Analyse 3

Notes de cours

André Giroux
Département de Mathématiques et Statistique
Université de Montréal
Mai 2004

Table des matières

Table des figures

1 INTRODUCTION

L'analyse mathématique est l'étude approfondie du calcul différentiel et intégral. Ce cours porte sur le calcul différentiel des fonctions de plusieurs variables. On commence par y établir les propriétés algébriques, géométriques et topologiques de l'espace euclidien à n dimensions, l'espace $\mathbb{R}^n$. On y étudie ensuite le calcul différentiel des fonctions numériques de plusieurs variables, les fonctions $\mathbb{R}^n \to \mathbb{R}$. On y analyse enfin les transformations différentiables des espaces euclidiens, les fonctions $\mathbb{R}^n \to \mathbb{R}^m$, avec en particulier une démonstration du théorème des fonctions inverses et une de celui des fonctions implicites. Comme application, on présente les méthodes classiques du calcul différentiel pour l'optimisation d'une fonction $f(x_1, x_2, \ldots, x_n)$, avec ou sans contrainte sur les variables $x_1, x_2, \ldots, x_n$.

L'étudiant est réputé être familier avec le calcul différentiel des fonctions d'une variable, les fonctions $\mathbb{R} \to \mathbb{R}$. Rappelons quelques résultats importants de ce calcul.

a) Le critère de Cauchy.

Une suite $\{x_n\}_{n\in\mathbb{N}}$ numérique est convergente si et seulement si

$$\lim_{n,m\to+\infty} |x_n - x_m| = 0.$$

b) Le théorème de Bolzano-Weierstrass.

Toute suite $\{x_n\}_{n\in\mathbb{N}}$ de points d'un intervalle compact $[a, b]$ contient une suite partielle convergeant vers un point de cet intervalle.

c) La propriété des valeurs extrêmes.

L'image d'un intervalle compact par une fonction continue est un intervalle compact.

d) Le théorème des accroissements finis.

Si $f : [a, b] \to \mathbb{R}$ est dérivable sur $]a, b[$ et continue sur $[a, b]$, il existe un point $c \in]a, b[$ tel que

$$f(b) - f(a) = f'(c)(b - a).$$

e) Le théorème de Taylor.

Si f est $n + 1$ fois dérivable dans un intervalle ouvert I contenant x_0, on peut écrire, pour $x \in I$,

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \frac{f^{(n+1)}(\xi)}{(n + 1)!}(x - x_0)^{n+1}$$

avec ξ entre x_0 et x.

f) Les fonctions convexes.

Une fonction dérivable $f : (a,b) \rightarrow \mathbb{R}$ est convexe sur (a,b) si et seulement si sa dérivée est croissante sur (a,b). Elle satisfait alors les inégalités

$$f(x_3) \leq \frac{x_2 - x_3}{x_2 - x_1} f(x_1) + \frac{x_3 - x_1}{x_2 - x_1} f(x_2)$$

quelques soient $a < x_1 < x_3 < x_2 < b$ et

$$f(x) \geq f(x_0) + f'(x_0)(x - x_0)$$

quelques soient $x, x_0 \in (a,b)$.

L'étudiant est aussi supposé connaître les concepts fondamentaux de l'algèbre linéaire : matrices, déterminants, écriture matricielle d'un système d'équations linéaires, règle de Cramer pour le résoudre, vecteurs, transformations linéaires $\mathbb{R}^n \rightarrow \mathbb{R}^m$ et théorème des axes principaux.

1.1 Exercices

Justifier ses réponses.

1. Calculer
$$\limsup_{n \to +\infty} \sin n\frac{\pi}{2} \cos n\pi \ , \ \liminf_{n \to +\infty} \sin n\frac{\pi}{2} \cos n\pi.$$

2. Déterminer, sans calculatrice, le plus grand des deux nombres π^e et e^π.

3. Montrer que
$$1 - x \leq e^{-x} \leq 1 - x + \frac{x}{e}$$
lorsque $0 < x < 1$.

2 L'ESPACE EUCLIDIEN

Un point $\mathbf{x}$ de l'espace euclidien à n dimensions $\mathbb{R}^n$ est un n-tuplet :

$$\mathbf{x} = (x_1, x_2, \ldots, x_n).$$

L'**addition** et la **multiplication scalaire** y sont définies par

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$

et

$$\lambda\,\mathbf{x} = (\lambda\,x_1, \lambda\,x_2, \ldots, \lambda\,x_n)$$

respectivement. On peut donc écrire

$$\mathbf{x} = \sum_{j=1}^{n} x_j \mathbf{e}_j^{(n)}$$

si

$$\mathbf{e}_j^{(n)} = (0, 0, \ldots, 1, \ldots, 0)$$

(le 1 occupant la $j^{i\grave{e}me}$ position). Pour utiliser l'écriture matricielle, il sera quelquefois commode d'identifier le point $\mathbf{x}$ avec son vecteur position, l'élément de $\mathbb{R}^{n\times 1}$ (la matrice $n \times 1$, le vecteur colonne) dont les entrées sont les nombres x_j. Alors $\mathbf{x}^T$ désignera l'élément de $\mathbb{R}^{1\times n}$ (la matrice $1 \times n$, le vecteur ligne) obtenu par transposition matricielle :

$$\mathbf{x}^T = (x_1\,x_2\,\ldots\,x_n).$$

(Remarquer que les entrées d'une matrice $1 \times n$ ne sont pas séparées par des virgules.)

2.1 Propriétés algébriques

Une somme

$$\sum_{k=1}^{N} \lambda_k \mathbf{x}_k$$

est une **combinaison linéaire** des points $\mathbf{x}_k$, une **combinaison affine** si $\sum_{k=1}^{N} \lambda_k = 1$ et une **combinaison convexe** si, de plus, $\lambda_k \geq 0$ pour tout k. Un ensemble $E \subseteq \mathbb{R}^n$ est un **ensemble convexe** s'il contient toute combinaison convexe de ses points.

Exemple.

La **droite** passant par $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ ($\mathbf{a} \neq \mathbf{b}$) est l'ensemble des combinaisons affines de $\mathbf{a}$ et $\mathbf{b}$:

$$\mathbf{x} = (1-\lambda)\mathbf{a} + \lambda\mathbf{b} = \mathbf{a} + \lambda(\mathbf{b}-\mathbf{a}), \quad \lambda \in \mathbb{R}.$$

Dans le cas générique où $a_k \neq b_k$ pour tout k, cela impose les $n-1$ contraintes suivantes sur les coordonnées du point $\mathbf{x}$ qui la parcourt :

$$\frac{x_1 - a_1}{b_1 - a_1} = \frac{x_2 - a_2}{b_2 - a_2} = \cdots = \frac{x_n - a_n}{b_n - a_n}.$$

Le **segment** $[\mathbf{a}, \mathbf{b}]$ est l'ensemble des combinaisons convexes de $\mathbf{a}$ et $\mathbf{b}$

$$\mathbf{x} = (1-\lambda)\mathbf{a} + \lambda\mathbf{b}, \quad 0 \leq \lambda \leq 1.$$

Exemple.

Soient $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ $m+1$ points tels que les m vecteurs

$$\mathbf{x}_1 - \mathbf{x}_0, \mathbf{x}_2 - \mathbf{x}_0, \ldots, \mathbf{x}_m - \mathbf{x}_0$$

soient linéairement indépendants. Le **polyèdre** (le polytope)

$$[\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m]$$

de **sommets** $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ est l'ensemble des combinaisons convexes de ces points. C'est un ensemble convexe. Lorsque $m = 2$, on obtient un **triangle** dont les **côtés** sont les segments $[\mathbf{x}_0, \mathbf{x}_1]$, $[\mathbf{x}_1, \mathbf{x}_2]$ et $[\mathbf{x}_2, \mathbf{x}_0]$. Lorsque $m = 3$, on obtient un **tétraèdre** dont les **faces** sont les triangles $[\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2]$, $[\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_3]$, $[\mathbf{x}_0, \mathbf{x}_2, \mathbf{x}_3]$ et $[\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3]$ et les **arêtes** sont les côtés $[\mathbf{x_0}, \mathbf{x_1}]$, $[\mathbf{x_0}, \mathbf{x_2}]$, $[\mathbf{x_0}, \mathbf{x_3}]$, $[\mathbf{x_1}, \mathbf{x_2}]$, $[\mathbf{x_1}, \mathbf{x_3}]$ et $[\mathbf{x_2}, \mathbf{x_3}]$ de ces triangles.

Exemple.

Un **pavé** (un parallélépipède rectangle) P est défini par n inégalités strictes ou larges :

$$P = (a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n)$$

(dans $\mathbb{R}$, $[a, b]$ désigne un intervalle fermé, $]a, b[$, un intervalle ouvert et (a, b), un intervalle quelconque).

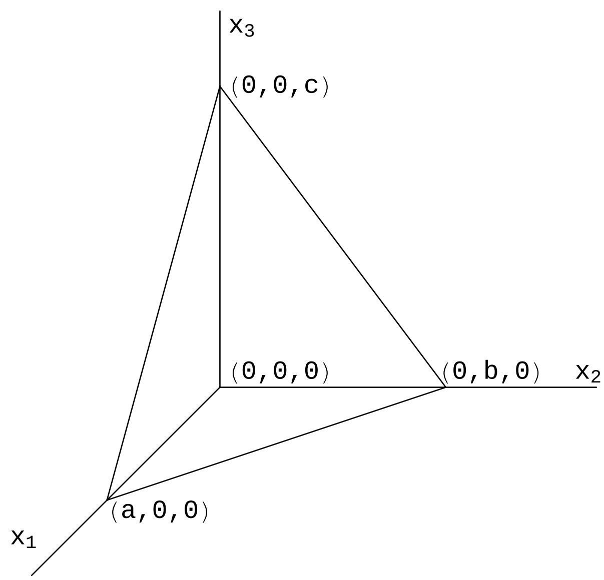

FIG. 1 – Un tétraèdre dans $\mathbb{R}^3$

Lorsque $n = 2$, il est possible définir un produit $\mathbf{x}\,\mathbf{y}$ qui prolonge à $\mathbb{R}^2$ la structure de corps qui existe sur $\mathbb{R}$. Identifions $x\,\mathbf{e}_1^{(1)} \in \mathbb{R}$ avec $x\,\mathbf{e}_1^{(2)} \in \mathbb{R}^2$. Il suffit de définir

$$\mathbf{e}_1^{(2)}\,\mathbf{e}_1^{(2)} = \mathbf{e}_1^{(2)},\ \mathbf{e}_1^{(2)}\,\mathbf{e}_2^{(2)} = \mathbf{e}_2^{(2)}\,\mathbf{e}_1^{(2)} = \mathbf{e}_2^{(2)},\ \mathbf{e}_2^{(2)}\,\mathbf{e}_2^{(2)} = -\mathbf{e}_1^{(2)}$$

et de postuler la distributivité de ce produit sur l'addition et sa commutativité avec la multiplication scalaire ; on obtient :

$$\mathbf{x}\,\mathbf{y} = (x_1y_1 - x_2y_2)\,\mathbf{e}_1^{(2)} + (x_1y_2 + x_2y_1)\,\mathbf{e}_2^{(2)}.$$

Si $x_1\,\mathbf{e}_1^{(2)} + x_2\,\mathbf{e}_2^{(2)} \neq \mathbf{0}$,

$$\frac{1}{x_1\,\mathbf{e}_1^{(2)} + x_2\,\mathbf{e}_2^{(2)}} = \frac{x_1}{x_1^2 + x_2^2}\,\mathbf{e}_1^{(2)} + \frac{-x_2}{x_1^2 + x_2^2}\,\mathbf{e}_2^{(2)}.$$

Puisque $(\mathbf{e}_2^{(2)})^2 = -1$, ce corps ne peut pas être ordonné. Il s'agit en fait du corps des nombres complexes $\mathbb{C}$.

Le produit précédent ne peut pas être prolongé à $\mathbb{R}^3$. Supposons en effet le contraire. Identifions $x_1\,\mathbf{e}_1^{(2)} + x_2\,\mathbf{e}_2^{(2)} \in \mathbb{R}^2$ avec $x_1\,\mathbf{e}_1^{(3)} + x_2\,\mathbf{e}_2^{(3)} \in \mathbb{R}^3$ et posons

$$\mathbf{e}_2^{(3)}\,\mathbf{e}_3^{(3)} = u_1\,\mathbf{e}_1^{(3)} + u_2\,\mathbf{e}_2^{(3)} + u_3\,\mathbf{e}_3^{(3)}.$$

Alors on devra avoir

$$\begin{aligned} -\mathbf{e}_3^{(3)} &= (\mathbf{e}_2^{(3)})^2\, \mathbf{e}_3^{(3)} = \mathbf{e}_2^{(3)}\, (\mathbf{e}_2^{(3)}\, \mathbf{e}_3^{(3)}) \\ &= \mathbf{e}_2^{(3)}\, (u_1 \mathbf{e}_1^{(3)} + u_2 \mathbf{e}_2^{(3)} + u_3 \mathbf{e}_3^{(3)}) \\ &= u_1\, \mathbf{e}_2^{(3)} - u_2\, \mathbf{e}_1^{(3)} + u_3\, (u_1\, \mathbf{e}_1^{(3)} + u_2\, \mathbf{e}_2^{(3)} + u_3\, \mathbf{e}_3^{(3)}) \end{aligned}$$

de telle sorte que

$$\mathbf{0} = (-u_2 + u_1 u_3)\, \mathbf{e}_1^{(3)} + (u_1 + u_2 u_3)\, \mathbf{e}_2^{(3)} + (1 + u_3^2)\, \mathbf{e}_3^{(3)}$$

et en particulier

$$1 + u_3^2 = 0$$

ce qui est absurde.

2.2 Propriétés géométriques

Le **produit scalaire** est défini par

$$\mathbf{x} \cdot \mathbf{y} = \sum_{j=1}^{n} x_j y_j = \mathbf{x}^T \mathbf{y}$$

et la **norme d'un vecteur** par

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{\sum_{j=1}^{n} x_j^2}.$$

Avec ces notations, l'**inégalité de Cauchy-Schwarz** s'écrit

$$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\|$$

avec égalité si et seulement si les vecteurs $\mathbf{x}$ et $\mathbf{y}$ sont linéairement dépendants et l'**inégalité du triangle** devient

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$

avec égalité précisément lorsque les vecteurs $\mathbf{x}$ et $\mathbf{y}$ sont des multiples positifs l'un de l'autre.

Exemple.

La **boule ouverte** de centre $\mathbf{x}_0$ et de rayon $r > 0$ est définie par une inégalité stricte

$$B(\mathbf{x}_0, r) = \{\mathbf{x} \mid \|\mathbf{x} - \mathbf{x}_0\| < r\}$$

et la **sphère** de mêmes centre et rayon est

$$S(\mathbf{x}_0, r) = \{\mathbf{x} \mid \|\mathbf{x} - \mathbf{x}_0\| = r\}.$$

Une boule est un ensemble convexe.

Observons que l'inégalité de Cauchy-Schwarz est équivalente à

$$\|\mathbf{y}\| = \sup\left\{\frac{|\mathbf{x}\cdot\mathbf{y}|}{\|\mathbf{x}\|} \mid \mathbf{x} \neq \mathbf{0}\right\} = \sup\{|\mathbf{x}\cdot\mathbf{y}| \mid \|\mathbf{x}\| = 1\}.$$

La **norme d'une transformation linéaire A** : $\mathbb{R}^n \to \mathbb{R}^m$ est définie par

$$\|\mathbf{A}\|_\infty = \sup\left\{\frac{\|\mathbf{A}(\mathbf{x})\|}{\|\mathbf{x}\|} \mid \mathbf{x} \neq \mathbf{0}\right\} = \sup\{\|\mathbf{A}(\mathbf{x})\| \mid \|\mathbf{x}\| = 1\}.$$

Si

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \in \mathbb{R}^{m\times n}$$

est sa matrice relativement aux bases canoniques, c'est-à-dire si

$$a_{i,j} = \mathbf{e}_i^{(m)} \cdot \mathbf{A}(\mathbf{e}_j^{(n)})$$

donc

$$\mathbf{A}(\mathbf{x}) = \mathbf{A}\left(\sum_{j=1}^{n} x_j \mathbf{e}_j^{(n)}\right) = \sum_{j=1}^{n} x_j \mathbf{A}(\mathbf{e}_j^{(n)})$$

$$= \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{i,j}\mathbf{e}_i^{(m)} = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{i,j}x_j\right)\mathbf{e}_i^{(m)},$$

on a

$$\|\mathbf{A}\|_\infty \leq \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{i,j}^2} = \|\mathbf{A}\|.$$

Cela suit en effet de l'inégalité de Cauchy-Schwarz :

$$\|\mathbf{A}(\mathbf{x})\|^2 = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{i,j} x_j \right)^2 \leq \sum_{i=1}^{m} \sum_{j=1}^{n} a_{i,j}^2 \sum_{j=1}^{n} x_j^2 = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{i,j}^2 \right) \|\mathbf{x}\|^2.$$

On a donc

$$\|\mathbf{A}(\mathbf{x})\| \leq \|\mathbf{A}\|_\infty \|\mathbf{x}\| \leq \|\mathbf{A}\| \, \|\mathbf{x}\|.$$

Le nombre $\|\mathbf{A}\|$ est bien sûr plus facile à calculer que le nombre $\|\mathbf{A}\|_\infty$ mais ce dernier est indépendant de la base choisie pour représenter $\mathbf{A}$ et se généralise plus facilement au cas de transformations linéaires entre espaces de dimension infinie. (Dans ce cours, nous n'utilisons que la base canonique et nous identifions, quand cela est commode, la transformation linéaire avec sa matrice).

Exemple.
Si la matrice $\mathbf{A}$ de l'opérateur linéaire $\mathbf{A} : \mathbb{R}^n \to \mathbb{R}^n$ est diagonale,

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & 0 & \cdots & 0 \\ 0 & a_{2,2} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & a_{n,n} \end{pmatrix},$$

on a

$$\|\mathbf{A}\| = \sqrt{\sum_{j=1}^{n} a_{j,j}^2}$$

et

$$\|\mathbf{A}\|_\infty = \sup\{|a_{j,j}| \mid 1 \leq j \leq n\}.$$

Théorème 1 *Soient* $\mathbf{L}, \mathbf{M} : \mathbb{R}^n \to \mathbb{R}^m$ *et* $\mathbf{N} : \mathbb{R}^m \to \mathbb{R}^p$ *des transformations linéaires et* $\lambda \in \mathbb{R}$ *un nombre. Alors*

1. $\|\lambda \mathbf{L}\|_\infty = |\lambda| \, \|\mathbf{L}\|_\infty$ *;*

2. $\|\mathbf{L} + \mathbf{M}\|_\infty \leq \|\mathbf{L}\|_\infty + \|\mathbf{M}\|_\infty$ *;*

3. $\|\mathbf{N} \circ \mathbf{L}\|_\infty \leq \|\mathbf{N}\|_\infty \, \|\mathbf{L}\|_\infty$.

Démonstration.

1. On a

$$\begin{aligned}\|\lambda\,\mathbf{L}\|_\infty &= \sup\{\|\lambda\,\mathbf{L}(\mathbf{x})\| \mid \|\mathbf{x}\| = 1\} = \sup\{|\lambda|\,\|\mathbf{L}(\mathbf{x})\| \mid \|\mathbf{x}\| = 1\}\\ &= |\lambda| \sup\{\|\mathbf{L}(\mathbf{x})\| \mid \|\mathbf{x}\| = 1\} = |\lambda|\,\|\mathbf{L}\|_\infty.\end{aligned}$$

2. On a

$$\begin{aligned}\|\mathbf{L}+\mathbf{M}\|_\infty &= \sup\{\|\mathbf{L}(\mathbf{x})+\mathbf{M}(\mathbf{x})\| \mid \|\mathbf{x}\| = 1\} \leq \sup\{\|\mathbf{L}(\mathbf{x})\| + \|\mathbf{M}(\mathbf{x})\| \mid \|\mathbf{x}\| = 1\}\\ &\leq \sup\{\|\mathbf{L}(\mathbf{x})\| \mid \|\mathbf{x}\| = 1\} + \sup\{\|\mathbf{M}(\mathbf{x})\| \mid \|\mathbf{x}\| = 1\} = \|\mathbf{L}\|_\infty + \|\mathbf{M}\|_\infty.\end{aligned}$$

3. On a

$$\begin{aligned}\|\mathbf{N}\circ\mathbf{L}\|_\infty &= \sup\{\|\mathbf{N}(\mathbf{L}(\mathbf{x}))\| \mid \|\mathbf{x}\| = 1\} \leq \sup\{\|\mathbf{N}\|_\infty\,\|\mathbf{L}(\mathbf{x})\| \mid \|\mathbf{x}\| = 1\}\\ &= \|\mathbf{N}\|_\infty \sup\{\|\mathbf{L}(\mathbf{x})\| \mid \|\mathbf{x}\| = 1\} = \|\mathbf{N}\|_\infty\,\|\mathbf{L}\|_\infty.\end{aligned}$$

C.Q.F.D.

L'**angle** formé par les vecteurs $\mathbf{x}$ et $\mathbf{y}$ est défini par

$$\angle(\mathbf{x},\mathbf{y}) = \arccos\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}.$$

On a donc $0 \leq \angle(\mathbf{x},\mathbf{y}) \leq \pi$, l'une de ces inégalités ne devenant une égalité que si les vecteurs $\mathbf{x}$ et $\mathbf{y}$ sont linéairement dépendants. Les vecteurs $\mathbf{x}$ et $\mathbf{y}$ sont **orthogonaux** lorsque

$$\mathbf{x}\cdot\mathbf{y} = 0.$$

Exemple.

