Université de Bordeaux
Licence 2, Mathématiques Fondamentales

TMQ302U
Fonctions des plusieurs variables

S. Kupin

Notes du cours
Version 1

10 octobre 2018

Table des matières

Introduction

Ces notes contiennent la liste succincte des définitions et de résultats principaux du cours TMQ302U ”Fonctions de plusieurs variables” donné à l'Université de Bordeaux. Ce document est non-exhaustif et on se gardera de le considérer comme un ”support principal” du travail sur ce cours ; il vaudrait mieux le regarder comme un supplément au contenu du cours magistral et travaux dirigés.

Rappelons que les DS et DST de ce module se baseront sur le programme officiel du cours et le matériel présenté lors des séances de CM et de TD, en non pas seulement sur le contenu de ce polycopié.

Remerciements. Je remercie chaleureusement Christophe Bavard, Duc-Manh Nguyen, Pïerre Parent et Guillaume Ricotta pour avoir m'apporté leur aide précieux à la rédaction de ces notes.

Chapitre 1

Topologie de $\mathbb{R}^d$

1.1 Rappel : les espaces vectoriels $\mathbb{R}^d$ et $\mathbb{C}^d$

Dans cette sous-section, nous rappelons la notion d'espace vectoriel de dimension finie. Ce sujet a été traité en détail lors du cours d'algèbre linéaire de L1.

Définition 1.1.1. *L'ensemble* $\mathbb{R}^d, d \geq 1$, *est défini comme*

$$\mathbb{R}^d = \left\{ x = \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix} \ : \ x_i \in \mathbb{R}, \ i = 1, \dots, d \right\}.$$

On introduit les opérations suivantes sur $\mathbb{R}^d$:

1. L'addition, " $+$ " : $\mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d, \quad (x, y) \in \mathbb{R}^d \times \mathbb{R}^d \mapsto x + y \in \mathbb{R}^d$, ou, plus précisément :

$$x + y = \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix} + \begin{bmatrix} y_1 \\ \vdots \\ y_d \end{bmatrix} \overset{déf}{=} \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_d + y_d \end{bmatrix}, \qquad \forall x, y \in \mathbb{R}^d.$$

2. La multiplication par un scalaire (appartenant à $\mathbb{R}$), " $\cdot$ " : $\mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d, \quad (\lambda, x) \in \mathbb{R} \times \mathbb{R}^d \mapsto \lambda \cdot x \in \mathbb{R}^d$, ou bien

$$\lambda \cdot x = \lambda x = \lambda \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix} \overset{déf}{=} \begin{bmatrix} \lambda x_1 \\ \vdots \\ \lambda x_d \end{bmatrix}, \qquad \forall \lambda \in \mathbb{R}, x \in \mathbb{R}^d.$$

Il est facile de vérifier que ces opérations satisfont un certain nombre de propriétés. La première série de ces propriétés porte sur l'addition :

- (commutativité de l'addition) $\forall x, y \in \mathbb{R}^d : x + y = y + x$,
- (associativité de l'addition) $\forall x, y, z \in \mathbb{R}^d : (x + y) + z = x + (y + z)$,
- (existence de l'élément neutre) $x + 0_d = 0_d + x = x$. Ici,

$$0_d = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^d.$$

Pour garder des notations raisonnablement simples, on note 0_d par 0.

- (existence de l'élément inverse) $x + (-x) = (-x) + x = 0$.

La deuxième série porte sur l'opération de multiplication par les scalaires et la compatibilité avec l'addition :

- (associativité de multiplication) $\forall \alpha, \beta \in \mathbb{R},\ x \in \mathbb{R}^d : (\alpha\beta)x = \alpha(\beta x)$,
- (distributivité) $\forall \alpha, \beta \in \mathbb{R},\ x, y \in \mathbb{R}^d : (\alpha + \beta)x = \alpha x + \beta x$ et $\alpha(x + y) = \alpha x + \beta y$,
- (élément neutre par rapport à la multiplication) $\forall x \in \mathbb{R}^d : 1 \cdot x = x \cdot 1 = x$.

On dit que l'ensemble $\mathbb{R}^d$ avec les opérations " + " et " · " satisfaisant les relations données ci-dessus est *un espace vectoriel* ; on le note $(\mathbb{R}^d, " + ", " \cdot ")$ si on veut mettre l'accent sur sa structure (ou on écrit $\mathbb{R}^d$ pour plus de simplicité).

Par la dimension de l'espace vectoriel on entend le nombre de vecteurs de base de cet espace. Il est clair que le nombre de vecteurs de base ne change pas si on passe d'une base à une autre ; la dimension de l'espace est donc bien définie.

Quant à $\mathbb{R}^d$, nous allons souvent travailler avec la base standard $(e^i)_{i=1,\dots,d}$ de l'espace, *i.e.,*

$$e^1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix},\ e^2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix},\ \dots, e^d = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Il vient facilement que $\dim \mathbb{R}^d = d, d \geq 1,$ et

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix} = \sum_{i=1}^{d} x_i e^i.$$

La définition de l'espace vectoriel complexe $\mathbb{C}^d$ (et des opérations là-dessus) est complètement analogue, *i.e.,*

Définition 1.1.2. *L'ensemble* $\mathbb{C}^d, d \geq 1,$ *est défini comme*

$$\mathbb{C}^d = \left\{ z = \begin{bmatrix} z_1 \\ \vdots \\ z_d \end{bmatrix} \ :\ z_i \in \mathbb{R},\ i = 1, \dots, d \right\}.$$

Pour l'essentiel de ce cours, nous nous intéresserons aux propriétés de sous-ensembles et au calcul d'applications sur $\mathbb{R}^d$; nous ferons toutefois quelques remarques à propos de $\mathbb{C}^d$ par souci de complétude.

1.2 La structure euclidienne de $\mathbb{R}^d$: produit scalaire, norme et distance

Dans cette section, on introduit le produit scalaire sur $\mathbb{R}^d$. Comme nous allons voir, ce produit scalaire donne lieu à une norme, qui définit à son tour une distance sur $\mathbb{R}^d$. Ces objets définissent *la structure euclidienne* de l'espace (contrairement à d'autres normes, cf. la discussion ci-dessous).

Soient

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix},\ y = \begin{bmatrix} y_1 \\ \vdots \\ y_d \end{bmatrix} \in \mathbb{R}^d.$$

Définition 1.2.1. *Le produit scalaire* $(.,.) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ *est défini de la manière suivante :*

$$(x, y) = \sum_{i=1}^{d} x_i y_i.$$

On peut écrire aussi $(x, y) = y^t\, x$, ou bien

$$(x, y) = [y_1, \ldots, y_d] \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix}.$$

Ici, y^t est la transposée du vecteur y.
On peut voir que :

1. (positivité) $\forall x \in \mathbb{R}^d : (x, x) \geq 0$ et $(x, x) = 0$ ssi (= si et seulement si) $x = 0$, (PS1)
2. (symétrie) $\forall x, y \in \mathbb{R}^d : (x, y) = (y, x)$, (PS2)
3. (linéarité) $\forall \alpha, \beta \in \mathbb{R}, \forall x, y, z \in \mathbb{R}^d : (\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z)$. (PS3)

La linéarité du produit scalaire par rapport au deuxième argument vient des deux dernières propriétés.

Proposition 1.2.2 (inégalité de Cauchy-Schwarz)**.** *Soient* $x, y \in \mathbb{R}^d$*, alors*

$$|(x, y)| \leq \sqrt{(x, x)}\sqrt{(y, y)}.$$

Le cas d'égalité dans cette inégalité est possible ssi les vecteurs x, y *sont liés.*

La norme sur $\mathbb{R}^d$ est l'analogue de la fonction valeur absolue sur $\mathbb{R}$; elle se définit de la manière suivante.

Définition 1.2.3. *La norme euclidienne sur* $\mathbb{R}^d$ *est donnée par*

$$||x|| = ||x||_2 \stackrel{déf}{=} \sqrt{(x, x)} = \left(\sum_{i=1}^{d} x_i^2 \right)^{1/2}.$$

Pour simplifier l'écriture, nous allons souvent désigner cette norme par $||.||$, et non pas par $||.||_2$.
La norme possède les propriétés suivantes :

1. (positivité) $\forall x \in \mathbb{R}^d : ||x|| \geq 0$ et $||x|| = 0$ ssi $x = 0$, (N1),
2. (homogénéité) $\forall \lambda \in \mathbb{R}, \forall x \in \mathbb{R}^d : ||\lambda x|| = |\lambda| ||x||$, (N2),
3. (inégalité triangulaire) $\forall x, y \in \mathbb{R}^d : ||x + y|| \leq ||x|| + ||y||$. (N3)

L'inégalité triangulaire (N3) peut s'écrire sous une forme équivalente : $\forall x, y \in \mathbb{R}^d : |||x|| - ||y||| \leq ||x - y||$ (N3').
A son tour, la norme (euclidienne) donne lieu à la distance (euclidienne elle aussi). L'analogue de la distance sur $\mathbb{R}$ est la fonction $d(x, y) = |x - y|,\ x, y \in \mathbb{R}$.

Définition 1.2.4. *La distance euclidienne entre les points* $x, y \in \mathbb{R}^d$ *est*

$$d(x, y) = d_2(x, y) \stackrel{déf}{=} ||x - y|| = \left(\sum_{i=1}^{d} (x_i - y_i)^2 \right)^{1/2}.$$

Comme pour la norme, on omettra le sous-indice 2 pour simplifier l'écriture, *i.e.,* on écrira $d(x,y)$ au lieu de $d_2(x,y)$.

Les propriétés suivantes de la distance découlent immédiatement de celles de la norme (N1)-(N3), cf. p. 5 :

1. (la positivité) $\forall x, y \in \mathbb{R}^d : d(x,y) \geq 0$ et $x = y$ ssi $d(x,y) = 0$, (D1),
2. (la symétrie) $\forall x, y \in \mathbb{R}^d : d(x,y) = d(y,x)$, (D2)
3. (l'inégalité triangulaire) $\forall x, y, z \in \mathbb{R}^d : d(x,y) \leq d(x,z) + d(z,y)$, ou bien $|d(x,z) - d(z,y)| \leq d(x,y)$. (D3)

1.2.1 Autres normes sur l'espace $\mathbb{R}^d$

Plus généralement, une norme sur $\mathbb{R}^d$ est une application $||.|| : \mathbb{R}^d \to \mathbb{R}_+$ satisfaisant les propriétés (N1)-(N3), cf. p. 5. D'une manière similaire, une distance sur $\mathbb{R}^d$ est une application $d(.,.) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+$ satisfaisant (D1)-(D3), cf. p. 6.

