

INSA Toulouse, cycle préparatoire (pour profils autres que mathématiques).

Mathématiques - Analyse 2

Fonctions de plusieurs variables

Calcul différentiel, extrema, changements de variables, intégrales multiples

*Ressources issues de nombreux enseignants-chercheurs ou enseignants agrégés de mathématiques, en particulier de ceux
de
l'INSA Toulouse, département de mathématiques appliquées.*

*Version éditée par J. Monnier, enseignant-chercheur (professeur) INSA Toulouse.
jerome.monnier@insa-toulouse.fr*

Table des matières

Chapitre 1. Fonctions de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$: limite, continuité. Applications linéaires, quadratiques.	5
1. Rappels sur l'espace $\mathbb{R}^m$	5
2. Fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^p$	6
3. Applications linéaires et affines ; applications quadratiques	6
4. Concepts de limite et continuité dans $\mathbb{R}^n$	7
Chapitre 2. Concepts de dérivées dans $\mathbb{R}^n$: dérivée directionnelle, dérivées partielles, fonction différentielle.	9
1. Dérivées directionnelle et partielles	9
2. L'application différentielle	10
3. Opérateurs différentiels usuels du premier ordre	13
4. Dérivées partielles d'ordres supérieurs	16
5. Formule de Taylor à n variables	17
Chapitre 3. Extrema d'une fonction de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$ (premiers pas vers l'optimisation).	19
1. Condition nécessaire d'extremum	19
2. Condition suffisante d'extremum - Matrice Hessienne	19
Chapitre 4. Changements de variables	23
1. Définition	23
2. Principes d'un changement de variables	23
3. Exemple d'application : résolution d'une équation différentielle par changement de variables	25
4. Coordonnées cylindriques et sphériques	26
Chapitre 5. Intégrales multiples	27
1. Propriétés fondamentales	27
2. Intégrale multiple sur un compact élémentaire	28
3. Changement de variables dans une intégrale multiple (double, triple)	29

Fonctions de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$: limite, continuité. Applications linéaires, quadratiques.

On introduit les notions de limite et continuité de fonctions de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$ i.e. des fonctions à n variables et à p composantes.

Exemple d'une fonction à 3 variables et à 2 composantes. $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, $f(x, y, z) = (e^{2x+y} + z, x^2y/z)$.

Le cas particulier des applications *linéaires* est traité en fin de chapitre.

1. Rappels sur l'espace $\mathbb{R}^m$

Soit x un point de $\mathbb{R}^m$, on écrira : $x = (x_1, \dots, x_j, \dots, x_m)$. On a également :

$$x = \sum_{j=1}^m x_j e_j^m$$

où e_j^m est le j ième vecteur de la base canonique.

DÉFINITION 1.1. Une norme sur $\mathbb{R}^n$ est une application de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^+$ (valeurs positives) et satisfaisant les trois axiomes :

- i) $\forall x \in \mathbb{R}^n$, $\|x\| = 0$ implique $x = 0$.
- ii) $\forall x \in \mathbb{R}^n, \forall \lambda \in \mathbb{R}$, $\|\lambda x\| = |\lambda| \|x\|$.
- iii) $\forall (x, y) \in (\mathbb{R}^n)^2$, $\|x + y\| \leq \|x\| + \|y\|$

L'espace $\mathbb{R}^m$ est un espace vectoriel normé. Il existe de nombreuses normes sur $\mathbb{R}^m$.

Les normes usuelles sont : $\|\cdot\|_{1, \mathbb{R}^m}$, $\|\cdot\|_{2, \mathbb{R}^m}$ et $\|\cdot\|_{\infty, \mathbb{R}^m}$ qui sont définies comme suit. Pour $x \in \mathbb{R}^m$,

$$\|x\|_1 = \sum_{j=1}^m |x_j|, \quad \|x\|_2 = \sqrt{\sum_{j=1}^m x_j^2}, \quad \|x\|_{\infty} = \max_{j=1, \dots, m} |x_j|.$$

Dans $\mathbb{R}^n$, toutes les normes sont équivalentes.

On notera $\langle x, y \rangle$ le produit scalaire usuel de $\mathbb{R}^m$:

$$\langle x, y \rangle = \sum_{j=1}^m x_j y_j$$

La norme 2 (norme euclidienne) lui est associée : $\|x\|_2^2 = \langle x, x \rangle$.

On rappelle que la base canonique $\mathcal{B}_m$ est orthonormée pour ce produit scalaire, i.e. $\langle e_i^m, e_j^m \rangle = 1$ si $i = j$; $= 0$ sinon.

DÉFINITION 1.2 (Boules). On notera $B(x, r) = \{y \in \mathbb{R}^m / \|x - y\| < r\}$ la boule ouverte de centre $x \in \mathbb{R}^m$ de rayon $r > 0$ et $\bar{B}(x, r) = \{y \in \mathbb{R}^m / \|x - y\| \leq r\}$ la boule fermée de centre $x \in \mathbb{R}^m$ de rayon $r > 0$.

Par la suite, lorsque l'on parlera de voisinage d'un point a de $\mathbb{R}^m$, on pensera à une boule ouverte centrée en a .

EXERCICE 1.1. Dessiner la boule unité $B(0, 1)$ de $\mathbb{R}^2$ pour les trois normes définies ci-dessus.

Un sous-ensemble non vide Ω de $\mathbb{R}^m$ est un ouvert de $\mathbb{R}^m$ si : $\forall a \in \Omega, \exists r > 0, B(a, r) \subset \Omega$.

**

2. Fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^p$

Soient n et p deux entiers ≥ 1 . Soit f une fonction définie sur un sous-ensemble D_f de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$. f est une fonction à n variables, à p composantes.

Rappel : une fonction est une application numérique i.e. à valeurs dans $\mathbb{R}^p$ ou $\mathbb{C}^p$.

EXERCICE 1.2. Déterminer l'ensemble de définition des fonctions suivantes :

a) $f : \mathbb{R}^3 \rightarrow \mathbb{R}, f(x, y, z) = \frac{z}{x+y} + e^{z(x+y)}.$

b) $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, f(x, y) = (e^{2x+y}, \ln(x)).$

Notation

Pour $f(x)$, on utilisera souvent la notation $f(x_1, \dots, x_j, \dots, x_n)$.

Si $n = 2$, on écrira souvent $f(x, y)$ au lieu de $f(x_1, x_2)$, et si $n = 3$, $f(x, y, z)$ au lieu de $f(x_1, x_2, x_3)$.

Composantes

On note également :

$$f(x) = f(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_i(x_1, \dots, x_n), \dots, f_p(x_1, \dots, x_n)).$$

Les fonctions f_i sont appelées les *fonctions composantes* de f . Ce sont des fonctions à valeurs réelles.

Si $p = 1$, on dira que f est une fonction *scalaire*. Au contraire, si $p > 1$ on parlera de fonction *vectorielle*.

3. Applications linéaires et affines ; applications quadratiques

Applications linéaires. Les applications linéaires constituent un cas particulier important à distinguer.

DÉFINITION 1.3. Une application f de $\mathbb{R}^n$ dans $\mathbb{R}^p$ est dite *linéaire* si les deux axiomes suivants sont vérifiés :

i) $\forall (x, y) \in (\mathbb{R}^n)^2, f(x + y) = f(x) + f(y).$

ii) $\forall x \in \mathbb{R}^n, \forall \lambda \in \mathbb{R}, f(\lambda x) = \lambda f(x).$

Une conséquence de la linéarité est que l'image d'une combinaison linéaire est la combinaison linéaire des images :

$$f\left(\sum_i \lambda_i x_i\right) = \sum_i \lambda_i f(x_i)$$

Rappelons que les matrices de $M_{p,n}$ définissent des applications linéaires de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$.

On note $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$ l'ensemble des applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^p$. C'est un espace vectoriel sur $\mathbb{R}$.

EXERCICE 1.3. a) Soit $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3, f(x, y, z) = (2x + 3y + 5z, x - y + z, y - z).$

Montrer que f est linéaire de $\mathbb{R}^3$ dans $\mathbb{R}^3$.

b) Soit $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3, f(x, y) = (x^2, x - y).$

Montrer que f n'est pas linéaire.

On dit que f est une application affine de $\mathbb{R}^n$ dans $\mathbb{R}^p$ si :

$$\exists a \in \mathbb{R}^p \text{ et } \varphi \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p) \text{ tels que } \forall x \in \mathbb{R}^n, f(x) = a + \varphi(x).$$

Exemple : $f(x, y, z) = (2 - x + 3y + z, 1 + x - z) = (2, 1) + (-x + 3y + z, x - z).$

Applications bilinéaires. On dit que $\varphi : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \rightarrow \mathbb{R}^p$ est une application bilinéaire si pour tout $x \in \mathbb{R}^{n_1}$, $y \mapsto \varphi(x, y)$ est linéaire et pour tout $y \in \mathbb{R}^{n_2}$, $x \mapsto \varphi(x, y)$ est linéaire i.e. si φ est linéaire par rapport à chacune de ses variables. Soit : $\forall (\lambda_1, \lambda_2) \in \mathbb{R}^2$,

$$\forall (x, y_1, y_2), \varphi(x, \lambda_1 y_1 + \lambda_2 y_2) = \lambda_1 \varphi(x, y_1) + \lambda_2 \varphi(x, y_2)$$

$$\forall (x_1, x_2, y), \varphi(\lambda_1 x_1 + \lambda_2 x_2, y) = \lambda_1 \varphi(x_1, y) + \lambda_2 \varphi(x_2, y)$$

On note $Bil(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}, \mathbb{R}^p)$ l'ensemble de ces applications.

