{R}^n, \mathbb{R}^p)$.

We shall now see systematically how the definition of the derivative as a linear map actually includes the cases which we have been studied previously. We have three cases:

[1] For all $x \in K$ and for all $t > 0$, $tx \in K$.

We consider a map $f : J \to \mathbb{R}$ from an open interval J into $\mathbb{R}$. Then $Df(x)$ is the linear mapping from $\mathbb{R}$ to $\mathbb{R}$ define by $Df(x)(t) = f'(x)t$.
Let U be an open subset of $\mathbb{R}^n$ and let $f : U \to \mathbb{R}$ be a mapping, differentiable at a point $x \in U$. Then $Df(x)$ is the linear mapping from $\mathbb{R}^n$ to $\mathbb{R}$ define by $Df(x)(u) = \nabla f(x) \cdot u$.
Let J be an interval in $\mathbb{R}$, and let $f : J \to \mathbb{R}^p$ be a mapping. Then $Df(x)$ is the linear mapping from $\mathbb{R}$ to $\mathbb{R}^p$ define by $Df(x)(t) = tf'(x)$.

Theorem 8 (Maps with coordinates) *Let U be open in $\mathbb{R}^n$, let f be a mapping from U to $\mathbb{R}^{p_1} \times \ldots \times \mathbb{R}^{p_k}$. Let $(f_1, \ldots, f_k)$ be the coordinate mappings from U to $\mathbb{R}^{p_j}$, that is $f(x) = (f_1(x), \ldots, f_k(x))$. Then f is (Frechet)-differentiable at x if and only if each f_j is differentiable at x, and if this is the case, then $Df(x) = (Df_1(x), \ldots, Df_k(x))$.*

Theorem 9 *Let $\psi : \mathbb{R}^n \to \mathbb{R}^p$ be a linear mapping. Then ψ is (Frechet)-differentiable at every point of $\mathbb{R}^n$ and $D\psi(x) = \psi$ for every $x \in \mathbb{R}^n$.*

Let ϕ from $\mathbb{R}^n \times \mathbb{R}^p$ to $\mathbb{R}^k$ be a bilinear mapping, that is the partial mapping $\phi(x, \cdot)$ is linear for all x in $\mathbb{R}^n$ and $\phi(\cdot, y)$ is linear for all y in $\mathbb{R}^p$. Then ϕ is (Frechet)-differentiable at every point of $\mathbb{R}^n \times \mathbb{R}^p$ and for all $(u, v) \in \mathbb{R}^n \times \mathbb{R}^p$, $D\phi(x)(u, v) = \phi(u, y) + \phi(x, v)$ for every $(x, y) \in \mathbb{R}^n \times \mathbb{R}^p$.

1.3.2 The Jacobian matrix of a differentiable map

Theorem 10 *Let U be an open set of $\mathbb{R}^n$, and let $f : U \to \mathbb{R}^p$ be a mapping which is (Frechet)-differentiable at x. Then the continuous linear map $Df(x)$ is represented by the Jacobian matrix*

$$J_f(x) = \left(\frac{\partial f_i}{\partial x_j}(x)\right) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \ldots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & \vdots & \vdots \\ \frac{\partial f_p}{\partial x_1}(x) & \ldots & \frac{\partial f_p}{\partial x_n}(x) \end{pmatrix}$$

where f_i is the i-th coordinate function of f.

We see that if f is (Frechet)-differentiable at every point of U, then $x \mapsto J_f(x)$ is a mapping from U into the space of $p \times n$ matrices, which is a space of dimension pn.

Definition 14 We shall say that f is of class $\mathcal{C}^1$ on U, or is a $\mathcal{C}^1$ mapping, if f is (Frechet)-differentiable on U and if in addition the derivative $Df : U \to L(\mathbb{R}^n, \mathbb{R}^p)$ is continuous, which is equivalent to assume that pn partial derivatives $\frac{\partial f_i}{\partial x_j}$ are continuous.