On peut donner des normes et des distances différentes de la norme et de la distance euclidiennes, par exemple :

$$\begin{aligned} ||x||_p &= \left(\sum_{i=1}^{d} |x_i|^p \right)^{1/p}, \quad 1 \leq p < \infty, \\ ||x||_\infty &= \max_{i=1,\ldots,d} |x_i|, \quad p = \infty, \\ d_p(x,y) &= ||x - y||_p, \quad 1 \leq p \leq \infty \end{aligned}$$

où $x, y \in \mathbb{R}^d$. Notons que $||.||_2$ et $d_2(.,.)$ ($p = 2$) sont exactement la norme et la distance définies précedemment ; pour $p \neq 2$, la norme $||.||_p$ (et la distance $d_p(.,.)$) ne provient pas d'un produit scalaire.

Pour d'autres informations à ce sujet, voir la section 1.4.1.

1.2.2 La norme et la distance sur $\mathbb{C}^d$

Les applications admettant les propriétés (PS1)-(PS3) (une exception : il faut remplacer la propriété (PS2) par la propriété (PS2') (anti-symétrie) : $\forall x, y \in \mathbb{C}^d : (x,y) = \overline{(x,y)}$), (N1)-(N3), (D1)-(D3), cf. p. 5 définissent le produit scalaire, la norme et la distance sur $\mathbb{C}^d$, respectivement.

Ces objets sont donnés par les relations suivantes :

$$\begin{aligned} (x,y) &= \sum_{i=1}^{d} x_i \bar{y}_i, \\ ||x||_2 &= \left(\sum_{i=1}^{d} |x_i|^2 \right)^{1/2}, \\ d_2(x,y) &= \left(\sum_{i=1}^{d} |x_i - y_i|^2 \right)^{1/2}, \end{aligned}$$

où $x, y \in \mathbb{C}^d$. Plus généralement, pour $1 \leq p \leq \infty$,

$$\begin{aligned} ||x||_p &= \left(\sum_{i=1}^{d} |x_i|^p \right)^{1/p}, \quad 1 \leq p < \infty, \\ ||x||_\infty &= \max_{i=1,\ldots,d} |x_i|, \quad p = \infty, \\ d_p(x,y) &= ||x - y||_p, \quad 1 \leq p \leq \infty. \end{aligned}$$

1.3 Eléments de topologie : ouverts, fermés, intérieurs, adhérences, etc.

Dans la suite de ce chapitre, nous considérons l'espace vectoriel $\mathbb{R}^d$ équipé de la norme euclidienne.

Définition 1.3.1. *La boule ouverte centrée au point $x^0 \in \mathbb{R}^d$ de rayon $r \geq 0$ est donnée par*

$$B(x^0, r) = \{x \in \mathbb{R}^d : ||x - x^0|| = d(x, x^0) < r\}.$$

Notons que la définition coïncide avec la définition "intuitive" ; par ailleurs, $B(x^0, 0) = \emptyset$. La boule $B(x^0, r), r > 0$, est parfois appelée *un voisinage* (ou *un voisinage ouvert*) du point x^0.

On procède de la même manière pour *la boule fermée, i.e.,*

$$\bar{B}(x^0, r) = B_f(x^0, r) \stackrel{déf}{=} \{x \in \mathbb{R}^d : ||x - x^0|| = d(x, x^0) \leq r\}.$$

Les définitions suivantes sont centrales pour le premier chapitre du cours.

Définition 1.3.2. *L'ensemble $A \subset \mathbb{R}^d$ est borné, s'il existe un $R \geq 0$ tel que $A \subset B(0, R)$.*

Définition 1.3.3. *L'ensemble $O \subset \mathbb{R}^d$ est dit ouvert, si tout point $x \in O$ possède un voisinage ouvert dans O,* i.e.,

$$\forall x \in O \ \exists r = r(x) > 0 : \ B(x, r) \subset O.$$

Définition 1.3.4. *L'ensemble $F \subset \mathbb{R}^d$ est fermé, si F^c (*i.e., *le complémentaire de F) est ouvert.*

On constate qu'une boule ouverte $B(x, r)$ est un ensemble ouvert, et la boule fermée $\bar{B}(x, r) = B_f(x, r)$ est un ensemble fermé ; la terminologie introduite n'est donc pas contradictoire.

La famille d'ensembles ouverts sur $\mathbb{R}^d$ est notée τ_O (τ_F pour la famille d'ensembles fermés, respectivement) ; elles possèdent les propriétés suivantes :

1 les seuls ensembles ouverts et fermés en même temps sont $\emptyset$ et $\mathbb{R}^d$,

2 une réunion quelconque d'ouverts est ouverte, *i.e.,* pour toute famille $\{O_\alpha\}_{\alpha \in A}$, où tout O_α est ouvert, $\cup_{\alpha \in A} O_\alpha$ est ouverte,

2' (idem pour les fermés, avec la réunion remplacée par l'intérsection) une intersection quelconque de fermés est fermée, *i.e.,* pour toute famille $\{F_\beta\}_{\beta \in B}$, où tout F_β est fermé, $\cap_{\beta \in B} F_\beta$ est fermée,

3 une intersection finie d'ouverts est ouverte, *i.e.,* pour une famille $\{O_i\}_{i=1,\ldots,n}$, où tout O_i est ouvert, $\cap_{i=1}^n O_i$ est ouverte,

3' une réunion finie de fermés est fermée, *i.e.,* pour une famille $\{F_i\}_{i=1,\ldots,n}$, où tout F_i est fermé, $\cup_{\beta \in B} F_\beta$ est fermée.

Comme la famille d'ouverts τ_O satisfait les propriétés (1)-(3), on dit qu'elle définit *la topologie* sur $\mathbb{R}^d$ (idem pour la famille τ_F de fermés à propriétés (1), (2'), (3')).

On passe rapidement aux notions d'intérieur, d'adhérence et de frontière.

Définition 1.3.5. *Soit A un sous-ensemble de $\mathbb{R}^d$.*

1. *L'intérieur de A, noté A^o, est le plus grand ouvert contenu dans A,* i.e.,

$$A^o \stackrel{déf}{=} \bigcup_{\substack{O \text{ est ouvert,} \\ O \subset A}} O.$$

2. *L'adhérence de A, notée $\bar{A}$, est le plus petit fermé contenant A,* i.e.,

$$\bar{A} \overset{déf}{=} \bigcap_{\substack{F \text{ est fermé}, \\ A \subset F}} F.$$

3. *La frontière de A, notée ∂A (ou encore $Fr(A)$) est* $\partial A \overset{déf}{=} \bar{A} \backslash A^o.$

La liste des propriétés élémentaires suivantes en découle immédiatement :

- $A^o \subset A$, $A \subset \bar{A}$;
- A^o est toujours ouvert ; de plus $A = A^o$ ssi A est ouvert ;
- $\bar{A}$ est fermé ; en particulier, $A = \bar{A}$ ssi A est fermé, etc. etc.

D'autres caractérisations de l'intérieur, l'adhérence et de la frontière d'un ensemble sont données dans la section 1.4.

1.4 Notion de suite convergente et de limite

Considérons $(x^n)_n$ une suite d'éléments de $\mathbb{R}^d$ (notation : $(x^n)_n \subset \mathbb{R}^d$). Chaque élément x^n de la suite est un vecteur (ou un point) de $\mathbb{R}^d$, *i.e.,*

$$x^n = \begin{bmatrix} x_1^n \\ x_2^n \\ \vdots \\ x_d^n \end{bmatrix}.$$

On attire l'attention sur le fait que dans l'écriture $x_j^n, n = 0, \dots, \infty,\ j = 1, \dots, d$, l'indice supérieur n désigne le numéro du vecteur x^n dans la suite $(x^n)_n$, et l'indice inférieur j note le numéro de la composante du vecteur x^n.

Voici encore une définition importante (à comparer avec la définition analogue dans le cas $d = 1$).

Définition 1.4.1. *On dit que la suite $(x^n)_n \subset \mathbb{R}^d$ admet la limite $y \in \mathbb{R}^d$ (ou bien que la suite $(x^n)_n$ converge vers y, ou que la suite $(x^n)_n$ converge ; notation : $y = \lim_{n\to+\infty} x^n$) ssi*

$$\forall \varepsilon > 0\ \exists N = N(\varepsilon)\ \forall n \geq N : \quad d(x^n, y) = ||x^n - y|| < \varepsilon.$$

D'une manière équivalente, ceci revient à dire que $(x^n)_{n\geq N} \subset B(y, \varepsilon)$.

Informellement, on dira que $y = \lim_{n\to+\infty} x^n$ revient à $||x^n - y|| \to 0$ $(n \to +\infty)$.

Il est aisé de voir que la convergence d'une suite $(x^n)_n \subset \mathbb{R}^d$ équivaut à la convergence de d suites $(x_j^n)_n,\ j = 1, \dots, \infty$ dans $\mathbb{R}$, *i.e.,*

Proposition 1.4.2. *Soit $(x^n)_n \subset \mathbb{R}^d$. Nous avons $y = \lim_{n\to+\infty} x^n$ ssi*

$$y_j = \lim_{n\to+\infty} x_j^n,\ j = 1, \dots, d.$$

Ci-dessus, $j = 1, \dots, d$, est arbitraire et fixé, et $n \to +\infty$.

Il est clair que les résultats habituels sur les opérations avec des limites de suites de $\mathbb{R}$ (les limites de la somme/différence de deux suites, multiplication par un scalaire, etc. etc.) se transfèrent au cas vectoriel.

On peut maintenant donner des caractérisations utiles de l'intérieur, de l'adhérence et de la frontière d'un ensemble introduits précédemment, cf. section 1.3.

Proposition 1.4.3. *Soit $A \subset \mathbb{R}^d$ un ensemble. Alors :*

1.
$$A^o = \{x \in A : \exists r = r(x) > 0,\ B(x,r) \subset A\},$$
soit en mots : l'intérieur de A est l'ensemble des $x \in A$ qui appartiennent à A avec un certain voisinage.

2.
$$\bar{A} = \{y \in \mathbb{R}^d : \exists (x^n)_n \subset A, \quad x^n \to y\ (n \to \infty)\},$$
soit en mots : l'adhérence de A est l'ensemble des points y que l'on peut réaliser comme limites de suites $(x^n)_n$ appartenant à A.

3.
$$\begin{aligned}\partial A &= \{y \in \mathbb{R}^d : \forall r > 0\ A \cap B(y,r) \neq \emptyset,\ A^c \cap B(y,r) \neq \emptyset\}\\ &= \{y \in \mathbb{R}^d : \exists (x^n)_n \subset A,\ x^n \to y(n \to \infty),\ \exists (z^n)_n \subset A^c,\ z^n \to y(n \to \infty)\},\end{aligned}$$
en mots : la frontière de A est l'ensemble des points y, que l'on peut approcher par des suites appartenant à A, et par les suites appartenant à A^c.