L'**hyperplan** H passant par $\mathbf{a}$ et orthogonal à la direction déterminée par le vecteur $\mathbf{n}$ (la direction **normale**) est

$$H = \{\mathbf{x} \mid \mathbf{n}\cdot(\mathbf{x}-\mathbf{a}) = 0\}.$$

Il induit deux **demi-espaces ouverts**

$$H_+ = \{\mathbf{x} \mid \mathbf{n}\cdot(\mathbf{x}-\mathbf{a}) > 0\}$$

et

$$H_- = \{\mathbf{x} \mid \mathbf{n}\cdot(\mathbf{x}-\mathbf{a}) < 0\}.$$

Lorsque $n = 3$, le **produit vectoriel** est défini par

$$\mathbf{x}\times\mathbf{y} = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1).$$

On a

$$\mathbf{x} \times \mathbf{y} \cdot \mathbf{x} = \mathbf{x} \times \mathbf{y} \cdot \mathbf{y} = 0.$$

Si $\mathbf{x}_1$, $\mathbf{x}_2$ et $\mathbf{x}_3$ sont trois points du plan H tels que les vecteurs déplacements $\mathbf{x}_2 - \mathbf{x}_1$ et $\mathbf{x}_3 - \mathbf{x}_1$ sont linéairement indépendants, H peut être décrit par l'équation

$$(\mathbf{x}_2 - \mathbf{x}_1) \times (\mathbf{x}_3 - \mathbf{x}_1) \cdot (\mathbf{x} - \mathbf{x}_1) = 0.$$

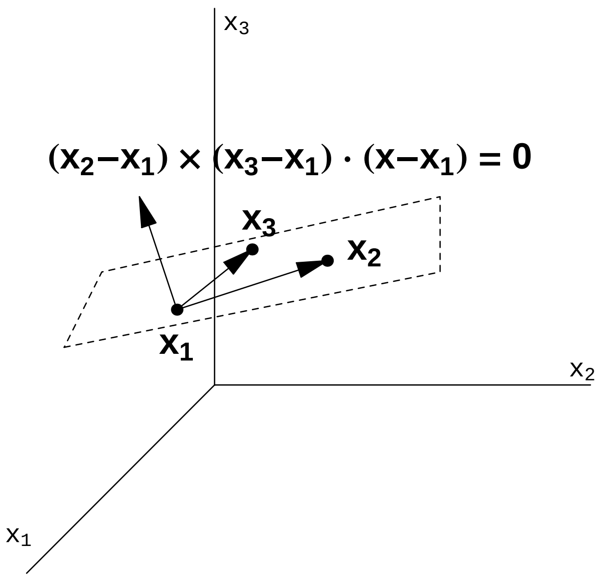

FIG. 2 – Un plan dans $\mathbb{R}^3$

2.3 Propriétés topologiques

La **distance** entre $\mathbf{x}$ et $\mathbf{y}$ est $\|\mathbf{x} - \mathbf{y}\|$. Elle satisfait l'inégalité

$$\|\mathbf{x} - \mathbf{y}\| \leq \|\mathbf{x} - \mathbf{z}\| + \|\mathbf{z} - \mathbf{y}\|$$

pour tout $\mathbf{z}$.

Un ensemble $E \subseteq \mathbb{R}^n$ est **ouvert** si à chaque point $\mathbf{x}_0 \in E$ correspond $r > 0$ tel que

$$B(\mathbf{x}_0, r) \subseteq E.$$

Exemple.

Une boule ouverte $B(\mathbf{a}, R)$ est ouverte : si $\mathbf{x}_0 \in B(\mathbf{a}, R)$, soit $\|\mathbf{x}_0 - \mathbf{a}\| = \rho < R$. Alors si $\|\mathbf{x} - \mathbf{x}_0\| < R - \rho$, $\mathbf{x} \in B(\mathbf{a}, R)$ car

$$\|\mathbf{x} - \mathbf{a}\| \leq \|\mathbf{x} - \mathbf{x}_0\| + \|\mathbf{x}_0 - \mathbf{a}\| < R - \rho + \rho = R.$$

Exemple.

Un demi-espace ouvert

$$H_+ = \{\mathbf{x} \mid \mathbf{n} \cdot (\mathbf{x} - \mathbf{a}) > 0\}$$

est ouvert. Si $\mathbf{x}_0 \in H_+$, soit $\mathbf{n} \cdot (\mathbf{x}_0 - \mathbf{a}) = \delta > 0$. Alors si $\|\mathbf{x} - \mathbf{x}_0\| < \delta/\|\mathbf{n}\|$, $\mathbf{x} \in H_+$ car

$$\mathbf{n} \cdot (\mathbf{x} - \mathbf{a}) = \mathbf{n} \cdot (\mathbf{x} - \mathbf{x}_0) + \mathbf{n} \cdot (\mathbf{x}_0 - \mathbf{a}) \geq -\|\mathbf{n}\| \, \|\mathbf{x} - \mathbf{x}_0\| + \delta > 0.$$

Toute réunion, toute intersection finie d'ensembles ouverts est encore un ensemble ouvert.

Exemple.

Dans $\mathbb{R}$, les ensembles ouverts sont précisément les ensembles qui peuvent s'écrire comme une réunion finie ou dénombrable d'intervalles ouverts disjoints.

Un ensemble $E \subseteq \mathbb{R}^n$ est **fermé** si son complémentaire $E^c = \mathbb{R}^n \setminus E$ est ouvert.

Exemple.

Un hyperplan H est fermé puisque son complémentaire $H_+ \cup H_-$ est ouvert.

Exemple.

Un pavé P défini par des inégalités larges,

$$P = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n],$$

est fermé. Si $\mathbf{x}_0 \notin P$, il faut que, par exemple, on ait $x_{0,1} - b_1 = \delta > 0$. Alors si $\|\mathbf{x} - \mathbf{x}_0\| < \delta$, on a $(x_1 - x_{0,1})^2 < \delta^2$ donc

$$x_1 > x_{0,1} - \delta = b_1$$

et $\mathbf{x} \in P$.

Il faut remarquer qu'un ensemble n'est pas nécessairement ouvert ou fermé et que, à l'opposé, l'espace $\mathbb{R}^n$ tout entier est à la fois ouvert et fermé.

Un ensemble $E \subseteq \mathbb{R}^n$ est **borné** s'il est existe $R > 0$ tel que

$$E \subseteq B(\mathbf{0}, R).$$

Exemple.
Un polyèdre et un pavé sont bornés, un hyperplan ne l'est pas.

Un ensemble $E \subseteq \mathbb{R}^n$ est **compact** s'il est fermé et borné.

Exemple.
Un pavé fermé est compact.

La suite $\{\mathbf{x}_k\}_{k\in\mathbb{N}}$ **converge** vers $\mathbf{a}$ si

$$\lim_{k\to+\infty} \|\mathbf{x}_k - \mathbf{a}\| = 0.$$

Les inégalités

$$\sup\{|x_{k,j} - a_j| \mid 1 \leq j \leq n\} \leq \|\mathbf{x}_k - \mathbf{a}\| \leq \sum_{j=1}^{n} |x_{k,j} - a_j|$$

montrent que

$$\lim_{k\to+\infty} \mathbf{x}_k = \mathbf{a}$$

si et seulement si

$$\lim_{k\to+\infty} x_{k,j} = a_j \quad \text{pour} \quad 1 \leq j \leq n.$$

En conséquence, le **critère de convergence de Cauchy** est valable : la suite $\{\mathbf{x}_k\}_{k\in\mathbb{N}}$ admet une limite si et seulement si

$$\lim_{k,p\to+\infty} \|\mathbf{x}_k - \mathbf{x}_p\| = 0.$$

Et, corollaire immédiat, toute série **normalement convergente**, c'est-à-dire telle que

$$\sum_{k=0}^{+\infty} \|\mathbf{x}_k\| < +\infty,$$

est convergente :

$$\left|\sum_{k=N+1}^{M} \mathbf{x}_k\right| \leq \sum_{k=N+1}^{M} \|\mathbf{x}_k\|.$$

Théorème 2 *Soit $E \subseteq \mathbb{R}^n$. Alors E est fermé si et seulement si la limite de toute suite convergente $\{\mathbf{x}_k\}_{k\in\mathbb{N}}$ de points de E est dans E.*

Démonstration.

La condition est nécessaire. Supposons E est fermé et soit

$$\mathbf{a} = \lim_{k\to+\infty} \mathbf{x}_k \ , \mathbf{x}_k \in E.$$

Si $\mathbf{a}$ appartenait à l'ensemble ouvert E^c, on pourrait trouver $r > 0$ tel que

$$B(\mathbf{a}, r) \subseteq E^c.$$

Or cela est impossible puisqu'il existe un indice k_r tel que $k > k_r$ implique

$$\|\mathbf{x}_k - \mathbf{a}\| < r.$$

Donc $\mathbf{a} \in E$.

La condition est suffisante. Supposons que E n'est pas fermé. L'ensemble E^c n'étant pas ouvert, on pourrait trouver un point $\mathbf{a} \in E^c$ tel que toute boule ouverte centrée en $\mathbf{a}$, $B(\mathbf{a}, r)$, coupe E. Soit donc

$$\mathbf{x}_k \in E \cap B(\mathbf{a}, 1/k) \ \text{ pour tout } \ k \in \mathbb{N}.$$

Alors, par l'hypothèse, on devrait avoir $\mathbf{a} = \lim_{k\to+\infty} \mathbf{x}_k \in E$ ce qui est absurde. C.Q.F.D.

Exemple.

Le **cône positif fermé** défini par

$$\mathbb{R}^n_+ = \{\mathbf{x} \mid x_j \geq 0 \ , \ 1 \leq j \leq n\}$$

est un ensemble fermé.

Une **boule fermée**

$$\overline{B}(\mathbf{x}_0, r) = \{\mathbf{x} \mid \|\mathbf{x} - \mathbf{x}_0\| \leq r\}$$

et une sphère

$$S(\mathbf{x}_0, r) = \{\mathbf{x} \mid \|\mathbf{x} - \mathbf{x}_0\| = r\}$$

sont des ensembles fermés. Ces deux derniers ensembles sont donc compacts.

Théorème 3 *Soit $E \subseteq \mathbb{R}^n$. Alors E est compact si et seulement si toute suite $\{\mathbf{x}_k\}_{k\in\mathbb{N}}$ de points de E contient une suite partielle $\{\mathbf{x}_{k_j}\}_{j\in\mathbb{N}}$ qui converge vers un point de E.*

Démonstration.

La condition est nécessaire. Supposons que E est compact. Les n suites numériques $\{x_{k,j}\}_{k\in\mathbb{N}}$ sont alors toutes bornées. On peut donc de la suite donnée extraire une suite partielle telle que les premières coordonnées $x_{k,1}$ des points qui la composent admettent une limite a_1. De cette suite partielle, on peut en extraire une autre telle que les deuxièmes coordonnées des points qui la composent admettent aussi une limite a_2. Ainsi de suite. Après n étapes, on obtient une suite partielle $\{\mathbf{x}_{k_p}\}_{p\in\mathbb{N}}$ de la suite originale qui converge vers $\mathbf{a} = (a_1, a_2, \ldots, a_n)$. L'ensemble E étant fermé, $\mathbf{a} \in E$.

La condition est suffisante. E est fermé puisque si

$$\mathbf{a} = \lim_{k\to+\infty} \mathbf{x}_k,$$

toute les suites partielles possibles de la suite $\{\mathbf{x}_k\}_{k\in\mathbb{N}}$ convergent vers $\mathbf{a}$ qui doit donc appartenir à E. E est borné. S'il ne l'était pas, on pourrait trouver des points $\mathbf{x}_k \in E$ tels que

$$\|\mathbf{x_{k+1}}\| > \|\mathbf{x}_k\| + 1$$

et, toute suite convergente étant bornée, cette suite n'admettrait aucune suite partielle convergente, contrairement à l'hypothèse. C.Q.F.D.

Exemple.

Un polyèdre

$$P = [\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_m]$$

est compact. Soit en effet

$$\mathbf{y}_k = \sum_{i=0}^{m} \lambda_{k,i}\, \mathbf{x}_i \;,\; k \in \mathbb{N}$$

une suite de points de P. Les points

$$\boldsymbol{\lambda}_k = (\lambda_{k,0}, \lambda_{k,1}, \ldots, \lambda_{k,m}) \in \mathbb{R}^{m+1}$$

appartiennent au sous-ensemble compact

$$E = [0,1]^{m+1} \cap \{\boldsymbol{\lambda} \mid \lambda_0 + \lambda_1 + \ldots + \lambda_m = 1\}$$

de $\mathbb{R}^{m+1}$. On peut donc en extraire une suite partielle convergeant vers un point de E, soit

$$\overline{\boldsymbol{\lambda}} = \lim_{j\to+\infty} \boldsymbol{\lambda}_{k_j}.$$

Alors

$$\mathbf{y} = \sum_{i=0}^{m} \overline{\lambda}_i \, \mathbf{x}_i \in P$$

et

$$\|\mathbf{y}_{k_j} - \mathbf{y}\| \leq \sum_{i=0}^{m} |\lambda_{k_j,i} - \overline{\lambda}_i| \, \|\mathbf{x}_i\|$$

donc

$$\mathbf{y} = \lim_{j \to +\infty} \mathbf{y}_{k_j}.$$

L'**intérieur** $\overset{\circ}{E}$ d'un ensemble $E \subseteq \mathbb{R}^n$ est la réunion de tous les ensembles ouverts contenus dans E — il peut être vide. C'est donc le plus grand ensemble ouvert contenu dans E.

L'**adhérence** $\overline{E}$ d'un ensemble $E \subseteq \mathbb{R}^n$ est l'intersection de tous les ensembles fermés qui contiennent E. C'est donc le plus petit ensemble fermé qui contienne E.

La **frontière** ∂E d'un ensemble $E \subseteq \mathbb{R}^n$ est définie par

$$\partial E = \overline{E} \cap \overline{E^c}.$$

Théorème 4 *Un point* $\mathbf{x}$ *appartient à* $\overline{E}$ *si et seulement si il est la limite d'une suite de points de* E*.*

Démonstration.

La condition est nécessaire. S'il existait une boule ouverte $B(\mathbf{x}, r) \subseteq E^c$, le complémentaire de cette boule serait un ensemble fermé contenant E et $\mathbf{x}$ ne serait pas dans l'intersection de tels fermés, c'est-à-dire dans $\overline{E}$. Les boules $B(\mathbf{x}, 1/k)$ contenant donc chacune au moins un point de E, $\mathbf{x}$ est la limite d'une suite de points de E.

La condition est suffisante. Si $\mathbf{x}$ est la limite d'une suite de points de E et F est un ensemble fermé contenant E, $\mathbf{x}$ est la limite d'une suite de points de F et, cet ensemble étant fermé, $\mathbf{x} \in F$. F étant arbitraire, $\mathbf{x} \in \overline{E}$. C.Q.F.D.

2.4 Exercices

Justifier ses réponses.

1. Les nombres λ_i dans la représentation

$$\mathbf{x} = \sum_{i=0}^{m} \lambda_i \mathbf{x}_i$$

d'un point du polyèdre $[\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_m]$ sont ses **coordonnées barycentriques** et le **barycentre** du polyèdre est le point dont toutes les coordonnées barycentriques sont égales.

– Montrer que les coordonnées barycentriques sont uniques (on suppose les vecteurs $\mathbf{x}_1-\mathbf{x}_0, \mathbf{x}_2-\mathbf{x}_0, \dots, \mathbf{x}_m-\mathbf{x}_0$ linéairement indépendants).

– Montrer que le barycentre d'un triangle coïncide avec l'intersection de ses médianes (les droites joignant les sommets aux milieux des côtés opposés).

2. Montrer qu'un ensemble est convexe si et seulement si il contient tous les segments admettant deux de ses points pour extrémités.

En déduire qu'une fonction $f : \mathbb{R} \to \mathbb{R}$ est convexe si et seulement si son **épigraphe**,

$$E_f = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 \geq f(x_1)\},$$

est convexe.

3. Soient

$$\|\mathbf{x}\|_1 = \sum_{j=1}^{n} |x_j|$$

et

$$B_1(\mathbf{a}, r) = \{\mathbf{x} \mid \|\mathbf{x} - \mathbf{a}\|_1 < r\}.$$

– Montrer que

$$\|\mathbf{x} + \mathbf{y}\|_1 \leq \|\mathbf{x}\|_1 + \|\mathbf{y}\|_1.$$

– Montrer que

$$\lim_{k \to +\infty} \|\mathbf{x_k} - \mathbf{x}\| = 0 \Leftrightarrow \lim_{k \to +\infty} \|\mathbf{x_k} - \mathbf{x}\|_1 = 0.$$

– Montrer que $B_1(\mathbf{a}, r)$ est convexe.

4. Mêmes questions pour
$$\|\mathbf{x}\|_\infty = \sup\{|x_j| \mid 1 \leq j \leq n\}$$
et
$$B_\infty(\mathbf{a}, r) = \{\mathbf{x} \mid \|\mathbf{x} - \mathbf{a}\|_\infty < r\}.$$
5. La distance entre le point $\mathbf{x}_0$ et l'ensemble E est
$$d(\mathbf{x}_0, E) = \inf\{\|\mathbf{x}_0 - \mathbf{x}\| \mid \mathbf{x} \in E\}.$$
Utiliser l'inégalité de Cauchy-Schwarz pour déterminer la distance entre le point $\mathbf{x}_0$ et l'hyperplan H d'équation $\mathbf{n} \cdot (\mathbf{x} - \mathbf{a}) = 0$ ainsi que le point $\mathbf{x}_m \in H$ où elle est atteinte.
6. Soit $\mathbf{L} : \mathbb{R}^n \to \mathbb{R}^n$ un opérateur linéaire inversible. Montrer que
$$\frac{1}{\|\mathbf{L}^{-1}\|_\infty} \leq \frac{\|\mathbf{L}\mathbf{x}\|}{\|\mathbf{x}\|} \leq \|\mathbf{L}\|_\infty.$$
7. Soit $\mathbf{A} : \mathbb{R}^3 \to \mathbb{R}^3$ l'opérateur linéaire dont la matrice relativement à la base canonique est
$$\mathbf{A} = \begin{pmatrix} \cos\theta & -\cos\phi\,\sin\theta & \sin\phi\,\cos\theta \\ \sin\theta & \cos\phi\,\cos\theta & -\sin\phi\,\sin\theta \\ 0 & \sin\phi & \cos\phi \end{pmatrix}.$$
Calculer $\|\mathbf{A}\|_\infty$ et $\|\mathbf{A}\|$.
8. Soient $\lambda \in \mathbb{R}$, $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$ et $\mathbf{C} \in \mathbb{R}^{p\times m}$. Montrer que
   –
   $$\|\lambda\mathbf{A}\| = |\lambda|\,\|\mathbf{A}\|.$$
   –
   $$\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|.$$
   –
   $$\|\mathbf{C}\mathbf{A}\| \leq \|\mathbf{C}\|\,\|\mathbf{A}\|.$$
9. Soit $\mathbf{A} : \mathbb{R}^n \to \mathbb{R}^m$ un opérateur linéaire. Montrer que
   –
   $$\|\mathbf{A}\|_\infty \leq \|\mathbf{A}\| \leq \sqrt{m \times n}\,\|\mathbf{A}\|_\infty.$$
   En déduire que, si $\mathbf{A}_k : \mathbb{R}^n \to \mathbb{R}^m$ est une suite d'opérateur linéaire, on a
   –
   $$\lim_{k\to+\infty} \|\mathbf{A}_k - \mathbf{A}\|_\infty = 0 \Leftrightarrow \lim_{k\to+\infty} \|\mathbf{A}_k - \mathbf{A}\| = 0.$$

10. Vrai ou faux ?
    - Toute réunion d'ensembles convexes est convexe.
    - Toute intersection d'ensembles convexes est convexe.
11. Mêmes questions en remplaçant « convexe » par « ouvert », par « fermé », par « borné », par « compact ».
12. Montrer que le pavé $[0,1[^n \subseteq \mathbb{R}^n$ n'est ni fermé, ni ouvert.
13. Déterminer l'intérieur de chacun des ensembles suivants : la boule $\{\mathbf{x} \mid \|\mathbf{x}-\mathbf{a}\| < 1\}$, l'hyperplan $\{\mathbf{x} \mid \mathbf{n}\cdot(\mathbf{x}-\mathbf{a}) = 0\}$ et le pavé $[0,1[^n$.
14. Déterminer l'adhérence de chacun des ensembles suivants : la boule $\{\mathbf{x} \mid \|\mathbf{x}-\mathbf{a}\| \leq 1\}$, le demi-espace $\{\mathbf{x} \mid \mathbf{n}\cdot(\mathbf{x}-\mathbf{a}) > 0\}$ et le pavé $[0,1[^n$.
15. Déterminer la frontière de chacun des ensembles suivants : la boule pointée $\{\mathbf{x} \mid 0 < \|\mathbf{x}-\mathbf{a}\| < 1\}$, l'hyperplan $\{\mathbf{x} \mid \mathbf{n}\cdot(\mathbf{x}-\mathbf{a}) = 0\}$, $\mathbb{Q}^n$ et le cône positif fermé $\{\mathbf{x} \mid x_j \geq 0\ ,\ 1 \leq j \leq n\}$.
16. Vrai ou faux ?
    - $E = \mathring{\overline{E}}$
    - $E = \overline{\mathring{E}}$
    - $E = \mathring{E} \cup \partial E$.
17. On considère la suite des points $\mathbf{x}_k$ de $\mathbb{R}^2$ définie par

$$x_{k+1,1} = \sqrt{x_{k,1}\, x_{k,2}}\,,\ x_{k+1,2} = \frac{x_{k,1} + x_{k,2}}{2}.$$

    Montrer que, quels que soient $x_{0,1} > 0$ et $x_{0,2} > 0$, elle converge et que sa limite est un point situé sur la droite $x_2 = x_1$.
18. Déterminer les valeurs de $\mathbf{x} \in \mathbb{R}^n$ pour lesquelles la série

$$\sum_{k=1}^{+\infty} \frac{\mathbf{x}}{\|\mathbf{x}\|^k}$$

    converge et calculer sa somme.
19. Soit $\mathbf{A} \in \mathbb{R}^{n\times n}$. Montrer que la série

$$\sum_{k=0}^{+\infty} \mathbf{A}^k$$

    converge si $\|\mathbf{A}\| < 1$ et calculer sa somme.

20. On considère une suite décroissante d'ensembles compacts non vides :

$$E_1 \supseteq E_2 \supseteq E_3 \supseteq \cdots$$

Montrer que l'intersection

$$\bigcap_{k \geq 1} E_k$$

est non vide. La conclusion tient-elle si l'on remplace « compacts » par « fermés » ?

3 FONCTIONS NUMÉRIQUES CONTINUES

Les fonctions continues $\mathbb{R}^n \to \mathbb{R}$ jouissent de propriétés analogues à celles des fonctions continues $\mathbb{R} \to \mathbb{R}$.

3.1 Définition

Théorème 5 *Soient $E \subseteq \mathbb{R}^n$ un ensemble, $\mathbf{x}_0 \in E$ un de ses points et $f : E \to \mathbb{R}$ une fonction numérique définie sur E. Les énoncés suivants sont équivalents :*

1. *Pour toute suite $\{\mathbf{x}_k\}_{k\in\mathbb{N}}$ de points de E distincts de $\mathbf{x}_0$, on a*

$$\lim_{k\to+\infty} \mathbf{x}_k = \mathbf{x}_0 \quad \textit{implique} \quad \lim_{k\to+\infty} f(\mathbf{x}_k) = L;$$

2. *à chaque $\epsilon > 0$ correspond $\delta > 0$ tels que pour tout $\mathbf{x} \in E$,*

$$0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta \quad \textit{implique} \quad |f(\mathbf{x}) - L| < \epsilon.$$

Démonstration.