On dit que φ est une *forme bilinéaire* si $p = 1$.

Exemple : $\varphi(x, y) = (xy, xy)$ est bilinéaire de $\mathbb{R} \times \mathbb{R}$ dans $\mathbb{R}^2$ ($\varphi \in Bil(\mathbb{R} \times \mathbb{R}, \mathbb{R}^2)$), mais $\varphi(x, y) = (xy, y)$ ne l'est pas. (Notons d'ailleurs que cette fonction n'est pas linéaire par rapport à $\mathbf{x} = (x, y) \dots$)

Formes quadratiques.

Une forme quadratique q sur $\mathbb{R}^n$ est une application de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$ tel qu'il existe φ bilinéaire ($\varphi \in Bil(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R})$) et symétrique, telle que :

$$q(x) = \varphi(x, x) \text{ soit } q(x) = q(x_1, \dots, x_j, \dots, x_n) = \sum_{j=1}^n a_{jj} x_j^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j$$

Exemple : $q : \mathbb{R}^3 \rightarrow \mathbb{R}$, $q(x, y, z) = 2x^2 - 3y^2 + 2xy - 3yz$.

**** POUR ALLER PLUS LOIN.**

$\varphi : \mathbb{R}^{n_1} \times \dots \times \mathbb{R}^{n_k} \rightarrow \mathbb{R}^p$ est une *application multilinéaire* si elle est linéaire par rapport à chacune de ses k variables.

Exemple : Le déterminant d'une matrice est une application multi-linéaire.

4. Concepts de limite et continuité dans $\mathbb{R}^n$

Dans la suite, Ω désigne un ouvert non vide de $\mathbb{R}^n$ et f désigne une fonction définie sur Ω , à valeurs dans $\mathbb{R}^p$, où n et p sont deux entiers ≥ 1 .

Définitions. Du point de vue conceptuel, la notion de limite dans $\mathbb{R}^n$ est semblable à celle dans $\mathbb{R}$ (cf la définition ci-dessous), même si certaines subtilités peuvent exister en terme de chemins d'approche au point a .

DÉFINITION 1.4. Soit $a \in \Omega$, $b \in \mathbb{R}^p$. On dit que f tend vers b quand x tend vers a si

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x \in \Omega, \quad 0 < \|x - a\| < \eta \implies \|f(x) - b\| < \varepsilon$$

ou encore,

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x \in \Omega \setminus \{a\}, \quad x \in B_{\mathbb{R}^n}(a, \eta) \implies f(x) \in B_{\mathbb{R}^p}(b, \varepsilon).$$

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

**** POUR ALLER PLUS LOIN.**

DÉFINITION 1.5. On dit que f est (uniformément) continue au point a si

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x \in \Omega, \quad \|x - a\| < \eta \implies \|f(x) - f(a)\| < \varepsilon,$$

ou encore,

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x \in \Omega, \quad x \in B_{\mathbb{R}^n}(a, \eta) \implies f(x) \in B_{\mathbb{R}^p}(f(a), \varepsilon).$$

**

Autrement dit, f est continue au point $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ si

$$\lim_{\|h\| \rightarrow 0} f(a_1 + h_1, \dots, a_n + h_n) = f(a_1, \dots, a_n)$$

Noter que $\|h\| \rightarrow 0$ peut être vu pour n'importe quelle des 3 normes classiques de $\mathbb{R}^n$ vu précédemment.

PROPOSITION 1.1. Soient f et g deux fonctions continues en un point a de $\mathbb{R}^n$. Alors,

- $(f + g)$ est continue au point a ,
- pour tout $\lambda \in \mathbb{R}$, λf est continue au point a .

De plus si f et g deux fonctions scalaires (i.e. $p = 1$), alors $f \cdot g$ et $\frac{f}{g}$ sont deux fonctions scalaires qui vérifient :

- $(f \cdot g)$ est continue au point a ,
- $\frac{f}{g}$ est continue au point a si $g(a) \neq 0$.

Continuité des applications linéaires.

THÉORÈME 1.1. Toute application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$ est continue.

De plus, il existe M tel que pour tout x , $\|f(x)\| \leq M\|x\|$.

Concepts de dérivées dans $\mathbb{R}^n$: dérivée directionnelle, dérivées partielles, fonction différentielle.

On étend la notion de dérivée ' au cas à n variables. Pour cela, on introduit les notions de : dérivées directionnelles, dérivées partielles, et application-fonction différentielle. Ces différentes notions ont des liens entre elles qui sont présentés ci-dessous.

Soient n et p deux entiers ≥ 1 . Soit f une fonction définie de $\mathbb{R}^n$ (ou sur un sous-ensemble D_f) à valeurs dans $\mathbb{R}^p$. f est une fonction à n variables et p composantes.

1. Dérivées directionnelle et partielles

1.1. Dérivée directionnelle.

DÉFINITION 2.1. Soit $v \in \mathbb{R}^n / \{0\}$ fixé.
 Si $\frac{f(a+tv) - f(a)}{t}$ a une limite quand $t \rightarrow 0$, cette limite est appelée la dérivée de f en a dans la direction v .
 Cette dérivée est notée $D_v f(a)$,

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

1.2. Dérivées partielles.

DÉFINITION 2.2. Soit $a \in \Omega$ et $j \in \{1, \dots, n\}$. Si f est dérivable en a selon la direction e_j^n ($j^{\text{ème}}$ vecteur de la base canonique de $\mathbb{R}^n$), on dit que f admet une dérivée partielle par rapport à x_j en a .

Autrement dit, la $j^{\text{ème}}$ dérivée partielle de f est sa dérivée (usuelle) appliquée uniquement à sa $j^{\text{ème}}$ variable, les autres variables restant fixes.

On note $\frac{\partial f}{\partial x_j}(a)$ ($\partial_j f(a)$ en forme contractée) sa $j^{\text{ème}}$ dérivée partielle au point a .

La $j^{\text{ème}}$ dérivée partielle $\frac{\partial f}{\partial x_j}$ est une fonction définie de Ω dans $\mathbb{R}^p$.

EXERCICE 2.1. Calculer les dérivées partielles de la fonction suivante :
 $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $f(x, y) = (e^{2x+y}, x^2 y)$.

On a :

DÉFINITION 2.3. Soit $f : \Omega \rightarrow \mathbb{R}^p$, Ω ouvert de $\mathbb{R}^n$. Si f admet des dérivées partielles sur Ω et ses dérivées partielles sont continues sur Ω alors f est de classe $\mathcal{C}^1$ sur Ω .

EXERCICE 2.2. Soit $\alpha \in \mathbb{R}$. La fonction f_α suivante est-elle de classe $\mathcal{C}^1$? sur quel ensemble ? (Discuter selon les valeurs de α).

$$f_\alpha : (x, y) \mapsto (e^{2x+y}, x^2 y^\alpha)$$

2. L'application différentielle

2.1. Application différentielle et dérivées partielles.

THÉORÈME 2.1. Soit Ω un ouvert non vide de $\mathbb{R}^n$, $a \in \Omega$ et $f : \Omega \rightarrow \mathbb{R}^p$. Si f admet des dérivées partielles sur Ω et si celles-ci sont continues en a , alors f est différentiable en a . Sa différentielle au point a notée Df_a est l'application de $\mathbb{R}^n$ dans $\mathbb{R}^p$ définie ainsi :

$$(1) \quad h \mapsto Df_a(h) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(a) h_j$$

** POUR ALLER PLUS LOIN

2.2. Retour au cas une seule variable. Lorsque la variable x est scalaire, i.e. $f : \mathbb{R} \rightarrow \mathbb{R}^p$, la définition de la dérivée $f'(x) = l$ est donnée par $l = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$.

Cette définition peut aussi s'écrire

$$f(x+h) = f(x) + lh + h\varepsilon(h)$$

où $\lim_{h \rightarrow 0} \varepsilon(h) = 0$.

On remarque que l'application $\left| \begin{array}{l} \mathbb{R} \rightarrow \mathbb{R}^p \\ h \mapsto lh \end{array} \right.$ est linéaire. On pourrait donc donner la définition suivante.

DÉFINITION 2.4. $f : \Omega \subset \mathbb{R} \rightarrow \mathbb{R}^p$ est dérivable en x s'il existe $\lambda : \mathbb{R} \rightarrow \mathbb{R}^p$ linéaire, $\eta > 0$ et $\varepsilon :]-\eta, \eta[\rightarrow \mathbb{R}^p$ tels que pour tout $h \in]-\eta, \eta[$, $f(x+h) = f(x) + \lambda(h) + h\varepsilon(h)$, avec $\varepsilon(h) \xrightarrow{h \rightarrow 0} 0$.

2.3. Cas de plusieurs variables. Cette dernière définition peut se généraliser au cas d'une fonction à n variables, soit f de $\mathbb{R}^n$ dans $\mathbb{R}^p$.

