

Notes de Cours - Analyse 4

Hacen DIB

24 Mars 2020

Table des matières

1	Rappels de Topologie de $\mathbb{R}^n$	5
1.1	Normes	5
1.1.6	Normes équivalentes	7
1.2	Notion de compacité	7
1.3	Notion de continuité	8
2	Calcul différentiel dans $\mathbb{R}^n$	11
2.1	Notions de différentiabilité	11
2.2	Dérivées partielles d'ordre supérieur	15
2.3	Formules de Taylor	17
2.4	Différentiabilité de fonctions vectorielles	19
2.5	Inversion locale	21
2.6	Théorème des fonctions implicites	24
2.7	Extremums locaux	26
2.8	Extrema liés	30
3	Calcul Intégral dans $\mathbb{R}^n$	35
3.1	Le cas $n = 2$	35
3.2	Propriétés	38
3.3	Théorème de Fubini	39
3.4	Changement de variables	42
3.5	Le cas $n = 3$	44
3.5.1	Théorème de Fubini	44
3.5.4	Changement de variables	46

Chapitre 1

Rappels de Topologie de $\mathbb{R}^n$

Le présent chapitre contient les principales notions, déjà traitées en module de Topologie, qui nous seront utiles en Analyse 4.

1.1 Normes

Définition 1.1.1 *On appelle norme sur $\mathbb{R}^n$ toute application*

$$N : \mathbb{R}^n \longrightarrow \mathbb{R}^+$$

telle que

$$(i) \ N(x) = 0 \Leftrightarrow x = 0$$

$$(ii) \ N(\lambda x) = |\lambda|N(x), \quad \forall x \in \mathbb{R}^n \quad \forall \lambda \in \mathbb{R}$$

$$(iii) \ N(x + y) \leq N(x) + N(y), \quad \forall x, y \in \mathbb{R}^n \quad (\text{inégalité triangulaire})$$

On note parfois (et même souvent) une norme par $\|\cdot\|$. La distance associée à une norme est définie par : $d(x, y) = N(x - y)$.

Exemple 1.1.2 *En voici quelques exemples :*

$$1. \ \|x\|_1 = \sum_{i=1}^n |x_i|$$

$$2. \ \|x\|_\infty = \sup_{1 \leq i \leq n} |x_i|$$

$$3. \ \|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2} \quad (\text{norme euclidienne})$$

Essayons pour $n = 2$ de déterminer la distance entre les points $(1, 0)$ et $(0, 1)$ en utilisant les trois normes données précédemment. On aura :

$$\square d_1((1, 0), (0, 1)) = \|(1, -1)\|_1 = 2$$

$$\square d_2((1, 0), (0, 1)) = \|(1, -1)\|_2 = \sqrt{2}$$

$$\square d_\infty((1, 0), (0, 1)) = \|(1, -1)\|_\infty = 1$$

On remarque que seule d_2 exprime la distance intuitive, c'est la distance euclidienne.

Proposition 1.1.3 Soit $\|\cdot\|$ une norme sur $\mathbb{R}^n$. On a $\forall x, y \in \mathbb{R}^n$

$$|\|x\| - \|y\|| \leq \|x - y\| \quad (1.1)$$

Démonstration : Il suffit d'écrire

$$\|x\| = \|x - y + y\| \leq \|x - y\| + \|y\|$$

d'où

$$\|x\| - \|y\| \leq \|x - y\|$$

Puis échanger le rôle de x et y . Se souvenir que si $a \leq b$ et $-a \leq b$ alors $|a| \leq b$. ■

Définition 1.1.4 (boule ouverte, boule fermée, ouvert, fermé)

- (a) L'ensemble $B(a, r) = \{x \in \mathbb{R}^n / \|x - a\| < r\}$ s'appelle *boule ouverte* de centre a et de rayon r .
- (b) L'ensemble $\overline{B}(a, r) = \{x \in \mathbb{R}^n / \|x - a\| \leq r\}$ s'appelle *boule fermée* de centre a et de rayon r .
- (c) Un sous-ensemble $\mathcal{O} \subset \mathbb{R}^n$ est dit *ouvert* si $\forall a \in \mathcal{O} \exists r > 0$ tel que $B(a, r) \subset \mathcal{O}$.
- (d) $\mathcal{F} \subset \mathbb{R}^n$ est dit *fermé* si son complémentaire $\mathbb{R}^n \setminus \mathcal{F}$ est ouvert.

Exemple 1.1.5 1. Boules dans $\mathbb{R}^2$

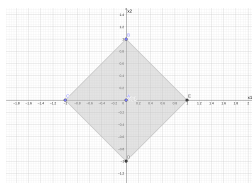


FIGURE 1.1 – $B_1((0, 0), 1)$

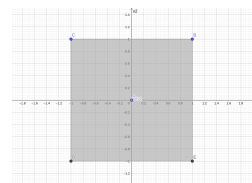


FIGURE 1.2 – $B_\infty((0, 0), 1)$

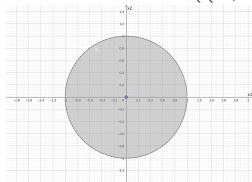


FIGURE 1.3 – $B_2((0, 0), 1)$

- 2. $\mathcal{P} = \{(x_1, x_2) \in \mathbb{R}^2 / x_2 > 0\}$ est un ouvert. (Avec n'importe quelle norme)
- 3. $A =]0, +\infty[\times \{0\}$ n'est ni ouvert ni fermé.

1.1.6 Normes équivalentes

Deux normes N_1 et N_2 sont dites équivalentes si et seulement si $\exists \alpha, \beta > 0$ tels que $\forall x \in \mathbb{R}^n$ on ait

$$\alpha N_2(x) \leq N_1(x) \leq \beta N_2(x) \quad (1.2)$$

Exemple 1.1.7 *On peut aisément établir les inégalités suivantes*

$$\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty$$

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty$$

Il est facile de remarquer que l'équivalence se traduit par des inclusions de boules de même centre (avec différents rayons) et ce pour les deux normes en question. Par exemple, dans un carré on peut toujours inscrire un disque et vice-versa. Nous montrerons plus loin que toutes les normes sur $\mathbb{R}^n$ sont équivalentes. Ceci permettra, en particulier, de ne pas faire grand cas de la norme quand on parlera de convergence de suites, de continuité, etc...

1.2 Notion de compacité

Nous adopterons la définition suivante de la compacité, appelée propriété de Bolzano-Weierstraß.

Définition 1.2.1 *Un sous-ensemble $K \subset \mathbb{R}^n$ (muni d'une norme donnée) est dit compact si et seulement si "de toute suite on peut extraire une sous-suite convergente".*

Exemple 1.2.2 $K = [-1, 1] \times [-1, 1] \times \cdots \times [-1, 1] \subset \mathbb{R}^n$ est compact.

En effet, fixons la norme $\|\cdot\|_\infty$. Après on pourra se passer de cette "restriction". Si $(x_k)_{k \geq 0}$ est une suite dans K , alors pour chaque indice $i = 1, 2, \dots, n$ la suite $(x_{i,k})_k$ est dans $[-1, 1]$. La compacité (propriété de Bolzano-Weierstraß) de $[-1, 1]$ (se reporter au cours d'Analyse 1) nous permet d'extraire de $(x_{1,k})_k$ une sous-suite convergente dans $[-1, 1]$, appelons-la $(x_{1,\phi_1(k)})_k$. On considère alors la suite $(x_{2,\phi_1(k)})_k$. On procède de la même façon. La deuxième extraction permet de dire que les deux premières composantes $(x_{1,\phi_1(\phi_2(k))})_k$ et $(x_{2,\phi_1(\phi_2(k))})_k$ sont convergentes. On réitère ce procédé jusqu'à la dernière composante. Il faut opérer n extractions successives et "emboîtées".

Nous terminons cette section par un théorème de caractérisation des compacts dans $\mathbb{R}^n$, discuté en détail dans le cours de Topologie.

Théorème 1 *K est compact dans $\mathbb{R}^n$ si et seulement si K est fermé et borné.*

1.3 Notion de continuité

Définition 1.3.1 Soit U un ouvert de $\mathbb{R}^n$ et $a \in U$. Alors $f : U \longrightarrow \mathbb{R}$ est continue au point a si et seulement si :

$$\lim_{x \rightarrow a} f(x) = f(a).$$

Autrement dit

$$\forall \varepsilon > 0 \quad \exists \delta_\varepsilon > 0 \quad \forall x \in U \quad \|x - a\| < \delta_\varepsilon \implies |f(x) - f(a)| < \varepsilon \quad (1.3)$$

La continuité de f sur U signifie qu'elle est continue en tout point de U .

Proposition 1.3.2 Les propriétés suivantes sont identiques au cas d'une variable en énoncé et démonstration.

- A-** f est continue en a si et seulement si pour toute suite (x_m) convergente vers a , la suite $(f(x_m))$ converge vers $f(a)$.
- B-** Si f et g sont continues, alors $f \pm g$, $f \cdot g$ et $f \circ g$ le sont aussi. De plus f/g est continue (en a) si $g(a) \neq 0$.
- C-** Si f est continue sur un compact, alors elle est bornée et elle atteint son maximum et son minimum.

Démonstration : Nous donnerons la démonstration de la propriété **C** seulement car les deux premières sont faciles à établir. La bornitude de f s'écrit comme suit

$$\exists M > 0 \quad \text{telle que} \quad \forall x \in K \quad |f(x)| \leq M$$

Faisons un raisonnement par l'absurde, et supposons que pour tout $M > 0$ il existe $x_M \in K$ tel que $|f(x_M)| > M$. Prenons pour M un entier naturel j . Ainsi on dispose maintenant d'une suite (x_j) dans K qui vérifie que $|f(x_j)| > j$, en particulier $\lim_{j \rightarrow +\infty} |f(x_j)| = +\infty$. La compacité de K nous donne la possibilité d'extraire une sous-suite, qu'on appellera aussi (x_j) , qui converge vers une limite a appartenant à K car c'est un fermé. Mais la continuité de f , par la propriété **A**, permet de dire que $\lim_{j \rightarrow +\infty} |f(x_j)| = |f(a)|$, d'où la contradiction. Montrons à présent que f atteint son maximum, pour le minimum c'est pareil. Rappelons la ε -caractérisation de la borne supérieure. Posons $M_f = \sup_{x \in K} f(x)$. Alors on a

$$\forall \varepsilon > 0 \quad \exists x_\varepsilon \in K \quad \text{tel que} \quad M_f - \varepsilon < f(x_\varepsilon) \leq M_f$$

Prenons pour ε une suite tendant vers 0, par exemple $\varepsilon = 1/2^k$. Cela nous donnera une suite (x_k) dans K telle que

$$\lim_{k \rightarrow +\infty} f(x_k) = M_f$$

La compacité de K nous permet d'extraire une sous-suite, notée aussi (x_k) , convergente dans K vers une limite a . Par la continuité de f on a $M_f = f(a)$. ■

Remarque 1.3.3 Dans beaucoup de situations la propriété **A** est utilisée dans le sens négatif, c'est-à-dire pour montrer qu'une fonction n'est pas continue. Par exemple pour la fonction

$$f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 1 & \text{si } (x, y) = (0, 0) \end{cases}$$

on peut considérer la suite $(1/k, 1/k)$, manifestement convergente vers $(0, 0)$, qui donne $f(1/k, 1/k) = 1/2 \not\rightarrow 1$, c'est-à-dire que cette fonction n'est pas continue en $(0, 0)$.

Nous allons terminer ce chapitre par une application de ce qui précède.

Proposition 1.3.4 Toutes les normes sur $\mathbb{R}^n$ sont équivalentes.

Démonstration : Il suffit de montrer qu'une norme quelconque N est équivalente à la norme $\|\cdot\|_\infty$ (par exemple). Soit $\{e_i\}_{1 \leq i \leq n}$ la base canonique de $\mathbb{R}^n$. Alors on a

$$N(x) = N\left(\sum_{i=1}^n x_i e_i\right) \leq \sum_{i=1}^n |x_i| N(e_i) \leq \|x\|_\infty \sum_{i=1}^n N(e_i)$$

Autrement dit $N(x) \leq b\|x\|_\infty$ avec $b = \sum_{i=1}^n N(e_i)$. Remarquons que $b > 0$ car sinon on aurait $\forall i = 1, 2, \dots, n \quad N(e_i) = 0$ ce qui est absurde. L'utilisation de (1.1) donne

$$|N(x) - N(x')| \leq N(x - x') \leq b\|x - x'\|_\infty$$

Cette dernière inégalité montre bien que N en tant qu'application est continue de $(\mathbb{R}^n, \|\cdot\|_\infty)$ dans $(\mathbb{R}, |\cdot|)$ car lipschitzienne. Considérons la "sphère" unité

$$K = \{x \in \mathbb{R}^n \mid \|x\|_\infty = 1\} = [-1, 1]^n.$$

Nous avons déjà montré que c'est un compact. Donc $\exists a \in K$ tel que $N(a) = \min_{x \in K} N(x)$.

Il est clair que $N(a) \neq 0$ car a est dans K . Soit maintenant $x \neq 0$. Alors $\frac{x}{\|x\|_\infty} \in K$

et donc $N\left(\frac{x}{\|x\|_\infty}\right) \geq N(a)$, ce qui donne par l'homogénéité de la norme

$N(x) \geq N(a)\|x\|_\infty$. cette inégalité reste vraie pour $x = 0$ et c'est l'inégalité qui nous manquait pour l'équivalence. ■

Chapitre 2

Calcul différentiel dans $\mathbb{R}^n$

Nous allons donner d'abord les principales notions de "dérivation" d'une fonction de $\mathbb{R}^n \rightarrow \mathbb{R}$. La norme utilisée désormais sera la norme euclidienne sauf mention contraire.

2.1 Notions de différentiabilité

Définition 2.1.1 Soit U un ouvert de $\mathbb{R}^n$. Une fonction $f : U \rightarrow \mathbb{R}$ est dite différentiable en un point $a \in U$ s'il existe une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}$, notée df_a , telle que

$$\lim_{\|h\| \rightarrow 0} \frac{f(a+h) - f(a) - df_a(h)}{\|h\|} = 0 \quad (2.1)$$

L'application linéaire df_a s'appelle différentielle de f au point a , ou bien application linéaire tangente en a . La différentiabilité de f sur U signifie sa différentiabilité en tout point de U . En notant la fraction précédente par $\varepsilon(h)$, on peut réécrire (2.1) comme suit :

$$f(a+h) = f(a) + df_a(h) + \|h\|\varepsilon(h) \quad (2.2)$$

où $\lim_{h \rightarrow 0} \varepsilon(h) = 0$. Rappelons que la linéarité de df_a s'écrit ainsi

$$df_a(h) = df_a(h_1, h_2, \dots, h_n) = \lambda_1 h_1 + \lambda_2 h_2 + \dots + \lambda_n h_n = \langle \lambda, h \rangle$$

où les λ_i sont des constantes qui dépendent de a et $\langle \cdot, \cdot \rangle$ est le produit scalaire usuel de $\mathbb{R}^n$.