La première condition entraîne la seconde. Si cette dernière était fausse en effet, on pourrait trouver $\epsilon > 0$ tel que quel que soit $\delta > 0$, on ait

$$|f(\mathbf{x}_\delta) - L| \geq \epsilon$$

pour au moins un point $\mathbf{x}_\delta$ de E tel que

$$0 < \|\mathbf{x}_\delta - \mathbf{x}_0\| < \delta.$$

Les points associés à 1, $1/2$, $1/3$, $\ldots$ formeraient une suite convergeant vers $\mathbf{x}_0$ sans que leurs images par f ne convergent vers L et la première condition serait violée.

La seconde condition entraîne la première. Si la suite $\{\mathbf{x}_k\}_{k\in\mathbb{N}}$ de points de E distincts de $\mathbf{x}_0$ converge vers $\mathbf{x}_0$, il existe un indice k_δ à partir duquel

$$0 < \|\mathbf{x}_k - \mathbf{x}_0\| < \delta$$

donc à partir duquel

$$|f(\mathbf{x}_k) - L| < \epsilon$$

et

$$\lim_{k\to+\infty} f(\mathbf{x}_k) = L.$$

C.Q.F.D.

Lorsque les conditions du théorème précédent sont vérifiées, on écrit :

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = L$$

(lire : f tend vers L lorsque $\mathbf{x}$ tend vers $\mathbf{x}_0$).

La fonction $f : E \to \mathbb{R}$ est **continue** en $\mathbf{x}_0 \in E$ si

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0).$$

Elle est continue sur E si elle est continue en chaque en chaque point de E. L'addition, la soustraction, la multiplication, la division et la composition de fonctions continues donnent des fonctions continues. Le **graphe** d'une fonction continue est l'**hypersurface**

$$G_f = \{\mathbf{x} \in \mathbb{R}^{n+1} \mid x_{n+1} = f(x_1, x_2, \ldots, x_n)\}.$$

Exemple.
La norme $f(\mathbf{x}) = \|\mathbf{x}\|$ est continue sur $\mathbb{R}^n$ puisque

$$|\|\mathbf{x}\| - \|\mathbf{x}_0\|| \leq \|\mathbf{x} - \mathbf{x}_0\|.$$

Exemple.
Une **fonction linéaire** $f(\mathbf{x}) = \mathbf{a}^T\mathbf{x}$ est continue sur $\mathbb{R}^n$ puisque

$$|\mathbf{a}^T\mathbf{x} - \mathbf{a}^T\mathbf{x}_0| \leq \|\mathbf{a}\|\|\mathbf{x} - \mathbf{x}_0\|.$$

Son graphe est un hyperplan.

Exemple.
Une **fonction quadratique** $f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$, *où l'on peut toujours supposer que* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *est symétrique*, est continue sur $\mathbb{R}^n$ puisque

$$|\mathbf{x}^T\mathbf{A}\mathbf{x}-\mathbf{x}_0^T\mathbf{A}\mathbf{x}_0| = |\mathbf{x}^T\mathbf{A}(\mathbf{x}-\mathbf{x}_0)+\mathbf{x}_0^T\mathbf{A}(\mathbf{x}-\mathbf{x}_0)| \leq \|\mathbf{A}\|(\|\mathbf{x}\|+\|\mathbf{x}_0\|)\|\mathbf{x}-\mathbf{x}_0\|.$$

Exemple.
Les **projections** $f_j(\mathbf{x}) = x_j$ sont continues sur $\mathbb{R}^n$ puisque

$$|x_j - x_{0,j}| \leq \|\mathbf{x} - \mathbf{x}_0\|.$$

Un **polynôme de degré** N,

$$P_N(\mathbf{x}) = \sum_{\|\boldsymbol{\alpha}\|_1 \le N} a_{\boldsymbol{\alpha}}\, \mathbf{x}^{\boldsymbol{\alpha}} = \sum_{\alpha_1+\alpha_2+\cdots+\alpha_n \le N} a_{\alpha_1\alpha_2\cdots\alpha_n}\, x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$$

($\boldsymbol{\alpha}$ est un indice multiple : les α_j sont des entiers positifs), est continu sur $\mathbb{R}^n$. Une **fonction rationnelle**,

$$R(\mathbf{x}) = \frac{P_N(\mathbf{x})}{Q_M(\mathbf{x})},$$

est continue sur son domaine de définition, l'ensemble

$$\{\mathbf{x} \mid Q_M(\mathbf{x}) \neq 0\}.$$

Exemple.
Des **fonctions transcendantes** telles

$$e^{\mathbf{x}^T \mathbf{A} \mathbf{x}},\ \arctan \mathbf{a}^T \mathbf{x} \text{ ou } \log \|\mathbf{x}\|$$

sont continues (sur leur domaine de définition).

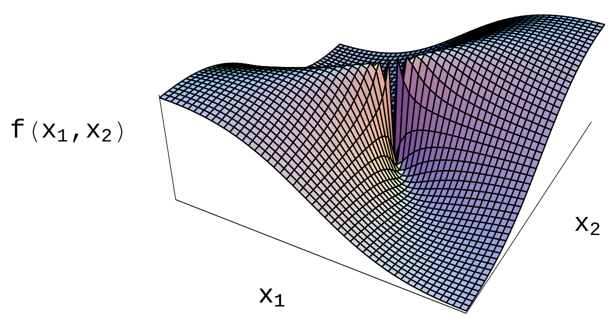

Fig. 3 – Une discontinuité à l'origine

Exemple.
La fonction $f : \mathbb{R}^2 \setminus \mathbf{0} \to \mathbb{R}$ définie par

$$f(x_1, x_2) = \frac{x_1 x_2}{x_1^2 + x_2^2},$$

ne peut pas être prolongée à une fonction continue sur $\mathbb{R}^2$ puisque

$$\lim_{\mathbf{x} \to \mathbf{0}} f(\mathbf{x})$$

n'existe pas :

$$f(x_1, \lambda x_1) = \frac{\lambda}{1+\lambda^2}.$$

Exemple.

La fonction $f : \mathbb{R}^2 \setminus \mathbf{0} \to \mathbb{R}$ définie par

$$f(x_1, x_2) = \begin{cases} \arctan \frac{x_2}{x_1} & \text{si } x_1 > 0 \\ \frac{\pi}{2} & \text{si } x_1 = 0,\, x_2 > 0 \\ \arctan \frac{x_2}{x_1} + \pi & \text{si } x_1 < 0,\, x_2 \geq 0 \\ -\frac{\pi}{2} & \text{si } x_1 = 0,\, x_2 < 0 \\ \arctan \frac{x_2}{x_1} - \pi & \text{si } x_1 < 0,\, x_2 < 0 \end{cases}$$

est discontinue sur $]-\infty, 0[$ puisque

$$\lim_{x_2 \to 0-} f(x_1, x_2) \neq f(x_1, 0)$$

quelque soit $x_1 < 0$.

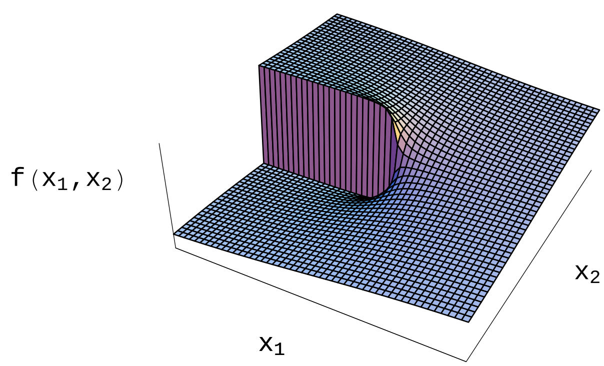

FIG. 4 – Une discontinuité le long d'un rayon

3.2 Propriétés

Théorème 6 *Soient $E \subseteq \mathbb{R}^n$ un ensemble ouvert et $f : E \to \mathbb{R}$ une fonction continue sur E. Alors l'image inverse d'un ensemble ouvert $O \subseteq \mathbb{R}^n$ par f est un ensemble ouvert.*

Démonstration.

Soit $\mathbf{x}_0 \in f^{-1}(O)$. Puisque O est ouvert, il existe $\epsilon > 0$ tel que

$$B(f(\mathbf{x}_0), \epsilon) \subseteq O.$$

Puisque que E est ouvert et que f est continue, il existe $\delta > 0$ tel que $B(\mathbf{x}_0, \delta) \subseteq E$ et que

$$B(\mathbf{x}_0, \delta) \subseteq f^{-1}(B(f(\mathbf{x}_0), \epsilon)) \subseteq f^{-1}(O).$$

C.Q.F.D.

Un ensemble $E \subseteq \mathbb{R}^n$ est **connexe** si on ne peut *pas* le représenter sous la forme

$$E = O_1 E \cup O_2 E,$$

les ensembles O_1 et O_2 étant des ouverts tels que $O_1 E \neq \emptyset$, $O_2 E \neq \emptyset$ et $O_1 O_2 E = \emptyset$.

Exemple.

Tout intervalle est un sous-ensemble connexe de $\mathbb{R}$.

Théorème 7 *Soient $E \subseteq \mathbb{R}^n$ un ensemble ouvert connexe et $f : E \to \mathbb{R}$ une fonction continue sur E. Alors $f(E)$ est un ensemble connexe.*

Démonstration.

Supposons au contraire que

$$f(E) = O_1 f(E) \cup O_2 f(E),$$

les ensembles O_1 et O_2 étant des ouverts tels que $O_1 f(E) \neq \emptyset$, $O_2 f(E) \neq \emptyset$ et $O_1 O_2 f(E) = \emptyset$. On aurait alors

$$E = f^{-1}(O_1) f^{-1}(f(E)) \cup f^{-1}(O_2) f^{-1} f((E)) = f^{-1}(O_1) E \cup f^{-1}(O_2) E,$$

les ensembles $f^{-1}(O_1)$ et $f^{-1}(O_2)$ étant des ouverts tels que $f^{-1}(O_1)E \neq \emptyset$, $f^{-1}(O_1)E \neq \emptyset$ et $f^{-1}(O_1) f^{-1}(O_2) E = \emptyset$ contrairement à l'hypothèse. C.Q.F.D.

Théorème 8 *Soient $E \subseteq \mathbb{R}^n$ un ensemble compact et $f : E \to \mathbb{R}$ une fonction continue sur E. Alors il existe $\mathbf{x}_m \in E$ et $\mathbf{x}_M \in E$ tels que*

$$f(\mathbf{x}_m) = \inf\{f(\mathbf{x}) \mid \mathbf{x} \in E\}\,,\; f(\mathbf{x}_M) = \sup\{f(\mathbf{x}) \mid \mathbf{x} \in E\}.$$

Démonstration.

Vérifions par exemple que f est bornée inférieurement et qu'elle atteint son minimum dans E — l'autre cas est semblable. Si f n'était pas bornée inférieurement dans E, on pourrait trouver une suite $\{\mathbf{x}_k\}_{k\in\mathbb{N}}$ de points de E telle que $f(\mathbf{x}_k) < -k$. L'ensemble E étant compact, cette suite devrait contenir une suite partielle convergeant vers un point de E, soit

$$\overline{\mathbf{x}} = \lim_{p\to+\infty} \mathbf{x}_{k_p}.$$

Par continuité, on devrait avoir

$$f(\overline{\mathbf{x}}) = \lim_{p\to+\infty} f(\mathbf{x}_{k_p}) = -\infty$$

ce qui est absurde.

Soit alors

$$\alpha = \inf\{f(\mathbf{x}) \mid \mathbf{x} \in E\}$$

et choisissons pour chaque $k \in \mathbb{N}$ un point $\mathbf{y}_k \in E$ tel que

$$\alpha < f(\mathbf{y}_k) < \alpha + \frac{1}{k}.$$

Par compacité, cette suite devra contenir une suite partielle convergeant vers un point de E, soit

$$\overline{\mathbf{y}} = \lim_{p\to+\infty} \mathbf{y}_{k_p}.$$

Par continuité, on aura

$$f(\overline{\mathbf{y}}) = \lim_{p\to+\infty} f(\mathbf{y}_{k_p}) = \alpha.$$

C.Q.F.D.

Théorème 9 *Soit $f : \mathbf{R}^n \to \mathbb{R}$ un fonction continue telle que*

$$\lim_{\|\mathbf{x}\|\to+\infty} f(\mathbf{x}) = +\infty.$$

Alors il existe $\mathbf{x}_m$ tel que

$$f(\mathbf{x}_m) = \inf\{f(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n\}.$$

Démonstration.

Par hypothèse, on peut trouver $R > 0$ tel que

$$f(\mathbf{x}) > f(\mathbf{0}) \quad \text{dès que } \|\mathbf{x}\| > R.$$

La boule $\overline{B}(\mathbf{0}, R)$ étant compacte, on peut y trouver un point $\mathbf{x}_m$ tel que

$$f(\mathbf{x}_m) \leq f(\mathbf{x}) \quad \text{si } \|\mathbf{x}\| \leq R.$$

Mais alors on aura en fait

$$f(\mathbf{x}_m) \leq f(\mathbf{x}) \quad \text{pour tout } \mathbf{x} \in \mathbb{R}^n.$$

C.Q.F.D.

Exemple.

Soit $\phi : [0, 1] \to \mathbb{R}$ une fonction continue et considérons le polynôme $f : \mathbb{R}^{n+1} \to \mathbb{R}$ suivant :

$$f(\mathbf{a}) = f(a_0, a_1, \ldots, a_n) = \int_0^1 \left(\phi(t) - \sum_{k=0}^n a_k t^k \right)^2 dt$$
$$= \int_0^1 \phi^2(t)\, dt - 2 \sum_{k=0}^n a_k \int_0^1 \phi(t) t^k\, dt + \sum_{k=0}^n \sum_{p=0}^n \frac{a_k a_p}{k+p+1}.$$

On a

$$\left| \sum_{k=0}^n a_k \int_0^1 \phi(t) t^k\, dt \right| \leq \sqrt{\sum_{k=0}^n \left(\int_0^1 \phi(t) t^k\, dt \right)^2} \sqrt{\sum_{k=0}^n a_k^2} = K_\phi \|\mathbf{a}\|.$$

D'autre part,

$$\sum_{k=0}^n \sum_{p=0}^n \frac{a_k a_p}{k+p+1} = \int_0^1 \left(\sum_{k=0}^n a_k t^k \right)^2 dt > 0 \text{ si } \mathbf{a} \neq \mathbf{0}$$

de telle sorte que, en posant

$$\mu = \inf \left\{ \sum_{k=0}^n \sum_{p=0}^n \frac{a_k a_p}{k+p+1} \mid \|\mathbf{a}\| = 1 \right\},$$

on a, la borne inférieure étant atteinte,

$$\mu > 0$$

et, par homogénéité,

$$\sum_{k=0}^{n}\sum_{p=0}^{n}\frac{a_k a_p}{k+p+1} \geq \mu\|\mathbf{a}\|^2 \text{ pour tout } \mathbf{a} \in \mathbb{R}^{n+1}.$$

Ainsi

$$f(\mathbf{a}) \geq \mu\|\mathbf{a}\|^2 - 2K_\phi\|\mathbf{a}\| - \int_0^1 \phi^2(t)\,dt$$

et

$$\lim_{\|\mathbf{a}\|\to+\infty} f(\mathbf{a}) = +\infty.$$

La fonction f atteint donc une valeur minimum : il existe au moins un polynôme $\sum_{k=0}^{n} a_k t^k$ de meilleure **approximation en moyenne quadratique** pour la fonction ϕ sur l'intervalle $[0, 1]$.

Exemple.

Soit $\phi : [0, 1] \to \mathbb{R}$ une fonction continue et considérons la fonction $f : \mathbb{R}^{n+1} \to \mathbb{R}$ suivante :

$$f(\mathbf{a}) = \sup\left\{\left|\phi(t) - \sum_{k=0}^{n} a_k t^k\right| \mid 0 \leq t \leq 1\right\}.$$

Comme

$$\|f(\mathbf{a}) - f(\mathbf{b})\| \leq \sup\left\{\left|\sum_{k=0}^{n} a_k t^k - \sum_{k=0}^{n} b_k t^k\right| \mid 0 \leq t \leq 1\right\} \leq \sqrt{n+1}\|\mathbf{a} - \mathbf{b}\|,$$

elle est continue. Il en est de même pour la fonction $g : \mathbb{R}^{n+1} \to \mathbb{R}$ associée à $\phi = 0$:

$$g(\mathbf{a}) = \sup\left\{\left|\sum_{k=0}^{n} a_k t^k\right| \mid 0 \leq t \leq 1\right\}.$$

Puisque $g(\mathbf{a}) > 0$ si $\mathbf{a} \neq \mathbf{0}$ et puisque la borne inférieure est atteinte, on a

$$\mu = \inf\{g(\mathbf{a}) \mid \|\mathbf{a}\| = 1\} > 0$$

et, par homogénéité,

$$g(\mathbf{a}) > \mu\|\mathbf{a}\| \text{ pour tout } \mathbf{a} \in \mathbb{R}^{n+1}.$$

Ainsi

$$f(\mathbf{a}) \geq g(\mathbf{a}) - \sup\{|\phi(t)| \mid 0 \leq t \leq 1\} \geq \mu\|\mathbf{a}\| - K_\phi$$

et

$$\lim_{\|\mathbf{a}\|\to+\infty} f(\mathbf{a}) = +\infty.$$

La fonction f atteint donc une valeur minimum : il existe au moins un polynôme $\sum_{k=0}^{n} a_k t^k$ de meilleure **approximation uniforme** pour la fonction ϕ sur l'intervalle $[0,1]$.

3.3 Exercices

Justifier ses réponses.

1. Soient $f_j : \mathbb{R} \to \mathbb{R}$ des fonctions continues. Montrer que la fonction

   $$f(\mathbf{x}) = f_1(x_1) f_2(x_2) \cdots f_n(x_n)$$

   est continue sur $\mathbb{R}^n$.

2. Déterminer l'ensemble des points de continuité des fonctions suivantes :
   - $f(x_1, x_2) = \operatorname{sgn} x_1 \operatorname{sgn} x_2$ (fonction signe) ;
   - $f(x_1, x_2) = x_1\, x_2\, \mathbb{I}_{\mathbb{Q}}(x_1)\, \mathbb{I}_{\mathbb{Q}}(x_2)$ (fonction indicatrice) ;
   - $f(x_1, x_2) = \operatorname{sgn}(x_1 - x_2) \operatorname{sgn}(x_1 + x_2)$.

3. Soient $E \subseteq \mathbb{R}^n$ un ensemble ouvert et $f : E \to \mathbb{R}$ une fonction. Montrer qu'elle est continue si et seulement si elle possède la propriété suivante : pour tout ensemble ouvert $O \subseteq \mathbb{R}^n$, l'ensemble $f^{-1}(O)$ est ouvert.

4. Vrai ou faux ?
   - Si $F \subseteq \mathbb{R}$ est fermé et $f : \mathbb{R}^n \to \mathbb{R}$ est continue, l'ensemble $f^{-1}(F)$ est fermé.
   - Si $F \subseteq \mathbb{R}^n$ est fermé et $f : \mathbb{R}^n \to \mathbb{R}$ est continue, l'ensemble $f(F)$ est fermé.

5. Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction continue. Que peut-on dire des ensembles suivants ?
   - $\{\mathbf{x} \mid f(\mathbf{x}) > \alpha\}$;
   - $\{\mathbf{x} \mid f(\mathbf{x}) \geq \alpha\}$;
   - $\{\mathbf{x} \mid f(\mathbf{x}) = \alpha\}$.

6. Soit $\mathbf{A} \in \mathbb{R}^{n\times n}$ une matrice carrée dont les entrées $a_{i,j} : \mathbb{R}^n \to \mathbb{R}$ sont des fonctions continues. Montrer que :
   - La norme $\|\mathbf{A}(\mathbf{x})\|$ est une fonction continue.
   - Le déterminant $\text{dét}(\mathbf{A}(\mathbf{x}))$ est une fonction continue.
   - L'ensemble $\{\mathbf{x} \mid \mathbf{A}(\mathbf{x}) \text{ est inversible}\}$ est un ensemble ouvert.

7. Montrer qu'un sous-ensemble connexe de $\mathbb{R}$ est nécessairement un intervalle.

8. Soient $E \subseteq \mathbb{R}^n$ un ensemble convexe et $f : E \to \mathbb{R}$ une fonction continue. Montrer que quels que soient $\mathbf{x}_1$, $\mathbf{x}_2 \in E$ et quel que soit y entre $f(\mathbf{x}_1)$ et $f(\mathbf{x}_2)$, il existe $\mathbf{x} \in [\mathbf{x}_1, \mathbf{x}_2]$ tel que $f(\mathbf{x}) = y$.
9. Soient $E \subseteq \mathbb{R}^n$ un ensemble compact et $f : E \to \mathbb{R}$ une fonction continue et strictement positive. Montrer qu'il existe un nombre μ strictement positif tel que $f(\mathbf{x}) \geq \mu$ pour tout $\mathbf{x} \in E$. La conclusion tient-elle si l'on remplace « compact » par « fermé » ?
10. Soient E, $F \subseteq \mathbb{R}^n$ des ensembles compacts. Vérifier que la fonction

$$f(\mathbf{x}) = \inf\{\|\mathbf{x} - \mathbf{y}\| \mid \mathbf{y} \in F\}$$

est continue. En déduire qu'il existe $\mathbf{x}_0 \in E$ et $\mathbf{y}_0 \in F$ tels que

$$\|\mathbf{x}_0 - \mathbf{y}_0\| \leq \|\mathbf{x} - \mathbf{y}\| \quad \text{pour tout } \mathbf{x} \in E,\, \mathbf{y} \in F.$$

11. Montrer que la fonction

$$f(\mathbf{x}) = \frac{\arctan \mathbf{a}^T \mathbf{x}}{1 + \|\mathbf{x}\|^2}$$

atteint son maximum et son minimum sur $\mathbb{R}^n$.

4 FONCTIONS NUMÉRIQUES DÉRIVABLES

Le calcul différentiel cherche à approximer localement une fonction quelconque par une fonction linéaire appropriée.