On transpose aussi la définition de suite de Cauchy (*e.g.,* la suite fondamentale) au cas de $\mathbb{R}^d$ à partir de $\mathbb{R}$.

Définition 1.4.4. *On dit que la suite $(x^n)_n \subset \mathbb{R}^d$ est une suite de Cauchy ssi*
$$\forall \varepsilon > 0\ \exists N = N(\varepsilon)\ \forall n, m \geq N : \ d(x^n, x^m) = ||x^n - x^m|| < \varepsilon.$$
D'une manière équivalente, ceci revient à dire que tous x^n, x^m avec $n, m \geq N$ possèdent la propriété $x^n \in B(x^m, \varepsilon)$ (ou bien $x^m \in B(x^n, \varepsilon)$).

Il est essentiel que la propriété de cette définition soit vérifiée pour *tous* $n, m \geq N$; on re-écrit la définition comme $||x^n - x^m|| \to 0$ pour $n, m \to +\infty$.

Proposition 1.4.5. *Soit $(x^n)_n \subset \mathbb{R}^d$ une suite donnée. Alors $(x^n)_n$ converge ssi elle est Cauchy.*

Observons que l'implication "Définition 1.4.1 $\Rightarrow$ Définition 1.4.4" est toujours vraie ; l'implication réciproque peut être fausse dans un espace vectoriel de dimension infinie. Si E est un espace vectoriel muni d'une norme, on dit que E est complet, si toute suite Cauchy dans l'espace est convergente. Nous avons donc

Proposition 1.4.6. *L'espace $\mathbb{R}^d$ ($\mathbb{C}^d$, respectivement) est complet.*

1.4.1 Les espaces vectoriels "en général"

Soit E un espace vectoriel (sur le corps $\mathbb{R}$ (ou $\mathbb{C}$)) ; l'expression "espace vectoriel" est quelquefois abrégée comme "ev". Supposons qu'il est muni d'une norme $||.||$ satisfaisant les propriétés (N1)-(N3), cf. p. 5 ; on dit alors que $(E, ||.||)$ est *un espace vectoriel normé* (un *evn* pour être bref). Si E est muni d'une distance $d(.,.)$ satisfaisant (D1)-(D3), cf. p. 6, on parle *d'un espace vectoriel métrique.* Notons que la structure d'un espace métrique en général ne requiert pas *a priori* la structure vectorielle de l'espace ; en particulier, les espaces métriques "généraux" ne font pas partie de ce cours.

Soit donc $(E, ||.||)$ un evn. Les constructions et les définitions des sections 1.3, 1.4 s'appliquent parfaitement à l'espace E et elles donnent la topologie (*i.e.,* les ouverts, fermés, etc.) et la notion de convergence des suites exactement comme présenté ci-dessus.

Par exemple, soit $\mathcal{P}_n[a,b]$ l'espace vectoriel des polynômes (réels) de degré $\leq n$, *i.e.,*
$$\mathcal{P}_n[a,b] = \{p(x) = \sum_{k=0}^{n} a_k x^k,\ a_k \in \mathbb{R}\}.$$

Il est clair que $\dim \mathcal{P}_n[a,b] = n+1$. On peut y introduire plusieurs normes, à savoir

$$\begin{aligned} ||p||_q &= \left(\sum_{k=0}^n |a_k|^q\right)^{1/q}, \quad 1 \le q \le +\infty, \\ ||p||'_\infty &= \max_{x\in[a,b]} |p(x)|, \quad ||p||'_1 = \int_a^b |p(x)|\,dx, \text{ etc. etc.} \end{aligned}$$

Pour donner un exemple d'un evn de dimension infinie, soit $E = C[a,b]$ et $F = C^1[a,b]$. Les normes sur ces ev peuvent être définies de façons suivantes : pour $f \in E$

$$||f||_\infty = \max_{x\in[a,b]} |f(x)|, \quad ||f||_1 = \int_a^b |f(x)|\,dx,$$

pour $g \in F$ on définit

$$\begin{aligned} ||g||' &= \max_{[a,b]}(|g(x)| + |g'(x)|), \qquad ||g||' = \max_{[a,b]} |g(x)| + \max_{[a,b]} |g'(x)|, \\ ||g||''' &= \int_a^b (|g(x)| + |g'(x)|)\,dx, \qquad ||g||^{IV} = \int_a^b |g(x)|\,dx + \max_{[a,b]} |g'(x)|, \end{aligned}$$

etc. etc. Comme on l'a déjà mentionné, dans ce cours on se concentre sur le cas de *dimension finie* et on fait quelques remarques sur le cas de dimension infinie (beaucoup plus compliqué) en passant.

On peut démontrer que tout ev E réel de dimension d est isomorphe à $\mathbb{R}^d$. De plus, les normes sur un evn E de dimension finie sont équivalentes deux-à-deux et elles définissent donc la même topologie. Cela montre donc que la topologie (ou la structure métrique) d'un evn réel de dimension d est exactement celle de $\mathbb{R}^d$ construite dans les sections 1.3, 1.4. La même discussion s'applique aux evn complexes de dimension d, qui seront tous "équivalents" à $\mathbb{C}^d$.

1.5 Compacité dans $\mathbb{R}^d$

Dans cette section, on introduit *les ensembles compacts* (ou *les compacts* tout court, pour être bref). On verra que ces ensembles possèdent un bon nombre des propriétés très particulières. Les compacts réapparaîtront en lien avec les applications continues dans le chapitre suivant, cf. Chapitre 2.

Définition 1.5.1. *Soit $K \subset \mathbb{R}^d$ un ensemble et $(O_\alpha)_{\alpha\in A}$ une famille quelconque d'ensembles ouverts $O_\alpha \subset \mathbb{R}^d$. On suppose que $(O_\alpha)_{\alpha\in A}$ recouvre K, c.à.d. $K \subset \cup_{\alpha\in A} O_\alpha$.*

On dit que K est compact *si la famille $(O_\alpha)_{\alpha\in A}$ admet une sous-famille* finie *$O_{\alpha_1}, \dots, O_{\alpha_n}$ telle que*

$$K \subset \cup_{i=1}^n O_{\alpha_i}.$$

Cette définition se lit "en mots" de la manière suivante : "K est compact, ssi chacun de ses recouvrements ouverts admet un sous-recouvrement ouvert fini".

On dit qu'un $A \subset \mathbb{R}^d$ admet un ε-réseau fini, s'il existent $(x^j)_{j=1,\dots,n}$, $n = n(\varepsilon)$ telles que

$$A \subset \cup_{j=1}^n B(x^j, \varepsilon).$$

Théorème 1.5.2 (sur les compacts de $\mathbb{R}^d$). *Soit $K \subset \mathbb{R}^d$ un ensemble. Les propriétés suivantes sont équivalentes (notation : LPSSE) :*

1. *K est compact,*
2. *K est fermé et borné,*

3. *toute suite $(x^n)_n \subset K$ admet une sous-suite $(x^{n_k})_k \subset (x^n)_n$ convergente,* i.e.,

$$x^{n_k} \to x^0 \quad (k \to \infty),$$

   et, de plus, $x^0 \in K$.

4. *pour tout $\varepsilon > 0$, K admet un ε-réseau fini.*

L'équivalence entre 1. et 2. dans $\mathbb{R}^d$ est souvent appelée "le théorème de Borel-Lebesgue" (ou celui de Heine-Borel). Le fait que 2. est équivalent à 3. dans $\mathbb{R}^d$ est appelé "le théorème de Bolzano-Weierstrass".

Notons que dans le cas d'evn de dimension infinie, on a 1. $\Rightarrow$ 2., mais la réciproque est fausse. Enfin, ajoutons que $A \subset \mathbb{R}^d$ est appelé pré-compact, ssi $\bar{A}$ est compact.

Chapitre 2

Applications sur $\mathbb{R}^d$ et leur continuité

2.1 Généralités sur les applications

Dans cette section, nous rappelons les définitions de base sur les applications (ou les fonctions) ; pour la suite du cours, nous faisons ce rappel surtout dans le cadre d'applications allant de $\mathbb{R}^d$ à $\mathbb{R}^{d_1}$.

Soient $A \subset \mathbb{R}^d$, $B \subset \mathbb{R}^{d_1}$ des ensembles. Une application $f : A \to B$ est un règle (ou une correspondance) qui associe l'unique $y \in B$ à chaque $x \in A$, c.à.d $f : x \mapsto y$, ou bien $y = f(x),\ x \in A, y \in B$. On dit quelquefois que A est l'ensemble de départ et B est l'ensemble d'arrivée pour f.

Comme $x \in \mathbb{R}^d$ et $y \in \mathbb{R}^{d_1}$, on peut détailler l'écriture $y = f(x)$ comme

$$\begin{bmatrix} y_1 \\ \vdots \\ y_{d_1} \end{bmatrix} = \begin{bmatrix} f_1(x) \\ \vdots \\ f_{d_1}(x) \end{bmatrix} = \begin{bmatrix} f_1(x_1, \dots, x_d) \\ \vdots \\ f_{d_1}(x_1, \dots, x_d) \end{bmatrix}.$$

Bien évidemment, on préférera l'écriture vectorielle $y = f(x)$ à celle d'avant (trop encombrante). Si $B \subset \mathbb{R}^1,\ d_1 = 1$, on dira que f est une fonction (de plusieurs variables). Dans le cas $d_1 > 1$, on dit que f est une application (de plusieurs variables) ; les fonctions $(f_j(.))_{j=1,\dots,d_1}$ s'appellent les composantes de f.

En pratique, l'application f est donnée par des formules (ou relations mathématiques).

Définition 2.1.1. *Le domaine de f (notation : D_f) est le plus grand ensemble de $x \in \mathbb{R}^d$ pour lequel $f(x)$ est bien défini.*

Clairement, $D_f \subset \mathbb{R}^d$ et

$$D_f = \cap_{j=1}^{d_1} D_{f_j}.$$

Définition 2.1.2. *L'image de f (notation : Δ_f ou encore $Im(f)$) est donné par :*

$$\Delta_f \overset{déf}{=} \{y \in \mathbb{R}^{d_1} : \exists x \in D_f : y = f(x)\} = f(D_f).$$

Plus généralement, on pose

$$f(A) \overset{déf}{=} \{y = f(x), x \in A\}$$

pour l'image (directe) de $A \subset D_f \subset \mathbb{R}^d$ et

$$f^{-1}(B) \overset{déf}{=} \{x \in D_f : \exists y \in B : y = f(x)\}$$

pour l'image réciproque de $B \subset \mathbb{R}^{d_1}$.

