Agrégation interne
2016/2017
Mathématiques

Calcul différentiel

1 Quelques rappels de topologie : normes, distances, etc.

Définition 1.1. *Soit E un espace vectoriel. Une application $N : E \to \mathbb{R}^+$ est une norme si*

1. *$N(x) = 0$ si et seulement si $x = 0$,*
2. *$N(\lambda x) = |\lambda| N(x)$ pour tout $\lambda \in \mathbb{R}$ et pour tout $x \in E$,*
3. *$N(x + y) \leq N(x) + N(y)$ pour tous $x, y \in E$.*

*Un espace vectoriel E muni d'une norme N, noté (E, N), est appelé un espace vectoriel normé (*evn*).*

Notation : Les normes sont souvent notées $\| \ \|$.

Exemple 1.1. *Soit $E = \mathbb{R}^n$. On note $x = (x_1, \dots, x_n)$ un élément de E. Les quantités ci-dessous définissent des normes sur E :*

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|, \quad \|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}, \quad \|x\|_\infty = \max_{i=1,\dots,n} |x_i|.$$

Définition 1.2. *Soit E un espace vectoriel et N_1, N_2 deux normes sur E. On dit que N_1 et N_2 sont équivalentes s'il existe $C, C' \in \mathbb{R}_+^*$ tels que*

$$\forall x \in E, \quad C N_1(x) \leq N_2(x) \leq C' N_1(x). \tag{1}$$

Théorème 1.1. *Si E est de dimension finie alors toutes les normes sur E sont équivalentes.*

Définition 1.3. *Soit E un ensemble. Une distance est une application $d : E \times E \to \mathbb{R}_+$ telle que*

1. *$d(x, y) = d(y, x)$ pour tous x, y dans E,*
2. *$d(x, y) = 0$ si et seulement si $x = y$,*
3. *$d(x, z) \leq d(x, y) + d(y, z)$ pour tous x, y, z dans E.*

Un ensemble E muni d'une distance d, noté (E, d), est appelé espace métrique, en abrégé e.m.

Exemple 1.2. *Si $(E, \| \ \|)$ est un evn, alors $d(x, y) := \|x - y\|$ définit une distance sur E. On dira que d est la distance associée à la norme $\| \ \|$. Si $A \subset E$ (pas nécessairement un sous-espace) est non vide alors (A, d) est un espace métrique.*

L'exemple ci-dessus fournit une classe importante d'espaces métriques, ce ne sont cependant pas les seuls.

Définition 1.4. *Soit E un ensemble et d_1, d_2 deux distances sur E. On dit que d_1 et d_2 sont équivalentes s'il existe $C, C' \in \mathbb{R}_+^*$ tels que*

$$\forall x, y \in E, \quad Cd_1(x,y) \leq d_2(x,y) \leq C'd_1(x,y).$$

Remarque 1.1. *Si E est un espace vectoriel et N_1, N_2 deux normes sur E, alors les distances d_1 et d_2 associées respectivement à N_1 et N_2 sont équivalentes ssi les normes N_1 et N_2 le sont.*

Définition 1.5. *Soit $x \in E$ et $r \geq 0$. On appelle boule ouverte, resp. fermée, de centre x et de rayon r l'ensemble $B(x,r) = \{y \in E \mid d(x,y) < r\}$, resp $\bar{B}(x,r) = \{y \in E \mid d(x,y) \leq r\}$.*

Remarque 1.2. *La notion de boule dépend du choix de la distance. Si d_1 et d_2 sont deux distances sur E, en général $B_{d_1}(x,r) \neq B_{d_2}(x,r)$. De même la notion de boule dépend de l'espace lui-même. Par exemple, dans $(\mathbb{R}, |\,|)$ on a $B(1,2) =]-1,3[$ alors que dans $(\mathbb{R}_+, |\,|)$ on a $B(1,2) = [0,3[$.*

Définition 1.6. *Soit $x \in E$. Un ensemble $V \subset E$ est appelé un voisinage de x s'il existe $r > 0$ tel que $B(x,r) \subset V$.*

Définition 1.7. *Un ensemble $O \subset E$ est dit ouvert si pour tout $x \in O$ il existe $r > 0$ tel que $B(x,r) \subset O$, i.e. si O est voisinage de chacun de ses points.*

Remarque 1.3. *L'ensemble vide est un ensemble ouvert ainsi que E lui même.*

Définition 1.8. *Un ensemble $F \subset E$ est dit fermé si son complémentaire $F^c = E \backslash F$ est ouvert.*

Proposition 1.2. *Toute réunion d'ensembles ouverts est un ouvert et toute intersection finie d'ensembles ouverts est un ouvert. Toute intersection de fermés est un fermé et toute réunion finie de fermés est un fermé.*

Tout comme pour la notion de boule, les notions de voisinage, ouvert, fermé dépendent a priori du choix de la distance (ainsi que de l'espace E lui-même).