Exemple 2.1.2 Examinons un exemple simple. Déterminons la différentielle de $f(x, y) = xy$, définie sur $\mathbb{R}^2$, en un point (a, b) . Il est facile de voir que

$$f(a+h_1, b+h_2) = (a+h_1)(b+h_2) = ab + bh_1 + ah_2 + h_1h_2$$

Deux candidats à la formule (2.2) s'offrent aisément : $df_{(a,b)}(h_1, h_2) = bh_1 + ah_2$ et

$\varepsilon(h_1, h_2) = \frac{h_1h_2}{\sqrt{h_1^2 + h_2^2}}$ si on choisit de travailler avec la norme euclidienne. Comme

$$\frac{|h_1h_2|}{h_1^2 + h_2^2} \leq \frac{1}{2}, \text{ alors } |\varepsilon(h_1, h_2)| \leq \frac{1}{\sqrt{2}} \sqrt{|h_1h_2|} \text{ qui assure que } \lim_{h \rightarrow 0} \varepsilon(h) = 0.$$

Définition 2.1.3 (*Dérivation dans le sens d'un vecteur*)

Soit $f : U \rightarrow \mathbb{R}$ et $a \in U$. Fixons un vecteur v dans $\mathbb{R}^n$. On appelle dérivée de f au point a dans le sens du vecteur v la limite suivante, si elle existe,

$$\lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t} = (D_v f)(a) \quad (2.3)$$

(remarquez que t est un réel).

Exemple 2.1.4 Examinons le cas de $f(x, y) = x^2 - y^2$ avec $a = (1, -1)$ et $v = (1, 1)$.
On a

$$f(1 + t, -1 + t) = (1 + t)^2 - (-1 + t)^2 = 4t \quad \text{et} \quad f(a) = 0$$

d'où $(D_v f)(a) = 4$.

Définition 2.1.5 (*Dérivation partielle*)

Soit $e_j = (0, 0, \dots, 0, 1, 0, \dots, 0)$ le jème vecteur de la base canonique de $\mathbb{R}^n$. On appelle dérivée partielle de f par rapport à x_j la dérivée de f dans le sens du vecteur e_j i.e.,

$$\frac{\partial f}{\partial x_j}(a) = (D_{e_j} f)(a) = \lim_{t \rightarrow 0} \frac{f(a_1, \dots, a_j + t, \dots, a_n) - f(a)}{t} \quad (2.4)$$

On appelle gradient de f le vecteur défini par

$$(\nabla f)(x) = (\vec{\text{grad}} f)(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

Les deux notations sont consacrées. Le symbole ∇ se lit "nabla".

Exemple 2.1.6 Considérons $f(x, y, z) = xy + xz + yz$. Alors

$$\frac{\partial f}{\partial x}(x, y, z) = y + z, \quad \frac{\partial f}{\partial y}(x, y, z) = x + z, \quad \frac{\partial f}{\partial z}(x, y, z) = x + y$$

Proposition 2.1.7 Si f est différentiable au point a , alors $\forall j = 1, \dots, n$; $\frac{\partial f}{\partial x_j}(a)$ existe, de plus

$$df_a(h) = \sum_{j=1}^n h_j \frac{\partial f}{\partial x_j}(a) = \langle h, \nabla f(a) \rangle \quad (2.5)$$

Démonstration : On a par hypothèse

$$f(a + h) = f(a) + \underbrace{\sum_{j=1}^n \lambda_j h_j}_{df_a(h)} + \|h\| \varepsilon(h)$$

Prenons alors $h = te_j$, alors

$$f(a + te_j) = f(a) + t\lambda_j + |t|\varepsilon(te_j)$$

Et d'après (2.4) on aura $\frac{\partial f}{\partial x_j}(a) = \lambda_j$. ■

Remarque 2.1.8 *La réciproque est en général fausse. Examinons l'exemple suivant :*

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

Il est facile de voir que $\frac{\partial f}{\partial x}(0, 0) = 0 = \frac{\partial f}{\partial y}(0, 0)$. Si f était différentiable, on aurait forcément $df_{(0,0)}(h) = 0$. Or

$$\frac{f(h_1, h_2) - f(0, 0) - df_{(0,0)}(h)}{\|h\|} = \frac{h_1 h_2}{h_1^2 + h_2^2}$$

ne tend manifestement pas vers 0 quand h tend vers $(0, 0)$, elle n'a d'ailleurs pas de limite.

Proposition 2.1.9 *Si f est différentiable en a , alors elle est continue en ce point. De plus elle est dérivable dans le sens de n'importe quel vecteur v et on a*

$$(D_v f)(a) = \langle (\nabla f)(a), v \rangle$$

Démonstration : La différentiabilité en a s'écrit

$$f(a + h) = f(a) + \langle (\nabla f)(a), h \rangle + \|h\|\varepsilon(h)$$

Ce qui donne aisément $\lim_{h \rightarrow 0} f(a + h) = f(a)$. Pour la deuxième affirmation, il suffit de prendre $h = tv$. ■

Dans la remarque précédente nous avons vu que l'existence des dérivées partielles premières n'implique pas en général la différentiabilité. Dans le prochain théorème, nous allons rajouter l'hypothèse de continuité de ces dérivées partielles afin d'obtenir la différentiabilité.

Théorème 2 *Si toutes les dérivées partielles $\frac{\partial f}{\partial x_j}(x)$ existent dans un voisinage de a et sont continues en a , alors f est différentiable au point a .*

Démonstration : Dans cette démonstration, on utilisera la norme $\|x\|_1 = \sum_{i=1}^n |x_i|$

pour sa commodité. Posons

$$g(x) = f(x) - \sum_{i=1}^n x_i \frac{\partial f}{\partial x_i}(a)$$

Nous allons montrer que $g(x+h) - g(a) = \|h\|_1 \varepsilon(h)$, c'est-à-dire que g est différentiable en a de différentielle nulle, ce qui donnera le résultat escompté pour f . On a d'abord

$$\frac{\partial g}{\partial x_j}(x) = \frac{\partial f}{\partial x_j}(x) - \frac{\partial f}{\partial x_j}(a)$$

La continuité des dérivées partielles prise comme hypothèse peut s'exprimer ainsi :

$$\forall \varepsilon > 0 \quad \exists \alpha_\varepsilon > 0 \quad \forall x \in B(a, \alpha_\varepsilon) \quad \forall j = 1, \dots, n \quad \left| \frac{\partial g}{\partial x_j}(x) \right| \leq \varepsilon.$$

Posons

$$\begin{aligned} y_0 &= (a_1, a_2, \dots, a_n) = a \\ y_1 &= (x_1, a_2, \dots, a_n) \\ y_2 &= (x_1, x_2, a_3, \dots, a_n) \\ &\vdots \\ y_n &= (x_1, x_2, \dots, x_n) = x \end{aligned}$$

Définissons n fonctions auxiliaires d'une variable par

$$\begin{aligned} g_k : [a_k, x_k] &\longrightarrow \mathbb{R} \\ t &\longrightarrow g_k(t) = g(x_1, x_2, \dots, x_{k-1}, t, a_{k+1}, \dots, a_n) \end{aligned}$$

L'intervalle de définition peut tout aussi être $[x_k, a_k]$. Remarquer que $g_1(t) = g(t, a_2, \dots, a_n)$ et $g_n(t) = g(x_1, x_2, \dots, x_{n-1}, t)$. D'après les données, les fonctions g_k vérifient les hypothèses du théorème classique des accroissements finis (à une variable). Donc on peut écrire

$$\begin{aligned} g_k(x_k) - g_k(a_k) &= (x_k - a_k) g'_k(c_k) = (x_k - a_k) \frac{\partial g}{\partial x_k}(x_1, x_2, \dots, x_{k-1}, c_k, a_{k+1}, \dots, a_n) \\ &\implies |g_k(x_k) - g_k(a_k)| \leq \varepsilon |x_k - a_k| \end{aligned}$$

On a par ailleurs $g_k(x_k) = g(y_k)$ et $g_k(a_k) = g(y_{k-1})$. De là

$$g(x) - g(a) = g(y_n) - g(y_0) = \sum_{k=1}^n g(y_k) - g(y_{k-1})$$

$$\implies |g(x) - g(a)| \leq \varepsilon \|x - a\|_1$$

ou encore

$$|g(a+h) - g(a)| \leq \varepsilon \|h\|_1$$

c'est-à-dire

$$\lim_{\|h\|_1 \rightarrow 0} \frac{g(a+h) - g(a)}{\|h\|_1} = 0$$

■

2.2 Dérivées partielles d'ordre supérieur

Les dérivées partielles premières $\frac{\partial f}{\partial x_j}$ sont à leur tour des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$, susceptibles d'être dérivables. Leurs dérivées partielles, notées $\frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) := \frac{\partial^2 f}{\partial x_i \partial x_j}$, s'appellent dérivées partielles secondes, et ainsi de suite. Une question naturelle se pose alors : peut-on avoir $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$ pour $i \neq j$? Avec seulement la dérivabilité à l'ordre 2 la réponse est non en général (voir exercice 4 de la série de T.D N°2). Mais avec la dérivabilité à l'ordre 2 et la continuité de ces dérivées deuxièmes la réponse est oui, c'est le théorème de Hermann Schwarz (1843-1921). On dira qu'une fonction est de classe C^p si elle admet des dérivées partielles continues jusqu'à l'ordre p inclu. Sans le théorème de Schwarz, le nombre de dérivées d'ordre p est n^p . Avec le théorème de Schwarz, ce nombre se réduit à $\frac{(n+p-1)!}{p!(n-1)!}$. Nous allons énoncer et démontrer le théorème de Schwarz pour des fonctions de deux variables, cela suffit pour le cas général.

Théorème 3 (de Schwarz)

Soit U un ouvert de $\mathbb{R}^2$, $a \in U$ et $f : U \rightarrow \mathbb{R}$ telle que $\frac{\partial^2 f}{\partial x \partial y}(x, y)$ et $\frac{\partial^2 f}{\partial y \partial x}(x, y)$ existent dans U et soient continues au point a . Alors

$$\frac{\partial^2 f}{\partial x \partial y}(a) = \frac{\partial^2 f}{\partial y \partial x}(a)$$

Démonstration : Soit $h = (h_1, h_2)$ de norme petite. Posons

$$A(h) = f(a_1 + h_1, a_2 + h_2) - f(a_1 + h_1, a_2) - f(a_1, a_2 + h_2) + f(a_1, a_2)$$

Considérons la fonction d'une variable $\varphi(t) = f(t, a_2 + h_2) - f(t, a_2)$. Alors par application du théorème des accroissements finis on a

$$A(h) = \varphi(a_1 + h_1) - \varphi(a_1) = h_1 \varphi'(a_1 + \theta_1 h_1), \quad \text{avec } 0 < \theta_1 < 1$$

et

$$\varphi'(a_1 + \theta_1 h_1) = \frac{\partial f}{\partial x}(a_1 + \theta_1 h_1, a_2 + h_2) - \frac{\partial f}{\partial x}(a_1 + \theta_1 h_1, a_2)$$

On peut appliquer encore une fois le théorème des accroissements finis pour la deuxième variable de la fonction $\frac{\partial f}{\partial x}(a_1 + \theta_1 h_1, \cdot)$. Cela donne

$$A(h) = h_1 h_2 \frac{\partial^2}{\partial y \partial x}(a_1 + \theta_1 h_1, a_2 + \theta_2 h_2), \quad \text{avec } 0 < \theta_2 < 1$$

On considère ensuite la fonction $\psi(t) = f(a_1 + h_1, t) - f(a_1, t)$ et on refait le même travail. On obtient

$$A(h) = h_1 h_2 \frac{\partial^2}{\partial x \partial y}(a_1 + \tilde{\theta}_1 h_1, a_2 + \tilde{\theta}_2 h_2)$$

et donc

$$\frac{\partial^2}{\partial y \partial x}(a_1 + \theta_1 h_1, a_2 + \theta_2 h_2) = \frac{\partial^2}{\partial x \partial y}(a_1 + \tilde{\theta}_1 h_1, a_2 + \tilde{\theta}_2 h_2)$$

On fait tendre alors h vers 0 et la continuité permet de conclure. ■

Exemple 2.2.1 *Considérons la fonction*

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

Il est facile de voir que f est continue partout. Calculons les dérivées partielles premières.

$$\frac{\partial f}{\partial x}(x, y) = \begin{cases} y \cos(\frac{x}{y}) & \text{si } y \neq 0 \\ 0 & \text{si } y = 0 \end{cases}$$

et

$$\frac{\partial f}{\partial y}(x, y) = \begin{cases} 2y \sin(\frac{x}{y}) - x \cos(\frac{x}{y}) & \text{si } y \neq 0 \\ 0 & \text{si } y = 0 \end{cases}$$

Le calcul des dérivées premières en $(x, 0)$ se fait en appliquant la définition (voir (2.4)). Toujours selon la définition on a

$$\frac{\partial^2 f}{\partial x \partial y}(0, 0) = \lim_{t \rightarrow 0} \frac{\frac{\partial f}{\partial y}(t, 0) - \frac{\partial f}{\partial y}(0, 0)}{t} = 0$$

et

$$\frac{\partial^2 f}{\partial y \partial x}(0, 0) = \lim_{t \rightarrow 0} \frac{\frac{\partial f}{\partial x}(0, t) - \frac{\partial f}{\partial x}(0, 0)}{t} = 1$$

D'autre part

$$\frac{\partial^2 f}{\partial y \partial x}(x, y) = \cos(\frac{x}{y}) + \frac{x}{y} \sin(\frac{x}{y}) \quad \text{si } y \neq 0$$

Il est évident que $\frac{\partial^2 f}{\partial y \partial x}$ n'est pas continue en $(0, 0)$ puisqu'elle n'a même pas de limite en ce point. Il n'est pas nécessaire d'examiner l'autre dérivée mixte. Ceci montre l'importance du théorème de Schwarz.

2.3 Formules de Taylor

Nous allons commencer par une formule de dérivation très utile.