DÉFINITION 2.5. Soit $x \in \Omega$. On dit que f est différentiable en x s'il existe $\lambda : \mathbb{R}^n \rightarrow \mathbb{R}^p$ linéaire, $\eta > 0$ et $\varepsilon : B_{\mathbb{R}^n}(0, \eta) \rightarrow \mathbb{R}^p$ telles que pour tout $h \in B_{\mathbb{R}^n}(0, \eta)$,

$$f(x+h) = f(x) + \lambda(h) + \|h\|\varepsilon(h), \quad \text{avec } \varepsilon(h) \xrightarrow{h \rightarrow 0} 0.$$

On appelle λ la différentielle de f en x ; on la note en fait Df_x .

Df_x peut aussi être notée $Df(x)$, $df(x)$ ou encore df_x selon les manuscrits.

2.4. Calcul d'une différentielle en pratique. Pour montrer que f est différentiable et calculer sa différentielle, on calcule $(f(x+h) - f(x))$ et on identifie dans l'expression les termes linéaires en h .

On a ainsi :

$$f(x+h) - f(x) = \lambda_x(h) + \text{reste}(h)$$

avec λ_x linéaire en h .

Si $\text{reste}(h) = o_0(\|h\|)$, on a alors montré la différentiabilité et exprimé sa différentielle en x qui est la partie linéaire en h i.e. $\lambda_x(h)$.

Si f est différentiable et en utilisant la notation Df_x , on a alors :

Si f est différentiable en $x \in \mathbb{R}^n$ alors : $\forall h \in \mathbb{R}^n$,

$$f(x+h) = f(x) + Df_x(h) + o_0(\|h\|)$$

où $o_0(\|h\|)$ peut s'écrire sous la forme : $\|h\|\varepsilon(h)$, $\varepsilon(h) \rightarrow_{h \rightarrow 0} 0$.

A x donné (x est fixé), Df_x est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$:
 $\forall (\lambda_1, \lambda_2) \in \mathbb{R} \times \mathbb{R}$, $\forall (h, k) \in \mathbb{R}^n \times \mathbb{R}^n$,

$$Df_x(\lambda_1 h + \lambda_2 k) = \lambda_1 Df_x(h) + \lambda_2 Df_x(k)$$

EXEMPLE 2.1. Différentielle de $\|\cdot\|_2^2$.

Notons $f : \mathbb{R}^n \rightarrow \mathbb{R}$; $x \mapsto f(x) = \|x\|_2^2 = \sum_{j=1}^n x_j^2$.

Pour tout $x, h \in \mathbb{R}^n$, on a :

$$f(x+h) = \sum_{j=1}^n (x_j + h_j)^2 = \sum_{j=1}^n x_j^2 + 2 \sum_{j=1}^n x_j h_j + \sum_{j=1}^n h_j^2.$$

A x fixé, l'application $h \mapsto 2 \sum_{j=1}^n x_j h_j$ est linéaire, et $\sum_{j=1}^n h_j^2 = \|h\|_2^2$ est négligeable par rapport à $\|h\|_2$ quand $h \rightarrow 0$.

Ainsi, f est différentiable en tout point $x \in \mathbb{R}^n$ et pour tout $x, h \in \mathbb{R}^n$,

$$D(\|x\|_2^2)(h) = 2 \sum_{j=1}^n x_j h_j = 2 \langle x, h \rangle$$

□

2.5. Autres propriétés de $Df_x(\cdot)$.

PROPOSITION 2.1.

1) (Unicité) Si f est différentiable en x , la différentielle de f en x est unique.

En effet on a : $\longrightarrow Df_x(h) = \lim_{t \rightarrow 0} \frac{f(x+th) - f(x)}{t}$.

2) (Continuité) Si f est différentiable en x alors f est continue en x .

3) (Linéarité) Si $f, g : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ sont différentiables en $x \in \Omega$, alors pour tous $(\lambda, \mu) \in \mathbb{R} \times \mathbb{R}$, $\lambda f + \mu g$ est différentiable en x et :

$$D(\lambda f + \mu g)(x) = \lambda Df_x + \mu Dg_x.$$

**

2.6. Interprétation de Df_x : linéarisation. Ré-écrivons le développement $f(a+h) = f(a) + Df_a(h) + \|h\|\varepsilon(h)$ avec $\lim_{h \rightarrow 0} \varepsilon(h) = 0$ comme suit :

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

Cette relation s'interprète comme suit : au voisinage de a , f peut être approchée (à l'ordre 1 en $(x-a)$) par l'application affine : $f(a) + Df_a(x-a)$.

On a linéarisé la fonction f au voisinage du point a .

On retrouve le même résultat que pour une fonction d'une seule variable réelle, dérivable au point a , pour laquelle la meilleure approximation à l'ordre 1 est donnée par la droite d'équation : $y(x) = f(a) + f'(a)(x-a)$.

Interprétation géométrique dans le cas d'une fonction scalaire à 2 variables : plan tangent.

REMARQUE 2.1. Cas scalaire : lien dérivée - application différentielle.

Si Ω est un ouvert de $\mathbb{R}$ ($n = 1$) et $f : \Omega \rightarrow \mathbb{R}^p$ i.e. f est une fonction vectorielle à une seule variable. f est différentiable en $x \in \Omega$ si et seulement si f est dérivable en x , et l'on a

$$Df_x(h) = h.f'(x) \text{ ou encore } f'(x) = Df_x(1).$$

En effet, pour tout réel h , on peut écrire $Df_x(h) = \lim_{t \rightarrow 0} \frac{f(x+th) - f(x)}{t} = \lim_{t \rightarrow 0} \frac{f(x+th) - f(x)}{th} h = f'(x)h$.

2.7. Cas particuliers. Soit Ω un ouvert de $\mathbb{R}^n$.

PROPOSITION 2.2. 1) Si $f : \Omega \rightarrow \mathbb{R}^p$ est constante, alors f est différentiable en tout point de Ω et $Df_x = 0$ pour tout $x \in \Omega$ (i.e. $Df_x : \mathbb{R}^n \rightarrow \mathbb{R}^p$ est l'application nulle).

2) Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ est linéaire, alors f est différentiable en tout point $x \in \mathbb{R}^n$ et $Df_x = f$ pour tout $x \in \mathbb{R}^n$. Autrement dit, la différentielle d'une application linéaire est elle-même.

Preuve :

1) Pour tout $x \in \Omega$ et h dans un voisinage de 0, $f(x+h) = f(x)$. D'où le résultat.

2) Pour tout $x \in \mathbb{R}^n$ et $h \in \mathbb{R}^n$, $f(x+h) = f(x) + f(h)$, f linéaire. D'où le résultat.

□

**** POUR ALLER PLUS LOIN**

On a

PROPOSITION 2.3. Si f est différentiable en un point a , alors pour tout $v \neq 0$, f admet une dérivée dans la direction v en a , avec $D_v f(a) = Df_a(v)$.

Attention. La dérivabilité selon tout vecteur en a n'entraîne pas nécessairement la différentiabilité de f en a . Nous savons construire des contre-exemples...

Par contre lorsque la fonction est suffisamment régulière (C^1) alors tout se passe bien.

3. Opérateurs différentiels usuels du premier ordre

3.1. Matrice jacobienne. Par la suite, pour ceux qui n'ont pas étudié la partie "Différentielles", vous pouvez remplacer dans les énoncés " f différentiable au point a " par " f de classe $\mathcal{C}^1$ au point a ".

DÉFINITION 2.6 (Jacobienne). Si $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ est différentiable en a , on appelle jacobienne de f au point a la matrice de Df_a dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^p$. On la note $J_f(a)$.

$J_f(a)$ est une matrice de taille $p \times n$,

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

La matrice jacobienne est la matrice de la différentielle Df_a de la fonction f . (Rappelons que la différentielle est par définition une application linéaire).

Exemples fondamentaux :

- Si f est une fonction scalaire, $J_f(a)$ est une matrice ligne.
- Si f est une application linéaire alors : $\forall a \in \mathbb{R}^n$, $Df_a = f$; et $\forall a \in \mathbb{R}^n$, $J_f(a) = \text{Mat}(f, \mathcal{B}_n, \mathcal{B}_p)$ (représentation de f dans les bases canoniques).

La matrice jacobienne intervient dans les changements d'équations mais aussi par exemple dans les développements limités.

EXERCICE 2.3. Soit $f : \mathbb{R}_*^+ \times]-\pi, \pi[\rightarrow \mathbb{R}^2$, $(\rho, \theta) \mapsto (x = \rho \cos \theta, y = \rho \sin \theta)$. Calculer la matrice jacobienne de f .

Noter que cette fonction f définit le changement de variable polaires.

3.2. Jacobien.

DÉFINITION 2.7 (Jacobian (cas $n = p$)).

Si $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ est différentiable au point $a \in \Omega$, on appelle jacobien de f en a le déterminant de la matrice jacobienne de f au point a . Il est noté :

$$\det J_f(a) = \frac{D(f_1, \dots, f_n)}{D(x_1, \dots, x_n)}(a).$$

Le jacobien apparaît lors des changements de variables dans les intégrales multiples (ie. à plusieurs variables).

EXERCICE 2.4. Soit $f : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $(\rho, \theta) \mapsto (x = \rho \cos \theta, y = \rho \sin \theta)$. Calculer le jacobien de f en tout point de Ω .