Exemple 1.3. *L'ensemble $[0,1[$ n'est pas un ouvert dans $(\mathbb{R}, |\,|)$ mais c'en est un dans $(\mathbb{R}_+, |\,|)$.*

Cependant, on a

Proposition 1.3. *Si d_1 et d_2 sont deux distances équivalentes alors les notions de voisinage, ouvert, fermé coïncident, i.e. tout ouvert pour d_1 est un ouvert pour d_2 et réciproquement, et de même pour les fermés, les voisinages etc.*

Proposition 1.4. *Soit (E,d) un espace métrique et $A \subset E$. On note $\overset{\circ}{A}$ le plus grand ouvert inclus dans A, appelé intérieur de A, et $\bar{A}$ est le plus petit fermé contenant A, appelé adhérence de A. On a par ailleurs,*

$$\overset{\circ}{A} := \{x \in E \,|\, \exists r > 0,\ B(x,r) \subset A\} \quad \text{et} \quad \bar{A} := \{x \in E \,|\, \forall r > 0,\ B(x,r) \cap A \neq \emptyset\}.$$

Définition 1.9. *Soit $u = (u_n)_n \in E^{\mathbb{N}}$ une suite d'éléments de E. On dit que la suite u converge s'il existe $\ell \in E$ tel que*

$$\forall \epsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \geq N \quad d(u_n, \ell) < \epsilon,$$

autrement dit si la suite de nombres réels $(d(u_n, \ell))_n$ tend vers 0 (dans $\mathbb{R}$).

Proposition 1.5. *Soit (E,d) un espace métrique et $F \subset E$. L'ensemble F est fermé si et seulement si, pour toute suite $(u_n)_n \in F^{\mathbb{N}}$, si $(u_n)_n$ converge alors sa limite est dans F.*

Définition 1.10. *Un ensemble $K \subset E$ est dit compact si toute suite $(u_n)_n$ d'éléments de K admet une sous-suite qui converge dans K.*

On rappelle qu'un ensemble A est borné s'il est inclus dans une boule (dans un evn celà revient à dire qu'il existe $M > 0$ tel que pour tout $x \in A$ on a $\|x\| \leq M$).

Théorème 1.6. *Si E est un evn de dimension finie alors un sous-ensemble K de E est compact ssi il est fermé et borné.*

2 Continuité des fonctions dans $\mathbb{R}^n$

Dans toute la suite on munira $\mathbb{R}^n$ et $\mathbb{R}^p$ d'une norme. Le choix de la norme sera sans importance puisque toutes les normes sont équivalentes en dimension finie.

Définition 2.1. *Soient $A \subset \mathbb{R}^n$, $f : A \to \mathbb{R}^p$ et $x \in A$. On dit que f est continue en x si*

$$\forall \epsilon > 0,\ \exists \delta > 0,\ \forall y \in A,\ \|y - x\| < \delta \Rightarrow \|f(y) - f(x)\| < \epsilon.$$

On dit que f est continue sur A si f est continue en tout point de A.

Proposition 2.1. *Soit $A \subset \mathbb{R}^n$ et $f : A \to \mathbb{R}^p$. Les propriétés suivantes sont équivalentes*

1. *f est continue sur A,*
2. *Pour tout ouvert U de F, $f^{-1}(U)$ est un ouvert de A,*
3. *Pour tout fermé V de F, $f^{-1}(V)$ est un fermé de A.*

On a bien entendu toutes les propriétés usuelles sur les fonctions continues : somme, composée, etc.

Théorème 2.2. *Soit $K \subset \mathbb{R}^n$ un compact et $f : K \to \mathbb{R}^p$ continue. Alors $f(K)$ est un compact de $\mathbb{R}^p$.*

Un cas particulier de ce théorème, lorsque $F = \mathbb{R}$, donne

Théorème 2.3. *Soit $K \subset \mathbb{R}^n$ un compact et $f : K \to \mathbb{R}$ continue. Alors f est bornée et atteint ses bornes, i.e. il existe $x_m, x_M \in K$ tels que pour tout $x \in K$ on ait $f(x_m) \leq f(x) \leq f(x_M)$.*

Parmi toutes les applications d'un espace vectoriel E dans un espace vectoriel F il y en a qui jouent un rôle important (voir la section suivante sur la différentiabilité), ce sont les applications linéaires. La proposition suivante est souvent très utile pour savoir si une application linéaire est continue.

Proposition 2.4. *Soient E et F deux evn et $f \in \mathcal{L}(E, F)$. Les assertions suivantes sont équivalentes*

1. *f est continue sur E,*
2. *f est continue en 0,*
3. *f est bornée sur la boule unité fermée de E, i.e. $\{f(x) \mid x \in E, \|x\|_E \leq 1\}$ est borné,*
4. *f est bornée sur la sphère unité de E, i.e. $\{f(x) \mid x \in E, \|x\|_E = 1\}$ est borné,*
5. *Il existe $C \geq 0$ tel que, pour tout $x \in E$, $\|f(x)\|_F \leq C\|x\|_E$.*

En dimension finie on a le résultat suivant.

Proposition 2.5. *Soient $(E, \|\ \|_E)$ et $(F, \|\ \|_F)$ deux evn. Si E est de dimension finie alors toute application linéaire de E dans F est continue. En particulier toutes les applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^p$ sont continues.*

Proposition 2.6. *Soit $f : E \to F$ une application linéaire. On a*

$$\sup_{x\in E,\, \|x\|_E=1} \|f(x)\|_F = \sup_{x\in E,\, \|x\|_E\leq 1} \|f(x)\|_F = \sup_{x\in E,\, x\neq 0} \frac{\|f(x)\|_F}{\|x\|_E}.$$

Ces quantités sont finies ssi f est continue et on note alors $\|f\|_{L(E,F)}$ leur valeur commune. L'application $f \mapsto \|f\|_{L(E,F)}$ définit une norme sur l'espace vectoriel des applications linéaires continues $L(E, F)$.