Proposition 2.3.1 *Soit $f : U \longrightarrow \mathbb{R}$ différentiable et $\varphi : I \longrightarrow U$ une application d'une variable t continûment dérivable sur un intervalle ouvert réel I*

$$\varphi(t) = (\varphi_1(t), \varphi_2(t), \dots, \varphi_n(t))$$

Posons $\psi(t) = f(\varphi(t))$. Alors

$$\psi'(t) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\varphi(t)) \cdot \varphi'_i(t) \quad (2.6)$$

Démonstration : La formule de Taylor-Lagrange appliquée à l'ordre 1 pour chaque φ_i permet d'écrire pour t fixé

$$\varphi(t + \tau) = \varphi(t) + \tau \varphi'(t) + \tau \theta_t(\tau)$$

où $\lim_{\tau \rightarrow 0} \theta_t(\tau) = 0$. Maintenant la différentiabilité de f donne

$$\psi(t + \tau) = \psi(t) + \langle (\nabla f)(\varphi(t)), \tau \varphi'(t) + \tau \theta_t(\tau) \rangle + \|\tau \varphi'(t) + \tau \theta_t(\tau)\| \varepsilon(\tau \varphi'(t) + \tau \theta_t(\tau))$$

puis

$$\frac{\psi(t + \tau) - \psi(t)}{\tau} = \langle (\nabla f)(\varphi(t)), \varphi'(t) + \theta_t(\tau) \rangle + \frac{|\tau|}{\tau} \|\varphi'(t) + \theta_t(\tau)\| \varepsilon(\tau \varphi'(t) + \tau \theta_t(\tau))$$

En passant à la limite quand τ tend vers 0, on obtient la formule annoncée. ■

Maintenant on est en mesure de donner les formules de Taylor avec reste de Lagrange et avec reste intégral.

Théorème 4 (*Formule de Taylor avec reste de Lagrange*)

Soit $U \subset \mathbb{R}^n$ un ouvert et $f : U \longrightarrow \mathbb{R}$ une application de classe C^{p+1} . Alors, pour h de norme petite, on a

$$\begin{aligned} f(x + h) = \sum_{k=0}^p \frac{1}{k!} \left(\sum_{i_1, \dots, i_k=1}^n h_{i_1} \dots h_{i_k} \frac{\partial^k f(x)}{\partial x_{i_1} \dots \partial x_{i_k}} \right) + \\ + \frac{1}{(p+1)!} \sum_{i_1, \dots, i_{p+1}=1}^n h_{i_1} \dots h_{i_{p+1}} \frac{\partial^{p+1} f(x + \theta h)}{\partial x_{i_1} \dots \partial x_{i_{p+1}}} \end{aligned} \quad (2.7)$$

avec $0 < \theta < 1$.

Démonstration : Définissons une fonction d'une variable t par $\eta(t) = f(x + th)$. Pour $\|h\|$ petite, on peut prendre t dans $[0, 1]$. On écrit tout simplement la formule de Taylor avec reste de Lagrange à l'ordre p pour la fonction η :

$$\eta(t) = \sum_{k=0}^p \frac{\eta^{(k)}(0)}{k!} t^k + \frac{t^{p+1}}{(p+1)!} \eta^{(p+1)}(\theta t)$$

La formule annoncée s'obtient en posant $t = 1$. Reste à évaluer les dérivées de η . On utilise la formule (2.6) :

$$\begin{aligned}\eta'(t) &= \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(x + th) \\ \eta''(t) &= \sum_{i,j=1}^n h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(x + th) \\ &\vdots \\ \eta^{(k)}(t) &= \sum_{i_1, \dots, i_k=1}^n h_{i_1} \dots h_{i_k} \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(x + th)\end{aligned}$$

Il suffit de prendre $t = 0$ quand $k = 0, 1, \dots, p$, puis $t = \theta$ pour $k = p + 1$ dans les expressions précédentes. ■

Théorème 5 (*Formule de Taylor avec reste intégral*)

Sous les mêmes hypothèses que le théorème précédent, on a

$$\begin{aligned}f(x + h) &= \sum_{k=0}^p \frac{1}{k!} \left(\sum_{i_1, \dots, i_k=1}^n h_{i_1} \dots h_{i_k} \frac{\partial^k f(x)}{\partial x_{i_1} \dots \partial x_{i_k}} \right) + \\ &\quad + \frac{1}{(p+1)!} \sum_{i_1, \dots, i_{p+1}=1}^n h_{i_1} \dots h_{i_{p+1}} \int_0^1 (1-\xi)^p \frac{\partial^{p+1} f(x + \xi h)}{\partial x_{i_1} \dots \partial x_{i_{p+1}}} d\xi \quad (2.8)\end{aligned}$$

La démonstration est pareille que la précédente en ce sens qu'elle utilise la formule de Taylor avec reste intégral pour la fonction d'une variable $\eta(\cdot)$. A titre indicatif, nous allons écrire la formule de Taylor-Lagrange à l'ordre 2 pour des fonctions de deux et trois variables. Pour le cas $n = 2, p = 2$, on a

$$\begin{aligned}f(x + h_1, y + h_2) &= f(x, y) + h_1 \frac{\partial f(x, y)}{\partial x} + h_2 \frac{\partial f(x, y)}{\partial y} + \\ &\quad + \frac{1}{2} \left\{ h_1^2 \frac{\partial^2 f(x, y)}{\partial x^2} + 2h_1 h_2 \frac{\partial^2 f(x, y)}{\partial x \partial y} + h_2^2 \frac{\partial^2 f(x, y)}{\partial y^2} \right\} + \mathcal{R}_{Lag}(h) \quad (2.9)\end{aligned}$$

avec

$$\begin{aligned}\mathcal{R}_{Lag}(h) &= \frac{1}{6} \left\{ h_1^3 \frac{\partial^3 f(x + \theta h_1, y + \theta h_2)}{\partial x^3} + 3h_1^2 h_2 \frac{\partial^3 f(x + \theta h_1, y + \theta h_2)}{\partial x^2 \partial y} + \right. \\ &\quad \left. + 3h_1 h_2^2 \frac{\partial^3 f(x + \theta h_1, y + \theta h_2)}{\partial x \partial y^2} + h_2^3 \frac{\partial^3 f(x + \theta h_1, y + \theta h_2)}{\partial y^3} \right\}\end{aligned}$$

Notez que nous avons utilisé le théorème de Schwarz. Notez aussi que la bornitude des dérivées troisièmes implique que $|\mathcal{R}_{Lag}(h)| \leq M \|h\|^3$.

Pour $n = 3, p = 2$, on aura

$$\begin{aligned} f(x + h_1, y + h_2, z + h_3) = & f(x, y, z) + h_1 \frac{\partial f(x, y, z)}{\partial x} + h_2 \frac{\partial f(x, y, z)}{\partial y} + h_3 \frac{\partial f(x, y, z)}{\partial z} \\ & + \frac{1}{2} \left\{ h_1^2 \frac{\partial^2 f(x, y, z)}{\partial x^2} + h_2^2 \frac{\partial^2 f(x, y, z)}{\partial y^2} + h_3^2 \frac{\partial^2 f(x, y, z)}{\partial z^2} + \right. \\ & \left. + 2h_1 h_2 \frac{\partial^2 f(x, y, z)}{\partial x \partial y} + 2h_1 h_3 \frac{\partial^2 f(x, y, z)}{\partial x \partial z} + 2h_2 h_3 \frac{\partial^2 f(x, y, z)}{\partial y \partial z} \right\} + \mathcal{R}_{Lag}(h) \end{aligned} \quad (2.10)$$

On peut condenser les écritures précédentes en introduisant la notion de *hessienne*.

Définition 2.3.2 On appelle *hessienne* la matrice des dérivées secondes définie par

$$Hess_f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

Remarquons que grâce au théorème de Schwarz la matrice hessienne est symétrique (par rapport à sa diagonale principale). On peut réécrire à présent la formule de Taylor-Lagrange à l'ordre 2 pour n quelconque d'ailleurs ainsi :

$$f(x + h) = f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2} \langle Hess_f(x)h, h \rangle + \mathcal{R}_{Lag}(h) \quad (2.11)$$

2.4 Différentiabilité de fonctions vectorielles

On entend par fonction vectorielle, une fonction à valeurs dans $\mathbb{R}^p$ i.e,

$$\begin{aligned} f : \mathbb{R}^n &\longrightarrow \mathbb{R}^p \\ x &\longrightarrow f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_p(x) \end{pmatrix} \end{aligned}$$

Définition 2.4.1 Une fonction $f : U \longrightarrow \mathbb{R}^p$ est dite différentiable en $a \in U$ s'il existe une application linéaire, notée $df_a : \mathbb{R}^n \longrightarrow \mathbb{R}^p$, telle que

$$\lim_{\|h\| \rightarrow 0} \frac{\|f(a + h) - f(a) - df_a(h)\|_{\mathbb{R}^p}}{\|h\|_{\mathbb{R}^n}} = 0 \quad (2.12)$$

On omettra les indices dans les normes quand il n'y a pas de confusion. On pourra écrire aussi la différentiabilité comme suit :

$$f(a+h) = f(a) + df_a(h) + \|h\|\varepsilon(h) \quad (2.13)$$

où $\varepsilon(\cdot)$ est une fonction vectorielle telle que $\lim_{\|h\| \rightarrow 0} \|\varepsilon(h)\| = 0$.

La linéarité de df_a s'exprime par (se référer au cours d'Algèbre Linéaire)

$$df_a(h) = \begin{pmatrix} \langle \nabla f_1(a), h \rangle \\ \langle \nabla f_2(a), h \rangle \\ \vdots \\ \langle \nabla f_p(a), h \rangle \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1(a)}{\partial x_1} & \frac{\partial f_1(a)}{\partial x_2} & \cdots & \frac{\partial f_1(a)}{\partial x_n} \\ \frac{\partial f_2(a)}{\partial x_1} & \frac{\partial f_2(a)}{\partial x_2} & \cdots & \frac{\partial f_2(a)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_p(a)}{\partial x_1} & \frac{\partial f_p(a)}{\partial x_2} & \cdots & \frac{\partial f_p(a)}{\partial x_n} \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix} \quad (2.14)$$

La matrice $p \times n$ précédente représentant df_a s'appelle *jacobienne* de f en a . Elle est aussi notée

$$J_f(a) = \left(\frac{\partial f_i(a)}{\partial x_j} \right)_{1 \leq i \leq p, 1 \leq j \leq n} \quad (2.15)$$

Exemple 2.4.2 Si $f(x, y, z) = \begin{pmatrix} xy + xz + yz \\ xyz \end{pmatrix}$ alors

$$J_f(x, y, z) = \begin{pmatrix} y + z & x + z & x + y \\ yz & xz & xy \end{pmatrix}$$

La différentielle (et par "ricochet" la jacobienne J_f) est linéaire par rapport à f i.e, $d(\lambda f + \mu g) = \lambda df + \mu dg$ (ou bien $J_{\lambda f + \mu g} = \lambda J_f + \mu J_g$). Ceci est le résultat évident de la linéarité de la dérivation partielle. On donne maintenant la règle, tout aussi utile, de la dérivation des fonctions composées.

Proposition 2.4.3 (*Dérivation des fonctions composées*)

Soient $f : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ et $g : \mathbb{R}^p \longrightarrow \mathbb{R}^q$ différentiables. Alors

$$d(g \circ f)_a = dg_{f(a)} \circ df_a \quad \text{ou bien} \quad J_{g \circ f}(a) = J_g(f(a)) \cdot J_f(a) \quad (2.16)$$

Pour ne pas se tromper dans la multiplication des matrices jacobienes, il faut se rappeler qu'on multiplie une matrice $q \times p$ (qui doit être écrite à gauche) par une matrice $p \times n$ (écrite à droite) pour obtenir une matrice $q \times n$.

Démonstration : Posons $y = f(x)$ et $z = g(y)$ pour distinguer les variables et les valeurs des deux fonctions. Alors pour $1 \leq i \leq q$ et $1 \leq j \leq n$

$$\frac{\partial (g \circ f)_i}{\partial x_j} = \frac{\partial}{\partial x_j} g_i(f_1(x), f_2(x), \dots, f_p(x)) = \sum_{k=1}^p \frac{\partial g_i(f(x))}{\partial y_k} \frac{\partial f_k(x)}{\partial x_n}$$

où nous avons fait usage de (2.6). C'est exactement la formule annoncée. ■

2.5 Inversion locale

Nous allons présenter une généralisation d'un résultat abordé en première année (Analyse 1), celui de l'inversibilité locale de fonctions C^1 . En effet, si pour une fonction d'une variable de classe C^1 on a $\varphi'(a) \neq 0$ alors elle est bijective d'un voisinage de a vers un voisinage de $\varphi(a)$. De plus la réciproque est de classe C^1 et $(\varphi^{-1})' = \frac{1}{\varphi' \circ \varphi^{-1}}$. La condition $\varphi'(a) \neq 0$ sera généralisée par l'inversibilité de la différentielle. L'outil essentielle dans la démonstration qui suivra est celui du théorème du point fixe de Banach. On peut l'énoncer dans le cas particulier de $\mathbb{R}^n$. Soit $T : \mathcal{F} \longrightarrow \mathcal{F}$ une application contractante d'un fermé $\mathcal{F}$ dans lui même i.e.,

$$\exists \rho \in [0, 1[\quad \text{tel que} \quad \forall x, x' \in \mathcal{F} \quad \|T(x) - T(x')\| \leq \rho \|x - x'\|$$

Alors T admet dans $\mathcal{F}$ un point fixe unique, c'est-à-dire il existe un unique point $x_* \in \mathcal{F}$ tel que $T(x_*) = x_*$. Ce résultat a été largement discuté dans le cours de Topologie.

Théorème 6 *Soit U un ouvert de $\mathbb{R}^n$ et soit $f : U \longrightarrow \mathbb{R}^n$ une application de classe C^1 . Si en un point a de U la différentielle df_a est inversible ($\det(J_f(a)) \neq 0$), alors il existe V_a un voisinage de a et $W_{f(a)}$ un voisinage de $f(a)$ tels que $f : V_a \longrightarrow W_{f(a)}$ est bijective. L'application réciproque f^{-1} est aussi de classe C^1 , de plus*

$$(df^{-1})_{f(x)} = (df_x)^{-1}$$

pour tout $x \in V_a$.