4.1 Définition

Soient $E \subseteq \mathbb{R}^n$ un ensemble ouvert, $\mathbf{x}_0 \in E$ un de ses points et $f : E \to \mathbb{R}$ une fonction numérique définie sur E. La fonction f est **dérivable** (ou différentiable) en $\mathbf{x}_0$ s'il existe une fonction linéaire $\mathbf{L} : \mathbb{R}^n \to \mathbb{R}$ telle que

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{f(\mathbf{x}) - f(\mathbf{x}_0) - \mathbf{L}(\mathbf{x}-\mathbf{x}_0)}{\|\mathbf{x}-\mathbf{x}_0\|} = 0$$

autrement dit si, dans un **voisinage** de $\mathbf{x}_0$ (un voisinage d'un point est un ensemble qui contient un ouvert qui contient le point), on a

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \mathbf{L}(\mathbf{x}-\mathbf{x}_0) + r(\mathbf{x})$$

avec

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{r(\mathbf{x})}{\|\mathbf{x}-\mathbf{x}_0\|} = 0.$$

La fonction linéaire $\mathbf{L}$ est unique puisque, choisissant $\mathbf{x} = \mathbf{x}_0 \pm h\mathbf{e}_j^{(n)}$, on a nécessairement

$$\mathbf{L}(\mathbf{e}_j^{(n)}) = \lim_{h\to 0} \frac{f(\mathbf{x}_0 + h\mathbf{e}_j^{(n)}) - f(\mathbf{x}_0)}{h}.$$

C'est la **dérivée** de f en $\mathbf{x}_0$, notée

$$\mathbf{L} = f'(\mathbf{x}_0) = Df(\mathbf{x}_0)$$

et les nombres

$$\mathbf{L}(\mathbf{e}_j^{(n)}) = \frac{\partial f}{\partial x_j}(\mathbf{x}_0) = D^{(j)}f(\mathbf{x}_0)$$

sont ses **dérivées partielles** en $\mathbf{x}_0$. Le calcul des dérivées partielles obéit aux mêmes règles que celui des dérivées « ordinaires ». La fonction f est dérivable sur E si elle est dérivable en chaque point de E.

Si f est dérivable en $\mathbf{x}_0$, elle est certainement continue en $\mathbf{x}_0$ puisqu'alors

$$f(\mathbf{x}) - f(\mathbf{x}_0) = \mathbf{L}(\mathbf{x}-\mathbf{x}_0) + r(\mathbf{x}).$$

Le **vecteur gradient** en $\mathbf{x}_0$ est

$$\mathbf{grad} f(\mathbf{x}_0) = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}_0), \frac{\partial f}{\partial x_2}(\mathbf{x}_0), \cdots, \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \right).$$

L'hyperplan de $\mathbb{R}^{n+1}$ dont l'équation est

$$x_{n+1} = f(\mathbf{x}_0) + \mathbf{grad} f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)$$

est **l'hyperplan tangent** au graphe

$$x_{n+1} = f(\mathbf{x})$$

en $\mathbf{x}_0$.

Remarque.
Avec l'identification habituelle, on a

$$f'(\mathbf{x}_0) \in \mathbb{R}^{1\times n} \quad \text{et} \quad \mathbf{grad} f(\mathbf{x}_0) \in \mathbb{R}^{n\times 1}.$$

Ces objets sont les transposés l'un de l'autre.

Exemple.
Une fonction linéaire $f(\mathbf{x}) = \mathbf{a}^T\mathbf{x}$ est partout dérivable et

$$f'(\mathbf{x}_0) = \mathbf{a}^T$$

pour tout $\mathbf{x}_0 \in \mathbb{R}^n$.

Exemple.
Une fonction quadratique $f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ est partout dérivable et

$$f'(\mathbf{x}_0) = 2\,\mathbf{x}_0^T\mathbf{A}$$

pour tout $\mathbf{x}_0 \in \mathbb{R}^n$. En effet,

$$\begin{gathered}\mathbf{x}^T\mathbf{A}\mathbf{x} - \mathbf{x}_0^T\mathbf{A}\mathbf{x}_0 - 2\,\mathbf{x}_0^T\mathbf{A}(\mathbf{x} - \mathbf{x}_0) \\ = \mathbf{x}^T\mathbf{A}(\mathbf{x} - \mathbf{x}_0) + \mathbf{x}_0^T\mathbf{A}(\mathbf{x} - \mathbf{x}_0) - 2\,\mathbf{x}_0^T\mathbf{A}(\mathbf{x} - \mathbf{x}_0) \\ = (\mathbf{x} - \mathbf{x}_0)^T\mathbf{A}(\mathbf{x} - \mathbf{x}_0)\end{gathered}$$

de telle sorte que

$$|\mathbf{x}^T\mathbf{A}\mathbf{x} - \mathbf{x}_0^T\mathbf{A}\mathbf{x}_0 - 2\,\mathbf{x}_0^T\mathbf{A}(\mathbf{x} - \mathbf{x}_0)| \leq \|\mathbf{A}\|\,\|\mathbf{x} - \mathbf{x}_0\|^2.$$

Exemple.

La fonction définie par

$$f(x_1,x_2)=\begin{cases}\dfrac{x_1x_2}{x_1^2+x_2^2} & \text{si } (x_1,x_2)\neq(0,0),\\ 0 & \text{sinon.}\end{cases}$$

n'est pas dérivable à l'origine puisqu'elle n'y est pas continue. Elle y admet néanmoins des dérivées partielles nulles :

$$\lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h}=\lim_{h\to 0}\frac{f(0,h)-f(0,0)}{h}=0.$$

On peut donc former le vecteur gradient à l'origine mais « il ne sert à rien ».

On a d'ailleurs, en utilisant les règles du calcul différentiel, que

$$\frac{\partial f}{\partial x_1}(x_1,x_2)=\begin{cases}\dfrac{x_2(x_2^2-x_1^2)}{(x_1^2+x_2^2)^2} & \text{si } (x_1,x_2)\neq(0,0),\\ 0 & \text{sinon}\end{cases}$$

et

$$\frac{\partial f}{\partial x_2}(x_1,x_2)=\begin{cases}\dfrac{x_1(x_1^2-x_2^2)}{(x_1^2+x_2^2)^2} & \text{si } (x_1,x_2)\neq(0,0),\\ 0 & \text{sinon.}\end{cases}$$

Les dérivées partielles sont ainsi partout définies et continues partout sauf à l'origine :

$$\frac{\partial f}{\partial x_1}(x_1,\lambda x_1)=-\frac{\partial f}{\partial x_2}(x_1,\lambda x_1)=\frac{\lambda(\lambda^2-1)}{x_1(1+\lambda^2)^2}\ \text{ si } x_1\neq 0.$$

4.2 Fonctions continûment dérivables

Soient $E\subseteq\mathbb{R}^n$ un ensemble ouvert et $f:E\to\mathbb{R}$ une fonction numérique définie sur E et admettant des dérivées partielles dans E. Si ces dérivées partielles admettent elles-même des dérivées partielles dans E, ces dernières, les **dérivées partielles d'ordre 2** de la fonction f, sont dénotées au point $\mathbf{x}$ par

$$D^{(i,j)}f(\mathbf{x})=D^{(i)}(D^{(j)}f)(\mathbf{x})$$

ou par

$$\frac{\partial^2 f}{\partial x_i\partial x_j}(\mathbf{x})=\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)(\mathbf{x})$$

et

$$\frac{\partial^2 f}{\partial x_j^2}(\mathbf{x}) = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_j}\right)(\mathbf{x}).$$

D'éventuelles **dérivées partielles d'ordre k** sont dénotées par

$$D^{\boldsymbol{\alpha}} f(\mathbf{x}) = D^{(\alpha_1,\alpha_2,\ldots,\alpha_n)} f(\mathbf{x}) = \frac{\partial^k f}{\partial x_1^{\alpha_1}\partial x_2^{\alpha_2}\cdots\partial x_n^{\alpha_n}}(\mathbf{x})$$

($\|\boldsymbol{\alpha}\|_1 = k$).

La fonction f est de **classe** $C^{(k)}$ **dans** E, $f \in C^{(k)}(E)$, si elle admet dans E des dérivées partielles continues jusqu'à l'ordre k. Elle est de **classe** $C^{(\infty)}$ **dans** E, $f \in C^{(\infty)}(E)$, si elle admet dans E des dérivées partielles continues de tous ordres. Elle est de **classe** $C^{(0)}$ **dans** E, $f \in C^{(0)}(E)$, si elle est continue dans E.

Théorème 10 *Soit $f \in C^{(1)}(E)$. Alors f est dérivable dans E.*

Démonstration.

Pour simplifier l'écriture, nous n'écrivons le raisonnement que pour le cas $n = 2$ et nous posons $x = x_1$, $y = x_2$. Le cas général est similaire.

Soit $(x_0, y_0) \in E$ un point quelconque. On a, en vertu du théorème des accroissements finis,

$$f(x,y) - f(x_0,y_0) - \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) - \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0)$$

$$= f(x,y) - f(x_0,y) + f(x_0,y) - f(x_0,y_0) - \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) - \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0)$$

$$= \left(\frac{\partial f}{\partial x}(x_1,y) - \frac{\partial f}{\partial x}(x_0,y_0)\right)(x-x_0) + \left(\frac{\partial f}{\partial x}(x_0,y_1) - \frac{\partial f}{\partial x}(x_0,y_0)\right)(y-y_0)$$

pour des nombres x_1 entre x et x_0 et y_1 entre y et y_0 appropriés. Donné $\epsilon > 0$, soient, en vertu de la continuité des dérivées partielles, $\delta_1 > 0$ et $\delta_2 > 0$ tels que

$$\left|\frac{\partial f}{\partial x}(\mathbf{x}) - \frac{\partial f}{\partial x}(\mathbf{x}_0)\right| < \frac{\epsilon}{2} \quad \text{dès que } \|\mathbf{x} - \mathbf{x}_0\| < \delta_1$$

et

$$\left|\frac{\partial f}{\partial y}(\mathbf{x}) - \frac{\partial f}{\partial y}(\mathbf{x}_0)\right| < \frac{\epsilon}{2} \quad \text{dès que } \|\mathbf{x} - \mathbf{x}_0\| < \delta_1.$$

Si $\|\mathbf{x}-\mathbf{x}_0\| < \inf\{\delta_1, \delta_2\}$, on a

$$\left| f(x,y) - f(x_0,y_0) - \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) - \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0) \right|$$
$$< \frac{\epsilon}{2}\,|x-x_0| + \frac{\epsilon}{2}\,|y-y_0| \leq \epsilon\,\|\mathbf{x}-\mathbf{x}_0\|.$$

C.Q.F.D.

Exemple.

Les dérivées partielles d'un polynôme P_N étant elles-même des polynômes, on a $P_N \in C^{(\infty)}(\mathbb{R}^n)$. De façon semblable, une fonction rationnelle R est de classe $C^{(\infty)}$ sur son domaine de définition (c'est un ensemble ouvert). De même, une fonction transcendante telle

$$e^{\mathbf{x}^T\mathbf{A}\mathbf{x}} \quad \text{ou} \quad \sin \mathbf{a}^T\mathbf{x}$$

est-elle de classe $C^{(\infty)}$ sur $\mathbb{R}^n$. Toutes ces fonctions sont donc dérivables (sur leur domaine de définition).

Théorème 11 *Soit $f \in C^{(2)}(E)$. Alors*

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

Démonstration.

Pour simplifier l'écriture, nous n'écrivons le raisonnement que pour le cas $n=2$ et nous posons $x = x_1$, $y = x_2$. Le cas général est similaire.

Soit $(x,y) \in E$ un point quelconque. Pour $u > 0$, posons

$$E(u) = \frac{f(x+u,y+u) - f(x+u,y) - f(x,y+u) + f(x,y)}{u^2}.$$

En utilisant le théorème des accroissements finis à deux reprises, on voit qu'il existe $x_1 \in]x, x+u[$ et $y_1 \in]y, y+u[$ tels que

$$E(u) = \frac{1}{u}\left(\frac{(f(x+u,y+u) - f(x+u,y)) - (f(x,y+u) - f(x,y))}{u}\right)$$
$$= \frac{1}{u}\frac{\partial}{\partial x}\left(f(x_1,y+u) - f(x_1,y)\right)$$
$$= \frac{1}{u}\left(\frac{\partial}{\partial x}f(x_1,y+u) - \frac{\partial}{\partial x}f(x_1,y)\right) = \frac{\partial^2}{\partial y \partial x}f(x_1,y_1).$$

Par continuité,

$$\lim_{u\to 0+} E(u) = \frac{\partial^2}{\partial y \partial x} f(x,y).$$

Mais, de façon symétrique, on a aussi

$$E(u) = \frac{1}{u}\frac{\partial}{\partial y}\left(f(x+u,y_2) - f(x,y_2)\right) = \frac{\partial^2}{\partial x \partial y} f(x_2,y_2)$$

et

$$\lim_{u\to 0+} E(u) = \frac{\partial^2}{\partial x \partial y} f(x,y).$$

C.Q.F.D.

Exemple.
Un fonction $f \in C^{(k)}(E)$ admet au plus

$$\binom{n+k-1}{k}$$

dérivées partielles $D^{\boldsymbol{\alpha}} f$ d'ordre k ($\|\boldsymbol{\alpha}\|_1 = k$) distinctes dans E.

Exemple.
La fonction définie par

$$f(x_1,x_2) = \begin{cases} \dfrac{x_1 x_2 (x_1^2 - x_2^2)}{x_1^2 + x_2^2} & \text{si } (x_1,x_2) \neq (0,0), \\ 0 & \text{sinon.} \end{cases}$$

est de classe $C^{(1)}$ sur $\mathbb{R}^n$. On a en effet

$$\frac{\partial f}{\partial x_1}(x_1,x_2) = \begin{cases} \dfrac{x_1^4 x_2 + 4\,x_1^2 x_2^3 - x_2^5}{(x_1^2 + x_2^2)^2} & \text{si } (x_1,x_2) \neq (0,0), \\ 0 & \text{sinon} \end{cases}$$

et

$$\frac{\partial f}{\partial x_2}(x_1,x_2) = \begin{cases} -\dfrac{x_1 x_2^4 + 4\,x_1^3 x_2^2 - x_1^5}{(x_1^2 + x_2^2)^2} & \text{si } (x_1,x_2) \neq (0,0), \\ 0 & \text{sinon.} \end{cases}$$

Ces dérivées partielles sont bien continues à l'origine puisque

$$\left|\frac{x_1^4 x_2 + 4\,x_1^2 x_2^3 - x_2^5}{(x_1^2 + x_2^2)^2}\right| \leq |x_2|\,\frac{x_1^4 + 4\,x_1^2 x_2^2 + x_2^4}{(x_1^2 + x_2^2)^2} \leq 2\,|x_2|$$

et que
$$\left|\frac{x_1x_2^4 + 4\,x_1^3x_2^2 - x_1^5}{(x_1^2+x_2^2)^2}\right| \le |x_1|\,\frac{x_2^4 + 4\,x_1^2x_2^2 + x_1^4}{(x_1^2+x_2^2)^2} \le 2\,|x_1|.$$

La fonction admet des dérivées partielles mixtes d'ordre 2. On a

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(x_1,x_2) = \begin{cases} \dfrac{x_1^8 + 10\,x_1^6x_2^2 - 10\,x_1^2x_2^6 - x_2^8}{(x_1^2+x_2^2)^4} & \text{si } (x_1,x_2) \neq (0,0), \\ -1 & \text{sinon} \end{cases}$$

et

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1,x_2) = \begin{cases} \dfrac{x_1^8 + 10\,x_1^6x_2^2 - 10\,x_1^2x_2^6 - x_2^8}{(x_1^2+x_2^2)^4} & \text{si } (x_1,x_2) \neq (0,0), \\ 1 & \text{sinon.} \end{cases}$$

Ces dérivées mixtes ne sont bien entendu pas continues à l'origine :

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(x_1,\lambda x_1) = \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1,\lambda x_1) = \frac{1 + 10\,\lambda^2 - 10\,\lambda^6 - \lambda^8}{(1+\lambda^2)^4} \quad \text{si } x_1 \neq 0.$$

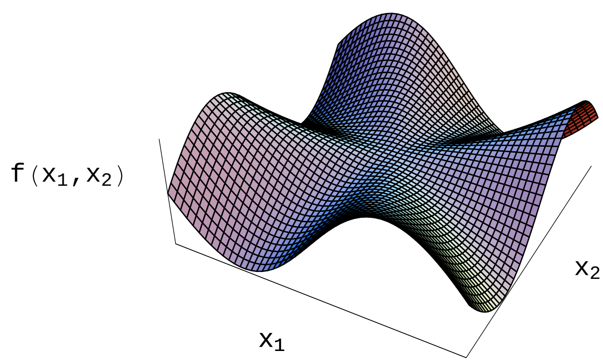

FIG. 5 – Une fonction continûment dérivable

4.3 Propriétés

Soient $E \subseteq \mathbb{R}^n$ un ensemble ouvert et $f : E \to \mathbb{R}$ une fonction numérique définie sur E.

Théorème 12 *Si f est dérivable dans E et si $[\mathbf{x}_1, \mathbf{x}_2] \subseteq E$, il existe $\mathbf{x}_3 \in [\mathbf{x}_1, \mathbf{x}_2]$ tel que*
$$f(\mathbf{x}_2) - f(\mathbf{x}_1) = f'(\mathbf{x}_3)(\mathbf{x}_2 - \mathbf{x}_1).$$

Démonstration.

Appliquons le théorème des accroissements finis à la fonction $g : [0,1] \to \mathbb{R}$ définie par la relation :

$$g(t) = f(\mathbf{x}_1 + t(\mathbf{x}_2 - \mathbf{x}_1)).$$

Il existe $t_3 \in]0,1[$ tel que

$$g(1) - g(0) = g'(t_3).$$

Puisque

$$\begin{aligned} g'(t_3) &= \lim_{h \to 0} \frac{g(t_3 + h) - g(t_3)}{h} \\ &= \lim_{h \to 0} \frac{f(\mathbf{x}_1 + t_3(\mathbf{x}_2 - \mathbf{x}_1) + h(\mathbf{x}_2 - \mathbf{x}_1)) - f(\mathbf{x}_1 + t_3(\mathbf{x}_2 - \mathbf{x}_1))}{h} \\ &= f'(\mathbf{x}_1 + t_3(\mathbf{x}_2 - \mathbf{x}_1))(\mathbf{x}_2 - \mathbf{x}_1), \end{aligned}$$

on a

$$f(\mathbf{x}_2) - f(\mathbf{x}_1) = f'(\mathbf{x}_3)(\mathbf{x}_2 - \mathbf{x}_1)$$

avec

$$\mathbf{x}_3 = \mathbf{x}_1 + t_3(\mathbf{x}_2 - \mathbf{x}_1).$$

C.Q.F.D.

Théorème 13 *Si $f \in C^{(k)}(E)$ et si $[\mathbf{x}_0, \mathbf{x}] \subseteq E$, il existe $\mathbf{y} \in [\mathbf{x}_0, \mathbf{x}]$ tel que*

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \sum_{\|\boldsymbol{\alpha}\|_1 < k} \frac{1}{\boldsymbol{\alpha}!} D^{\boldsymbol{\alpha}} f(\mathbf{x}_0)\,(\mathbf{x} - \mathbf{x}_0)^{\boldsymbol{\alpha}} + \sum_{\|\boldsymbol{\alpha}\|_1 = k} \frac{1}{\boldsymbol{\alpha}!} D^{\boldsymbol{\alpha}} f(\mathbf{y})\,(\mathbf{x} - \mathbf{x}_0)^{\boldsymbol{\alpha}}$$

où l'on a posé

$$\boldsymbol{\alpha}! = \alpha_1!\alpha_2! \cdots \alpha_n!.$$

Démonstration.

Calculons le développement limité d'ordre k à l'origine de la fonction $g : [0,1] \to \mathbb{R}$ définie par la relation :

$$g(t) = f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)).$$

Il existe $s \in]0,1[$ tel que

$$g(1) = g(0) + \sum_{p=1}^{k-1} \frac{1}{p!}\, g^{(p)}(0) + \frac{1}{k!}\, g^{(k)}(s).$$

Il suffit donc de vérifier que l'on a

$$g^{(p)}(t) = \sum_{\|\boldsymbol{\alpha}\|_1 = p} \frac{p!}{\boldsymbol{\alpha}!} D^{\boldsymbol{\alpha}} f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)) \, (\mathbf{x} - \mathbf{x}_0)^{\boldsymbol{\alpha}}.$$

Le résultat suivra avec $\mathbf{y} = \mathbf{x}_0 + s(\mathbf{x} - \mathbf{x}_0)$. Raisonnons par récurrence sur p. Si $p = 1$,

$$\begin{aligned}
g'(t) &= \lim_{h \to 0} \frac{g(t+h) - g(t)}{h} \\
&= \lim_{h \to 0} \frac{f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0) + h(\mathbf{x} - \mathbf{x}_0)) - f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0))}{h} \\
&= f'(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0))(\mathbf{x} - \mathbf{x}_0) \\
&= \sum_{\|\boldsymbol{\alpha}\|_1 = 1} \frac{1}{\boldsymbol{\alpha}!} D^{\boldsymbol{\alpha}} f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)) \, (\mathbf{x} - \mathbf{x}_0)^{\boldsymbol{\alpha}}.
\end{aligned}$$

Par récurrence sur p donc :

$$\begin{aligned}
g^{(p+1)}(t) &= \lim_{h \to 0} \frac{g^{(p)}(t+h) - g^{(p)}(t)}{h} \\
&= \sum_{\|\boldsymbol{\alpha}\|_1 = p} \frac{p!}{\boldsymbol{\alpha}!} (\mathbf{x} - \mathbf{x}_0)^{\boldsymbol{\alpha}} \lim_{h \to 0} \frac{D^{\boldsymbol{\alpha}} f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0) + h(\mathbf{x} - \mathbf{x}_0)) - D^{\boldsymbol{\alpha}} f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0))}{h} \\
&= \sum_{\|\boldsymbol{\alpha}\|_1 = p} \frac{p!}{\boldsymbol{\alpha}!} (\mathbf{x} - \mathbf{x}_0)^{\boldsymbol{\alpha}} \sum_{j=1}^{n} D^{(j)}(D^{\boldsymbol{\alpha}} f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)))(x_j - x_{0,j}) \\
&= \sum_{\|\boldsymbol{\alpha}\|_1 = p+1} \frac{(p+1)!}{\boldsymbol{\alpha}!} D^{\boldsymbol{\alpha}} f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)) \, (\mathbf{x} - \mathbf{x}_0)^{\boldsymbol{\alpha}}.
\end{aligned}$$

C.Q.F.D.