Remarque 2.1. *Bien qu'on ne l'indique pas explicitement pour ne pas trop alourdir la notation la quantité $\|f\|_{L(E,F)}$ dépend des normes choisies sur les espaces E et F.*

Remarque 2.2. *$\|f\|_{L(E,F)}$ est la plus petite constante C telle que $\|f(x)\|_F \leq C\|x\|_E$ pour tout $x \in E$.*

3 Différentiabilité - Différentielle

Dans toute cette section, Ω désignera un ouvert de $\mathbb{R}^n$.

Définition 3.1. *Soient $f : \Omega \to \mathbb{R}^p$, $a \in \Omega$ et $v \in \mathbb{R}^n$. On dit que f admet une dérivée au point a dans la direction v si la fonction de une variable $f_a^v : t \mapsto f(a + tv)$ est dérivable en $t = 0$. On appelle alors dérivée de f au point a dans la direction v la quantité $f_v'(a) = (f_a^v)'(0)$.*

Si $(e_1, \ldots, e_n)$ désigne la base canonique de $\mathbb{R}^n$ et si f admet une dérivée au point a dans la direction e_i on dira que f admet une dérivée partielle par rapport à la $i-$ème variable en a et on notera $\dfrac{\partial f}{\partial x_i}(a)$ la dérivée de f dans la direction e_i.

Remarque 3.1. *De façon équivalente, si $a = (a_1, \ldots, a_n)$, f admet une dérivée partielle par rapport à la $i-$ème variable en a si la fonction $f_i^a(t) = f(a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_n)$ est dérivable en a_i et on a $\dfrac{\partial f}{\partial x_i}(a) = (f_i^a)'(a_i)$.*

Remarque 3.2. *Attention l'existence de dérivées partielles n'entraine pas la continuité !*

Définition 3.2. *Soit $f : \Omega \to \mathbb{R}^p$ et $a \in \Omega$. On dit que f est différentiable en a s'il existe $L \in L(\mathbb{R}^n, \mathbb{R}^p)$ telle que*

$$\forall \epsilon > 0,\ \exists \delta > 0,\ \forall h \in \mathbb{R}^n,\ \|h\|_E < \delta \quad \Rightarrow \quad \|f(a + h) - f(a) - L(h)\| < \epsilon\|h\|.$$

De façon équivalente, f est différentiable en a s'il existe $L \in L(\mathbb{R}^n, \mathbb{R}^p)$ telle que

$$\lim_{h\to 0} \frac{1}{\|h\|}(f(a + h) - f(a) - L(h)) = 0, \tag{2}$$

ou encore telle que $f(a + h) = f(a) + L(h) + o(h)$.

Remarque 3.3. *Puisque $a \in \Omega$ et que Ω est ouvert, la fonction $h \mapsto f(a + h)$ est bien définie sur un voisinage de 0.*

Proposition 3.1. *Soit $f : \Omega \to \mathbb{R}^p$ et $a \in \Omega$. Si f est différentiable en a alors l'application L est unique. Elle est appelée différentielle de f en a et est notée $Df(a)$, ou encore $\mathrm{d}f(a)$, $\mathrm{d}f_a$ ou Df_a.*

Définition 3.3. *Soit $f : \Omega \to \mathbb{R}^p$. On dit que f est différentiable sur Ω si f est différentiable en tout $x \in \Omega$. On appelle alors différentielle de f l'application $Df : \left\{ \begin{array}{ccc} \Omega & \to & L(\mathbb{R}^n, \mathbb{R}^p) \\ x & \mapsto & Df(x) \end{array} \right.$.*

Proposition 3.2. *Soit $f : \Omega \to \mathbb{R}^p$ et $a \in \Omega$. Si f est différentiable en a alors f est continue en a.*

Proposition 3.3. *Si $f : \Omega \to \mathbb{R}^p$, $a \in \Omega$ et f est différentiable en a, alors f est dérivable en a dans n'importe quelle direction v et on a $f'_v(a) = Df(a)(v)$.*

Remarque 3.4. *La réciproque est fausse. Une fonction peut être dérivable en un point a dans toutes les directions sans être différentiable.*

Exemple 3.1. *Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = \frac{y^2}{x}$ si $x \neq 0$ et $f(0,y) = 0$. Montrer que f est dérivable en $(0,0)$ dans toutes les directions mais qu'elle n'y est pas différentiable.*

Dans la pratique, pour trouver la différentielle il suffit de connaitre sa valeur sur une base de $\mathbb{R}^n$ (c'est une application linéaire), par exemple la base canonique. Si f est différentiable, nécessairement $Df(a)(e_i) = \dfrac{\partial f}{\partial x_i}(a)$. On pourra donc utiliser la caractérisation suivante de la différentiabilité.

> f est différentiable en $a = (a_1, \ldots, a_n)$ ssi toutes les dérivées partielles de f au point a existent et si
> $$\lim_{h \to 0} \frac{1}{\|h\|} \left(f(a+h) - f(a) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) h_i \right) = 0,$$
> où $h = (h_1, \ldots, h_n)$. Dans ce cas $(Df)_a(h_1, \ldots, h_n) = \displaystyle\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) h_i$.

Remarque 3.5. *Si $E = \mathbb{R}$, tout élément $L \in L(\mathbb{R}, \mathbb{R}^p)$ est de la forme $L(x) = \ell x$ où $\ell \in \mathbb{R}^p$. La fonction f est donc différentiable en a s'il existe $\ell \in \mathbb{R}^p$ tel que*

$$\lim_{h \to 0} \frac{1}{|h|}(f(a+h) - f(a) - \ell h) = 0 \quad \iff \quad \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = \ell,$$

ce qui est précisément la définition de dérivable en a. On fera juste attention au fait que $\ell = f'(a)$ est le nombre dérivé de f en a tandis que $Df(a) \in L(\mathbb{R}, \mathbb{R}^p)$ est l'application $h \mapsto f'(a)h$.