Démonstration : Notons d'abord que dans ce théorème la jacobienne est une matrice carrée $n \times n$. C'est pourquoi son inversibilité a un sens et se vérifie par le fait que son déterminant est non nul. Notons par $\mathcal{M}_n(\mathbb{R})$ l'espace vectoriel des matrices réelles d'ordre n qui est de dimension n^2 . Étant de dimension finie, toutes les normes qu'on peut y définir sont équivalentes. Nous allons travailler ici avec la norme d'opérateur définie par

$$\forall A \in \mathcal{M}_n(\mathbb{R}), \quad \|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

car elle vérifie

$$\forall x \in \mathbb{R}^n, \quad \|Ax\| \leq \|A\| \|x\|$$

Nous avons utilisé la même notation pour la norme d'une matrice et d'un vecteur, mais il n'y a pas de confusion possible. Selon les hypothèses, on a pour la différentiabilité en a

$$\forall \alpha > 0, \exists \delta_1 > 0, \quad \|x - a\| \leq \delta_1 \implies \|f(x) - f(a) - df_a(x - a)\| \leq \alpha \|x - a\| \quad (2.17)$$

et pour la continuité des dérivées partielles premières

$$\forall \alpha > 0, \exists \delta_2 > 0, \quad \|x - a\| \leq \delta_2 \implies \|J_f(x) - J_f(a)\| \leq \alpha \quad (2.18)$$

On écrira indifféremment df_a ou $J_f(a)$. Considérons $\delta_3 = \min(\delta_1, \delta_2)$ afin que les deux inégalités précédentes soit vérifiées dans la même boule fermée de centre a . Pour montrer la bijection, nous allons montrer que pour y fixé proche de $f(a)$, il existe un unique x proche de a tel que $y = f(x)$. Définissons la fonction

$$g_y(x) = x + [J_f(a)]^{-1} (y - f(x))$$

Si $\|y - f(a)\| \leq \beta$, alors de

$$\|g_y(x) - a\| = \|[J_f(a)]^{-1}\| \{y - f(a) - (f(x) - f(a) - J_f(a)(x - a))\}$$

on déduit que

$$\|g_y(x) - a\| \leq \|[J_f(a)]^{-1}\| (\beta + \alpha\delta_3) \leq \delta_3$$

avec un choix adéquat des paramètres libres α et β . Par exemple $\alpha < \frac{1}{\|[J_f(a)]^{-1}\|}$ puis

$\beta \leq \delta_3 \left(\frac{1}{\|[J_f(a)]^{-1}\|} - \alpha \right)$. Cette dernière inégalité affirme que la fonction (manifestement continue) g_y envoie la boule fermée $\overline{B}(a, \delta_3)$ dans elle même. Montrons qu'on peut affiner le choix des paramètres pour qu'elle soit contractante. En effet

$$g_y(x) - g_y(x') = x - x' + [J_f(a)]^{-1} (f(x') - f(x))$$

peut-être arrangée comme suit

$$g_y(x) - g_y(x') = [J_f(a)]^{-1} \{ (f(x') - f(x) - J_f(x)(x' - x)) + (J_f(x) - J_f(a))(x' - x) \}$$

et donc

$$\|g_y(x) - g_y(x')\| \leq 2\alpha \|[J_f(a)]^{-1}\| \|x - x'\|$$

On n'a plus qu'à exiger en outre que

$$\rho = 2\alpha \|[J_f(a)]^{-1}\| < 1 \implies \alpha < \frac{1}{2\|[J_f(a)]^{-1}\|}$$

D'après le théorème du point fixe de Banach, il existe un unique $x_* \in \overline{B}(a, \delta_3)$ tel que $g_y(x_*) = x_* \implies y = f(x_*)$ i.e, f est bijective de $\overline{B}(a, \delta_3)$ vers $\overline{B}(f(a), \beta)$. Notons f^{-1} sa fonction réciproque. Prenons $y = f(a)$ dans la définition de g_y et deux points quelconques x, x' dans $\overline{B}(a, \delta_3)$. Alors

$$x - x' = g_{f(a)}(x) - g_{f(a)}(x') + [J_f(a)]^{-1} (f(x) - f(x'))$$

et donc

$$\|x - x'\| \leq \rho \|x - x'\| + \|[J_f(a)]^{-1}\| \|f(x) - f(x')\|$$

$$\implies \|x - x'\| \leq \frac{\|[J_f(a)]^{-1}\|}{1 - \rho} \|f(x) - f(x')\|$$

$$\implies \|f^{-1}(y) - f^{-1}(y')\| \leq \frac{\|[J_f(a)]^{-1}\|}{1 - \rho} \|y - y'\|$$

Ceci assure la continuité de f^{-1} . Reste à montrer sa différentiabilité.

Notons $\gamma = \frac{\| [J_f(a)]^{-1} \|}{1 - \rho}$. Rappelons qu'on peut s'inspirer des idées développées dans l'exercice 3 de la série de T.D.N°1, pour montrer que le choix précédent de α permet de dire que pour tout $x \in \overline{B}(a, \delta_3)$ la jacobienne $J_f(x)$ est inversible car elle est suffisamment proche de $J_f(a)$. Fixons $y = f(x)$ dans $\overline{B}(f(a), \beta)$. Prenons w de norme petite pour que $y + w \in \overline{B}(f(a), \beta)$ et posons $h = f^{-1}(y + w) - f^{-1}(y)$. Alors $\|h\| \leq \gamma \|w\|$. Aussi $y + w = f(x + h)$. D'où

$$\begin{aligned} f^{-1}(y + w) - f^{-1}(y) - (J_f(x))^{-1}(w) &= h - (J_f(x))^{-1}(w) \\ &= -(J_f(x))^{-1}(f(x + h) - f(x) - (J_f(x))(h)) \end{aligned}$$

On sait que

$$\|f(x + h) - f(x) - (J_f(x))(h)\| = \|h\| \|\varepsilon(h)\|$$

avec $\lim_{\|h\| \rightarrow 0} \|\varepsilon(h)\| = 0$. De là

$$\|f^{-1}(y + w) - f^{-1}(y) - (J_f(x))^{-1}(w)\| \leq \gamma \|(J_f(x))^{-1}\| \|w\| \|\varepsilon(h)\|$$

Ceci montre bien que f^{-1} est différentiable en y avec $(df^{-1})_y = (J_f(x))^{-1}$. ■

Exemple 2.5.1 *Considérons la fonction*

$$f(x_1, x_2) = \begin{pmatrix} x_1 + x_2 \\ x_1 x_2 \end{pmatrix}$$

qui est $C^\infty(\mathbb{R}^2)$. Il est facile de calculer sa jacobienne

$$J_f(x_1, x_2) = \begin{pmatrix} 1 & 1 \\ x_2 & x_1 \end{pmatrix}$$

et $\det(J_f(x_1, x_2)) = x_1 - x_2$. Si on se met au voisinage d'un point $a = (a_1, a_2)$ avec $a_1 \neq a_2$, par exemple une boule euclidienne de centre a et de rayon $r < \frac{|a_1 - a_2|}{\sqrt{2}}$ qui est la distance euclidienne de a à la droite Δ d'équation $x_1 - x_2 = 0$, alors f y est bijective. Contrairement au cas général où les calculs explicites sont très rares, on peut dans ce cas particulier pousser les calculs un peu plus loin. Par exemple si $y_1 = x_1 + x_2$ et $y_2 = x_1 x_2$, alors au voisinage de a avec $a_1 > a_2$, on aura

$$x_1 = \frac{y_1 + \sqrt{y_1^2 - 4y_2}}{2}, \quad x_2 = \frac{y_1 - \sqrt{y_1^2 - 4y_2}}{2}$$

et

$$J_{f^{-1}}(y_1, y_2) = (J_f)^{-1} = \frac{1}{x_1 - x_2} \begin{pmatrix} x_1 & -1 \\ -x_2 & 1 \end{pmatrix} = \frac{1}{\sqrt{y_1^2 - 4y_2}} \begin{pmatrix} \frac{y_1 + \sqrt{y_1^2 - 4y_2}}{2} & -1 \\ \frac{-y_1 + \sqrt{y_1^2 - 4y_2}}{2} & 1 \end{pmatrix}$$

2.6 Théorème des fonctions implicites

La question abordée dans la section précédente concerne la résolution de l'équation $f(x) = y$. C'est en réalité un système de n équations à n inconnues i.e,

$$\begin{cases} f_1(x_1, x_2, \dots, x_n) = y_1 \\ f_2(x_1, x_2, \dots, x_n) = y_2 \\ \vdots \\ f_n(x_1, x_2, \dots, x_n) = y_n \end{cases}$$

Nous allons dans cette section examiner des systèmes de p équations à n indéterminées avec $p < n$ i.e,

$$\begin{cases} f_1(x_1, x_2, \dots, x_n) = y_1 \\ f_2(x_1, x_2, \dots, x_n) = y_2 \\ \vdots \\ f_p(x_1, x_2, \dots, x_n) = y_p \end{cases}$$

Il est clair qu'il faudra choisir parmi les n variables x_i , p variables par rapport auxquelles il sera possible de résoudre, le reste seront considérées comme paramètres. La condition d'existence et d'unicité de la solution, comme on peut s'y attendre, est celle donnée dans la section précédente à savoir l'inversibilité de la sous-jacobienne $p \times p$ formées avec les dérivées des f_i par rapports aux variables choisies. Pour énoncer le résultat suivant, nous allons privilégier les variables $x_1, x_2, \dots, x_p$.

Théorème 7 *Identifions d'abord $\mathbb{R}^n$ à $\mathbb{R}^p \times \mathbb{R}^{n-p}$. Soit $f : \mathbb{R}^p \times \mathbb{R}^{n-p} \rightarrow \mathbb{R}^p$ de classe C^1 et $a = (a', a'') \in \mathbb{R}^p \times \mathbb{R}^{n-p}$ avec $b = f(a', a'')$. Supposons que la matrice carrée*

$$J_p^f = \left(\frac{\partial f_i(a)}{\partial x_j} \right)_{1 \leq i, j \leq p}$$

soit inversible. Alors il existe $U \subset \mathbb{R}^p$ voisinage de a' , $V \subset \mathbb{R}^{n-p}$ voisinage de a'' et $W \subset \mathbb{R}^p$ voisinage de b , et une fonction $\varphi : V \times W \rightarrow U$, tels que

$$\forall x'' \in V, \forall y \in W \quad x' = \varphi(x'', y)$$

est l'unique solution du système $f(x', x'') = y$; autrement dit $f(\varphi(x'', y), x'') = y$. On dit alors que $x' = \varphi(x'', y)$ est une fonction implicite définie par le système $f(x) = y$.

Démonstration : Il est d'abord clair, d'après les notations, que $x = (x', x'')$, c'est-à-dire $x' = (x_1, x_2, \dots, x_p)$ et $x'' = (x_{p+1}, x_{p+2}, \dots, x_n)$. Pour la démonstration, on se

ramène au théorème de l'inversion locale en complétant le système de la façon suivante :

$$\left\{ \begin{array}{l} f_1(x_1, x_2, \dots, x_n) = y_1 \\ f_2(x_1, x_2, \dots, x_n) = y_2 \\ \vdots \\ f_p(x_1, x_2, \dots, x_n) = y_p \\ \quad x_{p+1} = z_{p+1} \\ \quad x_{p+2} = z_{p+2} \\ \quad \vdots \\ \quad x_n = z_n \end{array} \right.$$

Les membres de gauche définissent une fonction $\tilde{f} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ de classe C^1 . Sa jacobienne s'écrit (triangulaire par blocs)

$$J_{\tilde{f}} = \begin{pmatrix} \left(\frac{\partial f_i}{\partial x_j} \right)_{1 \leq i, j \leq p} & \left(\frac{\partial f_i}{\partial x_j} \right)_{1 \leq i \leq p, p+1 \leq j \leq n} \\ 0 & I_{n-p, n-p} \end{pmatrix}$$

On a noté par $I_{n-p, n-p}$ la matrice identité $(n-p) \times (n-p)$. Il est facile de voir que $\det(J_{\tilde{f}}) = \det(J_p^f)$, et donc le théorème de l'inversion locale s'applique. La réciproque $\tilde{f}^{-1}$ existe dans un voisinage de (y, z) et $(x', x'') = \tilde{f}^{-1}(y, z)$. La fonction φ est formée des p premières composantes. Remarquez aussi que $z = x''$. ■

Exemple 2.6.1 *Considérons l'équation $x^2 + y^2 + z^2 = 1$ de la sphère euclidienne de centre $(0, 0, 0)$ et de rayon 1. Plaçons-nous par exemple au voisinage du "pôle nord" $(0, 0, 1)$. C'est une équation à trois inconnues i.e, $f(x, y, z) = x^2 + y^2 + z^2 = 1$. On a $f(0, 0, 1) = 1$ et $\frac{\partial f(0, 0, 1)}{\partial z} = 2z \mid_{(0, 0, 1)} = 2 \neq 0$. Donc dans un petit voisinage de $(0, 0, 1)$, z est une fonction de (x, y) . Dans ce cas particulier on peut même expliciter cette fonction : $z = \varphi(x, y) = \sqrt{1 - x^2 - y^2}$. Remarquons qu'au voisinage du même point x (ou bien y) n'est pas fonction de (y, z) . En effet, il existe deux solutions $x = \pm \sqrt{1 - y^2 - z^2}$ qui sont toutes les deux possibles, ce qui exclut la bijection (l'unicité de la solution). On peut aussi le vérifier par $\frac{\partial f(0, 0, 1)}{\partial x} = 2x \mid_{(0, 0, 1)} = 0$, non inversibilité de la sous-jacobienne.*

Exemple 2.6.2 *Regardons le système*

$$\begin{cases} f_1(x, y, z) = x^2 - y^2 + z^2 = 1 \\ f_2(x, y, z) = xyz = 1 \end{cases}$$

au voisinage du point $(1, 1, 1)$ qui est solution particulière. La jacobienne est

$$J_f(x, y, z) = \begin{pmatrix} 2x & -2y & 2z \\ yz & xz & xy \end{pmatrix}$$

Si on privilégie les variables (x, y) , il faut considérer la sous-jacobienne

$$J_{(x,y)}^f = \begin{pmatrix} 2x & -2y \\ yz & xz \end{pmatrix}$$

dont le déterminant est $\det(J_{(x,y)}^f) = 2z(x^2 + y^2) = 4$ au point $(1, 1, 1)$. Donc au voisinage de ce point, x et y sont fonctions de z i.e.,

$$x = \varphi_1(z) \quad , \quad y = \varphi_2(z)$$

On laisse au lecteur la tâche (un peu lourde !) de calculer les expressions de ces deux fonctions.

2.7 Extremums locaux

Dans ce qui suit nous allons examiner la question des maximums et minimums (extremums) locaux de fonctions de $\mathbb{R}^n \rightarrow \mathbb{R}$. Pour une variable ($n = 1$), il est bien connu que si f est C^1 et admet un extremum en $x = a$, alors $f'(a) = 0$. La réciproque est en général fautive, à savoir si a est un point critique ($f'(a) = 0$) alors on n'a pas forcément un extremum. Voir par exemple $f(x) = x^3$ qui n'a ni maximum, ni minimum en $x = 0$. Ce point s'appelle un point col (ou parfois point plat). Si le point critique n'est pas dégénéré ($f''(a) \neq 0$), alors dans ce cas on aura un maximum (respect. minimum) si $f''(a) < 0$ (respect. $f''(a) > 0$). Ce sont ces résultats que nous allons généraliser à $\mathbb{R}^n$.

Définition 2.7.1 Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$. On dit qu'elle admet un maximum local au point a , s'il existe V_a un voisinage de a tel que

$$\forall x \in V_a \quad f(x) \leq f(a)$$

Un maximum est dit strict si en plus on a $\forall x \in V_a \setminus \{a\} \quad f(x) < f(a)$.

Pour un minimum on a

$$\forall x \in V_a \quad f(x) \geq f(a)$$

Et pour le minimum strict, l'inégalité est stricte pour $x \neq a$.