3.3. Gradient. L'opérateur gradient peut se définir dans le cas d'une fonction à n variables mais à une seule composante ($p = 1$) ; soit : $f : \mathbb{R}^n \rightarrow \mathbb{R}$.

DÉFINITION 2.8. Soit $f : \Omega \rightarrow \mathbb{R}$ une fonction scalaire définie sur l'ouvert Ω de $\mathbb{R}^n$, $a \in \Omega$. Si f admet des dérivées partielles en a , on appelle gradient de f en a le vecteur de $\mathbb{R}^n$ ayant pour composantes les dérivées partielles de f au point a . On le note $\nabla f(a)$,

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right)$$

A noter que : $\nabla f(a) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(a) e_j^n$.

EXERCICE 2.5. Calculer le gradient de la fonction suivante :

$$f_\alpha : (x, y) \mapsto (x^2 + y^2)^\alpha$$

Et préciser son ensemble de définition.

Notez que : $f_{1/2} = \|(x, y)\|_2$.

REMARQUE 2.2. On peut écrire la différentielle de f en a à partir de la Jacobienne de f : pour tout $h \in \mathbb{R}^n$,

$$Df_a(h) = J_f(a) \cdot h$$

Ou bien à partir de son gradient si f est à une seule composante : $Df_a(h) = \langle \nabla f(a), h \rangle$.

3.3.1. Interprétation géométrique.

PROPOSITION 2.4. Soit $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, différentiable en a telle que $\nabla f(a) \neq 0$. Soit $u = \frac{1}{\|\nabla f(a)\|} \nabla f(a)$. La variation de f au voisinage de a est maximale suivant le vecteur u et minimale suivant le vecteur $-u$. i.e. pour tout vecteur v tel que $\|v\| = 1$,

$$-Df_a(u) \leq Df_a(v) \leq Df_a(u).$$

Exemples :

- Si f désigne un champ de température, alors sa variation autour du point a est la plus forte lorsqu'on se déplace dans la direction du gradient en a .
- Sur une carte d'isobares météo, le gradient de pression en un point est orthogonal à la ligne isobare. (Faire un dessin).

3.4. Divergence (cas $n = p$).

DÉFINITION 2.9 (Opérateur divergence (cas $n = p$)).

Soit $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ différentiable au tout point $x \in \Omega$. On appelle divergence de f , la fonction scalaire suivante :

$$\operatorname{div}(f)(x) = \frac{\partial f_1}{\partial x_1}(x) + \dots + \frac{\partial f_n}{\partial x_n}(x)$$

Les opérateurs divergence, gradient apparaissent dans tous les modèles mathématiques de physique et mécanique au sens large (fluides, structures, ondes, électromagnétisme etc).

Notez bien que l'opérateur divergence (d'un champ vectoriel) est une fonction scalaire tandis que l'opérateur gradient (d'un champ scalaire) est une fonction vectorielle. Dans la littérature anglo-saxonne, div est noté $\nabla \cdot$.

EXERCICE 2.6. Soit

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}^3, (x, y, z) \mapsto (x^3 + y - 1/z, x^2 - \exp(y), 1/x + y^2)$$

Calculer la divergence de f , et préciser son ensemble de définition.

** Pour aller plus loin

Rotationnel (cas $n = p = 3$). Dans le cas d'un champ vectoriel tridimensionnel avec $n = p = 3$, on peut définir son rotationnel.

DÉFINITION 2.10 (Opérateur rotationnel (cas $n = p = 3$)).

Soit $f : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$ différentiable tout point $x \in \Omega$. On appelle rotationnel de f , la fonction vectorielle suivante :

$$\operatorname{rot}(f)(x) = \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} \wedge \begin{pmatrix} f_x \\ f_y \\ f_z \end{pmatrix} (x) = \begin{pmatrix} \partial_y f_z - \partial_z f_y \\ \partial_z f_x - \partial_x f_z \\ \partial_x f_y - \partial_y f_x \end{pmatrix} (x)$$

L'opérateur rotationnel apparaît fréquemment dans les modèles mathématiques de mécanique des fluides mais aussi structures, ondes, électromagnétisme etc. Le rotationnel de f est également noté $\nabla \wedge f$. Dans la littérature anglo-saxonne, rot est noté curl .

Notez bien que l'opérateur rotationnel (d'un champ vectoriel 3D) est une fonction vectorielle (vecteur de $\mathbb{R}^3$).

EXERCICE 2.7. Soit

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}^3, (x, y, z) \mapsto (x^3 + y - 1/z, x^2 - \exp(y), 1/x + y^2)$$

Calculer le rotationnel de f , et préciser son ensemble de définition.

**

3.5. Fonctions composées. Soient n, p, q trois entiers ≥ 1 . Soient $f : \mathbb{R}^n \rightarrow \mathbb{R}^p, g : \mathbb{R}^p \rightarrow \mathbb{R}^q$ et $x \in \mathbb{R}^n$. Si f est de classe $\mathcal{C}^1$ en un point a , si g est de classe $\mathcal{C}^1$ au point $f(a)$, alors $(g \circ f)$ est de classe $\mathcal{C}^1$ au point a .

EXERCICE 2.8. Soient

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x, y) \mapsto (x + y^2, 2x) \text{ et } g : \mathbb{R}^2 \rightarrow \mathbb{R}, (u, v) \mapsto u + v$$

Calculer les dérivées partielles de $F(x, y) = (g \circ f)(x, y)$ au point $x = (x_0, y_0)$.

On établit le théorème de dérivation des fonctions composées suivant :

THÉORÈME 2.2. Soient $f : \mathbb{R}^n \rightarrow \mathbb{R}^p, g : \mathbb{R}^p \rightarrow \mathbb{R}^q$. On note $F = g \circ f : \Omega \rightarrow \mathbb{R}^q$. On note $(f_k)_{1 \leq k \leq p}$ les fonctions composantes de f .

On suppose que $f(x)$ et $g(y)$ sont de classe $\mathcal{C}^1$.

Alors, $F = g \circ f : \Omega \rightarrow \mathbb{R}^q$ est de classe $\mathcal{C}^1$, et pour tout $x \in \mathbb{R}^n$,

$$\forall j, 1 \leq j \leq n, \quad \frac{\partial F}{\partial x_j}(x) = \frac{\partial (g \circ f)}{\partial x_j}(x) = \sum_{k=1}^p \frac{\partial g}{\partial y_k}(f(x)) \cdot \frac{\partial f_k}{\partial x_j}(x)$$

En terme de jacobienne, cela s'écrit

$$J_{g \circ f}(x) = J_g(f(x)) \cdot J_f(x)$$

Cette formule est l'extension naturelle matricielle de la formule de dérivation de fonctions composées à une seule variable : $(g \circ f)'(x) = g'(f(x)) \cdot f'(x)$.

En appliquant le théorème précédent au cas où $g = f^{-1}$ (on a alors nécessairement $p = n$), on obtient :

COROLLAIRE 2.1. Si f est une application bijective de $\Omega \subset \mathbb{R}^n \rightarrow \Omega' \subset \mathbb{R}^n$ de classe $\mathcal{C}^1$ sur Ω et si sa réciproque f^{-1} est de classe $\mathcal{C}^1$ sur Ω' alors pour tout $a \in \Omega$,

$$J_{f^{-1}}(f(a)) = (J_f(a))^{-1}$$

EXERCICE 2.9. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x, y) \mapsto (x + y^2, 2x)$. Calculer $J_{f^{-1}}(y)$, $y = f(x)$.

4. Dérivées partielles d'ordres supérieurs

4.1. Dérivées d'ordres supérieurs. Soit $f : \Omega \rightarrow \mathbb{R}^p$, Ω ouvert de $\mathbb{R}^n$. La dérivée partielle $\frac{\partial f}{\partial x_j}$ est une application de Ω dans $\mathbb{R}^p$ (si f suffisamment régulière). Si cette application admet une dérivée partielle d'indice k en un point $x \in \Omega$, $\frac{\partial}{\partial x_k} \left(\frac{\partial f}{\partial x_j} \right) (x)$ sera notée $\frac{\partial^2 f}{\partial x_k \partial x_j} (x)$. C'est une dérivée partielle d'ordre 2.

Plus généralement, on définit par récurrence les dérivées partielles de f d'ordre k : $\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}} (x)$.

DÉFINITION 2.11. Soit $k \geq 1$. On dit que $f : \Omega \rightarrow \mathbb{R}^p$ est de classe $\mathcal{C}^k$ sur Ω si toutes ses dérivées partielles d'ordre k existent et sont continues sur Ω .

PROPOSITION 2.5. La composée de deux applications $\mathcal{C}^k$ est de classe $\mathcal{C}^k$.

4.2. Théorème de Schwarz. Ce théorème permet de commuter les indices de dérivation sous une condition de continuité.

THÉORÈME 2.3 (Théorème de Schwarz). Soit $f : \Omega \rightarrow \mathbb{R}^p$ une fonction définie sur un ouvert Ω de $\mathbb{R}^n$, à valeurs dans $\mathbb{R}^p$. Soit $a \in \Omega$.