Définition 3.4. *Si f est différentiable au point a on appelle matrice jacobienne de f au point a, notée $Jf(a)$, la matrice représentative de $Df(a)$ dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^p$.*

Si $f = (f_1, \ldots, f_p) : \Omega \to \mathbb{R}^p$ avec $f_i : \Omega \to \mathbb{R}$ et $h = (h_1, \ldots, h_n)$, on a alors

$$Df(a)(h) = \left(\sum_{j=1}^{n} \frac{\partial f_1}{\partial x_j}(a) h_j, \ldots, \sum_{j=1}^{n} \frac{\partial f_p}{\partial x_j}(a) h_j \right).$$

La matrice jacobienne $Jf(a)$ est donc la matrice $p \times n$

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

qui est en fait définie dès que f admet des dérivées partielles.

Proposition 3.4. *Soient $f, g : \Omega \to \mathbb{R}^p$ et $a \in \Omega$. Si f et g sont différentiables en a alors $f + g$ aussi et on a $D(f+g)(a) = Df(a) + Dg(a)$.*

Proposition 3.5. *Soient $\Omega \subset \mathbb{R}^n$, $\tilde{\Omega} \subset \mathbb{R}^p$ des ouverts, $f : \Omega \to \mathbb{R}^p$ telle que $f(\Omega) \subset \tilde{\Omega}$, $g : \tilde{\Omega} \to \mathbb{R}^d$ et $a \in \Omega$. Si f est différentiable en a et g est différentiable en $f(a)$. Alors la fonction $g \circ f : \Omega \to \mathbb{R}^d$ est différentiable en a et on a $D(g \circ f)(a) = Dg(f(a)) \circ Df(a)$.*

Remarque 3.6. *Si $E = F = \mathbb{R}$ dire que f et g sont différentiables et équivalent à dire qu'elles sont dérivables et on a alors $Df(a)(h) = f'(a)h$ et $Dg(f(a))(k) = g'(f(a))k$. On a donc $Dg(f(a)) \circ Df(a)(h) = g'(f(a))f'(a)h$. Le résultat de la proposition 3.5 se traduit ainsi par "$g \circ f$ est différentiable, donc dérivable, et $D(g \circ f)(a)(h) = g'(f(a))f'(a)h$, autrement dit $(g \circ f)'(a) = g'(f(a))f'(a)$". On retrouve bien la formule usuelle pour la dérivée de fonctions composées.*

La proposition précédente se traduit au niveau des matrices jacobiennes par

$$J\,(g \circ f)(a) = Jg(f(a)) \times Jf(a),$$

et au niveau des dérivées partielles par, si $f = (f_1, \ldots, f_p)$ et $g = (g_1, \ldots, g_d)$, pour tous $i = 1, \ldots, d$ et $j = 1, \ldots, n$,

$$\frac{\partial (g \circ f)_i}{\partial x_j}(a) = \sum_{k=1}^{p} \frac{\partial g_i}{\partial y_k}(f(a)) \times \frac{\partial f_k}{\partial x_j}(a).$$

4 Fonctions de classe C^1 - Accroissements finis

Définition 4.1. *Soit $f : \Omega \to \mathbb{R}^p$ une application différentiable. On dira que f est de classe C^1 si $Df : \Omega \to L(\mathbb{R}^n, \mathbb{R}^p)$ est continue.*

Remarque 4.1. *L'espace $L(\mathbb{R}^n, \mathbb{R}^p)$ est ici muni de la norme définie dans la Proposition 2.6.*

Tout comme pour la continuité et la différentiabilité, la somme et la composée de fonctions de classe C^1 est aussi de classe C^1.

Théorème 4.1. *Soient Ω un ouvert de $\mathbb{R}^n$ et $f : \Omega \to \mathbb{R}^p$. La fonction f est de classe C^1 sur Ω ssi*

1. *Pour tout $a \in \Omega$, et tout $i = 1, \ldots, n$, $\dfrac{\partial f}{\partial x_i}(a)$ existe.*
2. *Les n fonctions $\dfrac{\partial f}{\partial x_i} : \Omega \to \mathbb{R}^p$ sont continues.*

Théorème 4.2. *(Inégalité des accroissements finis) Soit Ω un ouvert convexe de $\mathbb{R}^n$ et $f : \Omega \to \mathbb{R}^p$ différentiable sur Ω, alors pour tous $a, b \in \Omega$ on a*

$$\|f(b) - f(a)\| \leq \sup_{x \in [a,b]} \|Df(x)\| \times \|b - a\|,$$

où $[a, b] = \{(1 - t)a + tb,\ t \in [0, 1]\}$.

Une conséquence importante de l'inégalité des accroissements finis est la suivante.

