Approfondissement en analyse

Licence de Mathématiques
Université Lyon 1

Dragoş Iftimie

Table des matières

1 Espaces normés

1.1 Norme

Soit E un espace vectoriel sur $\mathbb{R}$.

Définition. *On appelle norme sur E une application de E dans $\mathbb{R}_+$ habituellement notée $\| \ \|$ vérifiant pour tout $x, y \in E$ et tout $\lambda \in \mathbb{R}$*

i) $\|\lambda x\| = |\lambda| \, \|x\|$ *(homogénéité),*

ii) $(\|x\| = 0) \Rightarrow (x = 0)$ *(séparation),*

iii) $\|x + y\| \leq \|x\| + \|y\|$ *(inégalité triangulaire).*

Un espace vectoriel normé est un couple $(E, \| \ \|)$ où E est un espace vectoriel réel et $\| \ \|$ est une norme sur E.

Il est facile de voir que dans l'axiome de séparation **ii)** nous avons en fait l'équivalence.

Soit $(E, \| \ \|)$ un espace vectoriel normé. On définit la distance entre deux points x et y par $d(x, y) = \|x - y\|$.

Proposition 1. *Si $(E, \| \ \|)$ est un espace vectoriel normé alors la distance $d(x, y) = \|x - y\|$ vérifie les propriétés suivantes :*

a) $d \geq 0$ *;*

b) $d(x, y) = 0 \Leftrightarrow x = y$ *;*

c) $d(y, x) = d(x, y)$ *(symétrie) ;*

d) $\forall x, y, z \in E, \ d(x, z) \leq d(x, y) + d(y, z)$ *(inégalité triangulaire).*

e) $\forall x, y, z \in E, \ |d(x, y) - d(x, z)| \leq d(y, z)$ *(la distance entre les distances est plus petite que la distance).*

Exemples d'espaces normés.

a) La valeur absolue (ou le module) $| \ |$ est une norme sur $\mathbb{R}$.

b) La quantité

$$\|x\|_p = \begin{cases} \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}} & \text{pour } p < +\infty \\ \max_{i \in \{1, \ldots, n\}} |x_i| & \text{pour } p = +\infty. \end{cases}$$

définit une norme sur $\mathbb{R}^n$. Lorsque $p = 2$ nous obtenons la distance euclidienne $\| \cdot \|_2$.

c) La norme

$$\|x\|_p = \begin{cases} \left(\sum_{i=0}^{\infty} |x_i|^p \right)^{\frac{1}{p}} & \text{pour } p < +\infty \\ \sup_{i \in \mathbb{N}} |x_i| & \text{pour } p = +\infty. \end{cases}$$

sur l'espace des suites ℓ^p :

$$\ell^p = \{x = (x_i)_{i \geq 0} \ ; \ \|x\|_p < \infty\}.$$

d) Norme de la convergence uniforme sur $\mathscr{F}_b(X; \mathbb{R})$: Soit X un ensemble on munit $\mathscr{F}_b(X; \mathbb{R})$ de la norme

$$\forall f \in \mathscr{F}_b(X; \mathbb{R}), \|f\|_\infty = \sup_{x \in X} |f(x)|.$$

e) Soit I un intervalle fermé bornée de $\mathbb{R}$. La quantité

$$C^0(I,\mathbb{R}) \ni f \mapsto \|f\|_{L^p} = \left(\int_I |f|^p\right)^{\frac{1}{p}}$$

définit une norme sur $C^0(I,\mathbb{R})$ (= l'ensemble des fonctions continues de I dans $\mathbb{R}$).

Proposition 2 (Inégalité de Minkowski)**.**

a) Pour tout $x, y \in \mathbb{R}^n$ nous avons que

$$\|x+y\|_p \leq \|x\|_p + \|y\|_p.$$

b) Soit I un intervalle fermé bornée de $\mathbb{R}$. Pour tout $f, g \in C^0(I,\mathbb{R})$ nous avons que

$$\|f+g\|_{L^p} \leq \|f\|_{L^p} + \|f\|_{L^p}.$$

Démonstration. Voici les étapes de la preuve de la partie a).