En particulier, si f est de classe $C^{(k)}$ dans un voisinage de $\mathbf{x}_0$, elle y admet un **développement limité** : il existe $r > 0$ tel que

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \sum_{\|\boldsymbol{\alpha}\|_1 < k} \frac{1}{\boldsymbol{\alpha}!} D^{\boldsymbol{\alpha}} f(\mathbf{x}_0) \, (\mathbf{x} - \mathbf{x}_0)^{\boldsymbol{\alpha}} + \sum_{\|\boldsymbol{\alpha}\|_1 = k} \frac{1}{\boldsymbol{\alpha}!} D^{\boldsymbol{\alpha}} f(\mathbf{y}) \, (\mathbf{x} - \mathbf{x}_0)^{\boldsymbol{\alpha}}$$

pour tout $\mathbf{x}$ tel que $\|\mathbf{x} - \mathbf{x}_0\| < r$.

Exemple.

Le développement limité est unique. En particulier, les coefficients $a_{\boldsymbol{\alpha}}$ d'un polynôme

$$P_N(\mathbf{x}) = \sum_{\|\boldsymbol{\alpha}\|_1 \leq N} a_{\boldsymbol{\alpha}}\, \mathbf{x}^{\boldsymbol{\alpha}}$$

sont donnés par les **formules de Taylor** :

$$a_{\boldsymbol{\alpha}} = \frac{1}{\boldsymbol{\alpha}!}\, D^{\boldsymbol{\alpha}} f(\mathbf{0}).$$

Exemple.

Dans le cas fréquemment utilisé où $k = 2$, le développement limité peut s'écrire sous la forme

$$f(\mathbf{x}) = f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}\,(\mathbf{x} - \mathbf{x}_0)^T \mathbf{H}(\mathbf{y})(\mathbf{x} - \mathbf{x}_0)$$

où

$$\mathbf{H}(\mathbf{y}) = \begin{pmatrix} D^{(1,1)} f(\mathbf{y}) & D^{(1,2)} f(\mathbf{y}) & \cdots & D^{(1,n)} f(\mathbf{y}) \\ D^{(2,1)} f(\mathbf{y}) & D^{(2,2)} f(\mathbf{y}) & \cdots & D^{(2,n)} f(\mathbf{y}) \\ \vdots & \vdots & \cdots & \vdots \\ D^{(n,1)} f(\mathbf{y}) & D^{(n,2)} f(\mathbf{y}) & \cdots & D^{(n,n)} f(\mathbf{y}) \end{pmatrix} \in \mathbb{R}^{n \times n}$$

est la matrice (symétrique) des dérivées partielles d'ordre 2 (matrice de **Hesse** ou matrice **hessienne**).

4.4 Exercices

Justifier ses réponses.

1. Soient $f : \mathbb{R} \to \mathbb{R}$ et $g : \mathbb{R} \to \mathbb{R}$ deux fonctions dérivables. Montrer que la fonction
$$h(x_1, x_2) = f(x_1)\, g(x_2)$$
est dérivable.

2. Déterminer l'ensemble des points où la fonction
$$f(x_1, x_2) = |x_1|\, x_2^2$$
est dérivable.

3. Même question pour la fonction
$$f(x_1, x_2) = \sqrt[3]{x_1 x_2}.$$

4. Soit $p > 1$. Calculer $\|\mathbf{grad}(\|\mathbf{x}\|^p)\|$.
5. Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction dérivable. Calculer le gradient de la fonction $g : \mathbb{R}^n \to \mathbb{R}$ définie par $g(\mathbf{x}) = f(\|\mathbf{x}\|)$.
6. Soit
$$f_k(x) = \begin{cases} x^k \sin \frac{1}{x} & \text{si } x \neq 0, \\ 0 & \text{sinon.} \end{cases}$$
Montrer que
– f_0 n'est pas continue ;
– f_1 est continue mais non dérivable ;
– f_2 est dérivable mais non continûment dérivable ;
– f_3 est continûment dérivable mais non deux fois dérivable...
7. Soient $E \subseteq \mathbb{R}^n$ un ensemble ouvert, $\mathbf{x}_0 \in E$ un de ses points et $f : E \to \mathbb{R}$ une fonction numérique définie sur E. La **dérivée directionnelle** de f en $\mathbf{x}_0$ dans la direction du vecteur **unitaire** $\mathbf{e}$ ($\|\mathbf{e}\| = 1$) est
$$D_{\mathbf{e}} f(\mathbf{x}_0) = \lim_{h \to 0+} \frac{f(\mathbf{x}_0 + h\mathbf{e}) - f(\mathbf{x}_0)}{h}$$
(si la limite existe). Montrer qu'une fonction dérivable en $\mathbf{x}_0$ admet une dérivée directionnelle suivant toute direction $\mathbf{e}$ en $\mathbf{x}_0$ et que
$$D_{\mathbf{e}} f(\mathbf{x}_0) = \mathbf{grad} f(\mathbf{x}_0) \cdot \mathbf{e}.$$
8. Soient $E \subseteq \mathbb{R}^n$ un ensemble ouvert convexe et $f : E \to \mathbb{R}$ une fonction dérivable sur E. Montrer que, quels que soient $\mathbf{x}_1$ et $\mathbf{x}_2$ dans E, on a
$$|f(\mathbf{x}_2) - f(\mathbf{x}_1)| \leq \|\mathbf{x}_2 - \mathbf{x}_1\| \sup\{\|f'(\mathbf{x})\| \mid \mathbf{x} \in [\mathbf{x}_1, \mathbf{x}_2]\}.$$
9. Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction telle que
$$|f(\mathbf{x}_2) - f(\mathbf{x}_1)| \leq A \|\mathbf{x}_2 - \mathbf{x}_1\|^p$$
avec $p > 1$. Montrer qu'elle est constante.
10. Développer la fonction $f(x_1, x_2, x_3) = x_1 x_2 x_3$ au point $(1, 1, 1)$.
11. Soient $\phi \in C^{(k)}(\mathbb{R})$ et $f(\mathbf{x}) = \phi(\mathbf{a}^T \mathbf{x})$. Vérifier que le développement limité de cette fonction à l'origine peut s'écrire
$$f(\mathbf{x}) = f(\mathbf{0}) + \sum_{p=1}^{k-1} \phi^{(p)}(0) \sum_{\|\boldsymbol{\alpha}\|_1 = p} \frac{1}{\boldsymbol{\alpha}!} (a_1 x_1)^{\alpha_1} (a_2 x_2)^{\alpha_2} \cdots (a_n x_n)^{\alpha_n} + r_k$$
et préciser la forme du reste r_k.
12. Soient $\phi \in C^{(2)}(\mathbb{R})$ et $f(\mathbf{x}) = \phi(\mathbf{x}^T \mathbf{A} \mathbf{x})$. Calculer $f'(\mathbf{0})$ et $\mathbf{H}(\mathbf{0})$.

5 OPTIMISATION

Nous considérons le problème d'optimiser une fonction $f(\mathbf{x})$ de n variables.

5.1 Extremums locaux

Soient $E \subseteq \mathbb{R}^n$ un ensemble ouvert, $\mathbf{x}_0 \in E$ un de ses points et $f : E \to \mathbb{R}$ une fonction numérique définie sur E. La fonction f admet un **extremum local** (ou relatif) en $\mathbf{x}_0$ s'il existe $r > 0$ tel que $\operatorname{sgn}(f(\mathbf{x}) - f(\mathbf{x}_0))$ est constant dans la boule $B(\mathbf{x}_0, r)$ — un maximum si ce signe reste négatif, un minimum s'il reste positif.

Théorème 14 *Soient $E \subseteq \mathbb{R}^n$ un ensemble ouvert, $\mathbf{x}_0 \in E$ un de ses points et $f : E \to \mathbb{R}$ une fonction numérique définie sur E. Supposons que f admette des dérivées partielles dans E. Une condition nécessaire pour que f admette un extremum local en $\mathbf{x}_0$ est que*

$$\mathbf{grad} f(\mathbf{x}_0) = \mathbf{0}.$$

Démonstration.

Considérons par exemple le cas d'un maximum local. Pour $1 \leq j \leq n$, on a d'une part

$$0 \geq \lim_{h \to 0+} \frac{f(\mathbf{x}_0 + h\mathbf{e}_j^{(n)}) - f(\mathbf{x}_0)}{h} = \frac{\partial f}{\partial x_j}(\mathbf{x}_0)$$

et d'autre part

$$\frac{\partial f}{\partial x_j}(\mathbf{x}_0) = \lim_{h \to 0-} \frac{f(\mathbf{x}_0 + h\mathbf{e}_j^{(n)}) - f(\mathbf{x}_0)}{h} \geq 0.$$

C.Q.F.D.

Les points où le gradient s'annule sont les **points critiques** (ou stationnaire) de la fonction.

Théorème 15 *Soient $f \in C^{(2)}(E)$ et $\mathbf{x}_0 \in E$ un point critique de f.*

– *Si pour tout $\mathbf{u} \neq \mathbf{0}$,*

$$\mathbf{u}^T \mathbf{H}(\mathbf{x}_0) \mathbf{u} < 0,$$

f admet un maximum relatif en $\mathbf{x}_0$.

– *Si pour tout* $\mathbf{u} \neq \mathbf{0}$,

$$\mathbf{u}^T \mathbf{H}(\mathbf{x}_0)\mathbf{u} > 0,$$

f admet un minimum relatif en $\mathbf{x}_0$.

Démonstration.

Considérons par exemple le cas d'un maximum. Il existe un nombre $\delta_1 > 0$ tel que pour $\|\mathbf{x} - \mathbf{x}_0\| < \delta_1$,

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \frac{1}{2}\,(\mathbf{x} - \mathbf{x}_0)^T \mathbf{H}(\mathbf{y})(\mathbf{x} - \mathbf{x}_0)$$

où $\mathbf{y} \in [\mathbf{x}_0, \mathbf{x}]$ de sorte que

$$\operatorname{sgn}\left(f(\mathbf{x}) - f(\mathbf{x}_0)\right) = \operatorname{sgn}\,\big((\mathbf{x} - \mathbf{x}_0)^T \mathbf{H}(\mathbf{y})(\mathbf{x} - \mathbf{x}_0)\big)\,.$$

D'autre part,

$$\begin{aligned}\big|(\mathbf{x} - \mathbf{x}_0)^T \mathbf{H}(\mathbf{y})(\mathbf{x} - \mathbf{x}_0) &- (\mathbf{x} - \mathbf{x}_0)^T \mathbf{H}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)\big| \\ &= \big|(\mathbf{x} - \mathbf{x}_0)^T (\mathbf{H}(\mathbf{y}) - \mathbf{H}(\mathbf{x}_0))(\mathbf{x} - \mathbf{x}_0)\big| \\ &\leq \|\mathbf{H}(\mathbf{y}) - \mathbf{H}(\mathbf{x}_0)\|\,\|\mathbf{x} - \mathbf{x}_0\|^2.\end{aligned}$$

Il suit de l'hypothèse de négativité qu'il existe $\mu > 0$ tel que

$$(\mathbf{x} - \mathbf{x}_0)^T \mathbf{H}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) \leq -\,\mu\,\|\mathbf{x} - \mathbf{x}_0\|^2$$

et il suit de l'hypothèse $f \in C^{(2)}(E)$ qu'il existe un nombre $\delta_2 > 0$ tel que

$$\|\mathbf{x} - \mathbf{x}_0\| < \delta_2 \;\text{ implique }\; \|\mathbf{H}(\mathbf{x}) - \mathbf{H}(\mathbf{x}_0)\| < \frac{\mu}{2}.$$

Si $\|\mathbf{x} - \mathbf{x}_0\| < \inf\{\delta_1, \delta_2\}$, on aura donc

$$\operatorname{sgn}\,\big((\mathbf{x} - \mathbf{x}_0)^T \mathbf{H}(\mathbf{y})(\mathbf{x} - \mathbf{x}_0)\big) < 0.$$

C.Q.F.D.

Exemple.

Soient $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ les valeurs propres de $\mathbf{H}(\mathbf{x}_0)$. Si $\lambda_n < 0$, f atteint un maximum relatif en $\mathbf{x}_0$ et si $\lambda_1 > 0$, f atteint un minimum relatif en $\mathbf{x}_0$ (en vertu du théorème des axes principaux de l'algèbre linéaire).

Exemple.

Il y a un maximum relatif en $\mathbf{x}_0$ pourvu que

$$(-1)^k \det \begin{pmatrix} D^{(1,1)}f(\mathbf{y}) & D^{(1,2)}f(\mathbf{y}) & \cdots & D^{(1,k)}f(\mathbf{y}) \\ D^{(2,1)}f(\mathbf{y}) & D^{(2,2)}f(\mathbf{y}) & \cdots & D^{(2,k)}f(\mathbf{y}) \\ \vdots & \vdots & \cdots & \vdots \\ D^{(k,1)}f(\mathbf{y}) & D^{(k,2)}f(\mathbf{y}) & \cdots & D^{(k,k)}f(\mathbf{y}) \end{pmatrix} > 0$$

pour $1 \leq k \leq n$ et il y a un minimum relatif si

$$\det \begin{pmatrix} D^{(1,1)}f(\mathbf{y}) & D^{(1,2)}f(\mathbf{y}) & \cdots & D^{(1,k)}f(\mathbf{y}) \\ D^{(2,1)}f(\mathbf{y}) & D^{(2,2)}f(\mathbf{y}) & \cdots & D^{(2,k)}f(\mathbf{y}) \\ \vdots & \vdots & \cdots & \vdots \\ D^{(k,1)}f(\mathbf{y}) & D^{(k,2)}f(\mathbf{y}) & \cdots & D^{(k,k)}f(\mathbf{y}) \end{pmatrix} > 0$$

pour $1 \leq k \leq n$ (en vertu du théorème sur les déterminants mineurs principaux de l'algèbre linéaire).

5.2 Fonctions convexes

Soient $E \subseteq \mathbb{R}^n$ un ensemble convexe et $f : E \to \mathbb{R}$ une fonction numérique définie sur E. La fonction f est **convexe** sur E si, quels que soient $\mathbf{x}$ et $\mathbf{y}$ dans E et quel que soit λ dans l'intervalle $[0, 1]$,

$$f((1-\lambda)\mathbf{x} + \lambda\mathbf{y}) \leq (1-\lambda)f(\mathbf{x}) + \lambda f(\mathbf{y}).$$

Elle est **concave** si $-f$ est convexe.

Exemple.

Quel que soit $p \geq 1$, la fonction $f(\mathbf{x}) = \|\mathbf{x}\|^p$ est convexe puisqu'on peut l'écrire comme $f = h \circ g$ où $g(\mathbf{x}) = \|\mathbf{x}\|$ est convexe et $h(t) = t^p$ est convexe croissante.

Théorème 16 *Soient $E \subseteq \mathbb{R}^n$ un ensemble ouvert convexe et $f : E \to \mathbb{R}$ une fonction numérique dérivable sur E. Alors f est convexe sur E si et seulement si quels que soient $\mathbf{x}_1$ et $\mathbf{x}_2$ dans E,*

$$f(\mathbf{x}_2) - f(\mathbf{x}_1) \geq f'(\mathbf{x}_1)(\mathbf{x}_2 - \mathbf{x}_1).$$

Démonstration.

La condition est nécessaire. Supposant f convexe, on aura

$$f(\mathbf{x}_2) - f(\mathbf{x}_1) \geq \frac{f(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)) - f(\mathbf{x}_1)}{\lambda}$$

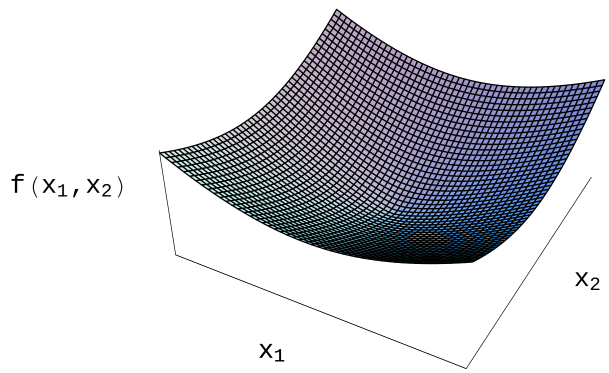

FIG. 6 – Une fonction convexe

pour tout $\lambda \in]0,1[$. Donc

$$f(\mathbf{x}_2) - f(\mathbf{x}_1) \geq f'(\mathbf{x}_1)(\mathbf{x}_2 - \mathbf{x}_1) + \frac{r(\lambda)}{\lambda}$$

où

$$\lim_{\lambda \to 0} \frac{r(\lambda)}{\lambda} = 0.$$

Laissant λ tendre vers 0,

$$f(\mathbf{x}_2) - f(\mathbf{x}_1) \geq f'(\mathbf{x}_1)(\mathbf{x}_2 - \mathbf{x}_1).$$

La condition est suffisante. Donnés $\mathbf{x}_1, \mathbf{x}_2 \in E$ et $\lambda \in]0,1[$, soit

$$\mathbf{x}_\lambda = (1-\lambda)\mathbf{x}_1 + \lambda \mathbf{x}_2.$$

Alors les inégalités

$$f(\mathbf{x}_2) - f(\mathbf{x}_\lambda) \geq f'(\mathbf{x}_\lambda)(\mathbf{x}_2 - \mathbf{x}_\lambda)$$

et

$$f(\mathbf{x}_\lambda) - f(\mathbf{x}_1) \leq f'(\mathbf{x}_\lambda)(\mathbf{x}_\lambda - \mathbf{x}_1)$$

entraînent

$$\frac{f(\mathbf{x}_2) - f(\mathbf{x}_\lambda)}{1-\lambda} \geq f'(\mathbf{x}_\lambda)(\mathbf{x}_2 - \mathbf{x}_1) \geq \frac{f(\mathbf{x}_\lambda) - f(\mathbf{x}_1)}{\lambda}$$

donc

$$f(\mathbf{x}_\lambda) \leq (1-\lambda) f(\mathbf{x}_1) + \lambda f(\mathbf{x}_2).$$

C.Q.F.D.

Théorème 17 *Soient $E \subseteq \mathbb{R}^n$ un ensemble ouvert convexe et $f \in C^{(2)}(E)$. Si pour tout $\mathbf{u} \in \mathbb{R}^n$ et pour tout $\mathbf{x} \in E$,*

$$\mathbf{u}^T \mathbf{H}(\mathbf{x})\mathbf{u} \geq 0,$$

la fonction f est convexe sur E.

Démonstration.
Soient $\mathbf{x}_1, \mathbf{x}_2 \in E$. Alors, il existe $\mathbf{y} \in [\mathbf{x}_1, \mathbf{x}_2]$ tel que

$$f(\mathbf{x}_2) - f(\mathbf{x}_1) - f'(\mathbf{x}_1)(\mathbf{x}_2 - \mathbf{x}_1) = \frac{1}{2}(\mathbf{x}_2 - \mathbf{x}_1)^T \mathbf{H}(\mathbf{y})(\mathbf{x}_2 - \mathbf{x}_1) \geq 0.$$

C.Q.F.D.

Exemple.
Soit $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A}\mathbf{x}$ une fonction quadratique. Alors $\mathbf{H}(\mathbf{x}) = 2\mathbf{A}$ pour tout $\mathbf{x}$ et

$$f(\mathbf{x}) = f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + f(\mathbf{x} - \mathbf{x}_0).$$

La fonction f est donc convexe si et seulement si elle est positive.

5.3 Exercices

Justifier ses réponses.

1. Soient $(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)$ des points du plan tels que

$$x_1 < x_2 < \cdots < x_N.$$

Déterminer la pente a et l'ordonnée à l'origine b de façon à ce que la droite $y = ax + b$ obtenue minimise la somme des carrés des écarts entre les points donnés (x_k, y_k) et les points calculés $(x_k, ax_k + b)$:

$$\sum_{k=1}^{N} (y_k - ax_k - b)^2.$$

(**droite des moindres carrés**)

2. Déterminer le minimum de l'expression

$$\int_{-1}^{1} (\sin \pi t - a - b t - c t^2)^2 dt$$

lorsque $a, b, c \in \mathbb{R}$.

3. Montrer qu'une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ qui est à la fois convexe et concave est nécessairement une fonction affine.
4. Montrer qu'une fonction convexe sur un polyèdre y atteint toujours son maximum en certains des sommets.
5. Montrer que la fonction

$$f(\mathbf{x}) = (1 + \mathbf{x}^T \mathbf{x})^{\mathbf{x}^T \mathbf{x}}$$

est convexe.

6 TRANSFORMATIONS DE L'ESPACE EUCLIDIEN

Soit $E \subseteq \mathbb{R}^n$ un ensemble. Une fonction $\mathbf{f} : E \to \mathbb{R}^m$ est déterminée par ses composantes, les m fonctions numériques $f_i : E \to \mathbb{R}$:

$$\mathbf{f} = (f_1, f_2, \ldots, f_m).$$

6.1 Exemples

Nous considérons d'abord quelques transformations de ce type.

Exemple.

Dans le cas où $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ est linéaire, $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$, les composantes f_i de $\mathbf{f}$ sont aussi des fonctions linéaires. L'image de $\mathbb{R}^n$ par $\mathbf{f}$ est un sous-espace vectoriel de $\mathbb{R}^m$ dont la dimension est égale au **rang** de $\mathbf{A}$, le nombre de vecteurs colonnes linéairement indépendants de $\mathbf{A}$ (c'est aussi le nombre de vecteurs lignes linéairement indépendants en vertu d'un théorème d'algèbre linéaire).

Exemple.

Lorsque $n = m = 3$, une fonction $\mathbf{f} : E \to \mathbb{R}^m$ est un **champ de vecteurs** dans E. Le champ

$$\mathbf{f}(\mathbf{x}) = \frac{\mathbf{x}}{\|\mathbf{x}\|^3}$$

est ainsi associé à la gravitation newtonienne dans $\mathbb{R}^3 \setminus \{\mathbf{0}\}$.

Pour visualiser une transformation $\mathbb{R}^2 \to \mathbb{R}^2$, $(x_1, x_2) \mapsto (y_1, y_2)$, on peut tracer dans le plan y_1y_2 les images des droites $x_1 = c_1$ et $x_2 = c_2$ par la transformation.

Exemple.

Pour la fonction

$$f_1(x_1, x_2) = e^{x_1} \cos x_2 \ , \ f_2(x_1, x_2) = e^{x_1} \sin x_2 \ ,$$

les droites $x_1 = c_1$ ont pour images les cercles $y_1^2 + y_2^2 = e^{2c_1}$ et les droites $x_2 = c_2$ ont pour images les droites $y_2 = \tan c_2 \, y_1$.

Exemple.