Nous avons quelques propriétés élémentaires suivantes :

$$\begin{aligned}
&A_1 \subset A_2 \Rightarrow f(A_1) \subset f(A_2),\\
&f(A_1 \cup A_2) = f(A_1) \cup f(A_2),\\
&f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2),\ A \subset f^{-1}(f(A)),\ \text{etc.}\\
&B_1 \subset B_2 \Rightarrow f^{-1}(B_1) \subset f^{-1}(B_2),\\
&f^{-1}(B_1 \cup B_2) = f(B_1) \cup f(B_2),\\
&f(B_1 \cap B_2) = f(B_1) \cap f(B_2),\ f(f^{-1}(B)) = B \cap \Delta_d,\ \text{etc.}
\end{aligned}$$

2.2 Limite d'une application

Soit $A \subset \mathbb{R}^d$ un ensemble, $f : A \to \mathbb{R}^{d_1}$ une application, et $x^0 \in \bar{A}$.

Définition 2.2.1. *On dit que $l \in \mathbb{R}^{d_1}$ est* la limite *de l'application f pour $x \to x^0$ (le long de A) (notation : $l = \lim_{x\to x^0,\, x\in A} f(x)$), ssi*

$$\forall \varepsilon > 0\ \exists \delta = \delta(\varepsilon) > 0\ \forall x \in A : ||x - x^0|| < \delta \Rightarrow ||f(x) - l|| < \varepsilon.$$

Ceci revient à dire que $f(B_{\mathbb{R}^d}(x^0, \delta) \cap A) \subset B_{\mathbb{R}^{d_1}}(l, \varepsilon)$.

En mots : pour tout voisinage ouvert V du point $l \in \mathbb{R}^{d_1}$, il existe un voisinage (suffisamment petit) du point $x^0 \in \bar{A}$ tel que son image appartient à V.

D'une manière équivalente, cette définition stipule que, pour toute la suite $(x^n)_n \subset A$ telle que $x^0 = \lim_{n\to+\infty} x^n$, on a

$$\lim_{n\to+\infty} f(x^n) = l.$$

Cette dernière relation sous-entend que : a) pour *toute* suite $(x^n)_n$ avec ces propriétés, la $\lim_{n\to+\infty} f(x^n)$ existe ; b) pour toute $(x^n)_n$, cette limite est *toujours* égale à l.

Il va de soi que la définition 2.2.1 est "intéressante" s'il y a des suites $(x^n)_n,\ x^n \to x^0$ non-triviales (*i.e.,* $x^n \neq x^0$ pour un nombre infini d'indices n), ou bien

$$\forall \delta > 0 :\ (B(x^0, \delta)\backslash\{x^0\}) \cap A \neq \emptyset.$$

Dans le cas contraire (*i.e.,* pour toute suite $(x^n)_n \subset A$, on a $x^n = x^0$ pour $n \geq N$), cette définition affirme seulement que $x^0 \in A$ et $f(x^0) = l$.

Nous avons les mêmes résultats à propos des opérations algébriques avec les limites que dans le cas de $\mathbb{R}^1$.

Proposition 2.2.2. *Soient $f, g : A \to \mathbb{R}^{d_1}$ des applications et $x^0 \in \bar{A}$. Supposons que*

$$l_1 = \lim_{x\to x^0,\ x\in A} f(x), \quad l_2 = \lim_{x\to x^0,\ x\in A} g(x).$$

Alors :

1. *pour tous $\lambda, \mu \in \mathbb{R}$, on a*
$$\lim_{x\to x^0,\ x\in A} (\lambda f(x) + \mu g(x)) = \lambda l_1 + \mu l_2.$$
2. *Soit maintenant $d_1 = 1$,* i.e., *$f, g : A \to \mathbb{R}$,*
   - $\lim_{x\to x^0,\ x\in A} f(x) \cdot g(x) = l_1 l_2,$
   - *si $l_2 \neq 0$, on a* $\lim_{x\to x^0,\ x\in A} f(x)/g(x) = l_1/l_2.$

Pour résumer, les limites d'applications commutent avec les opérations algébriques habituelles "tant que ces opérations sont bien définies" (*i.e.,* elles ont du sens).

2.3 Les applications continues et leur proriétés élémentaires

Soit $A \subset \mathbb{R}^d$ un ensemble, $f : A \to \mathbb{R}^{d_1}$ une application, et $x^0 \in A$. Informellement parlant, l'application f est *continue* au point x^0, si la limite

$$l = \lim_{x \to x^0,\, x \in A} f(x)$$

est égale à la "valeur naturelle" au point x^0, à savoir $l = f(x^0)$.

Plus précisément,

Définition 2.3.1. *1. L'application f est continue au point x^0 (au long de A) (notation : $f \in \mathcal{C}_A(x^0)$), ssi*

$$\lim_{x \to x^0,\, x \in A} f(x) = f(x^0).$$

D'une manière plus détaillée,

$$\forall \varepsilon > 0 \ \exists \delta = \delta(\varepsilon) > 0 \ \forall x \in A : ||x - x^0|| < \delta \Rightarrow ||f(x) - f(x^0)|| < \varepsilon,$$

i.e., *nous avons $l = f(x^0)$ dans la définition 2.2.1. Comme avant, ceci équivaut à dire que $f(B_{\mathbb{R}^d}(x^0, \delta) \cap A) \subset B_{\mathbb{R}^{d_1}}(f(x^0), \varepsilon)$.*

2. L'application f est continue sur A (notation : $f \in \mathcal{C}(A)$, ou $f \in \mathcal{C}(A, \mathbb{R}^{d_1})$ si on a envie d'indiquer explicitement l'espace d'arrivée), ssi f est continue en tout point de A (au long de A), i.e.,

$$\forall x^0 \in A : f \in \mathcal{C}_A(x^0).$$

Vu la proposition 2.2.2, la continuité d'applications s'accorde bien avec les opérations algébriques et la composition.

Proposition 2.3.2. *1. Soit $A \subset \mathbb{R}^d$ un ensemble et $f_1, f_2 : A \to \mathbb{R}^{d_1}$ des applications continues, $f_i \in \mathcal{C}(A, \mathbb{R}^{d_1})$, $i = 1, 2$. Alors pour tous $\lambda, \mu \in \mathbb{R}$, $\lambda f_1 + \mu f_2 \in \mathcal{C}(A, \mathbb{R}^{d_1})$.*

2. Soit maintenant $d_1 = 1$, i.e., *$f_i \in \mathcal{C}(A, \mathbb{R})$, $i = 1, 2$.*

- *On a $f_1 \cdot f_2 \in \mathcal{C}(A, \mathbb{R})$.*
- *Soit $f_2 \neq 0$ sur A. Alors $f_1/f_2 \in \mathcal{C}(A, \mathbb{R})$.*

3. Soient $A \subset \mathbb{R}^d$, $B \subset \mathbb{R}^{d_1}$ et $C \subset \mathbb{R}^{d_2}$. Ensuite $f \in \mathcal{C}(A, B), g \in \mathcal{C}(B, C)$, c.à.d.

$$A \xrightarrow{f} B \xrightarrow{g} C.$$

Alors l'application composée $F = g \circ f$ est continue, i.e., *$F = g \circ f \in \mathcal{C}(A, C)$.*

Nous avons aussi la caractérisation topologique suivante d'applications continues.

Théorème 2.3.3. *Soit $A \subset \mathbb{R}^d$ et $f : A \to \mathbb{R}^{d_1}$ une application. Les propriétés suivantes sont équivalentes :*

1. *$f \in \mathcal{C}(A, \mathbb{R}^{d_1})$,*
2. *pour tout ouvert $O_1 \subset \mathbb{R}^{d_1}$, il existe un ouvert $O \subset \mathbb{R}^d$ tel que $f^{-1}(O_1) = O \cap A$,*
3. *pour tout fermé $F_1 \subset \mathbb{R}^{d_1}$, il existe un fermé $F \subset \mathbb{R}^d$ tel que $f^{-1}(F_1) = F \cap A$.*

Ceci est bien évidemment faux si on remplace l'image réciproque $f^{-1}(.)$ par l'image directe $f(.)$.

2.3.1 Le prolongement d'une application par la continuité

Soit, comme d'habitude, $A \subset \mathbb{R}^d$ un ensemble et $f : A \to \mathbb{R}^{d_1}$, $f \in \mathcal{C}(A, \mathbb{R}^{d_1})$ une application continue. Soit $x^0 \in \bar{A}$. On souhaite savoir quand il est possible de prolonger l'application f au point x^0 "d'une manière naturelle", c.à.d. en sorte que l'application "prolongée" $\tilde{f}$ soit continue sur $\tilde{A} = A \cup \{x^0\}$.

La réponse à cette question est donnée par la proposition suivante. Nous disons que l'application f a la propriété $(*)_{x^0}$, ssi pour toute suite $(x^n)_n \subset A$ telle que $\lim_{n\to+\infty} x^n = x^0$, on a

$$\lim_{n\to+\infty} f(x^n) = l,$$

et la valeur $l \in \mathbb{R}^{d_1}$ ne dépend pas de la suite $(x^n)_n$ que l'on a choisie. On peut dire que l est la même pour toutes les suites $(x^n)_n \subset A$ possibles. D'une façon équivalente, il existe

$$\lim_{x\to x^0,\, x\in A} f(x) = l.$$

Proposition 2.3.4 (prolongement d'une application par continuité)**.** *Soient l'application f et l'ensemble A comme ci-dessus, $f \in \mathcal{C}(A, \mathbb{R}^{d_1})$, $x_0 \in \bar{A}$ et f satisfait la propriété $(*)_{x^0}$. Soit $\tilde{A} = A \cup \{x^0\}$. Alors il existe $\tilde{f} : \tilde{A} \to \mathbb{R}^{d_1}$ satisfaisant les points suivants :*

- $\tilde{f} \in \mathcal{C}(\tilde{A}, \mathbb{R}^{d_1})$,
- *$\tilde{f}$ restreinte sur A coïncide avec f,* i.e., $\tilde{f}\big|_A = f$.

L'application $\tilde{f}$ de cette proposition s'appelle *l'extension* (ou *le prolongement*) de f par continuité à $A \cup \{x^0\}$. Elle est donnée par

$$\tilde{f}(x) = \begin{cases} f(x) & , \quad x \in A, \\ l & , \quad x = x^0. \end{cases}$$

Exemple 2.3.5. *Soit*

$$f(x, y) = \frac{xy}{x^2 + y^2}, \ D_f = \mathbb{R}^2 \backslash \{(0, 0)\}.$$

Il est impossible de prolonger f au point $x^0 = (0, 0)$ par continuité, car la limite

$$\lim_{(x,y)\to 0,\,(x,y)\neq 0} f(x, y)$$

n'existe pas.