- On montre dans un premier temps que pour $a, b \in \mathbb{R}_+$ et $\frac{1}{p} + \frac{1}{q} = 1$ on a

$$ab \leq \frac{1}{p}a^p + \frac{1}{q}a^q.$$

  Ceci est une conséquence de la convexité de la fonction $x \to e^x$.
- Inégalité de Hölder : On montre ensuite que pour $(a_1, \ldots, a_n)$, $(b_1, \ldots b_n) \in \mathbb{R}^n$, on a

$$\left|\sum_{i=1}^n a_i b_i\right| \leq \left(\sum_{i=1}^n |a_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^n |b_i|^q\right)^{\frac{1}{q}} \qquad \text{avec } \frac{1}{p} + \frac{1}{q} = 1.$$

  On commence par montrer cette relation dans le cas particulier où $\sum_i^n |a_i|^p = \sum_{i=1}^n |b_i|^q = 1$, puis on la déduit dans le cas général.
- En partant de $|a_i + b_i|^p \leq |a_i + b_i|^{p-1}(|a_i| + |b_i|)$, on montre enfin que pour $(a_1, \ldots, a_n), (b_1, \ldots, b_n) \in \mathbb{R}^n$ on a

$$\left(\sum_{i=1}^n |a_i + b_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^n |a_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |b_i|^p\right)^{\frac{1}{p}}.$$

La preuve de la partie b) est similaire. □

1.2 Boules. Ensembles ouverts, fermés

Boules

Définition. *Soit $(E, \|\cdot\|)$ un espace normé, soit $x \in X$ et soit $r > 0$. On appelle boule ouverte (resp. boule fermée) de centre x et de rayon r l'ensemble*

$$B(x,r) = \{y \in E,\ d(x,y) < r\}$$
$$\textit{resp.} \qquad \overline{B}(x,r) = \{y \in E,\ d(x,y) \leq r\}.$$

On appelle ***ouvert de*** *E toute partie O de E qui vérifie*

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

On appelle ***fermé*** *toute partie de E dont le complémentaire est un ouvert.*

Proposition 3. *Une boule ouverte est un ensemble ouvert.*

Démonstration. Soit $B(x_0, r_0)$ une boule ouverte de (X, d). Soit $x \in B(x_0, r_0)$. On a $d(x_0, x) < r_0$ et on pose $\rho = \frac{r_0 - d(x_0,x)}{2}$. Alors la boule $B(x, \rho)$ est incluse dans $B(x_0, r_0)$. En effet pour $y \in B(x, \rho)$ on a

$$d(x_0, y) \leq d(x_0, x) + d(x, y) < d(x_0, x) + \frac{r_0 - d(x_0, x)}{2} = \frac{r_0 + d(x_0, x)}{2} < r_0.$$

□

1.3 Suites convergentes

Définition. *Soit $(E, \|\cdot\|)$ un espace normé. Soit $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de E et soit $x \in E$. On dit que la suite $(x_n)_{n\in\mathbb{N}}$ tend vers x, et on note $x_n \to x$, quand n tend vers l'infini si $d(x_n, x) = \|x_n - x\| \to 0$ quand $n \to \infty$. De manière équivalente,*

$$\forall \varepsilon > 0, \quad \exists N_\varepsilon \in \mathbb{N}, \quad \forall n \geq N_\varepsilon, \quad d(x_n, x) \leq \varepsilon.$$

Nous avons la caractérisation suivante des fermés par les suites convergentes.

Proposition 4. *Un ensemble F est fermé si et seulement si il contient toute limite de suite d'éléments de F. Plus précisément : $x_n \in F$ et $x_n \to x$ implique $x \in X$.*

Corollaire 5. *Toute boule fermée est un fermé.*

1.4 Continuité

Définition. *Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces normés et $f : E \to F$ une application.*

- *On dit que f admet la limite b au point a, et on note $\lim\limits_{x\to a} f(x) = b$ si $x_n \to a$ entraîne que $f(x_n) \to b$.*
- *On dit que f est continue en a si $\lim\limits_{x\to a} f(x) = f(a)$.*
- *On dit que f est continue (sur E) si f est continue en tout point.*

Proposition 6. *Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces normés et $f : E \to F$ une application. Alors f est continue si et seulement si l'image inverse d'un ouvert et un ouvert. Ou encore si et seulement si l'image inverse d'un fermé est un fermé.*