Si f admet des dérivées partielles secondes $\frac{\partial^2 f}{\partial x_k \partial x_j}$ et $\frac{\partial^2 f}{\partial x_j \partial x_k}$ sur Ω et si, de plus, ces fonctions sont continues

en a , alors :

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

Cette dérivée seconde de f pourra être notée : $\partial_{jk}^2 f(a)$.

EXERCICE 2.10. Soit

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto y.x^3 - 1/y$$

Vérifier que le théorème de Schwarz s'applique bien (pour les points où f est de classe $\mathcal{C}^2$ bien entendu).

4.3. Opérateur différentiel classique : le Laplacien.

DÉFINITION 2.12 (Laplacien).

Soit $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ de classe $\mathcal{C}^2$, et x un point de $\mathbb{R}^n$. On appelle laplacien de f en a , la fonction suivante :

$$\boxed{\Delta f(x) = \frac{\partial^2 f}{\partial x_1^2}(x) + \dots + \frac{\partial^2 f}{\partial x_n^2}(x)}$$

L'opérateur laplacien est un opérateur différentiel du deuxième ordre qui peut être défini tant pour une fonction f scalaire ($p = 1$) que vectorielle.

L'opérateur Laplacien s'étend de manière naturelle à une fonction à p composantes.

L'opérateur laplacien représente / modélise notamment les phénomènes de diffusion et il apparaît dans de nombreux modèles mathématiques de mécanique des fluides, structures, ondes, électromagnétisme etc.

PROPOSITION 2.6. a) Soient $u : \mathbb{R}^n \rightarrow \mathbb{R}^n$ de classe $\mathcal{C}^1$ (ie. un champ vectoriel), et $f : \mathbb{R}^n \rightarrow \mathbb{R}$ de classe $\mathcal{C}^1$ (f champ scalaire). On a :

$$\operatorname{div}(uf) = u \cdot \nabla f + f \operatorname{div}(u)$$

b) Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ de classe $\mathcal{C}^2$, on a :

$$\operatorname{div}(\nabla f) = \Delta f$$

Typiquement, u désigne un champ de vitesse et f un champ scalaire comme une température ou une concentration.

EXERCICE 2.11. Montrer cette proposition.

PROPOSITION 2.7. Soient $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ et $u : \mathbb{R}^3 \rightarrow \mathbb{R}$, tous deux de classe $\mathcal{C}^2$. On a :

$$\operatorname{div}(\operatorname{rot}(f)) = 0 \text{ et } \operatorname{rot}(\nabla u) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

EXERCICE 2.12. Montrer cette proposition.

5. Formule de Taylor à n variables

Par souci de simplification, on considère une fonction $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, i.e. à $p = 1$ composante. f est supposée être de classe $\mathcal{C}^k$. Soient a et h tels que : $[a, a + h] \subset \Omega$.

On définit l'application à variable réelle suivante : $u : [0, 1] \rightarrow \mathbb{R}^p$, $u(t) = f(a + th)$ (u est appelée "fonction auxiliaire" à la fonction f).

On calcule ensuite le développement limité de Taylor pour la fonction $u(t)$ au point 0.

Calculons les dérivées successives de u . On a :

$$u'(t) = \sum_{j=1}^n h_j \frac{\partial f}{\partial x_j}(a + th).$$

Pour j fixé, on peut dériver $t \mapsto \frac{\partial f}{\partial x_j}(a + th)$ (fonction auxiliaire associée à $\frac{\partial f}{\partial x_j}$) :

$$u''(t) = \sum_{i=1}^n \sum_{j=1}^n h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a + th).$$

Par récurrence, on obtient pour tout $1 \leq l \leq k$,

$$u^{(l)}(t) = \sum_{i_1, \dots, i_l=1}^n h_{i_1} \dots h_{i_l} \frac{\partial^l f}{\partial x_{i_1} \dots \partial x_{i_l}}(a + th).$$

Ecrivons à présent la formule de Taylor-Young pour $u(t)$ en $t = 0$,

$$u(1) = u(0) + u'(0) + \frac{u''(0)}{2!} + \dots + \frac{u^{(k-1)}(0)}{(k-1)!} + \dots$$

Finalement, on obtient la formule de Taylor-Young à n variables, de f au point a :

$$f(a + h) = f(a) + \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(a) + \dots + \frac{1}{(k-1)!} \sum_{i_1, \dots, i_{k-1}=1}^n h_{i_1} \dots h_{i_{k-1}} \frac{\partial^{k-1} f}{\partial x_{i_1} \dots \partial x_{i_{k-1}}}(a) + o(\|h\|^k)$$

Par ailleurs, on définit

DÉFINITION 2.13. Soit $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$. On appelle Hessienne de f au point a , la matrice des dérivées partielles secondes de f en a . On la note $H_f(a)$,

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

Si f est de classe $\mathcal{C}^2$, par le théorème de Schwarz, $H_f(a)$ est une matrice *symétrique*.

(A noter que cela implique notamment que $H_f(a)$ est diagonalisable, et de valeurs propres réelles).

La formule de Taylor-Young à l'ordre 2 de f au point a , s'écrit alors :

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

avec $H_f(a)$ la matrice Hessienne de f au point a .

EXERCICE 2.13. Soit f la fonction définie par : $f(x, y) = x^2 + y^2 - 2x + 1$.
Écrire le D.L. de Taylor-Young de f à l'ordre 2 en zéro.

Extrema d'une fonction de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$ (premiers pas vers l'optimisation).

Dans ce chapitre, on considère des fonctions *scalaires* à n variables ; f est définie de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$ (i.e. $p = 1$).

L'objet de ce chapitre est d'établir des conditions nécessaires et suffisantes de minimum local d'une fonctionnelle. Cette question est pleinement reliée aux problèmes d'optimisation différentiable qui se formalisent ainsi : trouver x^* tel que $j(x^*) = \min_x j(x)$, avec j le critère à minimiser / optimiser.

DÉFINITION 3.1. Soient Ω un ouvert de $\mathbb{R}^n$ et $f : \Omega \rightarrow \mathbb{R}$. Soit $a \in \Omega$. On dit que f admet un *extremum local* (maximum ou minimum) au point a s'il existe un voisinage $B(a, r) \subset \Omega$ de a tel que :

$$\begin{aligned} \forall x \in B(a, r), \quad f(x) &\leq f(a) & (\text{maximum}), \\ \forall x \in B(a, r), \quad f(x) &\geq f(a) & (\text{minimum}). \end{aligned}$$

On dit que l'extremum est *strict* si les inégalités ci-dessus sont strictes.

1. Condition nécessaire d'extremum

PROPOSITION 3.1. Soit $f : \Omega \rightarrow \mathbb{R}$ et $a \in \Omega$.

Si f admet un extremum au point a et si f est de classe C^1 , alors $\nabla f(a) = 0$.

Attention : ceci est une condition nécessaire mais non suffisante !

Terminologie. On appelle *point critique* tout point $a \in \Omega$ tel que : $\nabla f(a) = 0$.

EXERCICE 3.1. Soit $f(x, y) = x^2 + y^2 - 2x + 1$.

- Montrer que f admet un unique point critique, que vous déterminerez.
- Montrer que ce point critique est un minimum global.

EXERCICE 3.2. Soit $f(x, y) = -x^2 + y^2$.

- Montrer que f admet un unique point critique, que vous déterminerez.
- Montrer que ce point critique n'est ni un minimum local ni un maximum local.
- Faites un dessin.

2. Condition suffisante d'extremum - Matrice Hessienne

Rappel : cas d'une fonction scalaire i.e. $n = 1$.

Reprenons le cas d'une variable réelle (cas $n = 1$). La condition $f'(a) = 0$ n'entraîne pas nécessairement que f admet un extremum en a ! C'est une condition nécessaire mais non suffisante. Par contre, si f est deux fois dérivable en a (f de classe C^2 en a), si $f'(a) = 0$ et $f''(a) > 0$, alors f admet un minimum local en a .

Ceci se démontre en écrivant le développement de Taylor de f à l'ordre 2 au point a .

Extension au cas n variables.

Pour une fonction à n variables ($f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$), " $f''(a)$ " va être remplacé par le deuxième terme du développement de Taylor-Young : la *matrice Hessienne* de f au point a .

Rappelons que la matrice Hessienne $H_f(a)$ est symétrique et donc diagonalisable et ses valeurs propres sont réelles.

Le développement limité de Taylor d'ordre 2 de f au point a s'écrit :

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

Etant donné a , on définit $Q_f(a)$ la forme quadratique suivante :

$$Q_f(a)(h) = \langle H_f(a)h, h \rangle \text{ pour tout } h \in \mathbb{R}^n$$

Pour un point critique a , on a alors :

$$f(a+h) - f(a) = \frac{1}{2} Q_f(a)(h) + o_a(\|h\|^2)$$

C'est à dire que localement (i.e. au voisinage de a), le signe de $(f(a+h) - f(a))$ est le même que celui de $Q_f(a)(h)$.

Etudier la présence éventuel d'un extremum local au point a revient alors à étudier $Q_f(a)$.

Rappelons qu'une matrice A matrice est définie positive si il existe $c > 0$ tel que : $\forall x \in \mathbb{R}^n, \langle Ax, x \rangle \geq c\|x\|^2$.