Proposition 2.7.2 Si f admet en a un extremum alors $(\nabla f)(a) = 0$, ou encore $\forall i = 1, 2, \dots, n \quad \frac{\partial f(a)}{\partial x_i} = 0$.

Démonstration : Une manière simple de le prouver est de se ramener à une variable. En effet, si on pose $\varphi(t) = f(a_1, \dots, a_{i-1}, a_i + t, a_{i+1}, \dots, a_n)$, alors φ admet en 0 un extremum, d'où $\varphi'(0) = 0$. C'est le résultat escompté. ■

Définition 2.7.3 Si $(\nabla f)(a) = 0$, on dit que a est un point critique, et qu'il est non dégénéré si $\det(\text{Hess}_f(a)) \neq 0$.

La précédente proposition dit en particulier qu'il faut chercher les extremums parmi les points critiques.

Quelques rappels d'algèbre linéaire

Une matrice carrée $A = (a_{ij})_{i,j=1}^n$ est dite symétrique si $\forall i, j = 1, \dots, n \quad a_{ij} = a_{ji}$. On a alors $\forall x, y \in \mathbb{R}^n \quad \langle Ax, y \rangle = \langle x, Ay \rangle$. Une matrice carrée A est dite positive si $\forall x \in \mathbb{R}^n \quad \langle Ax, x \rangle \geq 0$, et elle est dite définie positive si elle est déjà positive avec en plus $\langle Ax, x \rangle = 0 \implies x = 0$.

Une matrice A symétrique est diagonalisable dans une base orthogonale i.e,

$$\exists P \text{ telle que } P^t P = P P^t = I \quad \text{et} \quad A = P D P^t$$

où P^t désigne la matrice transposée de P . La condition sur P exprime le caractère orthogonal. La matrice D est diagonale

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

où les λ_i sont les valeurs propres de A . La positivité de A est équivalente à la positivité des valeurs propres. La définie positivité est, elle, équivalente à la stricte positivité des valeurs propres. On peut montrer, sans grande difficulté, que l'encadrement suivant a lieu

$$\lambda_1 \|x\|^2 \leq \langle Ax, x \rangle \leq \lambda_n \|x\|^2 \quad (2.19)$$

si $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. Vérifier qu'une matrice symétrique est définie positive à l'aide de la définition ou bien grâce aux valeurs propres peut s'avérer une tâche très difficile. C'est pourquoi le critère suivant est de loin le plus pratique pour cette vérification. Donnons avant, la définition des mineurs principaux d'une matrice. On appelle mineurs principaux d'une matrice $A = (a_{ij})_{i,j=1}^n$ les déterminants suivants :

$$\Delta_1 = a_{11} \text{ , } \Delta_2 = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \text{ , } \Delta_3 = \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \text{ , } \dots \text{ , } \Delta_n = \det(A)$$

Critère de définie positivité : une matrice symétrique A est définie positive si et seulement si tous ses mineurs principaux sont strictement positifs.

$$\text{Exemple : } A = \begin{pmatrix} 4 & -1 & 3 & 0 \\ -1 & 2 & -1 & 1 \\ 3 & -1 & 5 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \text{ est définie positive car}$$

$$\Delta_1 = 4 \text{ , } \Delta_2 = 7 \text{ , } \Delta_3 = 19 \text{ , } \Delta_4 = 8$$

Enfin une matrice symétrique est négative (resp. définie négative) si $-A$ est positive (resp. définie positive). Elle est non définie si elle n'est ni positive ni négative i.e, elle admet des valeurs propres positives et négatives.

Théorème 8 Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ de classe C^2 et a un point critique. Alors

1. Si f admet en a un minimum local, alors la matrice hessienne $Hess_f(a)$ est positive.
2. Si f admet en a un maximum local, alors la matrice hessienne $Hess_f(a)$ est négative

Inversement, en supposant de plus qu'il est non dégénéré,

1. Si $Hess_f(a)$ est définie positive, alors f admet en a un minimum strict.
2. Si $Hess_f(a)$ est définie négative, alors f admet en a un maximum strict.

Si la hessienne est non définie, on dit que a est un col.

Démonstration : Tout est basé sur la formule de Taylor-Young

$$f(a+h) = f(a) + \langle h, (\nabla f)(a) \rangle + \frac{1}{2} \langle Hess_f(a).h, h \rangle + \|h\|^2 \varepsilon(h)$$

Comme a est un point critique, alors en prenant $h = t.u$, avec u un vecteur unitaire ($\|u\| = 1$),

$$f(a+tu) - f(a) = t^2 \left(\frac{1}{2} \langle Hess_f(a).u, u \rangle + \varepsilon(t.u) \right)$$

Faisons la démonstration pour un minimum. Pour $|t|$ petite, on a $f(a+tu) - f(a) \geq 0$, donc $\frac{1}{2} \langle Hess_f(a).u, u \rangle + \varepsilon(t.u) \geq 0$. En faisant tendre t vers 0 on aura la positivité

de la hessienne (à partir de u on revient à n'importe quel vecteur en posant $u = \frac{\xi}{\|\xi\|}$).

Inversement, si la hessienne est définie positive, alors sa plus petite valeur propre est strictement positive, $\lambda_1 > 0$. En utilisant (2.19) on aura

$$f(a+h) - f(a) \geq \|h\|^2 \left(\frac{1}{2} \lambda_1 + \varepsilon(h) \right)$$

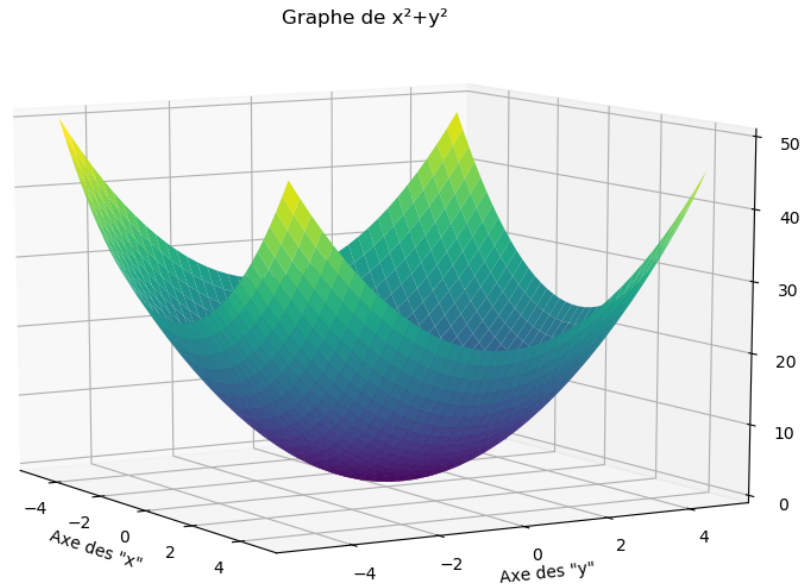
D'où le résultat en prenant $\|h\|$ suffisamment petite pour que $\frac{1}{2} \lambda_1 + \varepsilon(h) > 0$. ■

Exemple 2.7.4 Soit $f(x, y) = x^2 + y^2$. Alors

$$\nabla f(x, y) = \begin{pmatrix} 2x \\ 2y \end{pmatrix} \quad \text{et} \quad Hess_f(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Il y a un seul point critique $(0, 0)$ et il est non dégénéré puisque partout $\det(Hess_f) = 4 \neq 0$. Aussi la hessienne est définie positive. Donc en $(0, 0)$ il y a un minimum strict (ce qui est par ailleurs facilement détectable.)

Voici une illustration par le dessin du graphe dans un voisinage de $(0, 0)$

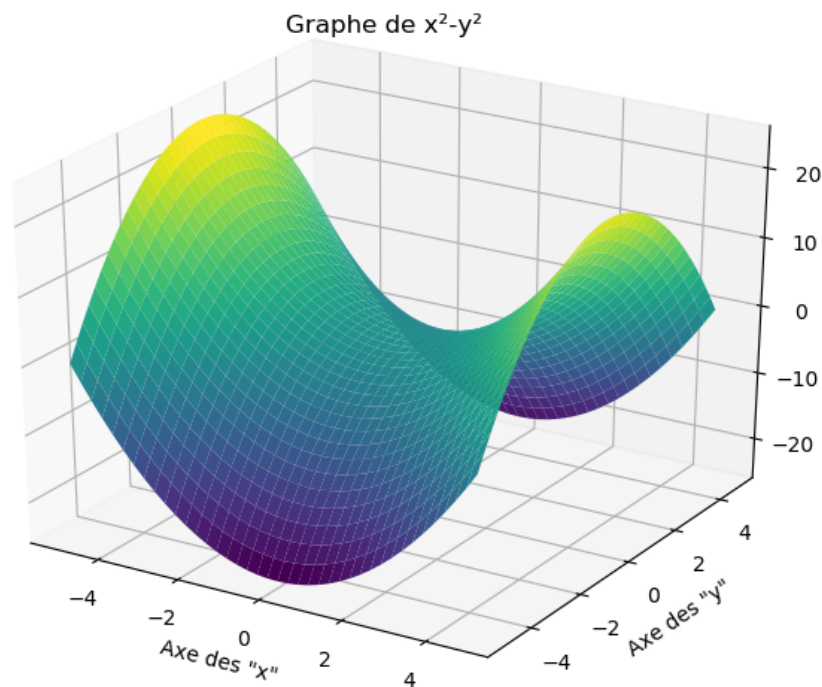


Exemple 2.7.5 Soit $g(x, y) = x^2 - y^2$. Alors

$$\nabla g(x, y) = \begin{pmatrix} 2x \\ -2y \end{pmatrix} \quad \text{et} \quad \text{Hess}_g(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

Il y a un seul point critique $(0, 0)$ et il est non dégénéré puisque partout $\det(\text{Hess}_g) = -4 \neq 0$. Dans ce cas la hessienne est non définie. Donc en $(0, 0)$ il y a un point "col" ou bien "selle".

Voici le graphe dans un voisinage de $(0, 0)$.



On voit que suivant un certain chemin sur cette surface $(0, 0, 0)$ apparaît comme un maximum, et suivant un autre chemin, il apparaît comme un minimum.

2.8 Extrema liés

Dans cette dernière section nous allons présenter la notion d'extrema liés (le mot extrema est une forme du pluriel d'extremum). C'est la recherche de maximum ou de minimum d'une fonction $f(x)$ avec en plus une condition (contrainte) sur les variables ($g(x) = 0$). On peut même envisager plusieurs contraintes. Le nombre de contraintes p est dans tous les cas strictement inférieur à n ($p < n$). Posons

$$\mathcal{C} = \{x \in \mathbb{R}^n \mid g(x) = 0\}$$

Le problème est d'étudier l'existence de $\min_{x \in \mathcal{C}} f(x)$ ou $\max_{x \in \mathcal{C}} f(x)$.

Proposition 2.8.1 *On suppose que les fonctions f et g sont de classe C^1 . Si f admet en a un extremum lié alors il existe un unique nombre réel λ tel que*

$$(\nabla f)(a) = \lambda(\nabla g)(a)$$

Ce nombre s'appelle multiplicateur de Lagrange.

Démonstration : Soit $\gamma : J \rightarrow \mathcal{C}$ une application différentiable définie sur un petit intervalle ouvert J contenant 0 et telle que $\gamma(0) = a$. D'après les hypothèses, l'application d'une variable réelle $f(\gamma(t))$ admet un extremum en 0, donc sa dérivée en 0 s'annule i.e.,

$$\langle (\nabla f)(a), \gamma'(0) \rangle = 0$$

D'autre part, $g(\gamma(t)) = 0$, d'où

$$\langle (\nabla g)(a), \gamma'(0) \rangle = 0$$

On apprend dans le cours de géométrie que le vecteur $\gamma'(0)$ est tangent à l'hyper-surface $\mathcal{C}$, et donc parcourt un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $n - 1$. Les deux vecteurs $(\nabla f)(a)$ et $(\nabla g)(a)$ sont donc orthogonaux à un même sous-espace de dimension $n - 1$. Il sont de ce fait colinéaires. ■

Posons $F(x, \lambda) = f(x) - \lambda g(x)$. D'après cette proposition chercher les extremums liés revient à résoudre le système

$$\begin{cases} (\nabla_x F)(x, \lambda) = 0 \\ \frac{\partial F}{\partial \lambda}(x, \lambda) = -g(x) = 0 \end{cases}$$

par rapport à x et λ . C'est un système de $n + 1$ équations à $n + 1$ inconnues.

Supposons que (a_*, λ_*) soit un point critique de F , c'est-à-dire une solution du système précédent. Si on suppose en plus que f et g sont de classe C^2 , alors la sous-hessienne de F au point (a_*, λ_*) qui concerne les variables x_i , renseignera sur la nature de ce point critique. Nous verrons plus clairement à travers les exemples qui suivront.

Exemple 2.8.2 *On veut résoudre par les méthodes développées ici, un problème bien connu en géométrie euclidienne plane. Il s'agit de trouver la distance euclidienne d'un*

point $A(a, b)$ à une droite (D) d'équation $\alpha x + \beta y + \gamma = 0$ ($\alpha^2 + \beta^2 \neq 0$). Soit $M(x, y) \in (D)$. La distance euclidienne de A à M est donnée par :

$$d(A, M) = \sqrt{(x - a)^2 + (y - b)^2}$$

Pour bénéficier du caractère C^∞ , nous travaillerons plutôt avec la fonction

$$f(x, y) = (x - a)^2 + (y - b)^2 = d(A, M)^2$$

car la distance est minimale si et seulement si f l'est. Ici la contrainte est manifestement $g(x, y) = \alpha x + \beta y + \gamma$. On a

$$F(x, y, \lambda) = (x - a)^2 + (y - b)^2 - \lambda(\alpha x + \beta y + \gamma)$$

et le système à résoudre

$$\begin{cases} \frac{\partial F}{\partial x} = 2(x - a) - \lambda\alpha = 0 \\ \frac{\partial F}{\partial y} = 2(y - b) - \lambda\beta = 0 \\ \frac{\partial F}{\partial \lambda} = -(\alpha x + \beta y + \gamma) = 0 \end{cases}$$

Des deux premières équations on aura $x = a + \frac{\lambda\alpha}{2}$ et $y = b + \frac{\lambda\beta}{2}$. En les remplaçant dans la troisième équation on a $\alpha a + \beta b + \gamma + \frac{\lambda}{2}(\alpha^2 + \beta^2) = 0$. La solution du système est donc

$$\begin{cases} \lambda_* = \frac{-2}{\alpha^2 + \beta^2}(\alpha a + \beta b + \gamma) \\ x_* = a - \frac{\alpha}{\alpha^2 + \beta^2}(\alpha a + \beta b + \gamma) \\ y_* = b - \frac{\beta}{\alpha^2 + \beta^2}(\alpha a + \beta b + \gamma) \end{cases}$$

Pour montrer que nous avons effectivement un minimum, il suffit de calculer la sous-hessienne

$$\tilde{H}_F(x, y) = \begin{pmatrix} \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial x \partial y} \\ \frac{\partial^2 F}{\partial x \partial y} & \frac{\partial^2 F}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Elle est définie positive. La valeur de ce minimum est $f(x_*, y_*) = \frac{(\alpha a + \beta b + \gamma)^2}{\alpha^2 + \beta^2}$. Et donc la distance est

$$d(A, (D)) = \frac{|\alpha a + \beta b + \gamma|}{\sqrt{\alpha^2 + \beta^2}}$$

formule bien connue par ailleurs.