Theorem 11 *Let U be an open set of $\mathbb{R}^n$, and let $f : U \to \mathbb{R}^p$ be a mapping. If the pn partial derivatives $\frac{\partial f_i}{\partial x_j}$ are defined on U and are continuous, then f is $\mathcal{C}^1$ on U.*

1.3.3 Basic Properties of the Derivative

Proposition 15 *Let U be open in $\mathbb{R}^n$. Let $f, g : U \to \mathbb{R}^p$ be two mappings which are (Frechet)-differentiable at $x \in U$. Then $f + g$ is (Frechet)-differentiable at x and $D(f + g)(x) = Df(x) + Dg(x)$. If c is real number, then cf is (Frechet)-differentiable at x and $D(cf)(x) = cDf(x)$.*

We recall that a bilinear mapping ψ from $\mathbb{R}^n \times \mathbb{R}^p$ to $\mathbb{R}^k$ is a mapping such that the partial mapping $\psi(x, \cdot)$ is linear for all x in $\mathbb{R}^n$ and $\psi(\cdot, y)$ is linear for all y in $\mathbb{R}^p$.

Proposition 16 *Let ψ be a bilinear mapping ψ from $\mathbb{R}^n \times \mathbb{R}^p$ to $\mathbb{R}^k$. Let U be an open subset of $\mathbb{R}^\ell$ and let $f : U \to \mathbb{R}^n$ and $g : U \to \mathbb{R}^p$ be two (Frechet)-differentiable mappings at $x \in U$. Then $\psi(f, g)$ is differentiable at x and for all $v \in \mathbb{R}^\ell$,*

$$D\psi(f, g)(x)(v) = \psi(Df(x)(v), g(x)) + \psi(f(x), Dg(x)(v))$$

Remark 8 If ψ is the inner product on $\mathbb{R}^n \times \mathbb{R}^n$, we have the following formula for all $v \in \mathbb{R}^\ell$,

$$D(f \cdot g)(x)(v) = Df(x)(v) \cdot g(x) + f(x) \cdot Dg(x)(v)$$

or

$$\nabla(f \cdot g)(x) = Df(x)^t(g(x)) + Dg(x)^t(f(x))$$

where $Df(x)^t$ is the transpose of the linear mapping $Df(x)$ and $Dg(x)^t$ is the transpose of the linear mapping $Dg(x)$.

Example: Let J be an open interval in $\mathbb{R}$ and let $t \mapsto A(t) = (a_{ij}(t))$ and $t \mapsto X(t)$ be two differentiable maps from J into the space of $p \times n$ matrices, and into $\mathbb{R}^n$ respectively. Thus for each t, $A(t)$ is an $p \times n$ matrix, and $X(t)$ is a column vector of dimension n. We can form the product $A(t)X(t)$, and thus the product map $t \mapsto A(t)X(t)$, which is differentiable. Our rule in this special case asserts that $\frac{\partial}{\partial t}A(t)X(t) = A'(t)X(t) + A(t)X'(t)$ where differentiation with respect to t is taken componentwise both on the matrix $A(t)$ and the vector $X(t)$. The product here is the product of a matrix and a vector.

1.3.4 Chain Rule

Proposition 17 (Chain Rule) *Let U be an open subset of $\mathbb{R}^n$ and let V be an open subset of $\mathbb{R}^p$. Let $f : U \to V$ and $g : V \to \mathbb{R}^k$ be two mappings. Let $x \in U$. Assume that f is (Frechet)-differentiable at x and g is (Frechet)-differentiable at $f(x)$. Then $g \circ f$ is (Frechet)-differentiable at x and $D(g \circ f)(x) = Dg(f(x)) \circ Df(x)$.*

Remark 9 $Df(x) : \mathbb{R}^n \to \mathbb{R}^p$ is a linear map, and $Dg(f(x)) : \mathbb{R}^p \to \mathbb{R}^s$ is a linear map, and so these linear maps can be composed, and the composite is a linear map, which is continuous because both $Dg(f(x))$ and $Df(x)$ are continuous. The composed linear map goes from $\mathbb{R}^n$ into $\mathbb{R}^s$, as it should.