Exemple 2.3.6. *Soit $\alpha > 1$ et*

$$g_\alpha(x, y) = \frac{(xy)^\alpha}{x^2 + y^2}, \ D_{g_\alpha} = \{(x, y) : x > 0, y > 0\}.$$

L'application g_α se prolonge au point $x^0 = (0, 0)$ par continuité, car la limite

$$l = \lim_{(x,y)\to 0,\, x>0, y>0} g_\alpha(x, y) = 0$$

existe et, par conséquent, la valeur réquise de $l = 0$.

2.4 Les applications continues, compacts et le théorème de Weierstrass

Dans cette section, on étudiera les propriétés des applications continues définies sur un compact. En particulier, on utilisera d'une manière essentielle les propriétés des compacts de la section 1.5.

Soit $A \subset \mathbb{R}^d$ et $f : A \to \mathbb{R}^{d_1}$ une application. Dans peu de temps, nous aurons besoin de la définition suivante.

Définition 2.4.1. *On dit que f est* uniformememnt continue *sur A, ssi*

$$\forall \varepsilon > 0 \; \exists \delta = d(\varepsilon) > 0 \; \forall x, y \in A : ||x - y|| < \delta \Rightarrow ||f(x) - f(y)|| < \varepsilon.$$

La continuité uniforme implique la continuité, le réciproque est faux en général.

Théorème 2.4.2 (de Weierstrass, I)**.** *Soit $K \subset \mathbb{R}^d$ un compact et $f : K \to \mathbb{R}^{d_1}$ une application continue là-dessus, $f \in \mathcal{C}(K, \mathbb{R}^{d_1})$. Alors $f(K)$ est aussi compact.*

Corollaire 2.4.3. *Sous les hypothèses de ce théorème, on a :*

1. *$f(K)$ est fermé et borné (cf. le théorème 1.5.2),*
2. *soit $d_1 = 1$,* i.e., *$f : K \to \mathbb{R}$. Alors $\sup_{x \in K} f(x)$ et $\inf_{x \in K} f(x)$ sont atteints, c.à.d il existent $x_{min}, x_{max} \in K$ tels que*

$$\sup_{x \in K} f(x) = \max_{x \in K} f(x) = f(x_{max}),$$
$$\inf_{x \in K} f(x) = \min_{x \in K} f(x) = f(x_{min}).$$

Ce résultat transfère le résultat analogue de $\mathbb{R}$ (démontré pour des fonctions continues sur des ségments $[a, b]$) dans le cadre beaucoup plus général.

Théorème 2.4.4 (de Weierstrass, II)**.** *Soit $K \subset \mathbb{R}^d$ un compact et $f : K \to \mathbb{R}^{d_1}$ une application continue, $f \in \mathcal{C}(K, \mathbb{R}^{d_1})$. Alors f est* uniformement continue *sur K.*

Donc, pour les applications définies sur un compact, la continuité et la continuité uniforme sont équivalentes.

Chapitre 3

Calcul différentiel

3.1 Les applications linéaires continues entre deux evn

Dans cette section, nous nous concentrons sur les applications linéaires allant de $\mathbb{R}^d$ à $\mathbb{R}^{d_1}$. En grande partie, ceci est le rappel du contenu du cours d'algèbre linéaire de L1. Il va de soi qu'une construction similaire a lieu pour les applications linéaires allant de F à E, où F, E sont des evn de dimension finie quelconques.

Définition 3.1.1. *Une application linéaire $A : \mathbb{R}^d \to \mathbb{R}^{d_1}$ est une application satisfaisant la relation*

$$A(\lambda x + \mu y) = \lambda Ax + \mu Ay$$

pour tous $x, y \in \mathbb{R}^d$ et tous $\lambda, \mu \in \mathbb{R}$.

L'ensemble d'applications linéaires de $\mathbb{R}^d$ à $\mathbb{R}^{d_1}$ est noté

$$\mathcal{L}(\mathbb{R}^d, \mathbb{R}^{d_1}) \stackrel{déf}{=} \{A : \mathbb{R}^d \to \mathbb{R}^{d_1} : \forall x, y \in \mathbb{R}^d \ \forall \lambda, \mu \in \mathbb{R} : \ A(\lambda x + \mu y) = \lambda Ax + \mu Ay\}.$$

Si on fixe des bases pour les espaces $\mathbb{R}^d$ et $\mathbb{R}^{d_1}$, on peut identifier l'ensemble $\mathcal{L}(\mathbb{R}^d, \mathbb{R}^{d_1})$ avec l'ensemble des matrices réelles $d_1 \times d$, *i.e.,*

$$\mathcal{L}(\mathbb{R}^d, \mathbb{R}^{d_1}) \simeq \mathcal{M}_{d_1,d}(\mathbb{R}) \stackrel{déf}{=} \left\{ \begin{bmatrix} a_{11} & \dots & a_{1d} \\ \vdots & \ddots & \vdots \\ a_{d_1 1} & \dots & a_{d_1 d} \end{bmatrix} : a_{ij} \in \mathbb{R}, \ i = 1, \dots, d_1; j = 1, \dots d \right\}.$$

Il est clair que l'ensemble $\mathcal{L}(\mathbb{R}^d, \mathbb{R}^{d_1})$ est un espace vectoriel. On va y introduire une norme (qu'on appelle *la norme induite*, ou encore une norme matricielle, ou une norme fonctionnelle), qui transformera l'espace $\mathcal{L}(\mathbb{R}^d, \mathbb{R}^{d_1})$ en evn.

Définition 3.1.2. *Une application $A \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^{d_1})$ est* bornée *ssi l'ensemble $A(B(0,1)) \subset \mathbb{R}^{d_1}$ est borné.*

Proposition 3.1.3. *Toute $A \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^{d_1})$ possède les propriétés suivantes :*

1. *A est bornée,* i.e., *la partie $A(B(0,1))$ est bornée,*
2. *$\forall x \in \mathbb{R}^d \ \forall r \geq 0 : A(B(x,r))$ est bornée,*
3. *A est continue au point $x^0 = 0$,*
4. *A est continue en un certain point $x^0 \in \mathbb{R}^d$,*
5. *A est continue sur $\mathbb{R}^d$.*

C.à.d., toute $A \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^{d_1})$ est bornée et continue (sur $\mathbb{R}^d$ tout entier).

La proposition précédente montre que pour tout $A \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^{d_1})$, la valeur suivante est bien définie

$$|||A||| \stackrel{déf}{=} \max_{x \in \bar{B}(0,1)} ||Ax|| = \max_{x \in \mathbb{R}^d:\ ||x|| \leq 1} ||Ax|| < \infty. \tag{3.1}$$

Il est facile de voir que $|||.|||$ définit une norme sur $\mathcal{L}(\mathbb{R}^d, \mathbb{R}^{d_1})$ (*i.e.*, l'application $|||.|||$ satisfait les propriétés (N1)-(N3), cf. p. 5). On appelle cette norme *la norme induite*. L'espace $(\mathcal{L}(\mathbb{R}^d, \mathbb{R}^{d_1}), |||.|||)$ est donc un evn.

Voici quelques informations supplémentaires sur $|||.|||$.

Proposition 3.1.4. *Soit $A \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^{d_1})$. Les égalités suivantes ont lieu :*

$$\begin{aligned} |||A||| &= \max_{||x|| \leq 1} ||Ax|| = \max_{||x||=1} ||Ax|| \\ &= \sup_{y \in \mathbb{R}^d,\ y \neq 0} \frac{||Ay||}{||y||} \\ &= \inf\{C : \forall x \in \mathbb{R}^d \quad ||Ax|| \leq C||x||\}. \end{aligned}$$

Proposition 3.1.5. *Soit $A \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^{d_1})$ une application identifiée avec sa matrice $[a_{ij}]_{i=1,\ldots,d_1; j=1,\ldots,d}$. Alors*

$$|||A|||^2 \leq \sum_{i,j} a_{ij}^2.$$

et, d'une manière triviale $\forall i, j : a_{ij}^2 \leq |||A|||^2$.

Supposons, par exemple, que l'application $F : D \to \mathcal{L}(\mathbb{R}^{d_1}, \mathbb{R}^{d_2})$ est continue, c.à.d. $F \in \mathcal{C}(D, \mathcal{L}(\mathbb{R}^{d_1}, \mathbb{R}^{d_2}))$. Ici, $D \subset \mathbb{R}^d$ est un ensemble. Cela revient à dire que $|||F(x) - F(x^0)||| \to 0$ pour $x \in D, x^0 \in D$ et $x \to x^0$, ou bien que pour tous $i = 1, \ldots, d_1; j = 1, \ldots, d$ on a $|F_{ij}(x) - F_{ij}(x^0)| \to 0$.

3.2 La différentielle d'une application et ses propriétés élémentaires

A partir de cette section, il est commode de supposer que le domaine d'une application est un ouvert. Soit donc $O \subset \mathbb{R}^d$ un ouvert et $f : O \to \mathbb{R}^{d_1}$ une application.

La définition suivante est centrale pour le reste du présent cours.

Définition 3.2.1. *Soit $x^0 \in O$ et $B(x^0, \delta) \subset O$ pour un $\delta > 0$. On dit que l'application f est* différentiable *au point x^0 (notation : $f \in \mathcal{D}(x^0)$), ssi pour tout $h \in B(0, \delta)$*

$$f(x^0 + h) = f(x^0) + A_{x^0} \cdot h + \alpha(x^0, f; h), \tag{3.2}$$

où $A_{x^0} \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^{d_1})$ et le reste $\alpha(x^0, f; h)$ satisfait la relation

$$\lim_{||h|| \to 0} \frac{||\alpha(x^0, f; h)||}{||h||} = 0.$$

Quelques mots à propos de la terminologie : le terme $f(x^0)$ à droite est le terme constant, le terme $A_{x^0} \cdot h$ est linéaire en h, et le terme $\alpha(x^0, f; h)$ est le reste (ou l'erreur). La relation limite précédente montre que l'erreur est négligeable par rapport au terme linéaire $A_{x^0} \cdot h$. La formule (3.2) affirme donc que $f \in \mathcal{D}(x^0)$ ssi, modulo un reste négligeable devant $||x - x^0||$, $f(x)$ se comporte comme le terme constant $f(x^0)$ plus un terme linéaire en $x - x^0$, *i.e.*,

$$f(x) = f(x^0) + A_{x^0} \cdot (x - x^0) + \alpha(x^0, f; (x - x^0)), \qquad (x - x^0) \in B(0, \delta).$$

Cette définition est à comparer avec la définition d'une fonction dérivable en un point sur $\mathbb{R}$.