On laisse le soin au lecteur de reprendre tous ces calculs dans le cas d'un point $A = (a, b, c)$ de $\mathbb{R}^3$ et d'un plan $\mathcal{P}$ d'équation $\alpha x + \beta y + \gamma z + \delta = 0$. La formule attendue est

$$d(A, \mathcal{P}) = \frac{|\alpha a + \beta b + \gamma c + \delta|}{\sqrt{\alpha^2 + \beta^2 + \gamma^2}}$$

Exemple 2.8.3 On se propose pour terminer ce chapitre, de traiter le problème de l'exemple précédent mais dans $\mathbb{R}^3$. Une droite dans l'espace est l'intersection de deux plans $(D) = \mathcal{P} \cap \mathcal{P}'$. Précisons les équations des deux plans

$$\mathcal{P} : \quad g_1(x, y, z) = \alpha x + \beta y + \gamma z + \delta = 0$$

$$\mathcal{P}' : \quad g_2(x, y, z) = \alpha' x + \beta' y + \gamma' z + \delta' = 0$$

Donnons d'abord la condition sur les paramètres pour que les deux plans ne soient ni confondus ni parallèles et distincts. Posons $\vec{n} = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$ et $\vec{n}' = \begin{pmatrix} \alpha' \\ \beta' \\ \gamma' \end{pmatrix}$ pour désigner

les vecteurs normaux à $\mathcal{P}$ et $\mathcal{P}'$ respectivement. L'intersection suivant une droite de ces deux plans est équivalente à l'indépendance linéaire des vecteurs normaux. Une manière de l'exprimer est d'utiliser le Gramien (déterminant de la matrice de Gram) :

$$G := \det \begin{pmatrix} \langle \vec{n}, \vec{n} \rangle & \langle \vec{n}, \vec{n}' \rangle \\ \langle \vec{n}, \vec{n}' \rangle & \langle \vec{n}', \vec{n}' \rangle \end{pmatrix} = \|\vec{n}\|^2 \|\vec{n}'\|^2 - \langle \vec{n}, \vec{n}' \rangle^2 \neq 0$$

En fait $G > 0$ par l'inégalité de Cauchy-Schwarz. Notons $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ et $A = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ et considérons

$$F(X, \lambda_1, \lambda_2) = \|X - A\|^2 - \lambda_1 g_1(X) - \lambda_2 g_2(X)$$

Remarquons qu'on a $g_1(X) = \langle X, \vec{n} \rangle + \delta$ et $g_2(X) = \langle X, \vec{n}' \rangle + \delta'$, d'où

$$F(X, \lambda_1, \lambda_2) = \|X - A\|^2 - \lambda_1 (\langle X, \vec{n} \rangle + \delta) - \lambda_2 (\langle X, \vec{n}' \rangle + \delta')$$

De là

$$\begin{aligned} (\nabla_X)F &= 2(X - A) - \lambda_1 \vec{n} - \lambda_2 \vec{n}' = 0 \\ \implies X &= X_* = A + \frac{1}{2} (\lambda_1 \vec{n} + \lambda_2 \vec{n}') \end{aligned}$$

Pour déterminer λ_1 et λ_2 il suffit de remplacer la valeur trouvée de X dans les contraintes $g_1(X) = 0$ et $g_2(X) = 0$. On obtient le système

$$\begin{pmatrix} \langle \vec{n}, \vec{n} \rangle & \langle \vec{n}, \vec{n}' \rangle \\ \langle \vec{n}, \vec{n}' \rangle & \langle \vec{n}', \vec{n}' \rangle \end{pmatrix} \begin{pmatrix} \frac{\lambda_1}{2} \\ \frac{\lambda_2}{2} \end{pmatrix} = \begin{pmatrix} -g_1(A) \\ -g_2(A) \end{pmatrix}$$

La solution est donnée par

$$\frac{1}{2} \begin{pmatrix} \lambda_{1,*} \\ \lambda_{2,*} \end{pmatrix} = \frac{1}{G} \begin{pmatrix} \langle \vec{n}', \vec{n}' \rangle & -\langle \vec{n}, \vec{n}' \rangle \\ -\langle \vec{n}, \vec{n}' \rangle & \langle \vec{n}, \vec{n} \rangle \end{pmatrix} \begin{pmatrix} -g_1(A) \\ -g_2(A) \end{pmatrix}$$

Il arrive que $X_* - A = \frac{\lambda_{1,*}}{2} \vec{n} + \frac{\lambda_{2,*}}{2} \vec{n}'$, et donc

$$\|X_* - A\|^2 = \left(\frac{\lambda_{1,*}}{2}\right)^2 \|\vec{n}\|^2 + \left(\frac{\lambda_{2,*}}{2}\right)^2 \|\vec{n}'\|^2 + 2 \left(\frac{\lambda_{1,*}}{2}\right) \left(\frac{\lambda_{2,*}}{2}\right) \langle \vec{n}, \vec{n}' \rangle$$

En remplaçant par les valeurs de $\frac{\lambda_{1,*}}{2}$ et $\frac{\lambda_{2,*}}{2}$, on obtient

$$\|X_* - A\|^2 = \frac{1}{G^2} \left[g_1^2(A) G \|\vec{n}'\|^2 + g_2^2(A) G \|\vec{n}\|^2 - 2g_1(A)g_2(A) \langle \vec{n}, \vec{n}' \rangle \right]$$

On en déduit en définitive

$$d(A, (D)) = \|X - A\| = \frac{1}{\sqrt{G}} \|g_1(A) \vec{n}' - g_2(A) \vec{n}\|$$

Chapitre 3

Calcul Intégral dans $\mathbb{R}^n$

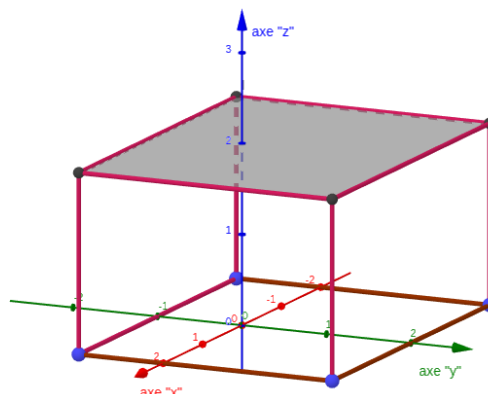
Le présent chapitre est une introduction au calcul intégral dans $\mathbb{R}^n$. Il s'agit d'une extension de la théorie de l'intégrale de Riemann développée en Analyse 1. Le but recherché en première instance étant de donner à l'étudiant le maximum d'outils pour réussir des calculs explicites. C'est pourquoi, les questions "trop théoriques" ne seront abordées que rarement. Nous présenterons les cas $n = 2$ et $n = 3$. Le cas général est similaire bien que hors programme.

3.1 Le cas $n = 2$

Les domaines de base sont les rectangles fermés $[a_1, a_2] \times [b_1, b_2]$, et les fonctions de base sont les fonctions constantes $f(x, y) = c$. L'intégrale est définie alors par

$$\iint_{[a_1, a_2] \times [b_1, b_2]} f(x, y) \, dx dy = c(a_2 - a_1)(b_2 - b_1)$$

Sa valeur absolue représente le *volume* du parallélépipède de base $[a_1, a_2] \times [b_1, b_2]$ et de hauteur $|c|$. Dans la figure suivante, on a représenté la fonction constante $f(x, y) = 2$ dans le rectangle $[-2, 2] \times [-1, 2]$



Pour définir l'intégrale double on considère une fonction $f : [a_1, a_2] \times [b_1, b_2] \rightarrow \mathbb{R}$ bornée i.e, $\exists m, M \in \mathbb{R}$ tels que $m \leq f(x, y) \leq M$, $\forall (x, y) \in Q = [a_1, a_2] \times [b_1, b_2]$. On considère aussi une subdivision pour chacun des deux intervalles :

$$a_1 = x_0 < x_1 < \cdots < x_{n-1} < x_n = a_2 \quad \text{et} \quad b_1 = y_0 < y_1 < \cdots < y_{n-1} < y_n = b_2$$

On aurait pu choisir un nombre différent de points pour chaque subdivision, mais on peut toujours revenir au même nombre en rajoutant des points à celle qui en a moins. Notons

$$\Delta_n = \{(x_i, y_j) \mid i, j = 0, 1, \dots, n\}$$

que nous appellerons *subdivision* du pavé $Q = [a_1, a_2] \times [b_1, b_2]$. On notera aussi $Q_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j]$ avec son aire $\mu(Q_{ij}) = (x_i - x_{i-1})(y_j - y_{j-1})$. Nous appellerons *pas* de la subdivision Δ_n le nombre

$$\rho_n = \max_{1 \leq i, j \leq n} \mu(Q_{ij})$$

On dira qu'une subdivision $\Delta'_{n'}$ est plus fine que Δ_n si $\Delta_n \subset \Delta'_{n'}$. En particulier, elle contient plus de points, dont les coordonnées se placeront forcément entre celles des points de Δ_n . D'où la diminution du pas : $\rho'_{n'} \leq \rho_n$. Posons

$$m_{ij} = \inf_{Q_{ij}} f(x, y) \quad , \quad M_{ij} = \sup_{Q_{ij}} f(x, y)$$

Les sommes de Darboux inférieure et supérieure sont définies par

$$s(f, \Delta_n) := \sum_{i,j=1}^n m_{ij} \mu(Q_{ij}) \quad , \quad S(f, \Delta_n) := \sum_{i,j=1}^n M_{ij} \mu(Q_{ij})$$

Une première série d'inégalités se déduit facilement

$$m\mu(Q) \leq s(f, \Delta_n) \leq S(f, \Delta_n) \leq M\mu(Q)$$

Une deuxième série d'inégalités fait intervenir deux subdivisions dont l'une est plus fine que l'autre, $\Delta_n \subset \Delta'_{n'}$

$$s(f, \Delta_n) \leq s(f, \Delta'_{n'}) \leq S(f, \Delta'_{n'}) \leq S(f, \Delta_n)$$

Ceci s'interprète en disant que, respectivement à l'ordre partiel défini par la finesse, les sommes de Darboux inférieures sont croissantes et les sommes de Darboux supérieures sont décroissantes. La démonstration repose sur les deux inégalités, faciles à établir

$$\text{si } A \subset B \text{ alors } \inf_B f(\xi) \leq \inf_A f(\xi) \quad \text{et} \quad \sup_A f(\xi) \leq \sup_B f(\xi)$$

On les applique à $Q'_{i'j'} \subset Q_{ij}$ pour obtenir les bonnes inégalités.

Proposition 3.1.1 Soient Δ_n^1 et Δ_n^2 deux subdivisions quelconques (non nécessairement comparables). Alors

$$s(f, \Delta_n^1) \leq S(f, \Delta_n^2)$$

Démonstration : Posons $\Delta_n = \Delta_n^1 \cup \Delta_n^2$. Alors cette subdivision est plus fine que Δ_n^1 et Δ_n^2 . D'où par ce qui précède

$$s(f, \Delta_n^1) \leq s(f, \Delta_n) \leq S(f, \Delta_n) \leq S(f, \Delta_n^2)$$

■

Notons $\mathfrak{D}_Q$ l'ensemble des subdivisions du pavé Q . On peut maintenant affirmer que les deux quantités suivantes existent et vérifient

$$\sup_{\Delta_n \in \mathfrak{D}_Q} s(f, \Delta_n) \leq \inf_{\Delta_n \in \mathfrak{D}_Q} S(f, \Delta_n)$$

Définition 3.1.2 On dit que la fonction f est Riemann-intégrable sur le pavé Q si et seulement si

$$\sup_{\Delta_n \in \mathfrak{D}_Q} s(f, \Delta_n) = \inf_{\Delta_n \in \mathfrak{D}_Q} S(f, \Delta_n)$$

Dans ce cas, la valeur commune s'appelle intégrale double de f sur Q et se note

$$\iint_Q f(x, y) \, dx dy$$

L'utilisation des sommes de Darboux est souvent réservée à des questions théoriques comme la non intégrabilité d'une fonction. Pour des problèmes pratiques, on a recours à la notion de somme de Riemann.

Définition 3.1.3 On appelle somme de Riemann associée à une subdivision Δ_n la somme

$$\mathfrak{R}(f, \Delta_n) = \sum_{1 \leq i, j \leq n} f(\xi_{ij}) \mu(Q_{ij}) \quad (3.1)$$

où $\xi_{ij} \in Q_{ij}$ sont des points arbitrairement choisis.

Proposition 3.1.4 Si f est Riemann-intégrable sur le pavé Q , alors on a

$$\lim_{\rho_n \rightarrow 0} \mathfrak{R}(f, \Delta_n) = \iint_Q f(x, y) \, dx dy$$

Démonstration : L'affirmation découle essentiellement du fait que $s(f, \Delta_n) \leq \mathfrak{R}(f, \Delta_n) \leq S(f, \Delta_n)$. ■

Une large classe de fonctions intégrables au sens de Riemann est formée des fonctions continues, et les fonctions continues par morceaux. La preuve est identique au cas d'une variable (revoir Analyse 1). C'est dans ces classes que nous allons travailler désormais.

3.2 Propriétés

Nous allons énumérer les premières propriétés, identiques par ailleurs au cas d'une variable.

Proposition 3.2.1 1. (*Linéarité*) Soient f, g deux fonctions continues sur un pavé Q , alors $\forall \alpha, \beta \in \mathbb{R}$ on a

$$\iint_Q (\alpha f(x, y) + \beta g(x, y)) \, dxdy = \alpha \iint_Q f(x, y) \, dxdy + \beta \iint_Q g(x, y) \, dxdy$$

$$2. \text{ (Positivité) Si } f \geq 0 \text{ alors } \iint_Q f(x, y) \, dxdy \geq 0$$

$$3. \text{ (Croissance) Si } f \leq g \text{ alors } \iint_Q f(x, y) \, dxdy \leq \iint_Q g(x, y) \, dxdy$$

$$4. \text{ On a } \left| \iint_Q f(x, y) \, dxdy \right| \leq \iint_Q |f(x, y)| \, dxdy$$

Démonstration :

1. La linéarité découle (par exemple) du fait que la somme de Riemann possède la même propriété par rapport à f i.e,

$$\mathfrak{R}(\alpha f + \beta g, \Delta_n) = \alpha \mathfrak{R}(f, \Delta_n) + \beta \mathfrak{R}(g, \Delta_n)$$

on passe à limite ensuite.