Propriétés de la forme quadratique $Q_f(a)$ On a les résultats suivants :

- (1) $Q_f(a)$ est positive définie (resp. négative définie) si et seulement si toutes les valeurs propres de $H_f(a)$ sont strictement positives (resp. négatives).
- (2) $Q_f(a)$ n'est ni positive ni négative si et seulement si $H_f(a)$ admet des valeurs propres positives et des valeurs propres négatives.
- (3) $Q_f(a)$ est positive (resp. négative) sans être définie si et seulement si toutes les valeurs propres de $H_f(a)$ sont positives (resp. négatives), et que 0 est valeur propre de $H_f(a)$.

EXERCICE 3.3. Dans le cas $n = 2$ -, faire un dessin correspondant au graphe type de $Q_f(a)$ pour chaque cas.

On a alors le

THÉORÈME 3.1 (Condition suffisante d'extremum local strict).

Soient $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ de classe $\mathcal{C}^2$, $a \in \Omega$ un point critique de f ($\nabla f(a) = 0$), et $H_f(a)$ la matrice hessienne de f au point a .

- i) Si toutes les valeurs propres de $H_f(a)$ sont strictement positives (resp. négatives) alors f admet un minimum (resp. maximum) local strict au point a .
- ii) Si $H_f(a)$ admet des valeurs propres positives et des valeurs propres négatives alors f admet un point selle (point col) en a . (f n'admet donc pas d'extremum en a).
- iii) Si $Q_f(a)$ est positive sans être définie, ou négative sans être définie, alors on ne peut rien affirmer.

**** POUR ALLER PLUS LOIN.**

Preuve : On écrit le développement de Taylor-Young au point a : pour h assez petit,

$$f(a+h) = f(a) + 0 + \frac{1}{2} Q_f(a)(h) + \|h\|^2 \varepsilon(h), \quad \varepsilon(h) \xrightarrow{h \rightarrow 0} 0.$$

i) $H_f(a)$ est définie positive. Il existe $c > 0$ tel que :

$$\forall x \in \mathbb{R}^n, \langle H_f(a)x, x \rangle \geq c\|x\|^2$$

Ainsi,

$$f(a+h) - f(a) \geq \frac{c}{2} \|h\|^2 + \|h\|^2 \varepsilon(h).$$

Il existe $\alpha > 0$ tel que si $\|h\| < \alpha$, $|\varepsilon(h)| \leq \frac{c}{4}$. D'où pour $\|h\| < \alpha$, $h \neq 0$,

$$f(a+h) - f(a) \geq \frac{c}{4} \|h\|^2 > 0.$$

f admet donc un minimum local strict en a .

iii) On note (λ_k, w_k) les éléments propres de H , $1 \leq k \leq n$: $Hw_k = \lambda_k w_k$.

On a : $\langle H_f(a)w_k, w_k \rangle = \lambda_k \|w_k\|^2$. A partir d'une renormalisation de w_k , on peut définir $x_1 \in \mathbb{R}^n$ et $x_2 \in \mathbb{R}^n$ tels que :

$$\langle H_f(a)x_1, x_1 \rangle = 1 \text{ et } \langle H_f(a)x_2, x_2 \rangle = -1$$

Prenant $h = tx_1$, t réel assez petit,

$$f(a + tx_1) = f(a) + \frac{t^2}{2} + t^2 \varepsilon(tx_1), \quad \varepsilon(t) \xrightarrow{t \rightarrow 0} 0.$$

Donc pour $t \neq 0$ assez petit, $f(a + tx_1) > f(a)$.

De même, pour $t \neq 0$ assez petit, $f(a + tx_2) < f(a)$. Donc tout voisinage de a contient des points x tels que $f(x) > f(a)$ et d'autres tels que $f(x) < f(a)$. f n'admet pas d'extremum en a .

ii) $H_f(a)$ est définie négative si et seulement si $-Hess_f(a)$ est définie positive... □

**

EXERCICE 3.4. Dans le cas $n = 2$, et pour chaque situation i), ii) et iii), tracer, dans la base des vecteurs propres, les 3 graphes illustrant le comportement local de f au voisinage du point critique a .

En pratique, étudier le comportement de $Q_f(a)$ revient à étudier les valeurs propres de la matrice Hessienne $H_f(a)$.

Signe des valeurs propres dans le cas $n = 2$ (voire $n = 3$).

Soit $A = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$ une matrice symétrique réelle. On a : $\det A = rt - s^2$.

Notons λ_1 et λ_2 les deux valeurs propres de A . Elles sont réelles car A est symétrique.

Si $\det A < 0$, alors $\lambda_1 \lambda_2 < 0$ donc A n'est ni positive ni négative.

Si $\det A > 0$, alors $\lambda_1 \lambda_2 > 0$ donc les deux valeurs propres ont le même signe.

De plus, $\text{tr} A = r + t = \lambda_1 + \lambda_2$, donc les valeurs propres sont du signe de $(r + t)$, i.e. du signe de r (car $rt > s^2 > 0$, donc r et t ont le même signe).

Cas $n = 3$. Dans le cas d'une matrice 3×3 , un raisonnement semblable peut encore être aisément mené.

EXERCICE 3.5. Reprenez les deux exercices précédents et déterminez la nature des points critiques en utilisant le théorème "conditions suffisantes".

Changements de variables

Les changements de variables sont requis dans tous les domaines de la modélisation et aussi par exemple pour la résolution d'équations différentielles. Les changements de variables sont également indispensables pour le calcul de nombreuses intégrales, cf le chapitre traitant d'intégration.

1. Définition

DÉFINITION 4.1. Soient Ω_1 et Ω_2 deux ouverts de $\mathbb{R}^n$. Un changement de variable défini de Ω_1 dans Ω_2 est une bijection "régulière", au sens où la fonction doit être de classe C^1 et d'inverse C^1 également.

2. Principes d'un changement de variables

On présente ici les principes d'un changement de variables au travers du cas classique du changement de coordonnées en variables polaires.

2.1. Coordonnées en variables polaires : définition et propriétés. Le changement de variable polaire est défini ainsi :

$$\varphi : (r, \theta) \mapsto (x, y) \text{ avec } \begin{cases} x &= r \cos \theta \\ y &= r \sin \theta \end{cases}$$

L'application φ est de classe C^1 (et même C^∞) sur $\mathbb{R}^2$. Elle définit un changement de variables de $\mathbb{R}^{+*} \times]-\pi, \pi]$ dans $\mathbb{R}^{*2}$.

On a :

$$J_\varphi(r, \theta) = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}$$

Son Jacobien est alors : $\frac{D(x, y)}{D(r, \theta)} = r$.

(A noter que ce changement de variable présente une singularité en $r = 0$, mais cela n'est en fait pas pénalisant pour la suite).

EXERCICE 4.1. *Faites un dessin.*

2.2. Expression du gradient après changement de variable : cas du changement en variables polaires. Etant donnée $f(x, y)$ une fonction définie sur $\mathbb{R}^2$. Pour une raison quelconque nous souhaiterions exprimer cette fonction en coordonnées polaires. On pose alors :

$$f^*(r, \theta) = f(x, y).$$

Soit : $f^* = f \circ \varphi$ puisque $f^*(r, \theta) = f(r \cos \theta, r \sin \theta)$.

Cherchons à exprimer $\nabla f(x, y)$ en fonction des variables polaires (r, θ) .

Pour cela on calcule les dérivées partielles de f comme suit (derivation de composées) :

$$(2) \quad \frac{\partial f}{\partial x}(x, y) = \frac{\partial f^*}{\partial r}(r, \theta) \frac{\partial r}{\partial x} + \frac{\partial f^*}{\partial \theta}(r, \theta) \frac{\partial \theta}{\partial x}$$

$$(3) \quad \frac{\partial f}{\partial y}(x, y) = \frac{\partial f^*}{\partial r}(r, \theta) \frac{\partial r}{\partial y} + \frac{\partial f^*}{\partial \theta}(r, \theta) \frac{\partial \theta}{\partial y}$$

Ce système s'écrit de manière équivalente ainsi :

$$(4) \quad \nabla f(x, y) = J_{\varphi^{-1}}(x, y) \nabla f^*(r, \theta)$$

avec

$$J_{\varphi^{-1}}(x, y) = \begin{pmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{pmatrix}$$

On a alors besoin des dérivées partielles du changement de variable inverse $\varphi^{-1}(x, y) = (r, \theta)$.

Dans le cas présent, on sait calculer ce changement de variable inverse φ^{-1} . (Notons que ceci n'est pas le cas pour tous les changements de variables!...).

On a : $r = (x^2 + y^2)^{1/2}$.

Concernant l'expression explicite de θ , pour raison de simplicité ($\tan$ est bijective sur $] -\pi/2, \pi/2[$), on considère φ^{-1} défini sur $(R_+^+)^2$ uniquement. On a alors : $\theta = \arctan\left(\frac{y}{x}\right)$..