L'application $A_{x^0} \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^{d_1})$ s'appelle *la différentielle* de f au point x^0. On la note $Df(x^0), \partial f(x^0)$ ou encore $df(x^0)$. Nous allons se conformer à la première possibilité, c.à.d.

$$Df(x^0) \stackrel{déf}{=} A_{x^0}.$$

Pour résumer, (3.2) se réécrit donc comme

$$f(x^0 + h) = f(x^0) + Df(x^0)h + \alpha(x^0, f; h).$$

Pour simplifier l'écriture, on notera le reste $\alpha(h)$ ou $\alpha(x^0, h)$.

Définition 3.2.2. *Soit $f : O \to \mathbb{R}^{d_1}$ une application. Elle est différentiable sur O (notation : $f \in \mathcal{D}(O)$, ou $f \in \mathcal{D}(O, \mathbb{R}^{d_1})$), ssi $\forall x^0 \in O :\ f \in \mathcal{D}(x^0)$.*

Donc si $f \in D(O, \mathbb{R}^{d_1})$, on a une application $Df : O \to \mathcal{L}(\mathbb{R}^d, \mathbb{R}^{d_1})$.

Définition 3.2.3. *Soit $f : O \to \mathbb{R}^{d_1}$ une application. Elle est* de classe $\mathcal{C}^1$ *sur O (notation : $f \in \mathcal{C}^1(O)$, ou $f \in \mathcal{C}^1(O, \mathbb{R}^{d_1})$), ssi $Df \in \mathcal{C}(O, \mathcal{L}(\mathbb{R}^d, \mathbb{R}^{d_1}))$.*

Autrement dit, l'application f est de classe $\mathcal{C}^1$ sur O, ssi $Df(x) \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^{d_1})$ existe pour tout $x \in O$, et

$$\forall x^0 \in O :\ |||Df(x) - Df(x^0)||| \to 0$$

pour $x \in O$ et $x \to x^0$, cf. la section 3.1 et notamment les propositions 3.1.4 et 3.1.5.

Voici quelques propriétés élémentaires des applications différentiables.

Proposition 3.2.4. *Soit $O \subset \mathbb{R}^d$ un ouvert, $f : O \to \mathbb{R}^{d_1}$ une application, et $x^0 \in O$.*

1. *Soit $f \in \mathcal{D}(x^0)$. Alors la différentielle $Df(x^0)$ est bien définie, c.à.d. s'il y a deux applications linéaires $A_{x^0}, B_{x^0} \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^{d_1})$ satisfaisant (3.2), alors $A_{x^0} = B_{x^0}$.*

2. *Si $f \in \mathcal{D}(x^0)$, alors $f \in \mathcal{C}(x^0)$, la réciproque est fausse en général.*

3. *Soient $f, g : O \to \mathbb{R}^{d_1}$, $f, g \in \mathcal{D}(x^0)$ et $\lambda, \mu \in \mathbb{R}$. Alors $\lambda f + \mu g \in \mathcal{D}(x^0)$ et*

$$D(\lambda f + \mu g)(x^0) = \lambda Df(x^0) + \mu Dg(x^0).$$

4. *Ecrivons l'application $f : O \to \mathbb{R}^{d_1}$ "composante par composante", c.à.d*

$$f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_{d_1}(x) \end{bmatrix}.$$

Alors $f \in \mathcal{D}(x^0)$ ssi $f_j \in \mathcal{D}(x^0)$, $j = 1, \ldots, d_1$ et, de plus,

$$Df(x^0) = \begin{bmatrix} Df_1(x) \\ \vdots \\ Df_{d_1}(x) \end{bmatrix}. \tag{3.3}$$

C'est le moment de dire quelques mots de plus à propos de la terminologie. Soit, comme toujours, $g : O \to \mathbb{R}$ une fonction, O un ouvert et $x^0 \in O$. Supposons que $g \in \mathcal{D}(x^0)$. Quelquefois (et surtout quand il s'agit d'équations aux dérivées partielles (= EDP)), on appelle la différentielle $Dg(x^0)$ le gradient de g, la notation :

$$Dg(x^0) = \operatorname{grad} g(x^0) = \nabla g(x^0).$$

La formule (3.3) se re-écrit comme

$$Df(x^0) = \begin{bmatrix} \nabla f_1(x) \\ \vdots \\ \nabla f_{d_1}(x) \end{bmatrix}. \tag{3.4}$$

Notons aussi que, pour la fonction g, la direction du gradient $\nabla g(x^0)$ indique la direction de la croissance de $g(x)$ la plus rapide (à partir du point x^0). Le vecteur $-\nabla g(x^0)$ donne la direction de la decroissance de $g(x)$ la plus rapide à partir du point x^0.

Pour finir cette section, voici une proposition sur la différéntiabilité d'application composée.

Proposition 3.2.5. *Soient $f : O \to O_1$, $g : O_1 \to \mathbb{R}^{d_2}$ des applications, et $O \subset \mathbb{R}^d$, $O_1 \subset \mathbb{R}^{d_1}$ des ouverts. Supposons que $f \in \mathcal{D}(O, O_1), g \in \mathcal{D}(O_1, \mathbb{R}^{d_2})$. Alors $F = g \circ f : O \to \mathbb{R}^{d_2}$ est différentiable sur O, et, de plus,*

$$DF(x) = Dg(f(x)) \cdot Df(x), \qquad x \in O.$$

Le "$\cdot$" designe la composition (ou la multiplication) d'applications linéaires données par les différentielles.

3.3 La dérivée directionnelle, les dérivées partielles d'une application

Soit $O \subset \mathbb{R}^d$ un ouvert, $f : O \to \mathbb{R}^{d_1}$, $x^0 \in O$ et $v \in \mathbb{R}^{d*}$ (*i.e.,* $v \in \mathbb{R}^d$ et $v \neq 0$).

Définition 3.3.1. *La dérivée directionnelle de f au point x^0 (par rapport au vecteur v) est donnée par*

$$D_v f(x^0) = \lim_{t\to 0} \frac{f(x^0 + tv) - f(x^0)}{t}.$$

Il est aisé de voir que la différentiabilité de f au point x^0 implique l'existence de $D_v f(x^0)$ pour tout $v \in \mathbb{R}^{d*}$. Notamment, nous avons

Proposition 3.3.2. *Soit $f \in \mathcal{D}(x^0)$, alors $\forall v \in \mathbb{R}^{d*}$*

$$D_v f(x^0) = Df(x^0) \cdot v.$$

Par contre l'implication réciproque est fausse en général, *i.e.,* on peut trouver des applications pour lesquelles $D_v f(x^0)$ existe pour tout $v \in \mathbb{R}^{d*}$, mais $f \notin \mathcal{D}(x^0)$.

Nous passons à la définition de la *dérivée partielle*. En pratique, les dérivées partielles permettent d'écrire la différentielle d'une application explicitement et elles sont donc très utiles.

Nous supposons que O, f, x^0 sont choisis comme d'habitude.

Définition 3.3.3. *Soit $j = 1, \dots, d$. La j-ième dérivée partielle de l'application f au point x^0 est donnée par*

$$\frac{\partial f}{\partial x_j}(x^0) \overset{déf}{=} D_{e^j} f(x^0) = \lim_{t\to 0} \frac{f(x^0 + te^j) - f(x^0)}{t}.$$

Ci-dessus, $(e^j)_{j=1,\dots,d}$ est la base (standard) de $\mathbb{R}^d$.

Par conséquent, si $f \in \mathcal{D}(O, \mathbb{R}^{d_1})$, alors $\frac{\partial f}{\partial x_j} : O \to \mathbb{R}^{d_1}$. C.à.d., $\frac{\partial f}{\partial x_j}$ est une application à valeurs vectorielles en général. Si $d_1 = 1$, $\frac{\partial f}{\partial x_j}$ est une fonction (*i.e.,* une application à valeurs scalaires).

Voici les autres notations utilisées pour la j-ième dérivée partielle de f :

$$\frac{\partial f}{\partial x_j}(x^0) = \partial_j f(x^0) = f'_{x_j}(x^0) = \dots.$$

Dans ce cours, nous allons utiliser la notation donnée dans la définition 3.3.3.

Voici une liste des remarques simples sur la dérivée partielle :

1. La proposition 3.3.2 implique immédiatement que, si $f \in \mathcal{D}(O, \mathbb{R}^{d_1})$, alors pour $x^0 \in O$

$$\frac{\partial f}{\partial x_j}(x^0) = Df(x^0) \cdot e^j.$$

   En particulier, la matrice de $Df(x^0)$ se représente comme

$$Df(x^0) = \left[\frac{\partial f}{\partial x_1}(x^0), \ldots, \frac{\partial f}{\partial x_d}(x^0) \right].$$

   Cette formule fournit la matrice $Df(x^0)$ "colonne par colonne" ; elle est à comparer avec la formule (3.4), qui donne l'écriture de $Df(x^0)$ "ligne par ligne".

2. L'existence de $\frac{\partial f}{\partial x_j}(x^0)$ pour tout $j = 1, \ldots, d$ contient *trop peu* d'informations sur f au voisinage de x^0. Elle n'implique ni la continuité de f au point x^0, ni la différentiabilité de f en ce point.

C'est le bon moment pour se poser la question suivante : quand une application f, définie sur un ouvert O, est-elle différentiable ? Comme l'étude de la différentiabilité à l'aide de la définition 3.2.1 est difficile en pratique, on aimerait bien avoir d'autres outils pour répondre à cette interrogation. Un moyen de le faire est donné par le théorème suivant.

Théorème 3.3.4 (sur l'appartenance de l'application à la classe C^1)**.** *Soit $O \subset \mathbb{R}^d$ un ouvert, $f : O \to \mathbb{R}^{d_1}$ une application.*

Alors $f \in \mathcal{C}^1(O, \mathbb{R}^{d_1})$ <u>ssi</u> toutes les dérivées partielles $\frac{\partial f}{\partial x_j}$ existent et sont continues sur O, c.à.d.

$$\frac{\partial f}{\partial x_j} \in \mathcal{C}(O, \mathbb{R}^{d_1}), \quad j = 1, \ldots, d.$$

Corollaire 3.3.5. *Si $\frac{\partial f}{\partial x_j} \in \mathcal{C}(O, \mathbb{R}^{d_1}), \quad j = 1, \ldots, d$, alors $f \in \mathcal{D}(O, \mathbb{R}^{d_1})$.*

Ce corollaire donne une condition suffisante pour la différentiabilité d'une application qui est beaucoup utilisée en pratique.