Remark 10 In terms of the Jacobian matrix we have: $J_{g\circ f}(x) = J_g(f(x))J_f(x)$, the multiplication being that of matrices.

Corollary 1 *Let U be an open subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}$ be a (Frechet)-differentiable mapping. Let J be an open interval of $\mathbb{R}$ and let $\varphi : J \to \mathbb{R}^n$ be a differentiable mapping such that $\varphi(t) \in U$ for $t \in J$. Then $f \circ \varphi$ is differentiable on J and $(f \circ \varphi)'(t) = \sum_{i=1}^n \varphi_i'(t)\frac{\partial f}{\partial x_i}(\varphi(t))$ for all $t \in J$.*
In particular if $\varphi(t) = \bar{x} + tu$ for some $\bar{x} \in U$ and $u \in \mathbb{R}^n$, we get $(f \circ \varphi)'(t) = \sum_{i=1}^n u_i \frac{\partial f}{\partial x_i}(\varphi(t)) = \nabla f(\bar{x} + tu) \cdot u$ for all $t \in J$.

1.3.5 The Mean Value Theorem

Theorem 12 *Let U be an open subset of $\mathbb{R}^n$ and f be a differentiable mapping from U to $\mathbb{R}$. Let $\underline{x}$ and $\bar{x}$ two elements of U such that the segment $[\underline{x}, \bar{x}] = \{(1-t)\underline{x} + t\bar{x} \mid t \in [0,1]\} \subset U$. Then, it exists $\xi \in]\underline{x}, \bar{x}[$ such that $f(\bar{x}) - f(\underline{x}) = Df(\xi)(\bar{x} - \underline{x})$.*

Remark 11 This theorem cannot be generalised to a differentiable mapping f taken its value in $\mathbb{R}^p$ with $p > 1$. Indeed, let f from $\mathbb{R}$ to $\mathbb{R}^2$ defined by $f(t) = (\cos t, \sin t)$. f is $\mathcal{C}^1$ on $\mathbb{R}$. We remark that $f(0) = f(2\pi)$. It does not exists $t \in]0, 2\pi[$ such that $f(2\pi) - f(0) = Df(t)(2\pi)$. Indeed, $Df(t) \neq 0$ for all $t \in \mathbb{R}$.

Nevertheless, we can obtain an upper bound of the norm $f(\bar{x}) - f(\underline{x})$ by using the norm as a linear mapping of $Df(\xi)$ for $\xi \in [\underline{x}, \bar{x}]$.

Theorem 13 *Let U be an open subset of $\mathbb{R}^n$ and f be a differentiable mapping from U to $\mathbb{R}^p$. Let $\underline{x}$ and $\bar{x}$ two elements of U such that the segment $[\underline{x}, \bar{x}] = \{(1-t)\underline{x} + t\bar{x} \mid t \in [0,1]\}$ is included in U. Then, it exists $\xi \in]\underline{x}, \bar{x}[$ such that $\|f(\bar{x}) - f(\underline{x})\|_p \leq \|Df(\xi)\|_{\mathcal{L}} \|\bar{x} - \underline{x}\|_n$.*

Corollary 2 *Let U be an open subset of $\mathbb{R}^n$ and f be a differentiable mapping from U to $\mathbb{R}^p$. Let $\underline{x}$ and $\bar{x}$ two elements of U such that the segment $[\underline{x}, \bar{x}] = \{(1-t)\underline{x} + t\bar{x} \mid t \in [0,1]\}$ is included in U. Then, $\|f(\bar{x}) - f(\underline{x})\|_p \leq \max\{\|Df(\xi)\|_{\mathcal{L}} \mid \xi \in [\underline{x}, \bar{x}]\} \|\bar{x} - \underline{x}\|_n$.*

A first consequence of this theorem is the fact that a differentiable mapping f such that $Df(x)$ is the nul linear mapping for every x is locally constant.