1.5 Applications linéaires et continues

Proposition 7. *Soit $(E, \|\ \|_E)$ et $(F, \|\ \|_F)$ deux espaces vectoriels normés. Une application linéaire $f : E \to F$ est continue si et seulement si il existe une constante $M > 0$ telle que*

$$\forall x \in E, \|f(x)\|_F \leq M \|x\|_E .$$

Démonstration. $\boxed{\Rightarrow}$ Si f est une application linéaire continue de E dans F, il existe $\rho > 0$ tel que la boule fermée de E, $B_{f,E}(0,\rho)$, soit incluse dans $f^{-1}(B_{f,F}(0,1))$. Pour $x \in E \setminus \{0\}$, le vecteur $\frac{\rho}{\|x\|_E} x$ appartient à $B_{f,E}(0,\rho)$ d'où

$$\frac{\rho}{\|x\|_E} \|f(x)\|_F = \left\| f\left(\frac{\rho}{\|x\|_E} x\right) \right\|_F \leq 1.$$

Il suffit de prendre $M = \frac{1}{\rho}$.
$\boxed{\Leftarrow}$ Si on a la constante M alors on a

$$\forall x, y \in E,\ \|f(y) - f(x)\|_F = \|f(y-x)\|_F \leq M \|y - x\|_E$$

et l'application linéaire f est continue. □

Exemples.

a) Si on munit $\mathscr{C}^0([0,1];\mathbb{R})$ de la norme $\|f\|_\infty = \sup\limits_{[0,1]} |f|$, alors la forme linéaire $f \to f(1/2)$ est continue.

b) Même exemple mais en remplaçant la norme $\|\cdot\|_\infty$ par la norme $\|f\|_1 = \int_0^1 |f(t)|\, dt$. La forme linéaire $f \to f(1/2)$ n'est pas continue.

Voici une application à l'équivalence des normes.

Définition (équivalence des normes). *Deux normes $\|\cdot\|, \|\cdot\|'$ sur un même espace vectoriel X sont dites équivalentes s'il existe $C, C' > 0$ tels que $C\|x\| \leq \|x\|' \leq C'\|x\|$ pour tout $x \in X$.*

Proposition 8. *Deux normes sur un espace vectoriel E définissent les mêmes ouverts si et seulement si elles sont équivalentes.*

Définition. *Les normes $\| \ \|_E$ et $\| \ \|_F$ étant fixées, on note $\mathscr{L}(E;F)$ l'espace des applications linéaires continues de E dans F. On appelle* ***dual topologique*** *de E et on note $E' = \mathscr{L}(E;\mathbb{R})$ l'espace des formes linéaires continues sur E.*

Proposition 9. *L'espace $\mathscr{L}(E;F)$ est un espace normé avec la norme triple suivante :*

$$\begin{aligned}|||f|||_{\mathscr{L}(E;F)} &= \sup_{x\neq 0} \frac{\|f(x)\|_F}{\|x\|_E} = \sup_{\|x\|_E=1} \|f(x)\|_F = \sup_{\|x\|_E\leq 1} \|f(x)\|_F = \sup_{\|x\|_E<1} \|f(x)\|_F \\ &= \inf\{M \ ; \ \forall x \in E, \|f(x)\|_F \leq M \|x\|_E\} = \min\{M \ ; \ \forall x \in E, \|f(x)\|_F \leq M \|x\|_E\}.\end{aligned}$$

Remarque. *Nous avons toujours l'inégalité*

$$\|f(x)\|_F \leq |||f|||_{\mathscr{L}(E;F)}\|x\|_E.$$

La norme triple est d'ailleurs la plus petite constante C avec la propriété que

$$\|f(x)\|_F \leq C\|x\|_E.$$

On termine avec une estimation de la norme de la composée de deux applications linéaire continues.

Proposition 10. *Si $(E, \| \ \|_E)$, $(F, \| \ \|_F)$ et $(G, \| \ \|_G)$ sont trois espaces vectoriels normés alors pour $f \in \mathscr{L}(E;F)$ et $g \in \mathscr{L}(F;G)$ la composée $g \circ f$ appartient à $\mathscr{L}(E;G)$ et on a*

$$|||g \circ f|||_{\mathscr{L}(E;G)} \leq |||g|||_{\mathscr{L}(F;G)}|||f|||_{\mathscr{L}(E;F)}.$$

Démonstration. Il est clair que la composée $g \circ f$ est linéaire et continue. De plus pour $x \in E$ on a

$$\|g \circ f(x)\|_G = \|g\,[f(x)]\|_G \leq \|g\| \, \|f(x)\|_F \leq \|g\| \, \|f\| \, \|x\|_E \,,$$

d'où la majoration de la norme de $g \circ f$. $\square$

1.6 Dimension finie

Les espaces normés de dimension finie ont un certain nombre de propriétés qui les distinguent des autres espaces normés. Nous les étudions dans cette partie.