3.4 Inégalité des accroissements finis

Il n'est pas difficile de voir que l'analogue "direct" du TAF (= le théorème des accroissements finis) n'a pas lieu en plusieurs dimensions.

C'est seulement sa version "faible" qui est vraie, et nous l'ennonçons ci-dessous.

Proposition 3.4.1. *Soient $O \subset \mathbb{R}^d$ un ouvert, et $f \in \mathcal{D}(O, \mathbb{R}^{d_1})$. Soient $x, y \in O$ en sorte que $[x, y] \subset O$. Alors :*

1.

$$||f(y) - f(x)|| \leq \sup_{z \in [x,y]} |||Df(z)||| \; ||y - x||.$$

2. Si $f \in \mathcal{C}^1(O, \mathbb{R}^d)$, on a

$$f(y) - f(x) = \int_0^1 (Df(x + t(y - x)))\, dt \,(y - x).$$

3.5 Rappel sur les formes bilinéaires

Ici, on rappelle quelques définitions et propriétés des formes bilinéaires. *A priori*, elles ont été traitées en détail dans le cours d'algèbre linéaire de L1.

Définition 3.5.1. *Un application $Q : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ s'apelle une forme bilinéaire, si elle est linéaire par rapport à ses deux arguments,* i.e.,

$$\begin{aligned} \forall \lambda, \mu \in \mathbb{R},\ x, x', y \in \mathbb{R}^d : &\qquad Q(\lambda x + \mu x', y) = \lambda Q(x, y) + \mu Q(x', y), \\ \forall \lambda, \mu \in \mathbb{R},\ x, y, y' \in \mathbb{R}^d : &\qquad Q(x, \lambda y + \mu y') = \lambda Q(x, y) + \mu Q(x, y'). \end{aligned}$$

L'ensemble de formes bilinéaires allant de $\mathbb{R}^d \times \mathbb{R}^d$ à $\mathbb{R}$ est noté par $\mathcal{BL}(\mathbb{R}^d, \mathbb{R}^d; \mathbb{R})$.

Il est clair que $\mathcal{BL}(\mathbb{R}^d, \mathbb{R}^d; \mathbb{R})$ est un espace vectoriel ; la proposition suivante montre que $\mathcal{BL}(\mathbb{R}^d, \mathbb{R}^d; \mathbb{R})$ est un evn qui est isométrique à $\mathcal{L}(\mathbb{R}^d, \mathbb{R}^d)$.

Proposition 3.5.2. *1. Soit $Q \in \mathcal{BL}(\mathbb{R}^d, \mathbb{R}^d; \mathbb{R})$. Le relations*

$$\begin{aligned} |||Q|||' &= \sup_{||x||, ||y|| \le 1} |Q(x, y)| \\ &= \inf\{C :\ |Q(x, y)| \le C||x||\,||y||\} < \infty \end{aligned} \tag{3.5}$$

définissent une norme sur l'espace en question. En particulier, toute forme bilinéaire est continue.

2. Pour tout $Q \in \mathcal{BL}(\mathbb{R}^d, \mathbb{R}^d; \mathbb{R})$, il existe $A \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^d)$ telle que

$$Q(x, y) = Q_A(x, y) = (Ax, y) = y^t\, Ax,$$

et, de plus, $|||Q|||' = |||A|||$.

Nous avons donc $\mathcal{BL}(\mathbb{R}^d, \mathbb{R}^d; \mathbb{R}^d) \simeq \mathcal{L}(\mathbb{R}^d, \mathbb{R}^d)$. On peut y rajouter encore un isomorphisme isométrique naturel ; ainsi

$$\mathcal{BL}(\mathbb{R}^d, \mathbb{R}^d; \mathbb{R}) \simeq \mathcal{L}(\mathbb{R}^d, \mathcal{L}(\mathbb{R}^d, \mathbb{R})) \simeq \mathcal{L}(\mathbb{R}^d, \mathbb{R}^d).$$

Passons maintenant aux formes bilinéaires symétriques, c.à.d., les formes Q pour lesquelles

$$Q(x, y) = Q(y, x), \quad \forall x, y \in \mathbb{R}^d.$$

Il est facile de voir que $Q = Q_A$ est symétrique ssi A l'est, *i.e.,* $A^t = A$ (en termes de la matrice correspondante). Soit $Q = Q_A$ symétrique. Nous adoptons la terminologie suivante :

1. on dit que la forme $Q = Q_A$ est *positive* (notation : $Q_A \ge 0$ ou $A \ge 0$), ssi
$$Q_A(x, x) = (Ax, x) \ge 0, \quad \forall x \in \mathbb{R}^d.$$
2. on dit que la forme $Q = Q_A$ est *strictement positive* (notation : $Q_A > 0$ ou $A > 0$), ssi
$$Q_A(x, x) = (Ax, x) > 0, \quad \forall x \in \mathbb{R}^{d*}.$$
3. on dit que la forme $Q = Q_A$ est *négative* (notation : $Q_A \le 0$ ou $A \le 0$), ssi
$$Q_A(x, x) = (Ax, x) \le 0, \quad \forall x \in \mathbb{R}^d.$$
4. on dit que la forme $Q = Q_A$ est *strictement négative* (notation : $Q_A < 0$ ou $A < 0$), ssi
$$Q_A(x, x) = (Ax, x) < 0, \quad \forall x \in \mathbb{R}^{d*}.$$

Pour voir si une matrice donnée $A \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^d)$ satisfait ces propriétés, on utilise le critère de Sylvester.

3.6 Dérivées partielles d'ordre supérieur

Pour discuter les différentielles d'ordre k ($k > 2$) de la même manière que nous l'avions fait pour la différentielle première, cf. la section 3.2, nous devrions utiliser les formes k-linéaires, qui sont hors programme pour ce cours. Pour cette raison, on choisit la version "réduite" de cette présentation et on se concentre surtout sur la discussion des dérivées partielles d'ordre $k \geq 2$.

Soit $O \subset \mathbb{R}^d$ un ouvert, $f : O \to \mathbb{R}^{d_1}$ une application et $x^0 \in O$. Supposons que les dérivées premières $\partial f/\partial x_j,\ j = 1, \ldots, d$, existent sur O. Commençons avec la définition des dérivées partielles d'ordre 2.

Définition 3.6.1. *1. Pour $k = 1, \ldots, d$, on définit la dérivée partielles $\partial^2 f/\partial x_k \partial x_j$ d'ordre 2 comme suit : pour $x^0 \in O$*

$$\frac{\partial^2 f}{\partial x_k \partial_j}(x^0) \overset{déf}{=} \frac{\partial}{\partial x_k}\left(\frac{\partial f}{\partial x_j}\right)(x^0) = \lim_{t\to 0} \frac{\partial f/\partial x_j(x^0 + te^k) - \partial f/\partial x_j(x^0)}{t}$$

(à condition que la limite existe et soit finie).

2. On dit que l'application f est de classe $\mathcal{C}^2(O)$, ssi toutes les dérivées partielles d'ordre 2 *existent et sont continues sur O,* i.e.,

$$\forall j, k = 1, \ldots, d,\ \frac{\partial^2 f}{\partial x_k \partial x_j} \in \mathcal{C}(O).$$

La définition des dérivées partielles d'ordre supérieur se fait par récurrence. Soient $i_1, \ldots, i_n$ des indices avec $i_k \in \{1, \ldots, d\}$ et $k = 1, \ldots, n$. Supposons que toutes les dérivées partielles d'ordre $n - 1$ existent, *i.e.,* les dérivées partielles

$$\frac{\partial}{\partial x_{i_{n-1}}}\left(\frac{\partial}{\partial x_{i_{n-2}}} \cdots \left(\frac{\partial f}{\partial x_{i_1}}\right) \cdots\right)$$

sont bien définies sur O.

Définition 3.6.2. *1. La dérivée partielle*

$$\frac{\partial^n f}{\partial x_{i_n} \partial x_{i_{n-1}} \ldots \partial x_{i_1}}$$

d'ordre n *est définie comme suit : pour $x^0 \in O$*

$$\begin{aligned}\frac{\partial^n f}{\partial x_{i_n} \partial x_{i_{n-1}} \ldots \partial x_{i_1}}(x^0) &\overset{déf}{=} \frac{\partial}{\partial x_{i_n}}\left(\frac{\partial}{\partial x_{i_{n-1}}} \cdots \left(\frac{\partial f}{\partial x_{i_1}}\right) \cdots\right)(x^0) \\ &= \lim_{t\to 0} \frac{\frac{\partial^{n-1} f}{\partial x_{i_{n-1}} \ldots \partial x_{i_1}}(x^0 + te^{i_n}) - \frac{\partial^{n-1} f}{\partial x_{i_{n-1}} \ldots \partial x_{i_1}}(x^0)}{t},\end{aligned}$$

sous la condition d'existence de la limite.

2. On dit que l'application f est de classe $\mathcal{C}^n(O)$, ssi toutes les dérivées partielles d'ordre n *existent et sont continues sur O,* i.e.,

$$\frac{\partial^n f}{\partial x_{i_n} \partial x_{i_{n-1}} \ldots \partial x_{i_1}} \in \mathcal{C}(O).$$

On se pose maintenant la question suivante : sous quelles conditions les dérivées partielles "commutent", c.à.d.

$$\frac{\partial^2 f}{\partial x_j \partial x_k} = \frac{\partial^2 f}{\partial x_k \partial x_j}, \quad j \neq k.$$

Bien sûr, on peut se poser la même question pour les dérivées partielles d'ordre n. Voici la réponse :

Théorème 3.6.3 (de Clairaut (ou de Schwarz)). *1. Soit $f : O \to \mathbb{R}^{d_1}$ une application, où O est un ouvert. Si $f \in \mathcal{C}^2(O)$ (*i.e., *les dérivées partielles d'ordre 2 existent et sont continues), alors*

$$\frac{\partial^2 f}{\partial x_j \partial x_k} = \frac{\partial^2 f}{\partial x_k \partial x_j}$$

sur O. On peut dire aussi que la valeur de la dérivée partielle ne dépend pas de l'ordre de calcul de dérivées partielles premières.

*2. (Idem pour les dérivées partielles d'ordre n ; nous donnons la formulation "en mots" et non à l'aide de formules, car l'écriture est assez encombrante) Si $f \in \mathcal{C}^n(O)$(*i.e., *les dérivées partielles d'ordre n existent et sont continues), alors la valeur de la dérivée partielle*

$$\frac{\partial^n f}{\partial x_{i_n} \partial x_{i_{n-1}} \ldots \partial x_{i_1}}$$

ne dépend pas de l'ordre de dérivation $i_1, i_2, \ldots, i_n$.