2. Idem pour la positivité
3. Pour la croissance, il suffit d'appliquer la positivité à $g - f$ et la linéarité ensuite.
4. Pour la dernière, on a bien $f \leq |f|$.

■

Définition 3.2.2 Notons $\mathbb{1}_A$ la fonction caractéristique d'un sous-ensemble A i.e,

$$\mathbb{1}_A(\xi) = \begin{cases} 1 & \text{si } \xi \in A \\ 0 & \text{si } \xi \notin A \end{cases}$$

Un sous-ensemble $A \subset \mathbb{R}^2$ borné est dit mesurable (ou quarrable) si et seulement si sa fonction caractéristique est intégrable. La bornitude implique l'existence d'au moins un pavé $Q \supset A$. La valeur de l'intégrale dans ce cas définit la surface de A i.e,

$$\mu(A) = \iint_Q \mathbb{1}_A(x, y) \, dxdy = \iint_A dxdy$$

Ces sous-ensembles bornés mesurables vont constituer une classe d'ensembles plus large que celle des pavés, et avec lesquels on pourra calculer des intégrales doubles de fonctions continues. En effet

$$\iint_A f(x, y) \, dx dy = \iint_Q f(x, y) \mathbb{1}_A(x, y) \, dx dy$$

Parmi les ensembles mesurables, il y a ceux qu'on appelle simples. Ils sont définis par

$$A = \{(x, y) \in \mathbb{R}^2 \mid a \leq x \leq b \text{ et } \varphi_1(x) \leq y \leq \varphi_2(x)\}$$

ou bien

$$A = \{(x, y) \in \mathbb{R}^2 \mid c \leq y \leq d \text{ et } \psi_1(y) \leq x \leq \psi_2(y)\}$$

où les fonctions $\varphi_1, \varphi_2, \psi_1, \psi_2$ sont au moins continues. Un exemple de tels ensembles le disque euclidien de centre $(0, 0)$ et de rayon r i.e.,

$$D((0, 0), r) = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq r^2\}$$

On peut écrire

$$\begin{aligned} D((0, 0), r) &= \{(x, y) \in \mathbb{R}^2 \mid -r \leq x \leq r \text{ et } -\sqrt{r^2 - x^2} \leq y \leq \sqrt{r^2 - x^2}\} \\ &= \{(x, y) \in \mathbb{R}^2 \mid -r \leq y \leq r \text{ et } -\sqrt{r^2 - y^2} \leq x \leq \sqrt{r^2 - y^2}\} \end{aligned}$$

Un ensemble de mesure (d'aire) nulle est dit négligeable. Par exemple un segment de droite. Il existe des ensembles négligeables plus sophistiqués. Cette notion va nous permettre d'énoncer la formule de Chasles.

Proposition 3.2.3 (*Formule de Chasles*)

Soient A, B deux ensembles bornés mesurables tels que $A \cap B$ soit négligeable. Alors

$$\iint_{A \cup B} f(x, y) \, dx dy = \iint_A f(x, y) \, dx dy + \iint_B f(x, y) \, dx dy$$

Comme conséquence immédiate, le calcul d'une intégrale sur un domaine pourra se faire en décomposant ce domaine en une réunion finie de domaines simples dont les intersections deux à deux sont négligeables.

3.3 Théorème de Fubini

Dans cette section nous allons aborder la première technique majeure pour calculer les intégrales doubles. Il s'agit de ramener ce calcul à celui de deux intégrales simples. L'ordre d'intégration sera sans importance grâce au théorème de Fubini (Guido FUBINI 1879-1943).

Théorème 9 (*de Fubini*)

Soit $f : Q \rightarrow \mathbb{R}$ une fonction continue sur un pavé Q . Alors

$$\iint_Q f(x, y) \, dx dy = \int_{a_1}^{a_2} \left(\int_{b_1}^{b_2} f(x, y) \, dy \right) dx = \int_{b_1}^{b_2} \left(\int_{a_1}^{a_2} f(x, y) \, dx \right) dy$$

Démonstration : Montrons par exemple la première égalité. Remarquons que la continuité (en deux variables) de f implique d'une part son intégrabilité double sur Q . Elle implique d'autre part que chacune des deux fonctions $f(x, \cdot)$ et $f(\cdot, y)$ est continue et donc intégrable sur son intervalle respectif. Notons $h(x) = \int_{b_1}^{b_2} f(x, y) dy$. L'outil essentiel dans cette démonstration est le fait qu'une fonction continue sur un compact (Q en l'occurrence) est uniformément continue i.e,

$$\forall \varepsilon > 0 \exists \delta_\varepsilon > 0 \quad |x - x'| + |y - y'| \leq \delta_\varepsilon \implies |f(x, y) - f(x', y')| \leq \varepsilon$$

(δ_ε ne dépend que de ε et pas des points $(x, y), (x', y')$). L'égalité sera établie si on démontre que l'intégrale double de f sur Q est égale à l'intégrale simple de h sur $[a_1, a_2]$. Nous allons le faire à l'aide des sommes de Riemann. Choisissons une subdivision Δ_n telle que $\forall i, j \quad |x_i - x_{i-1}| + |y_j - y_{j-1}| \leq \delta_\varepsilon$. Rappelons enfin que d'après les résultats (Analyse 3) sur les intégrales dépendants d'un paramètre, la fonction h est continue. Reprenons l'expressions de la somme de Riemann (3.1) où $\xi_{ij} = (u_i, v_j)$

$$\begin{aligned} \mathfrak{R}(f, \Delta_n) &= \sum_{1 \leq i, j \leq n} f(u_i, v_j)(x_i - x_{i-1})(y_j - y_{j-1}) \\ &= \sum_{i=1}^n (x_i - x_{i-1}) \left(\sum_{j=1}^n f(u_i, v_j)(y_j - y_{j-1}) \right) \end{aligned}$$

La dernière somme représente une somme de Riemann pour la fonction $f(u_i, \cdot)$. D'où pour n assez grand

$$\sum_{j=1}^n f(u_i, v_j)(y_j - y_{j-1}) = \int_{b_1}^{b_2} f(u_i, y) dy + \Theta_n(u_i) = h(u_i) + \Theta_n(u_i)$$

avec $\lim_{n \rightarrow +\infty} \Theta_n(u_i) = 0$. En fait cette limite est uniforme par rapport aux u_i . En effet,

$$\Theta_n(u_i) = \sum_{j=1}^n f(u_i, v_j)(y_j - y_{j-1}) - \int_{b_1}^{b_2} f(u_i, y) dy = \sum_{j=1}^n \int_{y_{j-1}}^{y_j} (f(u_i, v_j) - f(u_i, y)) dy$$

et donc

$$\max_{u_i} |\Theta_n(u_i)| \leq \varepsilon(b_2 - b_1)$$

Maintenant

$$\mathfrak{R}(f, \Delta_n) = \sum_{i=1}^n (x_i - x_{i-1})h(u_i) + \sum_{i=1}^n (x_i - x_{i-1})\Theta_n(u_i)$$

Remarquer que la première somme à droite est une somme de Riemann pour la fonction h , et donc converge vers l'intégrale simple de h . Enfin

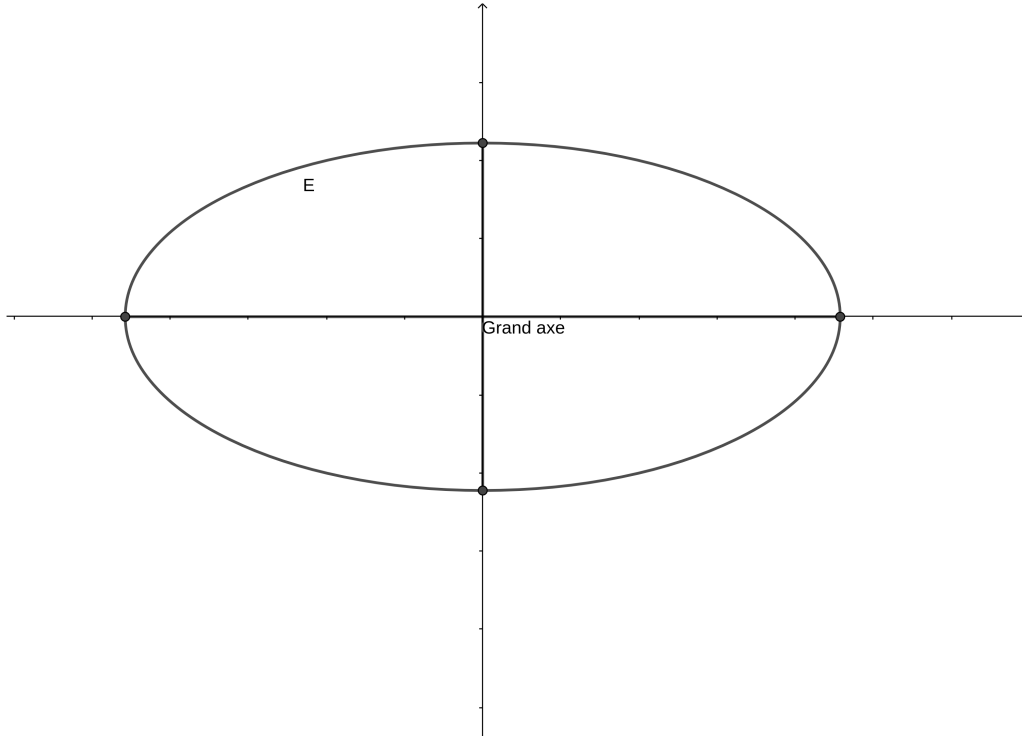
$$|\mathfrak{R}(f, \Delta_n) - \sum_{i=1}^n (x_i - x_{i-1})h(u_i)| \leq \varepsilon(b_2 - b_1)(a_2 - a_1)$$

Ceci montre que les deux sommes de Riemann tendent vers la même limite. ■

Le théorème de Fubini s'étend naturellement aux domaines simples. Nous omettrons la démonstration un peu technique. Donnons quelques exemples.

Exemple 3.3.1 Calculons l'aire d'une ellipse de grand axe $2a$ et de petit axe $2b$. Soit

$$\mathcal{E} = \left\{ (x, y) \in \mathbb{R}^2 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1 \right\}$$



En résolvant l'inégalité par rapport à y (par exemple) on obtient

$$\mathcal{E} = \left\{ (x, y) \in \mathbb{R}^2 \mid -a \leq x \leq a \quad \text{et} \quad -b\sqrt{1 - \frac{x^2}{a^2}} \leq y \leq \underbrace{b\sqrt{1 - \frac{x^2}{a^2}}}_{\varphi(x)} \right\}$$

D'où

$$\text{Aire}(\mathcal{E}) = \iint_{\mathcal{E}} dx dy = \int_{-a}^a \left(\int_{-\varphi(x)}^{\varphi(x)} dy \right) dx = 4b \int_0^a \sqrt{1 - \frac{x^2}{a^2}} dx$$

Il suffit alors de faire le changement $x = a \cos t$ avec $0 \leq t \leq \frac{\pi}{2}$ pour avoir

$$\text{Aire}(\mathcal{E}) = 4ab \int_0^{\frac{\pi}{2}} \sin^2 t dt = 2ab \int_0^{\frac{\pi}{2}} (1 - \cos(2t)) dt = \pi ab.$$

Pour le cas particulier $a = b = r$ on retrouve l'aire du disque.

Exemple 3.3.2 Soit à calculer l'intégrale double $I = \iint_D \cos(xy) dx dy$ où D est le domaine défini par

$$D = \left\{ (x, y) \in \mathbb{R}^2 \mid 1 \leq x \leq 2 \quad 0 \leq xy \leq \frac{\pi}{2} \right\}$$

La résolution par rapport à y donne $0 \leq y \leq \frac{\pi}{2x}$ et par le théorème de Fubini on aura

$$\begin{aligned} I &= \int_1^2 \left(\int_0^{\pi/2x} \cos(xy) \, dy \right) dx = \int_1^2 \left[\frac{\sin(xy)}{x} \right]_{y=0}^{y=\pi/2x} dx \\ &= \int_1^2 \frac{dx}{x} = [\ln(x)]_1^2 = \ln 2. \end{aligned}$$

3.4 Changement de variables

La deuxième technique importante pour calculer une intégrale double est celle du changement de variables.

Définition 3.4.1 Soient U et V deux ouverts de $\mathbb{R}^2$. On appelle *difféomorphisme* (ou *changement de variables*) entre U et V une application

$$\begin{aligned} \Phi : U &\longrightarrow V \\ (u, v) &\longrightarrow \Phi(u, v) = (x, y) \end{aligned}$$

de classe C^1 inversible, c'est-à-dire de jacobien non nul partout ($\det(J_\Phi) \neq 0$)

Rappelons que si $x = \varphi_1(u, v)$ et $y = \varphi_2(u, v)$ alors

$$J_\Phi = \begin{pmatrix} \frac{\partial \varphi_1}{\partial u} & \frac{\partial \varphi_1}{\partial v} \\ \frac{\partial \varphi_2}{\partial u} & \frac{\partial \varphi_2}{\partial v} \end{pmatrix}$$

Le théorème suivant sera donné sans démonstration. Le lecteur intéressé pourra la trouver par exemple dans "Mathématiques, tout-en-un, niveau L2" sous la direction de J.P.Ramis et A.Warusfel, Édition DUNOD 2007.

Théorème 10 Soit $\Phi : U \longrightarrow V$ un changement de variables entre deux ouverts bornés et soit $f : V \longrightarrow \mathbb{R}$ une application continue bornée. Alors

$$\iint_V f(x, y) \, dx dy = \iint_U (f \circ \Phi)(u, v) |\det J_\Phi(u, v)| \, du dv$$

Le premier exemple que nous allons présenter montre la pertinence de procéder par changement de variables.

Exemple 3.4.2 Soit à calculer

$$I = \iint_D \frac{1}{1 + x^2 + y^2} \, dx dy$$

où $D = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ est le disque unité de $\mathbb{R}^2$. Nous allons d'abord essayer d'utiliser le théorème de Fubini. Il est clair que $-1 \leq x \leq 1$ et $-\sqrt{1-x^2} \leq y \leq \sqrt{1-x^2}$, d'où (remarquer la parité en x et y)

$$I = \int_{-1}^1 \left[\frac{\arctan\left(\frac{y}{\sqrt{1+x^2}}\right)}{\sqrt{1+x^2}} \right]_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} dx = 4 \int_0^1 \frac{\arctan \sqrt{\frac{1-x^2}{1+x^2}}}{\sqrt{1+x^2}} dx$$

Il n'est pas évident d'imaginer un changement simple (en une variable) pour calculer cette dernière intégrale. Procédons à présent par le changement en coordonnées polaires :

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta \end{cases}$$

On a $J = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}$ et donc $\det J = r$. C'est donc un changement entre les ouverts suivants

$$U = \{(r, \theta) \in \mathbb{R}^2 \mid 0 < r < 1 \text{ et } 0 < \theta < 2\pi\}$$

et

$$V = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\} \setminus \{(x, 0) \mid 0 \leq x < 1\}$$

Remarquons que $\bar{V} = D$. De là on aura

$$I = \int_0^1 \int_0^{2\pi} \frac{r}{1+r^2} dr d\theta = 2\pi \left[\frac{1}{2} \ln(1+r^2) \right]_0^1 = \pi \ln 2$$

Exemple 3.4.3 On se propose de calculer l'intégrale de Gauss $\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$.