Proposition 4.3. *Soit $U \subset \mathbb{R}^n$ un ouvert connexe et $f : \Omega \to \mathbb{R}^p$ différentiable. Alors $Df = 0$ si et seulement si f est constante.*

Théorème 4.4. *Soit Ω un ouvert convexe de $\mathbb{R}^n$ et $f : \Omega \to \mathbb{R}$ différentiable sur Ω, alors pour tous $a, b \in \Omega$ il existe $c \in]a, b[$ tel que $f(b) - f(a) = Df(c)(b - a)$ où $]a, b[= \{(1 - t)a + tb,\ t \in]0, 1[\}$.*

Exemple 4.1. *Soit $f : \mathbb{R} \to \mathbb{R}^2$ définie par $f(t) = (\cos(t), \sin(t))$. La fonction f est différentiable (dérivable) et on a $Df(t)(h) = (-h\sin(t), h\cos(t))$ d'où $\|Df(t)\| = 1$ pour tout t. On peut vérifier qu'on a bien $\|f(t) - f(s)\| \leq |t - s|$ pour tous s, t.*

Par contre le Théorème 4.4 n'est plus vrai ! En effet, $f(2\pi) - f(0) = 0$ mais pour tout $t \in [0, 2\pi]$ on a $Df(t)(2\pi) = 2\pi(-\sin(t), \cos(t)) \neq (0, 0)$.

5 Différentielles d'ordre supérieur.

Définition 5.1. *Soit $f : \Omega \to \mathbb{R}^p$ une fonction différentiable. On dira que f est deux fois différentiable si $Df : \Omega \to L(\mathbb{R}^n, \mathbb{R}^p)$ est différentiable. La différentielle de Df au point $a \in \Omega$ est appelée différentielle seconde de f au point a, c'est un élément de $L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^p))$ noté $D^2f(a)$. On identifie en général $L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^p))$ avec $L_2(\mathbb{R}^n, \mathbb{R}^p)$ l'espace des applications bilinéaires de $\mathbb{R}^n$ dans $\mathbb{R}^p$ via l'isomorphisme φ défini par $(\varphi(f))(h, k) := (f(h))(k)$ et dont l'inverse est donné par $(\varphi^{-1}(g)(h))(k) := g(h, k)$.*

Par récurrence, on définit les fonctions k-fois différentiables ainsi que la différentielle d'ordre k que l'on identifie, de la même façon que ci-dessus, à une application k-linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$. Et on dit que la fonction f est de classe C^k si elle est k-fois différentiable et que la différentielle d'ordre k est continue.

On montre alors que f est deux fois différentiable ssi chacune des dérivées partielles $\dfrac{\partial f}{\partial x_j}$ est différentiable. On note alors $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(a)$ la dérivée partielle par rapport à la i-ème variable de $\dfrac{\partial f}{\partial x_j}$ au point a. De plus, si $h = (h_1, \ldots, h_n)$ et $k = (k_1, \ldots, k_n)$ on a

$$D^2f(a)(h, k) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i k_j.$$

Théorème 5.1 (Schwarz). *Soit Ω un ouvert de $\mathbb{R}^n$, $a \in \Omega$ et $f : \Omega \to \mathbb{R}^p$. Si f est deux fois différentiable au point a alors pour tous $i, j = 1, \ldots, n$ on a*

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

Autrment dit, la différentielle seconde $D^2f(a)$, vue comme application bilinéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$, est symétrique, i.e. pour tous $h, k \in \mathbb{R}^n$ on a $D^2f(a)(h, k) = D^2f(a)(k, h)$.

Si $f : \Omega \to \mathbb{R}$, la matrice $H_f(a) := \left(\dfrac{\partial^2 f}{\partial x_i \partial x_j}(a)\right)_{i,j=1,\ldots,n}$ est la matrice associée à la forme bilinéaire $D^2f(a)$ dans la base canonique de $\mathbb{R}^n$, i.e. pour tous $h, k \in \mathbb{R}^n$ on a $D^2f(a)(h, k) = {}^t h H_f(a) k$. Elle est appelée matrice hessienne de f au point a. Le théorème de Schwarz affirme que cette matrice est symétrique.

Théorème 5.2 (Taylor à l'ordre 2). *Si $f : \Omega \to \mathbb{R}^p$ est deux fois différentiable alors*

$$f(a + h) = f(a) + Df(a)(h) + \frac{1}{2} D^2f(a)(h, h) + o(\|h\|^2),$$

autrement dit

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

6 Fonctions à valeurs dans $\mathbb{R}$

On suppose dans toute cette section que f est à valeurs dans $\mathbb{R}$.

6.1 Vecteur gradient et plan tangent

Définition 6.1. *Si $f : \Omega \to \mathbb{R}$ est différentiable et $(e_1, \ldots, e_n)$ désigne la base canonique de $\mathbb{R}^n$, le vecteur $\mathrm{grad} f(a) = \nabla f(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) e_i$ est appelé vecteur gradient de f au point a. Pour tout $h \in \mathbb{R}^n$ on a alors $Df(a)(h) = \nabla f(a) \cdot h$ où $\cdot$ désigne le produit scalaire usuel sur $\mathbb{R}^n$.*

Définition 6.2. *Soit $f : \Omega \to \mathbb{R}$. Le graphe de f est l'ensemble $\{(x, f(x)) \in \mathbb{R}^n \times \mathbb{R} \,|\, x \in \Omega\}$. Si f est différentiable au point a l'hyperplan affine de $\mathbb{R}^{n+1}$ d'équation*

$$x_{n+1} = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) \iff x_{n+1} = f(a) + Df(a)(x - a)$$

est appelé hyperplan tangent au point a. C'est le plan passant par le point $(a, f(a))$ et de vecteur normal $(\nabla f(a), -1)$.