Soit maintenant $f : O \to \mathbb{R}$ (*i.e.,* $d_1 = 1$) et $f \in \mathcal{C}^2(O)$. Nous définissons la matrice

$$H_f(x) = [h_{ik}(x)]_{i,k=1,\ldots,d} = \left[\frac{\partial^2 f}{\partial x_i \partial x_k(x)}\right]_{i,k=1,\ldots,d}, \quad x \in O. \tag{3.6}$$

Cette matrice s'appelle la *matrice Hessienne* de f. La partie 1. du théorème 3.6.3 affirme que la matrice Hessienne de $f \in \mathcal{C}^2(O)$ est symétrique.

3.7 Formule de Taylor d'ordre 2

Soit, comme toujours $f : O \to \mathbb{R}$ ($d_1 = 1\,!$), et $O \subset \mathbb{R}^d$ un ouvert. Nous commençons cette section avec la définition de la deuxième différentielle de f sur O.

Rappelons que $f \in \mathcal{D}(O, \mathbb{R})$ ssi on a

$$f(x^0 + h) = f(x^0) + Df(x^0) \cdot h + \alpha(x^0; h), \quad x^0 \in O.$$

La définition suivante consiste à réécrire cette relation pour $Df(x)$ à la place de f.

Définition 3.7.1. *Soit f une application comme précisé ci-dessus. On dit que f est 2-fois différentiable sur O (notation : $f \in \mathcal{D}^2(O)$) ssi*

$$Df(x^0 + h) = Df(x^0) + h^t \cdot D^2 f(x^0) + \alpha_2(x^0; h),$$

à condition que les parties de cette relation aient un sense. Le reste $\alpha_2(x^0; h)$ est supposé satisfaire les propriétés habituelles, cf. la définition 3.2.1. Ici, h^t est la transposée du vecteur h.

Notons que l'écriture $h^t \cdot D^2 f(x^0)$ du terme linéaire à gauche vient du fait que $Df(x) \in \mathcal{L}(\mathbb{R}^d, \mathbb{R})$ et donc $Df(x)$ est *une ligne* à d composantes (c.à.d., une matrice de la taille $1 \times d$). Si on écrivait le terme linéaire comme $D^2 f(x^0) \cdot h$, on aurait eu une incompatibilité de dimension, car $D^2 f(x^0) \cdot h$ est *un vecteur* à d composantes (c.à.d., une matrice de la taille $d \times 1$).

Ensuite, comme $f : O \to \mathbb{R}$, on a $Df : O \to \mathcal{L}(\mathbb{R}^d, \mathbb{R})$ et donc

$$D^2 : O \to \mathcal{L}(\mathbb{R}^d, \mathcal{L}(\mathbb{R}^d, \mathbb{R})).$$

Or, le dernier evn a été étudié dans la section 3.5. Nous avons en particulier

$$\mathcal{L}(\mathbb{R}^d, \mathcal{L}(\mathbb{R}^d, \mathbb{R})) \simeq \mathcal{L}(\mathbb{R}^d, \mathbb{R}^d) \simeq \mathcal{BL}(\mathbb{R}^d, \mathbb{R}^d; \mathbb{R}).$$

Nous allons donc considérer la différentielle $D^2 f(x)$ comme l'élément de l'evn $\mathcal{BL}(\mathbb{R}^d, \mathbb{R}^d; \mathbb{R})$; voir le théorème suivant pour l'explication de cette convention. Notons aussi que la matrice de l'application $D^2 f(x)$ est exactement la matrice Hessienne $H_f(x)$, cf. section 3.6 et la relation (3.6).

Définition 3.7.2. *Soit $f \in D^2(O,\mathbb{R})$. On dit que f est de la classe $\mathcal{C}^2(O)$ (notation : $f \in \mathcal{C}^2(O)$), ssi $D^2 f(x)$ existe et*

$$D^2 f(x) \in \mathcal{C}^2(O),$$

où la continuité est entendue dans le sense de la norme $|||.|||$ sur $\mathcal{L}(\mathbb{R}^d,\mathbb{R}^d)$ (ou bien dans le sens de $|||.|||'$ sur $\mathcal{BL}(\mathbb{R}^d,\mathbb{R}^d;\mathbb{R})$), cf. les sections 3.1, 3.5.

Théorème 3.7.3 (formule de Taylor d'ordre 2)**.** *Soit $O \subset \mathbb{R}^d$ un ouvert, $f : O \to \mathbb{R}$ une fonction telle que $f \in \mathcal{C}^2(O,\mathbb{R})$. Soit aussi $x^0 \in O$ et $B(x^0,\delta) \subset O$ pour un $\delta > 0$. Alors, pour $h \in \mathbb{R}^d$, $||h|| < \delta$*

$$f(x^0+h) = f(x^0) + Df(x^0)\,h + \frac{1}{2}(D^2 f(x^0)h,h) + R_2(f,x^0;h), \tag{3.7}$$

où $R_2(f,x^0;h)$ est le reste intégral donné par

$$\begin{aligned} R_2(f,x^0;h) &= \int_0^1 (1-t)((D^2 f(x^0+th) - D^2 f(x^0))h,h)\,dt \\ &= h^t \left\{ \int_0^1 (1-t)((D^2 f(x^0+th) - D^2 f(x^0))\,dt \right\} h. \end{aligned}$$

En particulier, ceci implique que

$$\lim_{h\to 0} \frac{|R_2(f,x^0;h)|}{||h||^2} = 0.$$

On peut voir la relation (3.7) comme une "généralisation" de la formule (3.2) à l'ordre 2. Observons aussi que le troisième terme dans la partie de droite de (3.7) est donnée par la forme bilinéaire (quadratique) engendrée par $D^2 f(x^0)$, *i.e.,*

$$\frac{1}{2}(D^2 f(x^0)h,h) = \frac{1}{2}Q_{D^2 f(x^0)}(h,h).$$

Comme il était déjà mentionné, on peut songer à écrire la formule de Taylor d'ordre k pour des fonctions f de la classe $\mathcal{C}^k$. Cette construction ferait appel aux formes k-linéaires et elle est hors programme pour ce cours.

3.8 Étude de la nature des points extrémaux à l'aide de la formule de Taylor

Pour commencer, rappellons les définitions de points extrémaux. Comme d'habitude, nous avons $f : O \to \mathbb{R}$ et $x^0 \in O$.

Définition 3.8.1.

1. *Le point x^0 est le point de* maximum local*, s'il existe $B(x^0,\delta) \subset O$ telle que*

$$f(x) \le f(x^0), \quad \forall x \in B(x^0,\delta).$$

   Quelquefois, on dit que x^0 est un point de maximum local strict *s'il existe $B(x^0,\delta) \subset O$ telle que*

$$f(x) < f(x^0), \quad \forall x \in B(x^0,\delta)\backslash\{x^0\}.$$

2. *Le point x^0 est le point de* minimum local*, s'il y a $B(x^0,\delta) \subset O$ telle que*

$$f(x^0) \le f(x), \quad \forall x \in B(x^0,\delta).$$

   Quelquefois, on dit que x^0 est un point de minimum local strict *s'il existe $B(x^0,\delta) \subset O$ telle que*

$$f(x^0) < f(x), \quad \forall x \in B(x^0,\delta)\backslash\{x^0\}.$$

Les points de max. et de min. locaux s'appellent les extréma locaux.

3. Le point x^0 est le point de maximum global*, ssi*

$$f(x) \leq f(x^0), \quad \forall x \in O.$$

4. Le point x^0 est le point de minimum global*, ssi*

$$f(x^0) \leq f(x), \quad \forall x \in O.$$

Il y a une définition qui ne porte pas sur les points extrémaux, mais qui vient néanmoins dans la même liste : le point $x^0 \in O$ s'appelle un point-selle (un point de col), si pour tout $\delta > 0$ tel que $B(x^0, \delta) \subset O$, on a

$$\exists x \in B(x^0, \delta) : \ f(x) < f(x^0); \qquad \exists y \in B(x^0, \delta) : \ f(y) > f(x^0).$$

Proposition 3.8.2 (conditions nécessaires et conditions suffisantes pour les points extrémaux)**.** *Soit $O \subset \mathbb{R}^d$, $f : O \to \mathbb{R}$ et $f \in \mathcal{C}^2(O, \mathbb{R})$. De plus $x^0 \in O$.*

1. Si x^0 est un point extrémal, alors nécessairement x^0 est un point critique de f, c.à.d.

$$Df(x^0) = 0.$$

2. Si $Df(x^0) = 0$ et $D^2 f(x^0) > 0$, alors x^0 est un point de min. strict.

3. Si $Df(x^0) = 0$ et $D^2 f(x^0) < 0$, alors x^0 est un point de max. strict.

*4. Si $Df(x^0) = 0$ et $D^2 f(x^0) \gtrless 0$ (*i.e., *$D^2 f(x^0)$ est de signe indéfini), alors x^0 est le point-selle.*

Rappelons que la matrice de l'application $D^2 f(x)$ (ou bien de la forme bilinéaire correspondante) est la matrice Hessienne,

$$H_f(x) = \left[\frac{\partial^2 f}{{}_k\partial x_j}(x) \right]_{k,j=1,\dots,d}.$$

Cette forme est symétrique sous les hypothèses du théorème.

Vu les points 1. et 2. du théorème, on aurait la tendance à croire que si "$Df(x^0) = 0$ et $D^2 f(x^0) \geq 0$, alors x^0 est un point de min." (ainsi que "$Df(x^0) = 0$ et $D^2 f(x^0) \leq 0$, alors x^0 est un point de max."). Ceci est faux en général et ces situations nécessitent une étude spécifique cas par cas.

Fin

Bibliographie

[1] (sous direction de) Jean-Pierre Ramis, André Warusfel et al., Mathématiques Tout-en-un pour la Licence, Niveau L2 : Cours complets avec applications et 760 exercices corrigés, Editeur : Dunod (août 2007), Collection : Ramis, ISBN-10 : 2100507435, ISBN-13 : 978-2100507436.

[2] (sous direction de) Claude Deschamps, André Warusfel, Mathématiques tout-en-un 2e année MP-MP*, Editeur : Dunod (juin 2009), ISBN-10 : 2100534610, ISBN-13 : 978-2100534616.

[3] Jean-Marie Monier, Cours de mathématiques, tome 3 : Analyse, Cours et exercices corrigés, 2e année MP, PSI, PC, PT, Editeur : Dunod (août 2000), ISBN n'est pas disponible.

[4] Jean-Marie Monier, Analyse MP : Cours, méthodes et exercices corrigés, 783 pages, Editeur : Dunod ; Édition : 5e édition (mai 2007) Collection : J'intègre ISBN-10 : 2100510398 ISBN-13 : 978-2100510399.