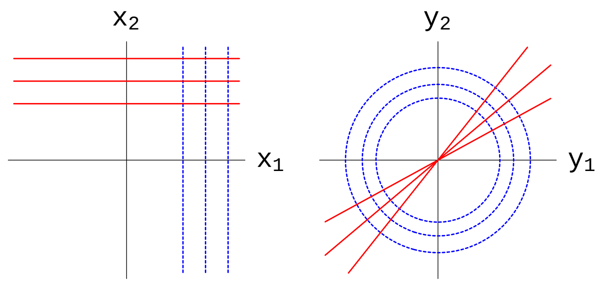

FIG. 7 – Une transformation du plan

Les **coordonnées sphériques** $r, \theta_1, \theta_2, \ldots, \theta_{n-1}$ sur $\mathbb{R}^n$ sont définies via les équations suivantes.

Lorsque $n = 2$ (**coordonnées polaires**),

$$x_1 = r\cos\theta_1\ ,\ x_2 = r\sin\theta_1$$

c'est-à-dire que pour, $\mathbf{x} \neq \mathbf{0}$,

$$r = \sqrt{x_1^2 + x_2^2}$$

et

$$\theta_1 = \begin{cases} \arctan\frac{x_2}{x_1} & \text{si } x_1 > 0 \\ \frac{\pi}{2} & \text{si } x_1 = 0,\ x_2 > 0 \\ \arctan\frac{x_2}{x_1} + \pi & \text{si } x_1 < 0,\ x_2 \geq 0 \\ -\frac{\pi}{2} & \text{si } x_1 = 0,\ x_2 < 0 \\ \arctan\frac{x_2}{x_1} - \pi & \text{si } x_1 < 0,\ x_2 < 0. \end{cases}$$

Lorsque $n = 3$,

$$\begin{aligned} x_1 &= r\cos\theta_1\sin\theta_2 \\ x_2 &= r\sin\theta_1\sin\theta_2 \\ x_3 &= r\cos\theta_2 \end{aligned}$$

avec

$$r > 0\ ,\ -\pi < \theta_1 \leq \pi\ ,\ 0 \leq \theta_2 \leq \pi.$$

On a donc

$$r = \sqrt{x_1^2 + x_2^2 + x_3^2}$$

$$\theta_1 = \begin{cases} \arctan \frac{x_2}{x_1} & \text{si } x_1 > 0 \\ \frac{\pi}{2} & \text{si } x_1 = 0,\ x_2 > 0 \\ \arctan \frac{x_2}{x_1} + \pi & \text{si } x_1 < 0,\ x_2 \geq 0 \\ -\frac{\pi}{2} & \text{si } x_1 = 0,\ x_2 < 0 \\ \arctan \frac{x_2}{x_1} - \pi & \text{si } x_1 < 0,\ x_2 < 0. \end{cases}$$

$$\theta_2 = \begin{cases} \arctan \dfrac{\sqrt{x_1^2 + x_2^2}}{x_3} & \text{si } x_3 > 0 \\ \frac{\pi}{2} & \text{si } x_3 = 0 \\ \arctan \dfrac{\sqrt{x_1^2 + x_2^2}}{x_3} + \pi & \text{si } x_3 < 0. \end{cases}$$

En général,

$$\begin{aligned} x_n &= r \cos \theta_{n-1} \\ x_{n-1} &= r \sin \theta_{n-1} \cos \theta_{n-2} \\ x_{n-2} &= r \sin \theta_{n-1} \sin \theta_{n-2} \cos \theta_{n-3} \\ &\cdots \\ x_3 &= r \sin \theta_{n-1} \sin \theta_{n-2} \cdots \cos \theta_2 \\ x_2 &= r \sin \theta_{n-1} \sin \theta_{n-2} \cdots \sin \theta_2 \sin \theta_1 \\ x_1 &= r \sin \theta_{n-1} \sin \theta_{n-2} \cdots \sin \theta_2 \cos \theta_1 \end{aligned}$$

(x_1 et x_2 dérogent à l'ordre « naturel » pour se conformer à l'usage courant dans le plan et l'espace). On a

$$r > 0\ ,\ -\pi < \theta_1 \leq \pi \ \text{ et } \ 0 \leq \theta_j \leq \pi \ \text{ pour } 2 \leq j \leq n-1.$$

6.2 Transformations continues

Soient $E \subseteq \mathbb{R}^n$ un ensemble, $\mathbf{x}_0 \in E$ un de ses points et $\mathbf{f} : E \to \mathbb{R}^m$ une fonction définie sur E. Elle est **continue** en $\mathbf{x}_0$ si à chaque $\epsilon > 0$ correspond $\delta > 0$ tel que

$$\mathbf{x} \in E \text{ et } \|\mathbf{x} - \mathbf{x}_0\| < \delta \quad \text{impliquent} \quad \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0)\| < \epsilon$$

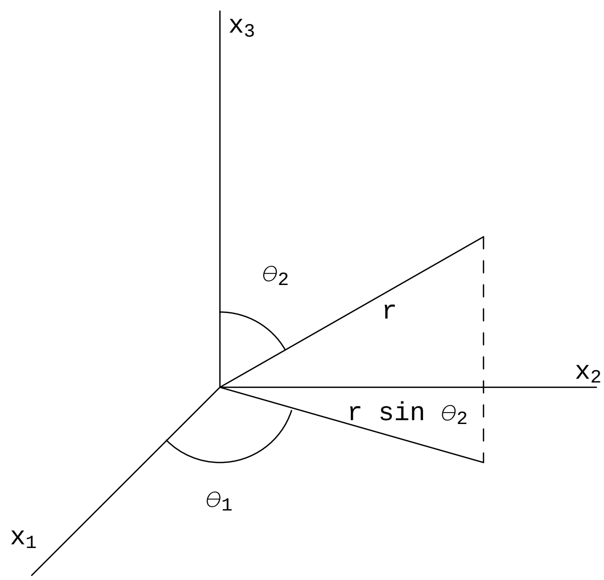

FIG. 8 – Les coordonnées sphériques dans $\mathbb{R}^3$

ou, de façon équivalente, si pour toute suite $\{\mathbf{x}_k\}_{k\in\mathbb{N}}$ de points de E,

$$\lim_{k\to+\infty} \mathbf{x}_k = \mathbf{x}_0 \quad \text{implique} \quad \lim_{k\to+\infty} \mathbf{f}(\mathbf{x}_k) = \mathbf{f}(\mathbf{x}_0).$$

La fonction $\mathbf{f}$ est continue en $\mathbf{x}_0$ si et seulement si chacune de ses composantes $f_i : E \to \mathbb{R}$ l'est. Elle est continue sur E si elle est continue en chaque point de E.

Exemple.
Une transformation linéaire $\mathbf{L} : \mathbb{R}^n \to \mathbb{R}^m$ est continue :

$$\|\mathbf{L}(\mathbf{x}) - \mathbf{L}(\mathbf{x}_0)\| \leq \|\mathbf{L}\|_\infty \|\mathbf{x} - \mathbf{x}_0\|.$$

Exemple.
La fonction $\mathbf{f} : \mathbb{R}^n \setminus \mathbf{0} \to \mathbb{R}^n$ donnée par

$$\mathbf{f}(\mathbf{x}) = \frac{\mathbf{x}}{\|\mathbf{x}\|}$$

est continue. Elle ne peut pas être prolongée à une fonction continue sur $\mathbb{R}^n$ tout entier puisque

$$\mathbf{f}(\lambda\, \mathbf{e}_1^{(n)}) = \operatorname{sgn} \lambda\, \mathbf{e}_1^{(n)}$$

ce qui n'admet pas de limite lorsque $\lambda \to 0$.

Exemple.

La fonction $\mathbf{f} :]0, +\infty[\times] - \pi, \pi] \times [0, \pi] \to \mathbb{R}^3$ donnée par

$$\mathbf{f}(r, \theta_1, \theta_2) = (r \cos \theta_1 \sin \theta_2, r \sin \theta_1 \sin \theta_2, r \cos \theta_2)$$

est continue.

Théorème 18 *Soient $E \subseteq \mathbb{R}^n$ un ensemble compact et $\mathbf{f} : E \to \mathbb{R}^m$ une fonction continue sur E. Alors l'ensemble $\mathbf{f}(E)$ est compact.*

Démonstration.

Soit $\{\mathbf{y}_k\}_{k\in\mathbb{N}}$ une suite de points de $\mathbf{f}(E)$. Il existe une suite $\{\mathbf{x}_k\}_{k\in\mathbb{N}}$ de points de E telle que

$$\mathbf{y}_k = \mathbf{f}(\mathbf{x}_k).$$

Par compacité, cette suite contient une suite partielle qui converge vers un point de E, soit

$$\overline{\mathbf{x}} = \lim_{p\to+\infty} \mathbf{x}_{k_p}.$$

En posant

$$\overline{\mathbf{y}} = \mathbf{f}(\overline{\mathbf{x}}),$$

on aura, par continuité,

$$\overline{\mathbf{y}} = \lim_{p\to+\infty} \mathbf{y}_{k_p}$$

et la suite donnée contient bien une suite partielle convergeant vers un point de $\mathbf{f}(E)$. C.Q.F.D.

Lorsque les conditions du théorème sont satisfaites, la fonction numérique $\mathbf{x} \mapsto \|\mathbf{f}(\mathbf{x})\|$ atteint son maximum et son minimum sur E.

Exemple.

Il existe un point $\mathbf{x}_0$ de la sphère $\|\mathbf{x}\| = 1$ où une transformation linéaire $\mathbf{L} : \mathbb{R}^n \to \mathbb{R}^m$ atteint sa norme :

$$\|\mathbf{L}\|_\infty = \|\mathbf{L}(\mathbf{x}_0)\|.$$

6.3 Transformations différentiables

Soient $E \subseteq \mathbb{R}^n$ un ensemble ouvert, $\mathbf{x}_0 \in E$ un de ses points et $\mathbf{f} : E \to \mathbb{R}^m$ une fonction définie sur E. La fonction $\mathbf{f}$ est **dérivable** (différentiable) en $\mathbf{x}_0$ s'il existe une transformation linéaire $\mathbf{L} : \mathbb{R}^n \to \mathbb{R}^m$ telle que

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0) - \mathbf{L}(\mathbf{x} - \mathbf{x}_0))}{\|\mathbf{x} - \mathbf{x}_0\|} = \mathbf{0}$$

autrement dit si, dans un voisinage de $\mathbf{x}_0$, on a

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + \mathbf{L}(\mathbf{x} - \mathbf{x}_0) + \mathbf{r}(\mathbf{x})$$

avec

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{\mathbf{r}(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}_0\|} = 0.$$

La fonction $\mathbf{f}$ est dérivable en $\mathbf{x}_0$ si et seulement si chacune de ses composantes f_i l'est et alors

$$L_i = f_i'(\mathbf{x}_0).$$

La transformation linéaire $\mathbf{L}$ est donc unique. C'est la **dérivée** de $\mathbf{f}$ en $\mathbf{x}_0$, notée

$$\mathbf{L} = \mathbf{f}'(\mathbf{x}_0) = D\mathbf{f}(\mathbf{x}_0).$$

Relativement aux bases canoniques, sa matrice est notée

$$\begin{pmatrix} D^{(1)}f_1(\mathbf{x}_0) & D^{(2)}f_1(\mathbf{x}_0) & \cdots & D^{(n)}f_1(\mathbf{x}_0) \\ D^{(1)}f_2(\mathbf{x}_0) & D^{(2)}f_2(\mathbf{x}_0) & \cdots & D^{(n)}f_2(\mathbf{x}_0) \\ \vdots & \vdots & \cdots & \vdots \\ D^{(1)}f_m(\mathbf{x}_0) & D^{(2)}f_m(\mathbf{x}_0) & \cdots & D^{(n)}f_m(\mathbf{x}_0) \end{pmatrix}$$

ou encore

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_2}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f_2}{\partial x_n}(\mathbf{x}_0) \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_m}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{pmatrix}.$$

C'est la matrice de **Jacobi** ou la matrice **jacobienne** de la fonction $\mathbf{f}$. En particulier, il suffit que les fonctions f_i soient toutes de classe $C^{(1)}$ dans E pour que la fonction $\mathbf{f}$ y soit dérivable. On dit que $\mathbf{f}$ est de **classe** $C^{(k)}$ **dans** E, $\mathbf{f} \in C^{(k)}(E)$, si toutes les fonction f_i le sont.

Lorsque $m = n$, le déterminant de la matrice de Jacobi est le **jacobien** de la transformation, noté

$$J\mathbf{f}(\mathbf{x}_0)$$

ou

$$\frac{\partial(f_1, f_2, \ldots, f_n)}{\partial(x_1, x_2, \ldots, x_n)}(\mathbf{x}_0).$$

Lorsque l'on écrit $y_i = f_i(x_1, x_2, \ldots, x_n)$, on écrit aussi

$$J\mathbf{f} = \frac{\partial(y_1, y_2, \ldots, y_n)}{\partial(x_1, x_2, \ldots, x_n)}.$$

Exemple.

Une transformation linéaire $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$, est sa propre dérivée. En tout point $\mathbf{x}_0$, on a

$$\mathbf{f}'(\mathbf{x}_0) = \mathbf{A}.$$

Exemple.

La fonction

$$\mathbf{f}(r, \theta_1, \theta_2) = (r\cos\theta_1 \sin\theta_2, r\sin\theta_1 \sin\theta_2, r\cos\theta_2)$$

admet pour matrice jacobienne

$$\begin{pmatrix} \cos\theta_1 \sin\theta_2 & -r\sin\theta_1 \sin\theta_2 & r\cos\theta_1 \cos\theta_2 \\ \sin\theta_1 \sin\theta_2 & r\cos\theta_1 \sin\theta_2 & r\sin\theta_1 \cos\theta_2 \\ \cos\theta_2 & 0 & -r\sin\theta_2 \end{pmatrix}$$

dans l'ensemble ouvert $E =]0, +\infty[\, \times\,]-\pi, \pi[\, \times\,]0, \pi[$. Les entrées de cette matrices étant toutes continues, la transformation est dérivable dans E. En fait, elle y est de classe $C^{(\infty)}$. Le jacobien de la transformation est

$$-r^2 \sin\theta_2.$$

Théorème 19 *Soient $E \subseteq \mathbb{R}^n$ un ensemble ouvert convexe et $\mathbf{f} : E \to \mathbb{R}^m$ une fonction dérivable. Quels que soient $\mathbf{x}_1, \mathbf{x}_2 \in E$, on a*

$$\|\mathbf{f}(\mathbf{x}_2) - \mathbf{f}(\mathbf{x}_1)\| \leq \sqrt{m\, n}\, \|\mathbf{x}_2 - \mathbf{x}_1\| \,\sup\{\|\mathbf{f}'(\mathbf{x})\| \mid \mathbf{x} \in [\mathbf{x}_1, \mathbf{x}_2]\}.$$

Démonstration.

En appliquant le théorème des accroissements finis (théorème (12)) à chaque composante f_i de $\mathbf{f}$, on obtient des points $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_m \in [\mathbf{x}_1, \mathbf{x}_2]$ tels que

$$\|\mathbf{f}(\mathbf{x}_2) - \mathbf{f}(\mathbf{x}_1)\|^2 = \sum_{i=1}^{m} (f_i(\mathbf{x}_2) - f_i(\mathbf{x}_1))^2$$

$$= \sum_{i=1}^{m} \left(f_i'(\mathbf{z}_i)(\mathbf{x}_2 - \mathbf{x}_1)\right)^2 \leq \|\mathbf{x}_2 - \mathbf{x}_1\|^2 \sum_{i=1}^{m} \|f_i'(\mathbf{z}_i)\|^2$$

$$= \|\mathbf{x}_2 - \mathbf{x}_1\|^2 \sum_{i=1}^{m} \sum_{j=1}^{n} \left(\frac{\partial f_i}{\partial x_j}(\mathbf{z}_i)\right)^2$$

$$\leq \|\mathbf{x}_2 - \mathbf{x}_1\|^2 n \sum_{i=1}^{m} \|\mathbf{f}'(\mathbf{z}_i)\|^2$$

$$\leq \|\mathbf{x}_2 - \mathbf{x}_1\|^2 \, n \times m \, \sup\{\|\mathbf{f}'(\mathbf{x})\| \mid \mathbf{x} \in [\mathbf{x}_1, \mathbf{x}_2]\}^2.$$

C.Q.F.D.

Théorème 20 *Soient $E \subseteq \mathbb{R}^n$ un ensemble ouvert et $\mathbf{f} : E \to \mathbb{R}^n$ une fonction admettant une matrice jacobienne dans E. Si $\mathbf{x} \mapsto \|\mathbf{f}(\mathbf{x})\|$ admet un maximum local en $\mathbf{x}_0 \in E$,*

$$J\mathbf{f}(\mathbf{x}_0) = 0.$$

Démonstration.

Si en effet $\|\mathbf{f}(\mathbf{x}_0)\| = 0$, il faut que $\|\mathbf{f}(\mathbf{x})\| = 0$ dans un voisinage de $\mathbf{x}_0$ donc que $J\mathbf{f}(\mathbf{x}_0) = 0$.

Si $\|\mathbf{f}(\mathbf{x}_0)\| \neq 0$, la fonction

$$\phi(\mathbf{x}) = \|\mathbf{f}(\mathbf{x})\|^2 = \sum_{i=1}^{n} (f_i(\mathbf{x}))^2$$

admettra un maximum local en $\mathbf{x}_0$ donc

$$\mathbf{grad}\phi(\mathbf{x}_0) = \mathbf{0}.$$

Comme

$$\frac{\partial \phi}{\partial x_j}(\mathbf{x}) = 2 \sum_{i=1}^{n} f_i(\mathbf{x}) \frac{\partial f_i}{\partial x_j}(\mathbf{x}),$$

le système linéaire homogène

$$\sum_{i=1}^{n} \frac{\partial f_i}{\partial x_j}(\mathbf{x}_0)u_i = 0 \quad , \quad 1 \leq j \leq n$$

dans les variables $u_1, u_2, \ldots, u_n$ admettra une solution non triviale $\mathbf{u} = \mathbf{f}(\mathbf{x}_0)$ et son déterminant $J\mathbf{f}(\mathbf{x}_0)$ devra s'annuler. C.Q.F.D.

6.4 Exercices

Justifier ses réponses.

1. Soit $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ une fonction. Vérifier qu'elle est bornée (en norme) sur l'ensemble E si et seulement si ses composantes $f_i : \mathbb{R}^n \to \mathbb{R}^m$ le sont. Posant

$$M = \sup\{\|\mathbf{f}(\mathbf{x})\| \mid \mathbf{x} \in E\}$$

et

$$M_i = \sup\{|f_i(\mathbf{x})| \mid \mathbf{x} \in E\},$$

montrer que

$$M_i \leq M \leq \sqrt{\sum_{i=1}^{m} M_i^2}.$$

2. Soit $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$ la fonction

$$\mathbf{f}(\mathbf{x}) = \begin{cases} \dfrac{(x_1^2 - x_2^2,\ 2x_1x_2)}{\sqrt{x_1^2 + x_2^2}} & \text{si } (x_1, x_2) \neq (0, 0), \\ 0 & \text{sinon.} \end{cases}$$

Vérifier que $\|\mathbf{f}(\mathbf{x})\| = \|\mathbf{x}\|$. Déterminer l'ensemble E_c des points où elle est continue puis l'ensemble E_d des points où elle est dérivable.

3. Calculer la dérivée de chacune des fonctions suivantes :

–
$$y_1 = e^{x_1} \cos x_2,\ y_2 = e^{x_1} \sin x_2$$

–
$$y_1 = x_1 + x_2 + x_3,\ y_2 = x_1^2 + x_2^2 + x_3^2$$

–
$$y_1 = x_1 + x_2,\ y_2 = x_1 - x_2,\ y_3 = \sqrt{x_1^2 + x_2^2}$$

–

$$y_1 = \cos x_1,\ y_2 = \sin x_1,\ y_3 = x_1.$$

4. Pour chacune des fonctions précédente, déterminer l'image $\mathbf{f}(\mathbb{R}^n)$ de $\mathbb{R}^n$ par la fonction $\mathbf{f}$.
5. Montrer, par un exemple approprié, que le théorème des accroissements finis,
$$\mathbf{f}(\mathbf{x}_2) - \mathbf{f}(\mathbf{x}_1) = \mathbf{f}'(\mathbf{x}_3)(\mathbf{x}_2 - \mathbf{x}_1),$$
est faux pour une fonction $\mathbb{R}^n \to \mathbb{R}^m$ si $m > 1$.
6. Soient $E \subseteq \mathbb{R}^n$ un ensemble ouvert convexe et $\mathbf{f} : E \to \mathbb{R}^m$ une fonction dérivable. Montrer que, si $\mathbf{f}'(\mathbf{x}) = \mathbf{0}$ pour tout $\mathbf{x} \in E$, la fonction $\mathbf{f}$ est constante dans E.
7. Vrai ou faux ?

   Soient $E \subseteq \mathbb{R}^n$ un ensemble ouvert et $\mathbf{f} : E \to \mathbb{R}^n$ une fonction admettant une matrice jacobienne dans E. Si $\mathbf{x} \mapsto \|\mathbf{f}(\mathbf{x})\|$ admet un minimum local en $\mathbf{x}_0 \in E$, $J\mathbf{f}(\mathbf{x}_0) = 0$.
8. On considère la transformation des coordonnées sphériques aux coordonnées cartésiennes sur $\mathbb{R}^3$:
$$\begin{aligned} x_1 &= r\cos\theta_1 \sin\theta_2 \\ x_2 &= r\sin\theta_1 \sin\theta_2 \\ x_3 &= r\cos\theta_2. \end{aligned}$$
Pour quelles valeurs de $\mathbf{x}$ la transformation inverse est-elle définie ? continue ? dérivable ? Quelle est sa dérivée ?

7 DÉRIVATION EN CHAÎNE

Nous considérons le problème de relier les dérivées partielles d'une fonction $f : \mathbb{R}^n \to \mathbb{R}$ relativement à deux systèmes de coordonnées différents sur $\mathbb{R}^n$.

7.1 Le théorème

Théorème 21 *Soient $E \subseteq \mathbb{R}^n$, $F \subseteq \mathbb{R}^m$ des ensembles ouverts et soient $\mathbf{f} : E \to \mathbb{R}^m$, $\mathbf{g} : F \to \mathbb{R}^p$ des fonctions telles que $\mathbf{f}(E) \subseteq F$. Si $\mathbf{f}$ est dérivable en $\mathbf{x}_0$ et si $\mathbf{g}$ est dérivable en $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$, la fonction composée $\mathbf{g} \circ \mathbf{f} : E \to \mathbb{R}^p$ est dérivable en $\mathbf{x}_0$ et*

$$(\mathbf{g} \circ \mathbf{f})'(\mathbf{x}_0) = \mathbf{g}'(\mathbf{y}_0) \circ \mathbf{f}'(\mathbf{x}_0).$$

Démonstration.