Commençons par définir ce que c'est un ensemble compact (notion non restreinte aux espaces de dimension finie).

Définition. *Soit E un espace normé et $A \subset E$. On dit que A est compact si de toute suite (x_n) de A un peut extraire une sous-suite (x_{n_k}) qui converge dans A : $x_{n_k} \to \ell$ et $\ell \in A$.*

Une caractérisation utile de la compacité en dimension finie est donnée par le théorème de Bolzano-Weierstrass qui affirme que les compacts en dimension finie sont les ensembles fermés et bornés. Commençons par le cas de la dimension 1.

Théorème 11 (Heine-Borel-Lebesgue). *Tout intervalle fermé borné de $\mathbb{R}$, $[a, b]$, $a, b \in \mathbb{R}$, est compact.*

Démonstration. Soit $(x_n)_{n\in\mathbb{N}}$ une suite de $[a,b]$, on va extraire une sous-suite convergente par une méthode de dichotomie. On construit la suite de couple $((a_k,b_k))_{k\in\mathbb{N}}$ par récurrence :

- $a_0 = a$ et $b_0 = b$
- On pose $c_k = \frac{a_k+b_k}{2}$. Si il existe un nombre infini d'indices n tels que $x_n \in [a_k, c_k]$ on prend $a_{k+1} = a_k$ et $b_{k+1} = c_k$. Sinon on prend $a_{k+1} = c_k$ et $b_{k+1} = b_k$.

On a formé ainsi deux suites adjacentes $(a_k)_{k\in\mathbb{N}}$ et $(b_k)_{k\in\mathbb{N}}$ telles que pour tout $k \in \mathbb{N}$ il existe $x_{n_k} \in [a_k, b_k]$. Une conséquence de la propriété de la borne supérieure est que deux suites adjacentes convergent et ont même limite. Cette limite est aussi limite de la sous-suite $(x_{n_k})_{k\in\mathbb{N}}$. □

Le résultat précédent implique le théorème de Bolzano-Weierstrass suivant :

Théorème 12 (Bolzano-Weierstrass)**.** *Dans $(\mathbb{R}^n, \|\cdot\|_\infty)$, les ensembles compacts sont les ensembles fermés et bornés.*

Lemme 13. *Une fonction continue sur un compact est bornée et atteint ses bornes.*

Voici un théorème qui regroupe plusieurs propriétés des espaces normés de dimension finie.

Théorème 14. *a) Dans un espace normé de dimension finie toutes les normes sont équivalentes.*

b) Dans un espace normé de dimension finie les ensembles compacts sont les ensembles fermés et bornés.

c) Dans un espace normé quelconque, tout sous-espace vectoriel de dimension finie est fermé.

d) Toute application linéaire définie sur un espace normé de dimension finie à valeurs dans un espace normé quelconque (pas forcément de dimension finie) est continue.

Lemme 15. *Soit E un espace normé quelconque et F un sous-espace vectoriel fermé strictement inclus dans E. Il existe $x \in E$ tel que $\|x\| = 1$ et $d(x, F) \geq \frac{1}{2}$.*

Le théorème suivant montre la réciproque du théorème de Bolzano-Weierstrass.

Théorème 16 (Riesz)**.** *Soit $(E, \|\cdot\|)$ un espace normé (pas forcément de dimension finie). La boule unité fermée est compacte si et seulement si E est de dimension finie.*

2 Calcul différentiel

Dans cette partie nous considérons $\mathbb{R}^n$ muni de la norme euclidienne $\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$. Mais toutes les normes étant équivalentes sur $\mathbb{R}^n$, les résultats qui suivent restent vrais avec toute autre norme.

2.1 Définitions et premières propriétés

Définition. *Soit U un ouvert de $\mathbb{R}^n$, $a \in U$ et $f : U \to \mathbb{R}^m$. On dit que f est différentiable en a s'il existe une application linéaire $L : \mathbb{R}^n \to \mathbb{R}^m$ telle que*

$$f(a+h) = f(a) + L(h) + o(\|h\|) \qquad \textit{quand } h \to 0.$$

(On a utilisé la notation habituelle $X = o(\|h\|)$ si $\frac{X}{\|h\|} \to 0$ quand $h \to 0$.)