D'où, après calculs :

$$J_{\varphi^{-1}}(x, y) = \begin{pmatrix} \cos \theta & \sin \theta \\ -\frac{\sin \theta}{r} & \frac{\cos \theta}{r} \end{pmatrix}$$

Ce qui donne :

$$(5) \quad \frac{\partial f}{\partial x}(x, y) = \frac{\partial f^*}{\partial r}(r, \theta) \cos \theta - \frac{\partial f^*}{\partial \theta}(r, \theta) \frac{\sin \theta}{r}$$

$$(6) \quad \frac{\partial f}{\partial y}(x, y) = \frac{\partial f^*}{\partial r}(r, \theta) \sin \theta + \frac{\partial f^*}{\partial \theta}(r, \theta) \frac{\cos \theta}{r}$$

Soit :

$$(7) \quad \boxed{\nabla f(x, y) = J_{\varphi^{-1}}^T(x, y) \cdot \nabla f^*(r, \theta)}$$

REMARQUE 4.1. On a également la relation :

$$\boxed{\nabla f^*(r, \theta) = J_{\varphi}^T(r, \theta) \cdot \nabla f(x, y)}$$

2.3. Résumé du calcul du gradient après changement de variables : cas général. Si on considère la fonction suivante (cas $n = 2$) : $\varphi(x, y) = (u(x, y), v(x, y))$, la jacobienne de φ au point (x, y) s'écrit :

$$J_{\varphi}(x, y) = \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix} (x, y)$$

$J_{\varphi}(x, y)$ se calcule toujours facilement car on dispose de $u(x, y)$ et $v(x, y)$.

Par contre, il est fréquent que l'on ne sache pas expliciter le changement de variables inverse φ^{-1} i.e. (x, y) en fonction de (u, v) ! ...

Par contre on peut facilement calculer les dérivées partielles de φ^{-1} (ici les fonctions $(x(u, v), y(u, v))$) en inversant la jacobienne $J_{\varphi}(x, y)$.

En effet, si $(u, v) = \varphi(x, y)$, on a : $J_{\varphi}(x, y) J_{\varphi^{-1}}(u, v) = I$; soit :

$$(8) \quad \boxed{J_{\varphi^{-1}}(u, v) = (J_{\varphi}(x, y))^{-1}}$$

On obtient ainsi les dérivées partielles de φ^{-1} sans avoir explicité φ^{-1} !

Dans le cas $\varphi(x, y) = (u(x, y), v(x, y))$, cela s'écrit :

$$\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} (u, v) = \left(\begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix} (x, y) \right)^{-1}$$

* RAPPEL. COMMENT INVERSER UNE MATRICE 2×2 ?

Rappelons que pour une matrice de taille 2×2 , voire 3×3 , on peut utiliser la formule de Cramer : $A^{-1} = \frac{1}{\det(A)} {}^T \text{co}A$. (Pour rappel, la comatrice de la matrice carrée A est la matrice constituée de ses cofacteurs).

Dans le cas 2×2 , si :

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

alors :

$$A^{-1} = \frac{1}{(ad-bc)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Dans le cas 3×3 , la formule générale s'applique relativement aisément. On a :

$$\text{co}A = (\text{cofac}_{i,j})_{1 \leq i,j \leq 3}$$

avec $\text{cofac}_{i,j} = (-1)^{i+j} \det(m)$, où m est la sous-matrice déduite de A en ayant enlevé la ligne i et la colonne j .

3. Exemple d'application : résolution d'une équation différentielle par changement de variables

Illustrons l'usage du changement de variable polaire pour résoudre une équation différentielle, ici une Equation aux Dérivées Partielles (EDP) linéaire à deux variables.

Considérons l'équation différentielle (de fait une EDP) suivante :

$$(9) \quad x \frac{\partial f}{\partial x}(x, y) + y \frac{\partial f}{\partial y}(x, y) = (x^2 + y^2)$$

L'objet est de calculer la/les fonctions f définies sur $\mathbb{R}^{++} \times \mathbb{R}$ à valeurs dans $\mathbb{R}^2$, f de classe $\mathcal{C}^1$, qui vérifient cette équation. En d'autres termes on cherche à résoudre cette équation sur $\Omega = \mathbb{R}^{++} \times \mathbb{R}$. Cette équation a-priori difficile à résoudre va en fait devenir bien plus simple après passage en variables polaires. (NB. Ca ne saute a-priori pas aux yeux).

1) Effectuons le changement en variables polaires. On définit : $f^*(r, \theta) = f(r \cos \theta, r \sin \theta)$ pour $(r, \theta) \in \mathbb{R}^{++} \times]-\frac{\pi}{2}, \frac{\pi}{2}[$.

En utilisant (4), l'équation s'écrit :

$$\begin{aligned} x \frac{\partial f}{\partial x}(x, y) + y \frac{\partial f}{\partial y}(x, y) &= r \cos \theta \left(\frac{\partial f^*}{\partial r} \cos \theta - \frac{\partial f^*}{\partial \theta} \frac{\sin \theta}{r} \right) (r, \theta) + r \sin \theta \left(\frac{\partial f^*}{\partial r} \sin \theta + \frac{\partial f^*}{\partial \theta} \frac{\cos \theta}{r} \right) (r, \theta) \\ &= r \frac{\partial f^*}{\partial r}(r, \theta) \end{aligned}$$

Ainsi on obtient l'équation différentielle plus simple suivante :

$$(10) \quad \frac{\partial f^*}{\partial r}(r, \theta) = r \quad , \text{ pour } (r, \theta) \in \mathbb{R}_*^+ \times]-\pi/2, \pi/2[$$

Cette nouvelle équation peut facilement se résoudre "à la main", et ce grâce au changement de variable précédent.

2) Résolution en variables (r, θ) .

En intégrant (10) on obtient que $f^*(r, \theta)$ est la somme de $\frac{r^2}{2}$ et d'une fonction qui ne dépend que de θ :

$$f^*(r, \theta) = \frac{r^2}{2} + c(\theta)$$

3) Retour aux variables (x, y) .

On peut désormais obtenir la solution de l'équation originelle en effectuant le changement de variable inverse. Comme $x > 0$, $\theta \in]-\frac{\pi}{2}, \frac{\pi}{2}[$, et on a $\theta = \arctan\left(\frac{y}{x}\right)$ ($\tan$ est bijective sur $] -\pi/2, \pi/2[$).

On peut donc réécrire $c(\theta) = \tilde{c}\left(\frac{y}{x}\right)$. On obtient finalement les solutions de l'EDP suivantes : $f(x, y) =$

$\frac{x^2 + y^2}{2} + \tilde{c}\left(\frac{y}{x}\right)$, pour toute fonction $\tilde{c}$ de classe $\mathcal{C}^1$ sur $\mathbb{R}$.

4. Coordonnées cylindriques et sphériques

L'extension des coordonnées polaires au 3D peut s'effectuer de deux manières :

- a) linéaire en z : cela donne les coordonnées cylindriques,
- b) angulaire en z : cela donne les coordonnées sphériques.

4.1. Coordonnées cylindriques. On pose

$$\Phi : U \rightarrow \mathbb{R}_*^3, (\rho, \theta, z) \mapsto (x = \rho \cos \theta, y = \rho \sin \theta, z)$$

avec $U = \mathbb{R}^{+*} \times]-\pi, \pi] \times \mathbb{R}$.

Φ définit le changement de variable cylindrique (Φ est une bijection de $\mathbb{R}^3$, de classe $\mathcal{C}^1$ et d'inverse $\mathcal{C}^1$).

EXERCICE 4.2. *Monter que son Jacobien est : $\det J_\Phi(\rho, \theta, z) = \rho$.*

Le calcul du Jacobien est indispensable pour un changement de variable dans une intégrale triple, cf chapitre suivant.

4.2. Coordonnées sphériques. On pose :

$$\Phi : U \rightarrow \mathbb{R}_*^3, (\rho, \varphi, \theta) \mapsto (x = \rho \sin \varphi \cos \theta, y = \rho \sin \varphi \sin \theta, z = \rho \cos \varphi)$$

avec $U = \mathbb{R}^{+*} \times]0, \pi] \times]-\pi, +\pi]$.

Φ définit le changement de variable sphérique (Φ est une bijection de $\mathbb{R}^3$, de classe $\mathcal{C}^1$ et d'inverse $\mathcal{C}^1$).

EXERCICE 4.3. *Monter que son Jacobien est : $\det J_\Phi(\rho, \theta, \varphi) = \rho^2 \sin \varphi$.*

(Valeur de Jacobien requise pour un changement de variable cylindrique dans une intégrale triple).

Intégrales multiples

Les intégrales multiples se retrouvent dans tous les domaines de la modélisation physique au sens large ainsi qu'en probabilités (et aussi dans bien d'autres domaines).

1. Propriétés fondamentales

EXERCICE 5.1.

Calculer les intégrales suivantes : a) $\int_{[0,1]^2} x \sin(x^2 + y) dx dy$.

b) $\int_{[0,1] \times [1,2]} e^{x+y} dx dy$.

Rem : Les ensembles $[0, 1]^2$ et $[0, 1] \times [1, 2]$ sont ce que l'on appelle des pavés de $\mathbb{R}^2$.

1.1. Théorème de Fubini. Enonçons ce théorème dans le cas particulier de pavés de $\mathbb{R}^n$.

THÉORÈME 5.1 (Intégrale sur un produit de pavé).

Soit P un pavé de $\mathbb{R}^n$ et Q un pavé de $\mathbb{R}^m$, $n, m \geq 1$.