Corollary 3 *Let U be an open subset of $\mathbb{R}^n$ and f be a differentiable mapping from U to $\mathbb{R}^p$. If $Df(x) = 0_{\mathcal{L}}$ for all $x \in U$, the, for all $\bar{x} \in U$, f is constant on the ball $B(\bar{x}, r)$ such that $B(\bar{x}, r) \subset U$.*

Another consequence is the fact that a $\mathcal{C}^1$ mapping is locally Lipschitz continuous.

Corollary 4 *Let U be an open subset of $\mathbb{R}^n$ and f be a continuously differentiable mapping from U to $\mathbb{R}^p$. Let $\bar{x} \in U$ and $r > 0$ such that the closed ball $\bar{B}(\bar{x}, r)$ is included in U. Then it exists $k \geq 0$ such that for all $(x, x') \in B(\bar{x}, r)^2$, $\|f(x') - f(x)\|_p \leq k\|x' - x\|_n$.*

Exercise 6 Compute the partial derivatives of the following mappings

1) $f(x, y) = x(2\ln(x+1) + y + 1) + e^{-y} + 2\ln(x+1) + y$;

2) $f(x, y, z) = x^\alpha y^\beta z^\gamma$; $\alpha > 0$, $\beta > 0$, $\gamma > 0$;

3) $f(x, y, z) = \sqrt{\alpha x + \beta y + \gamma z}$, $\alpha > 0$, $\beta > 0$, $\gamma > 0$;

4) $f(x, y, z) = y(x + x^{\frac{1}{2}} z^{\frac{1}{2}} + z)$;

5) $f(x, y, z) = (\alpha x^\rho + \beta y^\rho + \gamma z^\rho)^{\frac{1}{\rho}}$, $\alpha > 0$, $\beta > 0$, $\gamma > 0$ $\rho > 0$;

6) $f(x, y, z) = \frac{xyz}{x+y+z}$;

7) $f(x, y, z) = e^{\alpha x} e^{\beta y} e^{\gamma z}$;

8) $f(x, y, z) = \frac{z}{(\sqrt{x}+\sqrt{y})^2}$;

9) $f(x, y, z) = \ln(z) - \alpha\ln(x) - \beta\ln(y)$;

10) $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$;

11) $(x, y) \in \mathbb{R}^2 \mapsto f(x, y) = x^2 + (x + y - 1)^2 + y^2$;

12) $(x, y) \in \mathbb{R}^2 \mapsto f(x, y) = (x + y)^2 + x^4 + y^4$;

13) $(x, y) \in \mathbb{R}^2 \mapsto f(x, y) = 2x^4 - 3x^2 y + y^2$;

14) $f(x, y) = x^2 - xy + y^2 + x + y$, where $X = \{(x, y) \in \mathbb{R}^2, x \leq 0, y \leq 0, x + y \geq -3\}$;

15) $f(x, y) = x^2(1 + y)^3 + y^4$;

16) $f(x, y) = x^2 - y^2 + y^4/4$;

17) $f(x, y) = x^3 - 3x(1 + y^2)$;

18) $f(x, y) = \begin{cases} \frac{xy(1-x)(1-y)}{1-xy} & \text{if} \quad (x, y) \neq (1, 1) \\ 0 \quad \text{if} \quad (x, y) = (1, 1) \end{cases}$

Exercise 7 Let N be a norm on $\mathbb{R}^n$. Show that N is not differentiable at 0.

Exercise 8 Let f be a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}$. Show that f is differentiable on $\mathbb{R}^n$ and $Df(x) = f$ for all $x \in \mathbb{R}^n$.

Exercise 9 Let M be a $n \times p$ matrix. Let f be the mapping from $\mathbb{R}^n \times \mathbb{R}^p$ to $\mathbb{R}$ defined by:

$$f(x, y) = x \cdot My$$

1) Show that for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^p$, $f(x, y) \leq \|M\|_{\mathcal{L}} \|x\| \|y\|$.
2) Using the definition of the derivative show that f is differentiable on $\mathbb{R}^n \times \mathbb{R}^p$ and that the derivative is defined by :

$$Df(x, y)(h, k) = h \cdot My + x \cdot Mk$$

3) Deduce the derivative of the standard inner product on $\mathbb{R}^n$ as a mapping from $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}$.