On appelle L la différentielle de f au point a et on note $L = Df(a)$ ou encore $L = f'(a)$ lorsqu'il n'y a pas de risque de confusion avec la dérivée usuelle.

On dit que f est différentiable sur U si f est différentiable en tout point de U.

Remarques.

a) Pour pouvoir étudier la différentiabilité d'une fonction en un point il faut que la fonction soit définie au voisinage de ce point.

b) La notion de différentiabilité ne change pas quand on change les normes de $\mathbb{R}^n$ et $\mathbb{R}^m$. En effet, toutes les normes sur $\mathbb{R}^n$ et $\mathbb{R}^m$ sont équivalentes.

Proposition 17. *L'application linéaire L qui apparaît dans la définition de la différentiabilité est unique.*

Proposition 18. *Une fonction différentiable en a est continue en a.*

Exemples.

a) Toute application linéaire est différentiable en tout point et sa différentielle en un point arbitraire est elle-même.

b) L'application $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x) = x_1 x_2$ est différentiable en tout point et sa différentielle est donnée par

$$Df(a)(h) = a_1 h_2 + a_2 h_1.$$

c) Une fonction $f : \mathbb{R} \to \mathbb{R}$ est différentiable en un point a si et seulement si elle est dérivable. De plus, la différentielle $Df(a)$ est l'application linéaire donnée par la multiplication par la dérivée $f'(a)$:

$$Df(a)(h) = hf'(a).$$

2.2 Dérivées directionnelles

Définition. *Soit U un ouvert de $\mathbb{R}^n$, $a \in U$, $v \in \mathbb{R}^n$ et $f : U \to \mathbb{R}^m$. On dit que f admet au point a une dérivée directionnelle suivant la direction v, et on la note par $\frac{\partial f}{\partial v}(a)$, si la limite suivante existe :*

$$\frac{\partial f}{\partial v}(a) = \lim_{\varepsilon \searrow 0} \frac{f(a + \varepsilon v) - f(a)}{\varepsilon}.$$

En d'autres termes, la dérivée directionnelle $\frac{\partial f}{\partial v}(a)$ est la dérivée à droite en 0 de la fonction $t \mapsto f(a + tv)$.

Proposition 19. *Si f est différentiable en a alors f admet des dérivées directionnelles en a suivant toute direction et nous avons de plus que*

$$\frac{\partial f}{\partial v}(a) = Df(a)(v).$$

Exemples.

a) L'existence des dérivées directionnelles suivant toute direction n'entraîne pas forcément la différentiabilité de la fonction. Ni même la continuité. La fonction $f : \mathbb{R}^2 \to \mathbb{R}$

$$f(x, y) = \begin{cases} 1 & \text{si } y \geq x^2 \\ 1 & \text{si } y \leq 0 \\ 0 & \text{sinon} \end{cases}$$

admet des dérivées directionnelles en 0 qui sont nulles en toute direction. Mais f n'est pas continue en 0, et *a fortiori* n'est pas différentiable en 0.

b) Même si la fonction est continue et que toutes ses dérivées directionnelles existent en un point cela n'implique toujours pas que la fonction est différentiable. La fonction $f : \mathbb{R}^2 \to \mathbb{R}$

$$f(x,y) = \begin{cases} \frac{x(x^2-3y^2)}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$$

est continue en 0 et toutes ses dérivées directionnelles en 0 existent :

$$\frac{\partial f}{\partial v}(0) = f(v) \qquad \forall v \in \mathbb{R}^2.$$

Elle n'est cependant pas différentiable en 0 car les dérivées directionnelles en 0 ne sont pas linéaires en v.