Dans un voisinage de $\mathbf{x}_0$, on a :

$$\begin{aligned} &\mathbf{g}(\mathbf{f}(\mathbf{x})) - \mathbf{g}(\mathbf{y}_0) - \mathbf{g}'(\mathbf{y}_0)(\mathbf{f}'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)) \\ &= \mathbf{g}(\mathbf{f}(\mathbf{x})) - \mathbf{g}(\mathbf{y}_0) - \mathbf{g}'(\mathbf{y}_0)(\mathbf{f}(\mathbf{x}) - \mathbf{y}_0) \\ &+\mathbf{g}'(\mathbf{y}_0)(\mathbf{f}(\mathbf{x}) - \mathbf{y}_0 - \mathbf{f}'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)). \end{aligned}$$

Distinguons deux cas.

Cas où $\mathbf{g}'(\mathbf{y}_0) = \mathbf{0}$.

Donné $\epsilon > 0$, soit $\delta_g > 0$ tel que

$$\|\mathbf{g}(\mathbf{y}) - \mathbf{g}(\mathbf{y}_0)\| < \frac{\epsilon}{1 + \|\mathbf{f}'(\mathbf{x}_0)\|_\infty} \|\mathbf{y} - \mathbf{y}_0\|$$

dès que $\|\mathbf{y} - \mathbf{y}_0\| < \delta_g$. Soit aussi $\delta_1 > 0$ tel que

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{y}_0 - \mathbf{f}'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)\| < \|\mathbf{x} - \mathbf{x}_0\|$$

dès que $\|\mathbf{x} - \mathbf{x}_0\| < \delta_1$. En particulier, on a alors

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{y}_0\| < (1 + \|\mathbf{f}'(\mathbf{x}_0)\|_\infty)\|\mathbf{x} - \mathbf{x}_0\|.$$

Donc

$$\|\mathbf{x} - \mathbf{x}_0\| < \inf\{\delta_1, \frac{\delta_g}{1 + \|\mathbf{f}'(\mathbf{x}_0)\|_\infty}\}$$

implique

$$\|\mathbf{g}(\mathbf{f}(\mathbf{x})) - \mathbf{g}(\mathbf{y}_0)\| < \epsilon \, \|\mathbf{x} - \mathbf{x}_0\|$$

ce qui montre que
$$(\mathbf{g}\circ\mathbf{f})'(\mathbf{x}_0)=\mathbf{0}.$$

Cas où $\mathbf{g}'(\mathbf{y}_0)\neq\mathbf{0}$.

On a

$$\begin{gathered}\|\mathbf{g}(\mathbf{f}(\mathbf{x}))-\mathbf{g}(\mathbf{y}_0)-\mathbf{g}'(\mathbf{y}_0)(\mathbf{f}'(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0))\|\\ \leq\|\mathbf{g}(\mathbf{f}(\mathbf{x}))-\mathbf{g}(\mathbf{y}_0)-\mathbf{g}'(\mathbf{y}_0)(\mathbf{f}(\mathbf{x})-\mathbf{y}_0)\|\\ +\|\mathbf{g}'(\mathbf{y}_0)\|_\infty\|\mathbf{f}(\mathbf{x})-\mathbf{y}_0-\mathbf{f}'(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0)\|.\end{gathered}$$

Donné $\epsilon>0$, soit $\delta_f>0$ tel que

$$\|\mathbf{f}(\mathbf{x})-\mathbf{y}_0-\mathbf{f}'(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0\|<\frac{\epsilon}{2\,\|\mathbf{g}'(\mathbf{y}_0))\|_\infty}\|\mathbf{x}-\mathbf{x}_0\|$$

dès que $\|\mathbf{x}-\mathbf{x}_0\|<\delta_f$. Pour de tels $\mathbf{x}$, on aura en particulier que

$$\|\mathbf{f}(\mathbf{x})-\mathbf{y}_0\|<\left(\frac{\epsilon}{2\,\|\mathbf{g}'(\mathbf{y}_0))\|_\infty}+\|\mathbf{f}'(\mathbf{x}_0)\|_\infty\right)\|\mathbf{x}-\mathbf{x}_0\|.$$

Soit maintenant $\delta_g>0$ tel que $\|\mathbf{y}-\mathbf{y}_0\|<\delta_g$ implique

$$\begin{gathered}\|\mathbf{g}(\mathbf{y})-\mathbf{g}(\mathbf{y}_0)-\mathbf{g}'(\mathbf{y}_0)(\mathbf{f}(\mathbf{x})-\mathbf{y}_0)\|\\ <\frac{\epsilon}{2\left(\|\mathbf{f}'(\mathbf{x}_0)\|_\infty+\dfrac{\epsilon}{2\,\|\mathbf{g}'(\mathbf{y}_0))\|_\infty}\right)}\|\mathbf{y}-\mathbf{y}_0\|.\end{gathered}$$

Si
$$\|\mathbf{x}-\mathbf{x}_0\|<\inf\{\delta_f,\frac{\delta_g}{\|\mathbf{f}'(\mathbf{x}_0)\|_\infty+\dfrac{\epsilon}{2\,\|\mathbf{g}'(\mathbf{y}_0))\|_\infty}}\},$$

on aura

$$\|\mathbf{g}(\mathbf{f}(\mathbf{x}))-\mathbf{g}(\mathbf{y}_0)-\mathbf{g}'(\mathbf{y}_0)(\mathbf{f}'(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0))\|<\epsilon\,\|\mathbf{x}-\mathbf{x}_0\|.$$

C.Q.F.D.

Un cas particulier important du théorème précédent est celui où $\mathbf{g}=\mathbf{f}^{-1}$. On a alors
$$\left(\mathbf{f}^{-1}\right)'(\mathbf{y}_0)=\left(\mathbf{f}'(\mathbf{x}_0)\right)^{-1}.$$

Soulignons que cette formule suppose que l'on sait que la fonction inverse $\mathbf{f}^{-1}$ est dérivable en $\mathbf{y}_0=\mathbf{f}(\mathbf{x}_0)$. Le théorème entraîne alors que la transformation linéaire $\mathbf{f}'(\mathbf{x}_0)$ est inversible (donc que $m = n$) et que la relation précédente est valable.

En passant aux matrices des transformations linéaires $\mathbf{f}'(\mathbf{x}_0)$, $\mathbf{g}'(\mathbf{y}_0)$ et $(\mathbf{g} \circ \mathbf{f})'(\mathbf{x}_0)$, la relation

$$(\mathbf{g} \circ \mathbf{f})'(\mathbf{x}_0) = \mathbf{g}'(\mathbf{y}_0) \circ \mathbf{f}'(\mathbf{x}_0)$$

implique (en écrivant $\mathbf{y} = \mathbf{f}(\mathbf{x})$ et $\mathbf{z} = \mathbf{g}(\mathbf{y})$) :

$$\frac{\partial z_k}{\partial x_j} = \sum_{i=1}^{m} \frac{\partial z_k}{\partial y_i} \frac{\partial y_i}{\partial x_j}$$

pour $1 \leq k \leq p$ et $1 \leq j \leq n$. On écrit souvent ceci comme une relation entre opérateurs de dérivation :

$$\frac{\partial}{\partial x_j} = \sum_{i=1}^{m} \frac{\partial y_i}{\partial x_j} \frac{\partial}{\partial y_i}$$

ou, lorsque $n = 1$,

$$\frac{d}{dx} = \sum_{i=1}^{m} \frac{dy_i}{dx} \frac{\partial}{\partial y_i}.$$

La relation

$$\left(\mathbf{f}^{-1}\right)'(\mathbf{y}_0) = \left(\mathbf{f}'(\mathbf{x}_0)\right)^{-1}$$

quant à elle a comme conséquence supplémentaire la relation suivante entre les jacobiens :

$$\frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(y_1, y_2, \ldots, y_n)}(\mathbf{y}_0) = \frac{1}{\dfrac{\partial(y_1, y_2, \ldots, y_n)}{\partial(x_1, x_2, \ldots, x_n)}(\mathbf{x}_0)}.$$

7.2 Applications

Exemple.

Soient $E \subseteq \mathbb{R}^n$ un ensemble ouvert et $u, v : E \to \mathbb{R}$ deux fonctions numériques dérivables dans E. Alors les fonctions $u + v$, $u\,v$ et , si $v \neq 0$ dans E, u/v sont dérivables dans E et

$$(u + v)' = u' + v'$$
$$(u\,v)' = u'v + u\,v'$$
$$\frac{u'}{v} = \frac{u'v - u\,v'}{v^2}.$$

En effet, pour démontrer la première assertion, considérons les fonctions $\mathbf{f} : E \to \mathbb{R}^2$ et $g : \mathbb{R}^2 \to \mathbb{R}$ définies par

$$\mathbf{f}(\mathbf{x}) = (u(\mathbf{x}), v(\mathbf{x})) \text{ et } g(\mathbf{y}) = y_1 + y_2.$$

La fonction $\mathbf{f}$ est dérivable sur E et a pour matrice

$$\mathbf{f}'(\mathbf{x}) = \begin{pmatrix} D^{(1)}u(\mathbf{x}) & D^{(2)}u(\mathbf{x}) \\ D^{(1)}v(\mathbf{x}) & D^{(2)}v(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} u'(\mathbf{x}) \\ v'(\mathbf{x}) \end{pmatrix},$$

la fonction g est dérivable sur $\mathbf{R}^2$ et a pour matrice

$$g'(\mathbf{y}) = (1\ 1).$$

On a

$$(u+v)(\mathbf{x}) = (g \circ \mathbf{f})(\mathbf{x}).$$

Par suite,

$$(u+v)'(\mathbf{x}) = g'(\mathbf{f}(\mathbf{x}))\, \mathbf{f}(\mathbf{x}) = (1\ 1) \begin{pmatrix} u'(\mathbf{x}) \\ v'(\mathbf{x}) \end{pmatrix} = u'(\mathbf{x}) + v'(\mathbf{x})$$

ce qui est le résultat désiré.

Pour démontrer les deux autres assertions, il suffit de considérer les fonctions $g(\mathbf{y}) = y_1\, y_2$ et $g(\mathbf{y}) = y_1/y_2$ respectivement.

Exemple.

Considérons **l'opérateur différentiel linéaire d'ordre deux à coefficients constants** :

$$D = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j} \frac{\partial^2}{\partial x_1 \partial x_j}$$

où l'on peut supposer que *la matrice* $\mathbf{A}$ *est symétrique*. Il est toujours possible d'effectuer un changement linéaire de variables

$$\mathbf{y} = \mathbf{P}\mathbf{x}$$

de telle sorte que

$$D = \sum_{k=1}^{n} \lambda_k \frac{\partial^2}{\partial y_k^2},$$

les nombres $\lambda_1, \lambda_2, \ldots, \lambda_n$ étant les valeurs propres de la matrice $\mathbf{A}$. Il suffit en effet de choisir $\mathbf{P}$ telle que

$$\mathbf{PAP}^T = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

(en vertu du théorème des axes principaux de l'algèbre linéaire). On aura alors, puisque

$$y_k = p_{k,1}x_1 + p_{k,2}x_2 + \cdots + p_{k,n}x_n,$$

que

$$\frac{\partial}{\partial x_j} = \sum_{k=1}^{n} \frac{\partial y_k}{\partial x_j} \frac{\partial}{\partial y_k} = \sum_{k=1}^{n} p_{k,j} \frac{\partial}{\partial y_k}$$

puis que

$$\frac{\partial^2}{\partial x_i \partial x_j} = \sum_{k=1}^{n} p_{k,j} \frac{\partial}{\partial x_i}\left(\frac{\partial}{\partial y_k}\right) = \sum_{k=1}^{n} p_{k,j} \sum_{q=1}^{n} p_{q,i} \; \frac{\partial^2}{\partial y_q \partial y_k}.$$

Ainsi

$$\begin{aligned} D &= \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{q=1}^{n} p_{q,i}\, a_{i,j}\, p_{k,j} \frac{\partial^2}{\partial y_q \partial y_k} \\ &= \sum_{k=1}^{n}\sum_{q=1}^{n}\sum_{i=1}^{n}\sum_{j=1}^{n} p_{q,i}\, a_{i,j}\, p^T_{j,k} \frac{\partial^2}{\partial y_q \partial y_k} \\ &= \sum_{k=1}^{n}\sum_{q=1}^{n} (\mathbf{PAP}^T)_{k,q} \frac{\partial^2}{\partial y_q \partial y_k} = \sum_{k=1}^{n} \lambda_k \frac{\partial^2}{\partial y_k^2}. \end{aligned}$$

Lorsque $n = 2$, il n'y a que trois formes canoniques pour l'opérateur D, suivant que les deux valeurs propres sont de même signe, sont de signes opposés ou que l'une des deux est nulle. On obtient ainsi **l'équation de Laplace** (du potentiel)

$$\frac{\partial^2 \phi}{\partial x_1^2} + \frac{\partial^2 \phi}{\partial x_2^2} = 0,$$

l'équation d'onde

$$\frac{\partial^2 \phi}{\partial x_1^2} - \frac{\partial^2 \phi}{\partial x_2^2} = 0$$

et **l'équation de la chaleur** (de la diffusion)

$$\frac{\partial^2\phi}{\partial x_1^2} - \frac{\partial\phi}{\partial x_2} = 0.$$

Exemple.
En coordonnées polaires,

$$x_1 = r\cos\theta_1\ ,\ x_2 = r\sin\theta_1,$$

le **laplacien**

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}$$

devient (sauf à l'origine)

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\theta_1^2}$$

lorsqu'on l'applique à une fonction de classe $C^{(2)}$. En effet,

$$\frac{\partial}{\partial x_1} = \frac{\partial r}{\partial x_1}\frac{\partial}{\partial r} + \frac{\partial\theta_1}{\partial x_1}\frac{\partial}{\partial\theta_1} = \cos\theta_1\frac{\partial}{\partial r} - \frac{\sin\theta_1}{r}\frac{\partial}{\partial\theta_1}$$

et

$$\frac{\partial}{\partial x_2} = \frac{\partial r}{\partial x_2}\frac{\partial}{\partial r} + \frac{\partial\theta_1}{\partial x_2}\frac{\partial}{\partial\theta_1} = \sin\theta_1\frac{\partial}{\partial r} + \frac{\cos\theta_1}{r}\frac{\partial}{\partial\theta_1}.$$

Par suite,

$$\frac{\partial^2}{\partial x_1^2} = \cos\theta_1\frac{\partial}{\partial r}\left(\cos\theta_1\frac{\partial}{\partial r} - \frac{\sin\theta_1}{r}\frac{\partial}{\partial\theta_1}\right) - \frac{\sin\theta_1}{r}\frac{\partial}{\partial\theta_1}\left(\cos\theta_1\frac{\partial}{\partial r} - \frac{\sin\theta_1}{r}\frac{\partial}{\partial\theta_1}\right)$$
$$= \cos^2\theta_1\frac{\partial^2}{\partial r^2} - 2\,\frac{\cos\theta_1\,\sin\theta_1}{r}\frac{\partial^2}{\partial r\partial\theta_1} + \frac{\sin^2\theta_1}{r^2}\frac{\partial^2}{\partial\theta_1^2} + \frac{\sin^2\theta_1}{r}\frac{\partial}{\partial r} + 2\,\frac{\cos\theta_1\,\sin\theta_1}{r}\frac{\partial}{\partial\theta_1}.$$

De façon symétrique,

$$\frac{\partial^2}{\partial x_2^2} = \sin\theta_1\frac{\partial}{\partial r}\left(\sin\theta_1\frac{\partial}{\partial r} + \frac{\cos\theta_1}{r}\frac{\partial}{\partial\theta_1}\right) + \frac{\cos\theta_1}{r}\frac{\partial}{\partial\theta_1}\left(\sin\theta_1\frac{\partial}{\partial r} + \frac{\cos\theta_1}{r}\frac{\partial}{\partial\theta_1}\right)$$
$$= \sin^2\theta_1\frac{\partial^2}{\partial r^2} + 2\,\frac{\cos\theta_1\,\sin\theta_1}{r}\frac{\partial^2}{\partial r\partial\theta_1} + \frac{\cos^2\theta_1}{r^2}\frac{\partial^2}{\partial\theta_1^2} + \frac{\cos^2\theta_1}{r}\frac{\partial}{\partial r} - 2\,\frac{\cos\theta_1\,\sin\theta_1}{r}\frac{\partial}{\partial\theta_1}.$$

La formule suit en additionnant ces deux expressions.

Exemple.

En coordonnées sphériques,

$$x_1 = r\cos\theta_1 \sin\theta_2\ ,\ x_2 = r\sin\theta_1\sin\theta_2\ ,\ x_3 = r\cos\theta_2,$$

le laplacien

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}$$

s'écrit

$$\Delta = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta_2}\frac{\partial}{\partial\theta_2}\left(\sin\theta_2\frac{\partial}{\partial\theta_2}\right) + \frac{1}{r^2\sin^2\theta_2}\frac{\partial^2}{\partial\theta_1^2}.$$

Considérons en effet la transformation auxiliaire

$$x_1 = \rho\cos\theta_1\ ,\ x_2 = \rho\sin\theta_1.$$

En vertu de l'exemple précédent,

$$\Delta = \frac{\partial^2}{\partial\rho^2} + \frac{1}{\rho}\frac{\partial}{\partial\rho} + \frac{1}{\rho^2}\frac{\partial^2}{\partial\theta_1^2} + \frac{\partial^2}{\partial x_3^2}.$$

Soient ensuite

$$\rho = r\sin\theta_2\ ,\ x_3 = r\cos\theta_2.$$

Toujours en vertu du même calcul,

$$\begin{aligned}\Delta &= \frac{\partial^2}{\partial\rho^2} + \frac{\partial^2}{\partial x_3^2} + \frac{1}{\rho}\frac{\partial}{\partial\rho} + \frac{1}{\rho^2}\frac{\partial^2}{\partial\theta_1^2}\\ &= \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\theta_2} + \frac{1}{r\sin\theta_2}\frac{\partial}{\partial\rho} + \frac{1}{r^2\sin^2\theta_2}\frac{\partial^2}{\partial\theta_1^2}.\end{aligned}$$

D'autre part

$$\frac{\partial}{\partial\rho} = \frac{\partial r}{\partial\rho}\frac{\partial}{\partial r} + \frac{\partial\theta_2}{\partial\rho}\frac{\partial}{\partial\theta_2} = \sin\theta_2\frac{\partial}{\partial r} + \frac{\cos\theta_2}{r}\frac{\partial}{\partial\theta_2}.$$

Donc

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\theta_2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{\cos\theta_2}{r^2\sin\theta_2}\frac{\partial}{\partial\theta_2} + \frac{1}{r^2\sin^2\theta_2}\frac{\partial^2}{\partial\theta_1^2}.$$

ce qui est la formule annoncée développée.

7.3 Exercices

Justifier ses réponses.

1. Montrer que la composition $\mathbf{g} \circ \mathbf{f}$ de fonctions $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ et $\mathbf{g} : \mathbb{R}^m \to \mathbb{R}^p$ de classe $C^{(1)}$ est une fonction de classe $C^{(1)}$.
2. Calculer le jacobien
$$\frac{\partial(r, \theta_1, \theta_2)}{\partial(x_1, x_2, x_3)}$$
(l'exprimer en termes des variables x_1, x_2 et x_3).
3. Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction dérivable strictement positive et considérons la fonction $g : \mathbb{R}^n \to \mathbb{R}$ définie par
$$g(\mathbf{x}) = \sqrt{f(\mathbf{x})}.$$
Montrer qu'elle est dérivable et que
$$g'(\mathbf{x}) = \frac{f'(\mathbf{x})}{2\sqrt{f(\mathbf{x})}}.$$
4. Soient $f, g : \mathbb{R}^n \to \mathbb{R}$ deux fonctions dérivables et considérons la fonction $h : \mathbb{R}^n \to \mathbb{R}$ définie par
$$h(\mathbf{x}) = \log\left(1 + f^2(\mathbf{x}) + g^2(\mathbf{x})\right).$$
Montrer qu'elle est dérivable et que
$$h'(\mathbf{x}) = 2\,\frac{f(\mathbf{x})f'(\mathbf{x}) + g(\mathbf{x})g'(\mathbf{x})}{1 + f^2(\mathbf{x}) + g^2(\mathbf{x})}.$$
5. Soit $f \in C^{(2)}(\mathbb{R}^2)$ et considérons l'équation d'onde
$$\frac{\partial^2 f}{\partial x_1^2} = \frac{1}{c^2}\,\frac{\partial^2 f}{\partial x_2^2}.$$
Montrer que pour un changement linéaire approprié des variables,
$$\mathbf{y} = \mathbf{P}\mathbf{x},$$
elle devient
$$\frac{\partial^2 f}{\partial y_1 \partial y_2} = 0.$$
En déduire la forme générale de sa solution.

6. Vérifier que le laplacien

$$\Delta = \sum_{j=1}^{n} \frac{\partial^2}{\partial x_j^2}$$

est invariant sous une transformation orthogonale

$$\mathbf{y} = \mathbf{P}\mathbf{x} \ \text{ avec } \ \mathbf{P}^T = \mathbf{P}^{-1}$$

des variables.

7. Montrer qu'en deux dimensions, les solutions de classe $C^{(2)}$ de l'équation de Laplace (les **fonctions harmoniques**)

$$\Delta f = 0$$

qui ne dépendent que de $\sqrt{x_1^2 + x_2^2}$ sont de la forme

$$f(\mathbf{x}) = a \log \sqrt{x_1^2 + x_2^2} + b.$$

8. Montrer qu'en trois dimensions, les solutions de classe $C^{(2)}$ de l'équation de Laplace

$$\Delta f = 0$$

qui ne dépendent que de $\sqrt{x_1^2 + x_2^2 + x_3^2}$ sont de la forme

$$f(\mathbf{x}) = \frac{a}{\sqrt{x_1^2 + x_2^2 + x_3^2}} + b.$$

9. Soient $f : \mathbb{R}^n \to \mathbb{R}$ une fonction dérivable et $p \in \mathbb{R}$ un nombre réel. Montrer que les conditions suivantes sont équivalentes :
(i) quels que soient $\mathbf{x} \in \mathbb{R}^n$ et $t > 0$,

$$f(t\,\mathbf{x}) = t^p\, f(\mathbf{x});$$

(ii) quel que soit $\mathbf{x} \in \mathbb{R}^n$,

$$\sum_{j=1}^{n} x_j \frac{\partial f}{\partial x_j}(\mathbf{x}) = p\, f(\mathbf{x}).$$

(théorème d'Euler sur les **fonctions homogènes**).