Exercise 10 Let A be a $n \times n$ matrix, b, a vector in $\mathbb{R}^n$ and c a real number. Let f be the mapping from $\mathbb{R}^n$ to $\mathbb{R}$ defined by:

$$f(x) = x \cdot Ax + b \cdot x + c$$

1) Compute the partial derivatives of f on $\mathbb{R}^n$.
2) Show that the derivatives are continuous.
3) Provide the formula for the derivative of f at each point $\bar{x}$ of $\mathbb{R}^n$.

Exercise 11 Let f be the mapping from $\mathbb{R}^n$ to $\mathbb{R}$ defined by:

$$f(x) = \|x\|^2 = \sum_{i=1}^{n} x_i^2$$

1) Compute the partial derivatives of f at each point $\bar{x}$.
2) Show that f is differentiable at each point $\bar{x} \in \mathbb{R}^n$ and show that $Df(\bar{x})$ is defined by $Df(\bar{x})(h) = 2\bar{x} \cdot h$.

Exercise 12 Let f be a mapping from $\mathbb{R}^n$ to $\mathbb{R}$. We assume that there exists $c \in \mathbb{R}_+$ and $\alpha > 0$ such that for all $(x, y) \in (\mathbb{R}^n)^2$,

$$|f(y) - f(x)| \leq c\|y - x\|^{1+\alpha}$$

1) Show that the partial derivatives of f at each point of $\mathbb{R}^n$ are vanishing.
2) Deduce that f is constant.

Exercise 13 Let f be a differentiable mapping from $\mathbb{R}^3$ to $\mathbb{R}$. We assume that for all $(x, y, z) \in \mathbb{R}^3$, the three partial derivatives of f at (x, y, z) are non negative. Show that if (x', y', z') satisfies $x' \geq x$, $y' \geq y$ and $z' \geq z$, then $f(x', y', z') \geq f(x, y, z)$.

We now assume that for all $(x, y, z) \in \mathbb{R}^3$ the three partial derivatives of f at (x, y, z) are positive. Show that if (x', y', z') satisfies $x' \geq x$, $y' \geq y$ and $z' \geq z$ with one strict inequality among the three, then $f(x', y', z') > f(x, y, z)$.

Exercise 14 Let $\mathcal{M}_2$ be the space of dimension 4 of the 2×2 matrices. We consider the mapping "determinant" from $\mathcal{M}_2$ to $\mathbb{R}$.

For all $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $\det M = ad - bc$.

1) Compute the partial derivative of the mapping det.
2) Show that the mapping det is differentiable and give its derivative at $M \in \mathcal{M}_2$.
3) Show that $D\det(M) = 0_{\mathcal{L}}$ if and only if $M = 0$.

Exercise 15 Let f and g two differentiable mappings from $\mathbb{R}^n$ to $\mathbb{R}^p$. Let $\bar{x} \in \mathbb{R}^n$. We assume that $f(x) = g(x) + \|x - \bar{x}\|\varepsilon(x)$ where ε is a mapping from $\mathbb{R}^n$ to $\mathbb{R}^p$ satisfying $\lim_{x \to \bar{x}} \varepsilon(x) = 0_p$. Show that $f(\bar{x}) = g(\bar{x})$ and $Df(\bar{x}) = Dg(\bar{x})$.

Exercise 16 Let f be a differentiable mapping from an open subset U of $\mathbb{R}^n$ to $\mathbb{R}^p$. We assume that f is k Lipschitz continuous on U, i.e., $\exists k > 0$, $\forall x, y \in U^2$, $\|f(x) - f(y)\|_p \leq k\|x - y\|_n$. Show that for all x in U, $\|Df(x)\|_{\mathcal{L}} \leq k$.