2.3 Dérivées partielles

Définition. *Soit U un ouvert de $\mathbb{R}^n$, $a \in U$ et $f : U \to \mathbb{R}^m$. On dit que f admet au point a une dérivée partielle par rapport à la variable x_j, et on la note par $\frac{\partial f}{\partial x_j}(a)$, si la limite suivante existe :*

$$\frac{\partial f}{\partial x_j}(a) = \lim_{\varepsilon \to 0} \frac{f(a+\varepsilon e_j) - f(a)}{\varepsilon}$$

où on a noté par e_j le j-ème élément de la base canonique de $\mathbb{R}^n$: toutes les composantes de e_j sont nulles sauf la j-ème qui est égale à 1. En d'autres termes, la dérivée partielle $\frac{\partial f}{\partial x_j}(a)$ est la dérivée en a_j de la fonction $t \mapsto f(a_1, \ldots, a_{j-1}, t, a_{j+1}, \ldots, a_n)$ (on fige toutes les variables sauf x_j et on dérive par rapport à x_j).

La différentielle peut s'exprimer en fonction des dérivées partielles de la manière suivante.

Proposition 20. *Soit U un ouvert de $\mathbb{R}^n$, $a \in U$ et $f : U \to \mathbb{R}^m$ une fonction différentiable en a. Alors*

$$Df(a)(h) = \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a).$$

2.4 Matrice jacobienne, gradient

Une fonction f à valeurs dans $\mathbb{R}^m$ admet m composantes

$$f = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{pmatrix}$$

Comme la limite dans $\mathbb{R}^m$ se fait composante par composante et que les dérivées partielles sont définies via une limite, nous avons que f admet des dérivées partielles ssi chaque composante de f admet des dérivées partielles et la dérivée partielle de f s'obtient en prenant les dérivées partielles des composantes.

La différentielle est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$. Elle s'identifie donc à une matrice. On peut exprimer cette matrice en fonction des dérivées partielles des composantes de f.

Proposition 21. *Soit U un ouvert de $\mathbb{R}^n$, $a \in U$ et $f : U \to \mathbb{R}^m$ une fonction différentiable en a. La matrice de $Df(a)$ dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^m$ est donnée par la matrice suivante*

$$M_f(a) = \left(\frac{\partial f_i}{\partial x_j}(a)\right)_{\substack{1\leq i\leq m \\ 1\leq j\leq n}} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \ldots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \ldots & \frac{\partial f_2}{\partial x_n}(a) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & \ldots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix}$$

On appelle cette matrice la matrice jacobienne en a.

Définition. *Soit U un ouvert de $\mathbb{R}^n$, $a \in U$ et $f : U \to \mathbb{R}$ une fonction différentiable en a. On appelle gradient de f en a le vecteur ligne*

$$\nabla f(a) = \big(\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), \dots, \frac{\partial f}{\partial x_n}(a)\big).$$

Le gradient coïncide avec la matrice jacobienne. On peut facilement voir que le gradient est la direction où f augmente le plus vite.

2.5 Continuité des dérivées partielles et différentiabilité

Le critère le plus important pour la différentiabilité d'une fonction est celui de la continuité des dérivées partielles.

Théorème 22. *Soit U un ouvert de $\mathbb{R}^n$, $a \in U$ et $f : U \to \mathbb{R}^m$. On suppose que les dérivées partielles de f existent dans un voisinage de a et sont continues en a. Alors f est différentiable en a.*

Si la continuité des dérivées partielles est une condition suffisante de différentiabilité, ce n'est pas une condition nécessaire (seule l'existence des dérivées partielles est une condition nécessaire). En pratique, lorsqu'on veut décider de la différentiabilité d'une fonction concrète (qui a en général des points de singularité) on peut procéder de la manière suivante :

1. On étudie la continuité de la fonction. Si la fonction n'est pas continue elle n'est pas différentiable.
2. Si la fonction est continue, on étudie l'existence des dérivées partielles. Si au moins une des dérivées partielles n'existe pas, la fonction n'est pas différentiable.
3. Si la fonction est continue et toutes les dérivées partielles existent, on étudie la continuité des dérivées partielles. Si toutes les dérivées partielles sont continues alors la fonction est différentiable.
4. Si la fonction est continue et toutes les dérivées partielles existent mais certaines sont discontinues, il ne reste plus qu'à vérifier la définition de la différentiabilité. Mais l'application linéaire et continue L de la définition est connue (sa matrice est la matrice des dérivées partielles, la matrice jacobienne) donc la vérification de la définition est maintenant aisée.

Exemple. La fonction

$$f : \mathbb{R}^2 \to \mathbb{R}, \qquad f(x,y) = \begin{cases} x^2 \sin \frac{1}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$$

est continue en $(0,0)$, les dérivées partielles existent partout mais sont discontinues en $(0,0)$. La fonction est différentiable en $(0,0)$ de différentielle nulle.