Remarque 6.1. *Si $n = 1$ le graphe de f est une courbe C dans $\mathbb{R}^2$ et la droite tangente (hyperplan dans $\mathbb{R}^2$) au point $(a, f(a)) \in C$ est la droite d'équation*

$$y = f(a) + f'(a)(x - a).$$

Si $n = 2$ le graphe de f est une surface S dans $\mathbb{R}^3$ et le plan tangent au point $(a, b, f(a, b)) \in S$ est le plan d'équation

$$z = f(a) + \frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(y - b).$$

Proposition 6.1. *Soit $f : \Omega \to \mathbb{R}$ différentiable et $a \in \Omega$. Le vecteur gradient $\nabla f(a)$ indique la direction de plus grande pente de la surface représentative de f au point $(a, f(a))$.*

L'idée est la suivante :

1. on regarde l'intersection de la surface représentative de f et du plan affine P de $\mathbb{R}^{n+1}$ passant par le point $(a, f(a))$ et dirigé par les vecteurs e_{n+1} et v.
2. on obtient alors une courbe C dont l'équation dans le plan P est le graphe de la fonction $g(t) = f(a + tv)$.
3. lorsque $t = 0$ on est précisément au point $(a, f(a))$, la pente dans la direction v est donc donnée par $\frac{g'(0)}{\|v\|} = \frac{Df(a)(v)}{\|v\|} = \nabla f(a) \cdot \frac{v}{\|v\|}$.

L'inégalité de Cauchy-Schwartz montre que la pente maximale est donnée par $\|\nabla f(a)\|$ et que celle-ci est atteinte précisément quand v est colinéaire à $\nabla f(a)$.

6.2 Extrema locaux

Définition 6.3. *Soit $f : \Omega \to \mathbb{R}$ différentiable. On dit que $a \in \Omega$ est un point critique de f si $Df(a) = 0$ (l'application nulle de $\mathbb{R}^n$ dans $\mathbb{R}$), autrement dit si $\dfrac{\partial f}{\partial x_i}(a) = 0$ pour tout $i = 1, \ldots, n$.*

Proposition 6.2. *Soit $f : \Omega \to \mathbb{R}$ différentiable. Si f possède un extremum local en $a \in \Omega$ alors a est un point critique de f.*

On rappelle la notion de forme bilinéaire positive/négative.

Définition 6.4. *Soit $L : E \times E \to \mathbb{R}$ une forme bilinéaire symétrique. On dit que L est positive, resp. négative, si pour tout $h \in E$ on a $L(h,h) \geq 0$, resp. $L(h,h) \leq 0$. Si $L(h,h) > 0$, resp. $L(h,h) < 0$, pour tout $h \in E \backslash \{0\}$ on dit que L est définie positive, resp. définie négative.*

Proposition 6.3. *Soit $f : \Omega \to \mathbb{R}$ une fonction deux fois différentiable.*

- *Si a est un minimum, resp. maximum, local de f alors $Df(a) = 0$ et $D^2 f(a)$ est une forme bilinéaire positive, resp. négative.*
- *Si a est tel que $Df(a) = 0$ et $D^2 f(a)$ est définie positive, resp. définie négative, alors a est un minimum, resp. maximum, local strict de f.*

Attention, la différentielle seconde en un point a peut n'être ni positive ni négative, c'est le cas par exemple de la forme bilinéaire $f(h,k) = h_1 k_1 - h_2 k_2$. Un point critique a dont la différentielle seconde n'est ni positive ni négative est appelé point selle ou point col.
Pour étudier si la différentielle seconde est positive ou négative on peut utiliser le résultat suivant (appliqué à la matrice hessienne)

Proposition 6.4. *Soit A une matrice symétrique rélle. Elle est diagonalisable dans $\mathbb{R}$. De plus, A est (définie) positive, resp. négative, ssi toutes ses valeurs propres sont (strictement) positives, resp. négatives.*

En particulier en dimension 2, puisque la somme des valeurs propres est la trace de A et leur produit le déterminant de A, une matrice symétrique A est définie positive, resp. négative, si et seulement si $\det(A) > 0$ et $\mathrm{Tr}(A) > 0$, resp. $\mathrm{Tr}(A) < 0$. Si $\Omega \subset \mathbb{R}^2$ et si on note (notations de Monge)

$$r = \frac{\partial^2 f}{\partial x^2}(a), \quad s = \frac{\partial^2 f}{\partial x \partial y}(a) = \frac{\partial^2 f}{\partial y \partial x}(a) \quad \text{et} \quad t = \frac{\partial^2 f}{\partial y^2}(a),$$

on a alors le résultat suivant :

Proposition 6.5. *On pose $\Delta = rt - s^2$.*

1. *Si $\Delta > 0$ alors f admet un extremum local en a. Si $r > 0$ (ou $t > 0$) c'est un minimum local, et si $r < 0$ (ou $t < 0$) c'est un maximum local.*
2. *Si $\Delta < 0$ alors f n'admet pas d'extremum local en a, a est un point selle.*
3. *Si $\Delta = 0$ on ne peut rien dire a priori.*

7 Inversion locale et fonctions implicites

7.1 Les deux théorèmes

Définition 7.1. *Soient U, V deux ouverts de $\mathbb{R}^n$ et $f : U \to V$. On dit que f est un C^k-difféormorphisme (ou difféomorphisme de classe C^k) si f est bijective, de classe C^k et telle que f^{-1} est de classe C^k.*

Définition 7.2. *Soit $f : \Omega \to \mathbb{R}^n$ différentiable. Le déterminant de la matrice jacobienne $Jf(a)$ s'appelle le jacobien de f au point a et est noté* $\operatorname{Jac} f(a)$.