Soit $f : P \times Q \rightarrow \mathbb{R}$, $(x, y) \rightarrow f(x, y)$ une fonction continue. On a :

$$\int_{P \times Q} f(x, y) dx dy = \int_P \left(\int_Q f(x, y) dy \right) dx = \int_Q \left(\int_P f(x, y) dx \right) dy$$

COROLLAIRE 5.1 (Cas des fonctions à variables séparées).

Soit P un pavé de $\mathbb{R}^n$ et Q un pavé de $\mathbb{R}^m$, $n, m \geq 1$. Soit $f : P \rightarrow \mathbb{R}$ et $g : Q \rightarrow \mathbb{R}$ deux fonctions continues.

Alors,

$$\int_{P \times Q} f(x)g(y) dx dy = \left(\int_P f(x) dx \right) \cdot \left(\int_Q g(y) dy \right)$$

Dans ce cas (et si $n = m = 1$), l'intégrale double se ramène au produit de deux intégrales simples.

1.2. Théorème de la moyenne. Soit A un ensemble compact de $\mathbb{R}^n$ et $|A|$ sa mesure définie par : $|A| = \int_A dx$ (i.e. sa surface en 2D et son volume en 3D). Soit $f : A \rightarrow \mathbb{R}$ une fonction continue.

La valeur moyenne de f sur A est définie par :

$$\bar{f}_A = \frac{1}{|A|} \int_A f(x) dx$$

THÉORÈME 5.2 (Théorème de la moyenne). Soit A un ensemble compact de $\mathbb{R}^n$ et $f : A \rightarrow \mathbb{R}$ une fonction continue. On suppose A connexe. Alors il existe $x_0 \in A$, tel que :

$$f(x_0) = \bar{f}_A$$

Aussi pour tout $g : A \rightarrow \mathbb{R}$ continue et positive, il existe $x_0 \in A$ tel que :

$$\overline{(f \cdot g)}_A = f(x_0) \int_A g(x) dx = f(x_0) |A| \bar{g}_A.$$

2. Intégrale multiple sur un compact élémentaire

On montre ci-dessous comment calculer en pratique une intégrale multiple sur un ensemble qui peut s'écrire sous la forme dite de "compact élémentaire".

EXERCICE 5.2. *Quelle est la mesure des ensembles suivants ?*

- $A_1 = \{(x, y) \in \mathbb{R}^2, x + y \leq 1 \text{ et } x \geq 0 \text{ et } y \geq 0\}$
- $A_2 = \{(x, y) \in \mathbb{R}^2, x^2 + y \leq 0 \text{ et } y - x \geq 0\}$
- $A_3 = \{(x, y, z) \in \mathbb{R}^3, x + y + z \leq 1 \text{ et } x \geq 0 \text{ et } y \geq 0 \text{ et } z \geq 0\}$.

DÉFINITION 5.1. *On appelle compact élémentaire de $\mathbb{R}^n$ tout sous-ensemble compact de $\mathbb{R}^n$ défini par des inégalités "emboîtées" du type :*

$$\begin{aligned} a &\leq x_{\sigma(1)} \leq b \\ f_1(x_{\sigma(1)}) &\leq x_{\sigma(2)} \leq g_1(x_{\sigma(1)}) \end{aligned}$$

$$\vdots$$

$$f_{n-1}(x_{\sigma(1)}, \dots, x_{\sigma(n-1)}) \leq x_{\sigma(n)} \leq g_{n-1}(x_{\sigma(1)}, \dots, x_{\sigma(n-1)})$$

où $a, b \in \mathbb{R}$, $f_1, \dots, f_n, g_1, \dots, g_n$ sont continues, et σ est une permutation.

Le résultat suivant est alors central pour les calculs.

PROPOSITION 5.1 (Intégrale sur un compact élémentaire).

Soit A un compact élémentaire de $\mathbb{R}^n$ défini par les inégalités "emboîtées" suivantes :

$$\begin{aligned} a &\leq x_1 \leq b \\ f_1(x_1) &\leq x_2 \leq g_1(x_1) \end{aligned}$$

$$\vdots$$

$$f_{n-1}(x_1, \dots, x_{n-1}) \leq x_n \leq g_{n-1}(x_1, \dots, x_{n-1})$$

où $a, b \in \mathbb{R}$, et $f_1, \dots, f_n, g_1, \dots, g_n$ sont des fonctions continues.

Soit $f : A \rightarrow \mathbb{R}$ une fonction continue alors :

$$\int_A f(x) dx = \int_a^b \int_{f_1(x_1)}^{g_1(x_1)} \dots \left(\int_{f_{n-1}(x_1, \dots, x_{n-1})}^{g_{n-1}(x_1, \dots, x_{n-1})} f(x_1, \dots, x_n) dx_n \right) \dots dx_2 dx_1$$

Autrement dit, si l'ensemble A est un compact élémentaire, on peut calculer chaque intégrale élémentaire une à une, en respectant le bon ordre.

Tout ensemble A de $\mathbb{R}^n$ ne peut s'écrire sous forme de compact élémentaire.

En pratique, il faut savoir caractériser le domaine d'intégration Ω sous forme de compact élémentaire ; cela peut être plus ou moins facile...

3. Changement de variables dans une intégrale multiple (double, triple)

EXERCICE 5.3. Quelle est la mesure (i.e. l'aire) de $D = \left\{ (x, y) \in \mathbb{R}^2, \frac{1}{4} \leq x^2 + y^2 \leq 1 \right\}$?

3.1. Formule de changement de variables. On a la formule de changement de variables suivante dans une intégrale multiple.

THÉORÈME 5.3. Soit Ω un ouvert de $\mathbb{R}^n$ et $\Phi : \Omega \rightarrow \mathbb{R}^n$ un changement de variable (Φ bijection de $\mathbb{R}^n$, de classe $\mathcal{C}^1$ et d'inverse $\mathcal{C}^1$). Soient A un ensemble compact inclus dans Ω et $f : \Phi(A) \rightarrow \mathbb{R}$ continue. On a :

$$\int_{\Phi(A)} f(y) dy = \int_A (f \circ \Phi)(x) \cdot |Jac(\Phi)(x)| dx$$

où $Jac(\Phi)$ est le jacobien de $\Phi : Jac(\Phi)(x) = \det(D\Phi(x))$.

Si on note :

$$y = \Phi(x) \text{ i.e. } \begin{cases} y_1 = \Phi_1(x_1, \dots, x_n) \\ \vdots \\ y_n = \Phi_n(x_1, \dots, x_n) \end{cases}$$

La formule de changement de variables se ré-écrit :

$$\int_{\Phi(A)} f(y_1, \dots, y_n) dy_1 \dots dy_n = \int_A f(\Phi_1(x_1, \dots, x_n), \dots, \Phi_n(x_1, \dots, x_n)) \left| \frac{D(y_1, \dots, y_n)}{D(x_1, \dots, x_n)} \right| dx_1 \dots dx_n$$

REMARQUE 5.1.

1. Dans la formule de changement de variable, on peut retenir que :

$$dy_1 \dots dy_n = \left| \frac{D(y_1, \dots, y_n)}{D(x_1, \dots, x_n)} \right| dx_1 \dots dx_n.$$

2. Mais pourquoi n'avait-on pas une valeur absolue dans la formule de changement de variable en 1D (sur un intervalle $[a, b]$) ? En fait nous l'avons bien cette valeur absolue... sauf si l'on suppose implicitement que $a < b$...
3. La formule de changement de variable peut aussi s'écrire :

$$\int_A f(x) dx = \int_{\Phi(A)} (f \circ \Phi^{-1})(y) \cdot |Jac(\Phi^{-1})(y)| dy$$

et sachant que : $\det(D\Phi^{-1}(y)) = [\det(D\Phi(x))]^{-1}$.

Notons que si l'on considère un domaine élémentaire (i.e. local autour du point considéré), la mesure de l'image de ce domaine par f sera celui du domaine de départ multiplié par la valeur absolue du jacobien.

3.2. Changement de variable polaire. Pour tout ensemble A (ensemble compact) de $\mathbb{R}^2$, on a :

$$\int \int_A f(x, y) dx dy = \int \int_{\phi^{-1}(A)} f \circ \varphi(\rho, \theta) \rho d\rho d\theta$$

EXERCICE 5.4. Calculer la surface de la portion de disque de rayon R et d'"angle" $(0, 2\pi/3)$.

3.3. Changement de variable cylindrique. Pour tout ensemble A (ensemble compact) de $\mathbb{R}^3$, on a :

$$\iiint_A f(x, y, z) dx dy dz = \iiint_{\phi^{-1}(A)} f \circ \Phi(\rho, \theta, z) \rho d\rho d\theta dz$$

EXERCICE 5.5. *Calculer le volume de la portion de cylindre de révolution de rayon R , et de hauteur comprise entre 0 et h .*

3.4. Changement de variable sphérique. *Pour tout ensemble A (ensemble compact) de $\mathbb{R}^3$, on a :*

$$\iiint_A f(x, y, z) \, dx dy dz = \iiint_{\phi^{-1}(A)} f \circ \Phi(\rho, \theta, \varphi) \, \rho^2 \sin \varphi \, d\rho d\varphi d\theta$$

EXERCICE 5.6. *Calculer le volume de la boule de centre 0 et de rayon R dans $\mathbb{R}^3$.*