2.6 Composition et inverse

Nous avons que la composition de deux applications différentiables est différentiable et la différentielle de la composition est la composition des différentielles.

Proposition 23. *Soient U un ouvert de $\mathbb{R}^n$, V un ouvert de $\mathbb{R}^m$, $f : U \to \mathbb{R}^m$ et $g : V \to \mathbb{R}^k$. On suppose que f est différentiable en $a \in U$, que $f(a) \in V$ et que g est différentiable en $f(a)$. Alors $g \circ f$ est différentiable en a et*

$$D(g \circ f)(a) = Dg(f(a)) \circ Df(a).$$

Corollaire 24. *Si $f = f(y)$ est une fonction différentiable à valeurs réelles définie sur un ouvert de $\mathbb{R}^m$ et $g = g(x)$ est définie d'un ouvert de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^m$, alors*

$$\frac{\partial}{\partial x_i}\big(f(g_1, g_2, \ldots, g_m)\big) = \frac{\partial f}{\partial y_1}\frac{\partial g_1}{\partial x_i} + \frac{\partial f}{\partial y_2}\frac{\partial g_2}{\partial x_i} + \cdots + \frac{\partial f}{\partial y_m}\frac{\partial g_m}{\partial x_i}.$$

La différentielle de l'inverse est l'inverse de la différentielle. Plus précisément :

Théorème 25. *Soient U et V deux ouverts de $\mathbb{R}^n$ et $f : U \to V$ une bijection continue telle que f^{-1} est continue. On suppose que f est différentiable en un point $a \in U$ et que sa différentielle $Df(a)$ est inversible. Alors f^{-1} est différentiable en $f(a)$ et*

$$D(f^{-1})(f(a)) = \big(Df(a)\big)^{-1}.$$

2.7 Inégalité des accroissements finis

Commençons par montrer l'inégalité des accroissements finis dans un cas particulier.

Lemme 26. *Soit $f : [a, b] \to \mathbb{R}^n$ une fonction de classe C^1. Nous avons que*

$$\|f(a) - f(b)\| \le (b - a) \sup_{t\in[a,b]} \|f'(t)\|.$$

Voici maintenant le cas général :

Théorème 27. *Soit $f : U \to \mathbb{R}^m$ une fonction C^1 où U est un ouvert de $\mathbb{R}^n$. Soient $x, y \in U$ tels que le segment $[x, y] = \{tx + (1 - t)y \ ;\ t \in [0, 1]\}$ soit inclus dans U. Nous avons l'inégalité suivante*

$$\|f(x) - f(y)\| \le \|x - y\| \sup_{z\in[x,y]} \|Df(z)\|_{\mathscr{L}(\mathbb{R}^n,\mathbb{R}^m)}.$$

Remarque. *La norme de $Df(z)$ est calculée en tant qu'application linéaire et continue. Attention il faut utiliser les bonnes normes dans $\mathbb{R}^n$ et $\mathbb{R}^m$, les mêmes que celles utilisées pour $\|f(x) - f(y)\|$ et $\|x - y\|$.*

Corollaire 28. *Soit $f : U \to \mathbb{R}^m$ où U est un ouvert convexe de $\mathbb{R}^n$. On suppose que f est C^1 sur U et que sa différentielle est nulle en tout point. Alors f est constante.*

2.8 Dérivées d'ordre supérieur

On peut définir les fonctions de classe C^k en imposant que les dérivées partielles jusqu'à l'ordre k existent et soient continues.

Définition. *Soit $f : U \to \mathbb{R}^m$ où U est un ouvert de $\mathbb{R}^n$ et $k \ge 1$. On dit que f est de classe C^k si toutes les dérivées partielles d'ordre $\le k$ existent et sont continues.*

Un résultat très important affirme que les dérivées partielles commutent.

Théorème 29 (Schwarz)**.** *Soit f une fonction de classe C^2. Alors pour tout i et j nous avons que*

$$\frac{\partial}{\partial x_i}\Big(\frac{\partial f}{\partial x_j}\Big) = \frac{\partial}{\partial x_j}\Big(\frac{\partial f}{\partial x_i}\Big).$$

Nous noterons cette valeur commune par

$$\frac{\partial^2 f}{\partial x_i \partial x_j} := \frac{\partial}{\partial x_i}\Big(\frac{\partial f}{\partial x_j}\Big) = \frac{\partial}{\partial x_j}\Big(\frac{\partial f}{\partial x_i}\Big).$$

La formule de Taylor à l'ordre deux peut s'écrire sous la forme suivante.