Considérons $I_R = \int_{-R}^R e^{-x^2} dx$. A l'aide du théorème de Fubini on peut écrire

$$I_R^2 = \left(\int_{-R}^R e^{-x^2} dx \right) \left(\int_{-R}^R e^{-y^2} dy \right) = \iint_{Q_R} e^{-(x^2+y^2)} dx dy$$

où $Q_R = [-R, R] \times [-R, R]$. Posons $J_R = \iint_{D_R} e^{-(x^2+y^2)} dx dy$ où D_R est le disque de centre $(0, 0)$ et de rayon R . A l'aide du changement en coordonnées polaires on a

$$J_R = \int_0^R \int_0^{2\pi} e^{-r^2} r dr d\theta = 2\pi \left[\frac{-1}{2} e^{-r^2} \right]_0^R = \pi (1 - e^{-R^2})$$

Maintenant il suffit de voir qu'on peut "coincer" le carré Q_R entre deux disques i.e., $D_R \subset Q_R \subset D_{R\sqrt{2}}$. La positivité de la fonction à intégrer implique que

$$J_R \leq I_R^2 \leq J_{R\sqrt{2}} \iff \pi (1 - e^{-R^2}) \leq I_R^2 \leq \pi (1 - e^{-2R^2})$$

Le résultat s'en déduit par passage à la limite quand $R \rightarrow +\infty$.

3.5 Le cas $n = 3$

L'intégrale triple (dans $\mathbb{R}^3$) se construit de la même façon que l'intégrale double. D'ailleurs on peut étendre la construction à $\mathbb{R}^n$. Nous omettrons donc la répétition. Par contre on donnera (sans démonstration) le théorème de Fubini et la formule de changement de variables. On finira par quelques exemples.

3.5.1 Théorème de Fubini

Les ensembles mesurables bornés simples se définissent de la même façon, par exemple

$$A = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in D \subset \mathbb{R}^2 \text{ et } \varphi_1(x, y) \leq z \leq \varphi_2(x, y)\}$$

(si on résout par rapport à z) où D est la projection du domaine A dans le plan $(\vec{ox}, \vec{oy})$. Cela permettra de réduire l'intégrale triple à d'abord une intégrale simple (en z), puis une intégrale double (en (x, y)). Si on résout d'abord en y (par exemple), on opère *mutatis-mutandis*.

Théorème 11 (de Fubini)

Soit A un domaine borné mesurable simple et $f : A \longrightarrow \mathbb{R}$ une application continue bornée. Alors

$$\iiint_A f(x, y, z) \, dx dy dz = \iint_D \left(\int_{\varphi_1(x, y)}^{\varphi_2(x, y)} f(x, y, z) \, dz \right) dx dy$$

Exemple 3.5.2 Calculons le volume d'une boule sphérique de rayon R . Si on la centre en $(0, 0, 0)$ alors

$$\begin{aligned} B_R &= \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \leq R^2\} = \\ &= \left\{ (x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 \leq R^2 \text{ et } -\sqrt{R^2 - x^2 - y^2} \leq z \leq \underbrace{\sqrt{R^2 - x^2 - y^2}}_{\varphi(x, y)} \right\} \end{aligned}$$

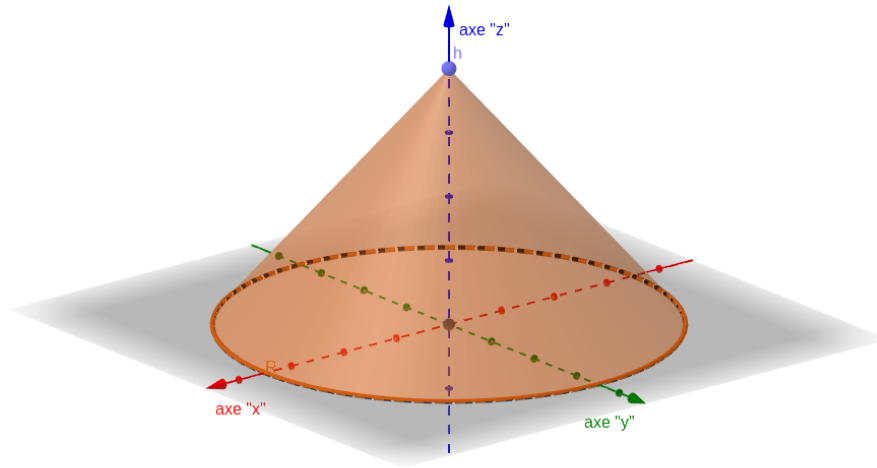
D'où

$$\text{Vol}(B_R) = \iiint_{B_R} dx dy dz = \iint_{x^2 + y^2 \leq R^2} \left(\int_{-\varphi(x, y)}^{\varphi(x, y)} dz \right) dx dy = 2 \iint_{x^2 + y^2 \leq R^2} \varphi(x, y) \, dx dy$$

En passant en coordonnées polaires on obtient

$$\text{Vol}(B_R) = 2 \int_0^R \int_0^{2\pi} \sqrt{R^2 - r^2} \, r dr d\theta = 2\pi \left[\frac{-2}{3} (R^2 - r^2)^{3/2} \right]_0^R = \frac{4\pi}{3} R^3$$

Exemple 3.5.3 On se propose de calculer le volume d'un cône droit de base circulaire de rayon R et de hauteur h .



Pour trouver l'équation de la surface latérale du cône, il suffit de trouver l'équation d'une génératrice. C'est une droite passant par le point $(0, 0, h)$ et un point du cercle (dans le plan XY) de centre $(0, 0)$ et de rayon R . La droite est dans le plan passant par ces deux points et l'origine. La projection sur le plan XY d'un point de cette génératrice a pour coordonnées $(x, y, 0)$ avec $x^2 + y^2 = r^2 \leq R^2$. L'équation de la génératrice (dans le plan indiqué) est donc

$$z = h \left(1 - \frac{r}{R} \right) = h \left(1 - \frac{\sqrt{x^2 + y^2}}{R} \right)$$

Le volume cherché est donc celui du domaine donné par

$$\mathcal{C} = \left\{ (x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 \leq R^2 \quad \text{et} \quad 0 \leq z \leq h \left(1 - \frac{\sqrt{x^2 + y^2}}{R} \right) \right\}$$

Donc, par le théorème de Fubini

$$\text{Vol}(\mathcal{C}) = \iiint_{\mathcal{C}} dx dy dz = \iint_{x^2 + y^2 \leq R^2} h \left(1 - \frac{\sqrt{x^2 + y^2}}{R} \right) dx dy$$

Et en passant en coordonnées polaires

$$\text{Vol}(\mathcal{C}) = h \int_0^R \int_0^{2\pi} \left(1 - \frac{r}{R} \right) r dr d\theta = 2\pi h \left[\frac{r^2}{2} - \frac{r^3}{3R} \right]_0^R = \frac{1}{3} \pi R^2 h.$$

3.5.4 Changement de variables

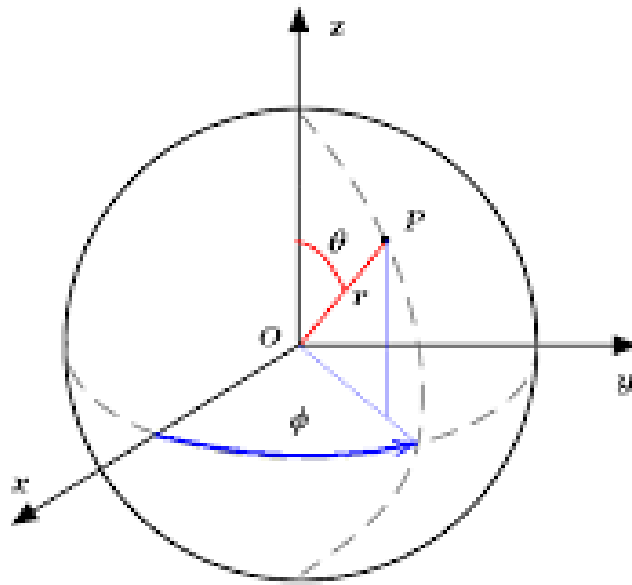
La formule de changement de variables est identique à celle du cas $n = 2$

$$\iiint_V f(x, y, z) \, dx dy dz = \iiint_U (f \circ \Phi)(u, v, w) |\det J_\Phi(u, v, w)| \, du dv dw$$

On donnera deux changements remarquables.

Coordonnées sphériques : ce sont les analogues en dimension 3 des coordonnées polaires.

$$\begin{cases} x = r \sin \theta \cos \varphi \\ y = r \sin \theta \sin \varphi \\ z = r \cos \theta \end{cases}$$



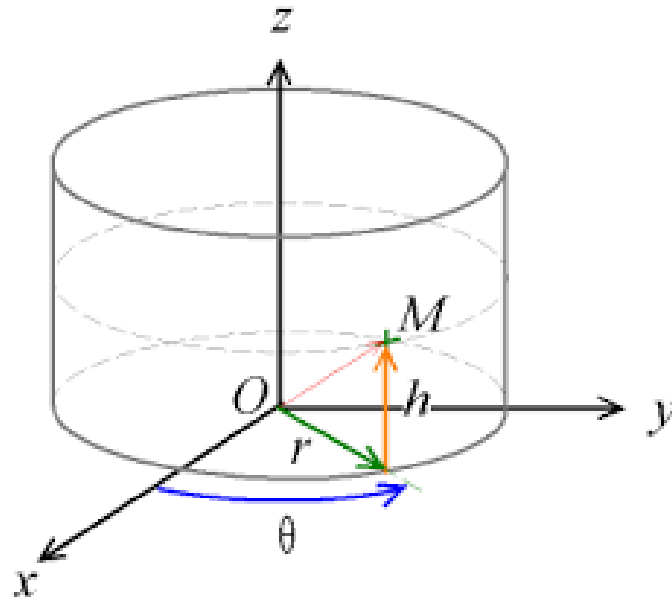
Pour décrire tout l'espace on doit prendre $r > 0$, $\theta \in [0, \pi]$, $\varphi \in [0, 2\pi[$. La matrice jacobienne se calcule aisément

$$J_\Phi = \begin{pmatrix} \sin \theta \cos \varphi & r \cos \theta \cos \varphi & -r \sin \theta \sin \varphi \\ \sin \theta \sin \varphi & r \cos \theta \sin \varphi & r \sin \theta \cos \varphi \\ \cos \theta & -r \sin \theta & 0 \end{pmatrix}$$

et $\det(J_\Phi) = r^2 \sin \theta$.

Coordonnées cylindriques : on fait opérer les coordonnées polaires sur XY et laisser z inchangé.

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta \\ z = h \end{cases}$$



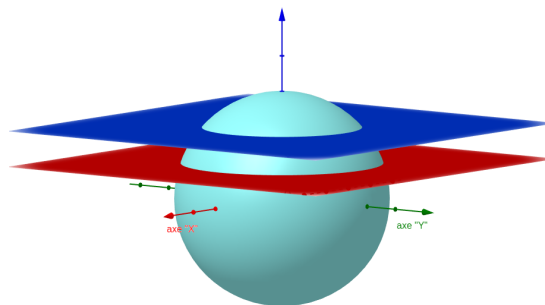
Pour décrire tout l'espace on doit prendre $r > 0$, $\theta \in [0, 2\pi[$, $h \in \mathbb{R}$. La matrice jacobienne se calcule facilement

$$J_{\Phi} = \begin{pmatrix} \cos \theta & -r \sin \theta & 0 \\ \sin \theta & r \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

et $\det(J_{\Phi}) = r$.

La plupart du temps les changements de variables sont dictés par la géométrie du domaine d'intégration. Les invariances par certaines transformations (rotation, homothétie,...) peuvent aider à trouver le bon changement. On peut aussi combiner changement de variables et théorème de Fubini.

Exemple 3.5.5 *Calculons le volume d'une portion de boule sphérique de rayon R et délimitée par deux plans parallèles au plan XOY .*



Appelons K cette portion. Alors

$$K = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \leq R^2 \text{ et } a \leq z \leq b\}$$

où $-R \leq a < b \leq R$. Contrairement à ce qu'on pourrait penser, les coordonnées cylindriques sont les plus adaptées. En effet, la paramétrisation de K devient

$$K = \{(r, \theta, h) \in \mathbb{R}^3 \mid r^2 + h^2 \leq R^2, \quad 0 \leq \theta < 2\pi, \quad a \leq h \leq b\}$$

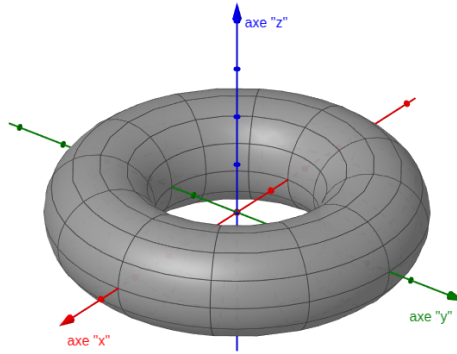
D'où

$$Vol(K) = \iint_K dx dy dz = \iint_K r dr d\theta dh = \int_0^{2\pi} \int_a^b \left(\int_0^{\sqrt{R^2-h^2}} r dr \right) d\theta dh$$

ou encore

$$Vol(K) = 2\pi \int_a^b \left[\frac{r^2}{2} \right]_0^{\sqrt{R^2-h^2}} dh = \pi \left[R^2 h - \frac{h^3}{3} \right]_a^b = \pi \left[R^2(b-a) - \frac{(b^3-a^3)}{3} \right]$$

Nous terminons par le calcul du volume d'un *tore* (une "chambre à air"). Géométriquement on fait tourner un cercle de centre $(R, 0, 0)$ et de rayon $a < R$ situé dans le plan XOZ autour de l'axe OZ pour engendrer un tore $\mathcal{T}$.



Une paramétrisation est donnée par

$$\begin{cases} x = (R + r \cos u) \cos v \\ y = (R + r \cos u) \sin v \\ z = r \sin u \end{cases} \quad \begin{matrix} 0 \leq r \leq a \\ 0 \leq u, v < 2\pi \end{matrix}$$

Le déterminant de la matrice jacobienne est $\det(J) = r(R + r \cos u)$. Ainsi

$$Vol(\mathcal{T}) = \int_0^a \int_0^{2\pi} \int_0^{2\pi} r(R + r \cos u) dr du dv = 2\pi^2 R a^2.$$