Théorème 7.1 (Inversion locale)**.** *Soient Ω un ouvert de $\mathbb{R}^n$, $f : \Omega \to \mathbb{R}^p$ de classe C^1 et $a \in \Omega$. Si $Df(a)$ est bijective (en tant qu'application linéaire), ce qui équivaut à $Jf(a)$ est inversible ou encore à $n = p$ et $\operatorname{Jac} f(a) \neq 0$, alors f est localement inversible au voisinage de a, i.e. il existe deux ouverts U et V de $\mathbb{R}^n$ avec $U \subset \Omega$ tels que $a \in U$, $f(a) \in V$ et $f : U \to V$ soit un C^1-difféomorphisme. De plus on a $Df^{-1}(f(a)) = (Df(a))^{-1}$. Si la fonction f est de classe C^k alors f^{-1} est un C^k-difféomorphisme.*

Remarque 7.1. *Contrairement au cas des fonctions de une variable, le fait que $Df(x)$ soit inversible pour tout $x \in \Omega$ n'implique pas que f est bijective de Ω sur $f(\Omega)$. La fonction $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $f(x, y) = (\mathrm{e}^x \cos(y), \mathrm{e}^x \sin(y))$ est de classe C^1. Pour tout $(x, y) \in \mathbb{R}^2$ on a $Jf(x, y) = \begin{pmatrix} \mathrm{e}^x \cos(y) & -\mathrm{e}^x \sin(y) \\ \mathrm{e}^x \sin(y) & \mathrm{e}^x \cos(y) \end{pmatrix}$, et donc $Jf(x, y)$ est inversible puisque $\operatorname{Jac} f(x, y) = \mathrm{e}^{2x} \neq 0$. Cependant f n'est pas injective (elle est 2π-périodique dans la variable y).*

Théorème 7.2 (Inversion globale)**.** *Soient Ω un ouvert de $\mathbb{R}^n$, $f : \Omega \to \mathbb{R}^p$ de classe C^k. Si f est injective et si $Df(x)$ est bijective pour tout $x \in \Omega$ alors $f(\Omega)$ est un ouvert de $\mathbb{R}^n$ et f est un C^k-difféomorphisme de Ω dans $f(\Omega)$.*

Remarque 7.2. *Ce que nous apprend le Théorème d'inversion globale ce n'est pas que f est bijective, on a supposé f injective donc elle est forcément bijective de Ω dans $f(\Omega)$. Il garantit d'une part que f^{-1} est de classe C^k mais aussi que l'image $f(\Omega)$ est un ouvert.*

Dans le théorème suivant on est amené à "découper" $\mathbb{R}^n$ sous la forme $\mathbb{R}^n \simeq \mathbb{R}^{n-p} \times \mathbb{R}^p$. On notera $m = n - p$ et un élément de $\mathbb{R}^n$ sera un couple (x, y) avec $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$ et $y = (y_1, \ldots, y_p) \in \mathbb{R}^p$.

Théorème 7.3 (Fonctions implicites)**.** *Soient $\Omega \subset \mathbb{R}^m \times \mathbb{R}^p$ un ouvert et $f := (f_1, \ldots, f_p) : \Omega \to \mathbb{R}^p$ de classe C^1. Soit $(a, b) \in \Omega$ tel que*

$$\begin{pmatrix} \frac{\partial f_1}{\partial y_1}(a,b) & \cdots & \frac{\partial f_1}{\partial y_p}(a,b) \\ \vdots & & \vdots \\ \frac{\partial f_p}{\partial y_1}(a,b) & \cdots & \frac{\partial f_p}{\partial y_p}(a,b) \end{pmatrix}$$

soit inversible. Alors il existe des voisinages V de a dans $\mathbb{R}^m$ et W de b dans $\mathbb{R}^p$ et une application $\varphi : V \to W$ de classe C^1 tels que

$$\forall (x, y) \in V \times W, \quad f(x, y) = f(a, b) \iff y = \varphi(x).$$

De plus on a

$$J\varphi(a) = -\begin{pmatrix} \frac{\partial f_1}{\partial y_1}(a,b) & \cdots & \frac{\partial f_1}{\partial y_p}(a,b) \\ \vdots & & \vdots \\ \frac{\partial f_p}{\partial y_1}(a,b) & \cdots & \frac{\partial f_p}{\partial y_p}(a,b) \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a,b) & \cdots & \frac{\partial f_1}{\partial x_m}(a,b) \\ \vdots & & \vdots \\ \frac{\partial f_p}{\partial x_1}(a,b) & \cdots & \frac{\partial f_p}{\partial x_m}(a,b) \end{pmatrix}.$$

Remarque 7.3. *Dans le cas $p = 1$ l'hypothèse s'écrit simplement $\frac{\partial f}{\partial x_n}(a, b) \neq 0$, et on a alors pour tout $i = 1, \ldots, n-1$ et en notant $x' = (x_1, \ldots, x_{n-1})$*

$$\frac{\partial \varphi}{\partial x_i}(x') = -\frac{\frac{\partial f}{\partial x_i}(x', \varphi(x'))}{\frac{\partial f}{\partial x_n}(x', \varphi(x'))}. \tag{3}$$

En particulier,

$$\frac{\partial \varphi}{\partial x_i}(a) = -\frac{\frac{\partial f}{\partial x_i}(a,b)}{\frac{\partial f}{\partial x_n}(a,b)}.$$

7.2 Applications

Application 1 : vecteur gradient et courbes de niveaux.