Théorème 30 (formule de Taylor)**.** *Soit f une fonction de classe C^2. Nous avons que*

$$f(a + h) = f(a) + \sum_i h_i \frac{\partial f}{\partial x_i}(a) + \frac{1}{2}\sum_{i,j} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a) + o(\|h\|^2)$$

quand $h \to 0$.

2.9 Extrema

Nous nous intéressons maintenant aux extrema d'une fonction, ou plus précisément comment les déterminer. Définissons d'abord les points d'extremum.

Définition. *Soit $f : A \subset \mathbb{R}^n \to \mathbb{R}$ une fonction.*

- *Un point d'extremum de f est un point a tel que f admet un maximum en a ($f(x) \leq f(a)$ pour tout $x \in A$) ou un minimum en a ($f(x) \geq f(a)$ pour tout $x \in A$).*
- *Un point d'extremum strict de f est un point a tel que f admet un maximum strict en a ($f(x) < f(a)$ pour tout $x \in A \setminus \{a\}$) ou un minimum strict en a ($f(x) > f(a)$ pour tout $x \in A \setminus \{a\}$).*
- *L'extremum est global si l'inégalité a lieu sur tout le domaine de définition de la fonction.*
- *L'extremum est local si l'inégalité a lieu sur une petite boule centrée en a.*

Un rôle important dans l'étude des extrema est joué par les points critiques.

Définition. *Un point critique d'une fonction f différentiable à valeurs réelles est un point a où le gradient de f s'annule : $\nabla f(a) = 0$.*

Nous avons que les points d'extrema sont des points critiques.

Proposition 31. *Tout point d'extremum local est un point critique.*

Il faut donc chercher les points d'extrema parmi les points critiques. Pour faire cela, nous nous appuyons sur la matrice hessienne.

Définition. *Soit $f : U \to \mathbb{R}$ où U est un ouvert de $\mathbb{R}^n$ une fonction de classe C^2. La matrice hessienne f en a, ou plus simplement la hessienne de f en a, est la matrice*

$$H_f(a) = \Big(\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\Big)_{1 \leq i,j \leq n}.$$

Par le théorème de Schwarz, la matrice hessienne est symétrique, donc diagonalisable. En fonction du signe de ses valeurs propres nous pouvons caractériser les points d'extrema.

Théorème 32. *Soit $f : U \to \mathbb{R}$ où U est un ouvert de $\mathbb{R}^n$ une fonction de classe C^2. Soit $a \in U$ un point critique.*

a) Si a est un point de minimum local, alors la matrice hessienne $H_f(a)$ est positive (de manière équivalente, ses valeurs propres sont toutes ≥ 0).

b) Si a est un point de maximum local, alors la matrice hessienne $H_f(a)$ est négative (de manière équivalente, ses valeurs propres sont toutes ≤ 0).

c) Si la matrice hessienne $H_f(a)$ est définie positive (de manière équivalente, si ses valeurs propres sont toutes strictement positives), alors a est un point de minimum local strict.

d) Si la matrice hessienne $H_f(a)$ est définie négative (de manière équivalente, si ses valeurs propres sont toutes strictement négatives), alors a est un point de maximum local strict.

e) Si la matrice hessienne $H_f(a)$ admet à la fois des valeurs propres strictement positives et strictement négatives, alors a n'est pas un point d'extremum. Dans ce cas, on appelle le point a un point selle.

En dimension deux, nous pouvons facilement trouver les signes des valeurs propres de la hessienne à partir du déterminant et de la trace sans avoir à calculer explicitement les valeurs propres. En effet, le déterminant est le produit des valeurs propres et la trace est la somme des valeurs propres. Nous obtenons alors la caractérisation suivante.

- Si $\det H_f(a) > 0$ et $\operatorname{tr} H_f(a) > 0$ alors a est un point de minimum local strict.
- Si $\det H_f(a) > 0$ et $\operatorname{tr} H_f(a) < 0$ alors a est un point de maximum local strict.
- Si $\det H_f(a) < 0$ alors a est un point selle (pas d'extremum dans ce point).