Définition 7.3. *Si $\Omega \subset \mathbb{R}^2$, $f : \Omega \to \mathbb{R}$ et $k \in \mathbb{R}$, on appelle courbe de niveau k l'ensemble $C_k = \{(x,y) \in \mathbb{R}^2 \,|\, x \in \Omega$ et $f(x,y) = k\}$.*

Proposition 7.4. *Soit $\Omega \subset \mathbb{R}^2$, $f : \Omega \to \mathbb{R}$ de classe C^1, $k \in \mathbb{R}$ et $(a,b) \in C_k$. Si $Df(a,b) \neq 0$, i.e. soit $\frac{\partial f}{\partial x}(a,b) \neq 0$ soit $\frac{\partial f}{\partial y}(a,b) \neq 0$, ou encore $\nabla f(a,b) \neq 0$, alors la courbe C_k admet une tangente au point (a,b) et le vecteur $\nabla f(a,b)$ est un vecteur normal à la courbe de niveau C_k au point (a,b) (et pointe dans la direction où f croit).*

Remarque 7.4. *Le résultat reste valable si $\Omega \subset \mathbb{R}^n$. On parle alors de surface de niveau et de plan (hyperplan) tangent.*

Application 2 : extrema liés.

Définition 7.4. *Soit $\Omega \subset \mathbb{R}^n$ un ouvert et $g : \Omega \to \mathbb{R}$. On dit que $\Gamma := \{x \in \Omega \,|\, g(x) = 0\}$ est une contrainte régulière si pour tout $x \in \Gamma$ il existe $i \in \{1, \ldots, n\}$ tel que $\frac{\partial g}{\partial x_i}(x) \neq 0$. En d'autres termes, si pour tout $x \in \Gamma$ on a $Dg(x) \neq 0$.*

Théorème 7.5. *Soit Ω un* ouvert *de $\mathbb{R}^n$, $f, g : \Omega \to \mathbb{R}$ de classe C^1 telles que $\Gamma := \{x \in \Omega \,|\, g(x) = 0\}$ soit une contrainte régulière. Si f possède un extremum local relatif à la contrainte Γ en $a \in \Gamma$, i.e. $f\lceil_\Gamma$ possède un extremum local, alors il existe (un unique) $\lambda \in \mathbb{R}$ tel que $Df(a) = \lambda Dg(a)$, i.e. $\frac{\partial f}{\partial x_i}(a) = \lambda \frac{\partial g}{\partial x_i}(a)$ pour tout $i = 1, \ldots, n$ ou encore $\nabla f(a) = \lambda \nabla g(a)$. Le nombre λ est appelé un multiplicateur de Lagrange.*

Remarque 7.5. *Il est important que f et g soient définies sur un ouvert ! Il ne faut cependant pas confondre l'ensemble Ω sur lequel sont définies f et g avec la contrainte Γ sur laquelle on cherche les extrema de f.*

Exemple 7.1. *Déterminer les points de la courbe $\Gamma = \{(x,y) \in \mathbb{R}^2 \,|\, x^4 + y^4 = 1\}$ les plus proches et les plus éloignés de l'origine. Il s'agit donc de trouver le minimum global et le maximum global de la fonction $f(x,y) = x^2 + y^2$ sur l'ensemble $\Gamma = \{(x,y) \in \mathbb{R}^2 \,|\, x^4 + y^4 = 1\}$. Ici $g(x,y) = x^4 + y^4 - 1$ et f sont bien de classe C^1 et définies sur l'ouvert $\Omega = \mathbb{R}^2$, par contre l'ensemble Γ n'est pas un ouvert de $\mathbb{R}^2$ mais un compact.*

7.3 Changement de variable dans les intégrales multiples

Théorème 7.6. *Soient $U, V \subset \mathbb{R}^n$ des ouverts, $\varphi : U \to V$ un C^1-difféomorphisme et $f : V \to \mathbb{C}$ une fonction continue.*

1) Si $f : V \to \mathbb{R}_+$ alors

$$\int_V f(y_1, \ldots, y_n)\, \mathrm{d}y_1 \cdots \mathrm{d}y_n = \int_U f \circ \varphi(x_1, \ldots, x_n) |\mathrm{Jac}\, \varphi(x_1, \ldots, x_n)| \mathrm{d}x_1 \cdots \mathrm{d}x_n,$$

l'égalité ayant lieu dans $\bar{\mathbb{R}}_+ = \mathbb{R}_+ \cup \{+\infty\}$.

2) Si f est intégrable alors la fonction $f \circ \varphi : U \to \mathbb{C}$ est intégrable et

$$\int_V f(y_1, \ldots, y_n)\,\mathrm{d}y_1 \cdots \mathrm{d}y_n = \int_U f \circ \varphi(x_1, \ldots, x_n) |\mathrm{Jac}\,\varphi(x_1, \ldots, x_n)| \mathrm{d}x_1 \cdots \mathrm{d}x_n.$$

Exemple 7.2. *Le changement de variables en coordonnées polaires est donné par l'application $\varphi(r, \theta) = (r\cos\theta, r\sin\theta) = (x, y)$ et on a $\mathrm{Jac}\,\varphi(r, \theta) = r$. Dans un changement de variable en coordonnées polaires on aura*

$$\int_V f(x, y)\,\mathrm{d}x\mathrm{d}y = \int_U f(r\cos\theta, r\sin\theta) r\,\mathrm{d}r\mathrm{d}\theta,$$

où U et V sont les ouverts correspondants. Par exemple, si $V = \mathbb{R}_+^* \times \mathbb{R}_+^*$ on a $U = \mathbb{R}_+^* \times]0, \frac{\pi}{2}[$ et φ est bien un C^1-difféomorphisme de U sur V.