8 FONCTIONS INVERSES

Nous considérons le problème de résoudre un système de n équations différentiables en n inconnues : $f(\mathbf{x}) = \mathbf{y}$.

8.1 Le théorème

Dans le cas où $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ est linéaire, le système $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ admet une solution unique $\mathbf{x} = \mathbf{f}^{-1}(\mathbf{y})$ quel que soit $\mathbf{y} \in \mathbb{R}^n$ si et seulement si $\mathbf{f}'$ est inversible.

Dans le cas $n = 1$, soit $f(x_0) = y_0$ où f est une fonction de classe $C^{(1)}$ dans un voisinage de x_0. On sait que si $f'(x_0) \neq 0$, il existe un voisinage ouvert U de x_0 et un voisinage ouvert V de y_0 tels que pour tout $y \in V$, l'équation $f(x) = y$ admette une solution unique x dans U, la fonction inverse $f^{-1} : V \to U$ ainsi définie étant elle aussi de classe $C^{(1)}$.

Par exemple, pour $f(x) = x^2$ et $x_0 > 0$, on a $U = V =]0, +\infty[$ et $f^{-1}(y) = \sqrt{y}$. Si $f(x) = x^3$ et $x_0 > 0$, on a encore $U = V =]0, +\infty[$ car la fonction inverse $f^{-1}(y) = \sqrt[3]{y}$ est définie partout mais n'est pas dérivable à l'origine.

Théorème 22 *Soient $E \subseteq \mathbb{R}^n$ un ensemble ouvert et $\mathbf{f} : E \to \mathbb{R}^n$ une fonction de classe $C^{(1)}$ dans E. Si $\mathbf{f}(\mathbf{x}_0) = \mathbf{y}_0$ et si $\mathbf{f}'(\mathbf{x}_0)$ est inversible, il existe un voisinage ouvert U de $\mathbf{x}_0$ et un voisinage ouvert V de $\mathbf{y}_0$ tels que l'équation $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ admette pour tout $\mathbf{y} \in V$ une solution $\mathbf{x} \in U$ unique. La fonction inverse $\mathbf{f}^{-1} : V \to U$ ainsi définie est de classe $C^{(1)}$ dans V.*

Démonstration.

Le raisonnement se divise en trois étapes : il existe un voisinage ouvert U de $\mathbf{x}_0$ dans lequel la fonction $\mathbf{f}$ est injective, son image $V = \mathbf{f}(U)$ par $\mathbf{f}$ est ouvert et la fonction inverse $\mathbf{f}^{-1} : V \to U$ est de classe $C^{(1)}$ dans V. Nous posons $\mathbf{A} = \mathbf{f}'(\mathbf{x}_0)$.

Étape 1. Pour montrer qu'il existe un voisinage ouvert U de $\mathbf{x}_0$ dans lequel la fonction $\mathbf{f}$ est injective, soit $R > 0$ tel que la dérivée $\mathbf{f}'(\mathbf{x})$ soit inversible dans la boule $B(\mathbf{x}_0, R) \subseteq E$. Si $\mathbf{x}_1$ et $\mathbf{x}_2$ sont dans $B(\mathbf{x}_0, R)$, on peut trouver, pour $1 \leq i \leq n$, des points $\mathbf{z}_i \in [\mathbf{x}_1, \mathbf{x}_2]$ tels que

$$f_i(\mathbf{x}_1) - f_i(\mathbf{x}_2) = f_i'(\mathbf{z}_i)(\mathbf{x}_1 - \mathbf{x}_2).$$

Donné

$$0 < \epsilon < \frac{1}{\|\mathbf{A}^{-1}\|_\infty},$$

soit, pour $1 \leq i \leq n$, $\delta_i > 0$ tel que $\mathbf{x} \in B(\mathbf{x}_0, \delta_i)$ implique

$$\|f_i'(\mathbf{x}) - f_i'(\mathbf{x}_0)\| < \frac{\epsilon}{\sqrt{n}}.$$

Soit $\delta = \inf\{\delta_1, \delta_2, \ldots, \delta_n\}$. Si $\mathbf{x}_1, \mathbf{x}_2 \in B(\mathbf{x}_0, \delta)$, on aura

$$\begin{aligned} &\|\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_2) - \mathbf{A}(\mathbf{x}_1 - \mathbf{x}_2)\|^2 \\ &= \sum_{i=1}^{n} \big(f_i(\mathbf{x}_1) - f_i(\mathbf{x}_2) - f_i'(\mathbf{x}_0)(\mathbf{x}_1 - \mathbf{x}_2)\big)^2 \\ &\leq \sum_{i=1}^{n} \|f_i'(\mathbf{z}_i) - f_i'(\mathbf{x}_0)\|^2 \|\mathbf{x}_1 - \mathbf{x}_2\|^2 \\ &< \epsilon^2 \, \|\mathbf{x}_1 - \mathbf{x}_2\|^2. \end{aligned}$$

En particulier,

$$\begin{aligned} \|\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_2)\| &\geq \|\mathbf{A}(\mathbf{x}_1 - \mathbf{x}_2)\|^2 - \epsilon \, \|\mathbf{x}_1 - \mathbf{x}_2\| \\ &> \left(\frac{1}{\|\mathbf{A}^{-1}\|_\infty} - \epsilon\right) \|\mathbf{x}_1 - \mathbf{x}_2\| > 0 \end{aligned}$$

si $\mathbf{x}_1 \neq \mathbf{x}_2$. La fonction $\mathbf{f}$ est injective dans la boule $U = B(\mathbf{x}_0, \delta)$.

Étape 2. Soient $V = \mathbf{f}(U)$ et $\mathbf{f}^{-1} : V \to U$ la fonction inverse de la restriction $\mathbf{f}/U$ de $\mathbf{f}$ à U. Montrons que V est un ensemble ouvert. Soit $\mathbf{y}_3 = \mathbf{f}(\mathbf{x}_3) \in V$. Choisissons $r_3 > 0$ tel que $\overline{B}(\mathbf{x}_3, r_3) \subseteq U$ et considérons l'ensemble compact $S_3 = \mathbf{f}(S(\mathbf{x}_3, r_3))$, image de la sphère compacte $S(\mathbf{x}_3, r_3)$ par la fonction continue $\mathbf{f}$. Puisque $\mathbf{y}_3 \notin S_3$,

$$d_3 = \inf\{\|\mathbf{y} - \mathbf{y}_3\| \mid \mathbf{y} \in S_3\} > 0.$$

Vérifions que $B(\mathbf{y}_3, d_3/2) \subseteq V$. Soit $\mathbf{y}_4 \in B(\mathbf{y}_3, d_3/2)$. Considérons la fonction $\phi : U \to \mathbb{R}$ définie par

$$\phi(\mathbf{x}) = \|\mathbf{f}(\mathbf{x}) - \mathbf{y}_4\|^2.$$

Il y a un point $\mathbf{x}_4 \in \overline{B}(\mathbf{x}_3, r_3)$ tel que

$$\phi(\mathbf{x}_4) = \inf\{\phi(\mathbf{x}) \mid \mathbf{x} \in \overline{B}(\mathbf{x}_3, r_3)\}.$$

En fait, $\mathbf{x}_4 \in B(\mathbf{x}_3, r_3)$ car si $\|\mathbf{x} - \mathbf{x}_3\| = r_3$, on a

$$\begin{aligned}\phi(\mathbf{x}) = \|\mathbf{f}(\mathbf{x}) - \mathbf{y}_4\|^2 &\geq (\|\mathbf{f}(\mathbf{x}) - \mathbf{y}_3\| - \|\mathbf{y}_3 - \mathbf{y}_4\|)^2 \\ &> \left(d_3 - \frac{d_3}{2}\right)^2 = \left(\frac{d_3}{2}\right)^2 = \phi(\mathbf{x}_3).\end{aligned}$$

Par suite, la fonction ϕ atteint un minimum local en $\mathbf{x}_4$ et donc

$$\frac{\partial \phi}{\partial x_j}(\mathbf{x}_4) = 2 \sum_{i=1}^{n} \left(f_i(\mathbf{x}_4) - y_{4,i}\right) \frac{\partial f_i}{\partial x_j}(\mathbf{x}_4) = 0$$

pour $1 \leq j \leq n$. Comme

$$\frac{\partial(f_1, f_2, \ldots, f_n)}{\partial(x_1, x_2, \ldots, x_n)} \neq 0,$$

la solution du système linéaire homogène

$$\sum_{i=1}^{n} \frac{\partial f_i}{\partial x_j}(\mathbf{x}_4)\, u_i = 0$$

pour $1 \leq j \leq n$ doit être la solution triviale $\mathbf{u} = \mathbf{0}$, c'est-à-dire que

$$\mathbf{y}_4 = \mathbf{f}(\mathbf{x}_4).$$

Étape 3. Montrons enfin que $\mathbf{f}^{-1} \in C^{(1)}(V)$. L'inégalité

$$\|\mathbf{f}^{-1}(\mathbf{y}_1) - \mathbf{f}^{-1}(\mathbf{y}_2)\| < \frac{1}{\frac{1}{\|\mathbf{A}^{-1}\|_\infty} - \epsilon} \|\mathbf{y}_1 - \mathbf{y}_2\|$$

obtenue à la fin de l'étape 1 et valable quels que soient $\mathbf{y}_1, \mathbf{y}_2 \in V$ montre déjà que $\mathbf{f}^{-1} \in C^{(0)}(V)$. Soit $\mathbf{y}_5 = \mathbf{f}(\mathbf{x}_5) \in V$ un point quelconque et vérifions que $\mathbf{f}^{-1}$ y est dérivable. Donné $\eta > 0$, soit $\mu > 0$ tel que

$$\|\mathbf{f}(\mathbf{x}_6) - \mathbf{f}(\mathbf{x}_5) - \mathbf{f}'(\mathbf{x}_5)(\mathbf{x}_6 - \mathbf{x}_5)\| < \eta\, \|\mathbf{x}_6 - \mathbf{x}_5\|$$

dès que $\|\mathbf{x}_6 - \mathbf{x}_5\| < \mu$. Soit maintenant $\Delta > 0$ tel que

$$\|\mathbf{x}_6 - \mathbf{x}_5\| = \|\mathbf{f}^{-1}(\mathbf{y}_6) - \mathbf{f}^{-1}(\mathbf{y}_5)\| < \mu$$

dès que $\|\mathbf{y}_6 - \mathbf{y}_5\| < \Delta$. Si $\|\mathbf{y}_6 - \mathbf{y}_5\| < \Delta$, on aura donc

$$\|\mathbf{f}'(\mathbf{x}_5)\left(\mathbf{f}'(\mathbf{x}_5)^{-1}(\mathbf{y}_6 - \mathbf{y}_5) - (\mathbf{f}^{-1}(\mathbf{y}_6) - \mathbf{f}^{-1}(\mathbf{y}_5))\right)\| < \eta\, \|\mathbf{f}^{-1}(\mathbf{y}_6) - \mathbf{f}^{-1}(\mathbf{y}_5)\|$$

d'où

$$\frac{1}{\|\mathbf{f}'(\mathbf{x}_5)^{-1}\|_\infty}\|\mathbf{f}^{-1}(\mathbf{y}_6)-\mathbf{f}^{-1}(\mathbf{y}_5)-\mathbf{f}'(\mathbf{x}_5)^{-1}(\mathbf{y}_6-\mathbf{y}_5)\|$$
$$<\eta\,\frac{1}{\dfrac{1}{\|\mathbf{A}^{-1}\|_\infty}-\epsilon}\|\mathbf{y}_6-\mathbf{y}_5\|$$

ce qui montre que $\mathbf{f}^{-1}$ est dérivable en $\mathbf{y}_5$ et que, bien sûr,

$$(\mathbf{f}^{-1})'(\mathbf{y}_5)=\mathbf{f}'(\mathbf{x}_5)^{-1}.$$

Cette dernière relation entraîne aussi que $\mathbf{f}^{-1}\in C^{(1)}(V)$ si $\mathbf{f}\in C^{(1)}(U)$ — et plus généralement que $\mathbf{f}^{-1}\in C^{(k)}(V)$ si $\mathbf{f}\in C^{(k)}(U)$. C.Q.F.D.

La démonstration précédente peut être grossièrement décrite de la façon suivante : on montre d'abord que la fonction $\mathbf{f}$ établit localement un isomorphe d'ensembles ($\mathbf{f}$ est localement une bijection), ensuite un isomorphisme topologique ($\mathbf{f}$ est localement un **homéomorphisme** — une bijection continue ainsi que son inverse) et enfin un isomorphisme différentiel ($\mathbf{f}$ est localement un **difféomorphisme** — une bijection dérivable ainsi que son inverse — de classe $C^{(1)}$ ou plus).

Sauf dans les cas les plus simples, il est illusoire d'espérer trouver explicitement les fonctions

$$x_j=\phi_j(y_1,y_2,\ldots,y_n)\ ,\ 1\leq j\leq n$$

solutions du système

$$f_i(x_1,x_2,\ldots,x_n)=y_i\ ,\ 1\leq i\leq n.$$

La règle de dérivation en chaîne permet quand même de calculer leurs dérivées partielles au point $\mathbf{y}_0$. Dérivant en effet les équations

$$f_i(x_1,x_2,\ldots,x_n)=y_i\ ,\ 1\leq i\leq n$$

par rapport à y_k, on obtient un système linéaire de n équations en n inconnues

$$\frac{\partial x_j}{\partial y_k}\ ,\ 1\leq j\leq n$$

qui s'écrit

$$\sum_{j=1}^{n}\frac{\partial f_i}{\partial x_j}\frac{\partial x_j}{\partial y_k}=0\ ,\ i\neq k$$

et

$$\sum_{j=1}^{n} \frac{\partial f_k}{\partial x_j} \frac{\partial x_j}{\partial y_k} = 1.$$

Comme son déterminant

$$\frac{\partial(f_1, f_2, \ldots, f_n)}{\partial(x_1, x_2, \ldots, x_n)}$$

est, par hypothèse non nul, il admet une solution unique qui peut, par exemple, s'écrire à l'aide de la règle de Cramer — c'est le théorème (21) explicité dans le cas particulier d'une fonction inversible.

8.2 Exemples

Exemple.

Le théorème des fonctions inverses a un caractère local. La fonction $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$ définie par

$$f_1(x_1, x_2) = e^{x_1} \cos x_2 \ , \ f_2(x_1, x_2) = e^{x_1} \sin x_2$$

n'est pas globalement inversible même si $\mathbf{f}'(\mathbf{x})$ est inversible partout. On a en effet

$$\frac{\partial(y_1, y_2)}{\partial(x_1, x_2)} = e^{2x_1} > 0.$$

et le système d'équations

$$y_1 = e^{x_1} \cos x_2 \ , \ y_2 = e^{x_1} \sin x_2$$

admet pour solutions

$$x_1 = \log \sqrt{y_1^2 + y_2^2} \ , \ x_2 = \arctan \frac{y_2}{y_1} + k\pi \ , \ k \in \mathbb{Z}.$$

Exemple.

Les équations

$$y_1 = x_1^2 - x_2^2 \ , \ y_2 = x_1 x_2$$

peuvent être résolues pour x_1 et x_2 au voisinage de tout point $\mathbf{x} \neq \mathbf{0}$ puisque

$$\frac{\partial(y_1, y_2)}{\partial(x_1, x_2)} = 2\left(x_1^2 + x_2^2\right).$$

On a effectivement

$$x_1 = \pm\sqrt{\frac{1}{2}\,y_1 + \sqrt{y_1^2 + 4y_2^2}}\,,\ x_2 = \pm\frac{y_2}{\sqrt{\frac{1}{2}\,y_1 + \sqrt{y_1^2 + 4y_2^2}}}.$$

Exemple.

Supposons que l'équation

$$x^3 + a_1x^2 + a_2x + a_3 = 0$$

admette trois racines distinctes $x_1 < x_2 < x_3$. Alors, si

$$(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2$$

est assez petit, l'équation voisine

$$x^3 + b_1x^2 + b_2x + b_3 = 0$$

admettra elle aussi trois racines distinctes. On a en effet

$$\begin{aligned} a_1 &= -(x_1 + x_2 + x_3) \\ a_2 &= x_1x_2 + x_2x_3 + x_3x_1 \\ a_3 &= -x_1x_2x_3 \end{aligned}$$

donc

$$\frac{\partial(a_1, a_2, a_3)}{\partial(x_1, x_2, x_3)} = -(x_1 - x_2)(x_2 - x_3)(x_3 - x_1) < 0.$$

Le système d'équation précédent établit donc une bijection entre un voisinage du point (x_1, x_2, x_3) (dont on peut supposer qu'il ne coupe aucun des plans $x_1 = x_2, x_2 = x_3$ et $x_3 = x_1$) et un voisinage du point (a_1, a_2, a_3).

Exemple.

Soit $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^n$ une fonction dérivable et telle que

$$\gamma = \sup\{\|\mathbf{g}'(\mathbf{x})\| \mid \mathbf{x} \in \mathbb{R}^n\} < \frac{1}{n}.$$

Alors la fonction $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ définie par

$$\mathbf{f}(\mathbf{x}) = \mathbf{x} + \mathbf{g}(\mathbf{x})$$

est bijective.

Dans le cas où $\mathbf{g}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ est linéaire, cela résulte du lemme suivant :

Lemme 1 *Si* $\|\mathbf{A}\| < 1$, $\mathbf{I} - \mathbf{A}$ *est inversible et*

$$(\mathbf{I} - \mathbf{A})^{-1} = \sum_{k=0}^{+\infty} \mathbf{A}^k.$$

Démonstration.

On a l'identité

$$\mathbf{I} - \mathbf{A}^N = (\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \cdots + \mathbf{A}^{N-1}).$$

Puisque $\|\mathbf{A}^k\| \leq \|\mathbf{A}\|^k$ et puisque la série

$$\sum_{k=0}^{+\infty} \|\mathbf{A}^k\|$$

est convergente, la série

$$\sum_{k=0}^{+\infty} \mathbf{A}^k$$

est elle aussi convergente et, laissant $N \to +\infty$ dans la relation précédente,

$$\mathbf{I} = (\mathbf{I} - \mathbf{A}) \sum_{k=0}^{+\infty} \mathbf{A}^k.$$

C.Q.F.D.

Dans le cas général, la fonction $\mathbf{f}$ est injective car la relation $\mathbf{f}(\mathbf{x}_1) = \mathbf{f}(\mathbf{x}_2)$ implique implique la relation

$$\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{g}(\mathbf{x}_2) - \mathbf{g}(\mathbf{x}_1)$$

et (théorème (19))

$$\|\mathbf{x}_1 - \mathbf{x}_2\| \leq n\gamma\|\mathbf{x}_1 - \mathbf{x}_2\| < \|\mathbf{x}_1 - \mathbf{x}_2\|$$

ce qui est absurde.

Pour montrer que $\mathbf{f}$ est surjective, soit $\mathbf{y} \in \mathbb{R}^n$ un point quelconque et considérons

$$\phi(\mathbf{x}) = \|\mathbf{f}(\mathbf{x}) - \mathbf{y}\|^2.$$

Alors

$$\begin{gathered}\sqrt{\phi(\mathbf{x})} = \|\mathbf{x} + \mathbf{g}(\mathbf{x}) - \mathbf{y}\| \\ \geq \|\mathbf{x}\| - \|\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{0})\| - \|\mathbf{g}(\mathbf{0}) - \mathbf{y}\| \\ \geq \|\mathbf{x}\| - n\gamma\|\mathbf{x}\| - \|\mathbf{g}(\mathbf{0}) - \mathbf{y}\|\end{gathered}$$

de telle sorte que

$$\lim_{\|\mathbf{x}\|\to+\infty} \phi(\mathbf{x}) = +\infty.$$

Il existe donc $\overline{\mathbf{x}} \in \mathbb{R}^n$ tel que

$$\phi(\overline{\mathbf{x}}) = \inf\{\phi(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n\}.$$

On a

$$\phi'(\overline{\mathbf{x}}) = \mathbf{0}$$

c'est-à-dire

$$0 = (\overline{x}_j + g_j(\overline{\mathbf{x}}) - y_j)\left(1 + \frac{\partial g_j}{\partial x_j}(\overline{\mathbf{x}})\right) + \sum_{i\neq j}(\overline{x}_i + g_i(\overline{\mathbf{x}}) - y_i)\,\frac{\partial g_i}{\partial x_j}(\overline{\mathbf{x}})$$

pour $1 \leq j \leq n$. Le système d'équations linéaires

$$\big(\mathbf{I} + \mathbf{g}'(\overline{\mathbf{x}})\big)\,\mathbf{u} = \mathbf{0}$$

admet donc pour solution

$$\mathbf{u} = \overline{\mathbf{x}} + \mathbf{g}(\overline{\mathbf{x}}) - \mathbf{y}.$$

Comme $\|\mathbf{g}'(\overline{\mathbf{x}})\| < 1$, $\mathbf{I} + \mathbf{g}'(\overline{\mathbf{x}})$ est inversible et la solution précédente doit être la solution triviale $\mathbf{u} = \mathbf{0}$:

$$\mathbf{y} = \mathbf{f}(\overline{\mathbf{x}}).$$

8.3 Exercices

Justifier ses réponses.

1. Déterminer les points $\mathbf{x}_0$ au voisinage desquels le système

$$y_1 = x_1^2 + x_2^2\ ,\ y_2 = x_1 x_2$$

   peut être résolu pour $\mathbf{x}$. Calculer explicitement cette solution.

2. Déterminer les points $\mathbf{x}_0$ au voisinage desquels le système

$$\begin{aligned} y_1 &= x_1 + x_2 + \cdots + x_n \\ y_2 &= x_1^2 + x_2^2 + \cdots + x_n^2 \\ &\cdots \\ y_n &= x_1^n + x_2^n + \cdots + x_n^n \end{aligned}$$

   peut être résolu pour $\mathbf{x}$.

3. Montrer que si la matrice

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

est inversible, le système d'équations

$$y_1 = ax_1 + \epsilon\, x_1^2 + bx_2 \ , \ y_2 = cx_1 + dx_2 + \epsilon\, x_2^2$$

peut être résolu pour $\mathbf{x}$ près de l'origine. Obtenir la solution explicite dans le cas où

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} .$$

Spécifier les points $\mathbf{y}$ pour lesquels la solution obtenue est valable. Qu'arrive-t-il lorsque ϵ tend vers 0 ?

4. Montrer que le système

$$y_1 = x_3 \cos x_1 x_2 \ , \ y_2 = x_3 \sin x_1 x_2 \ , \ y_3 = x_1 + x_3$$

peut être résolu pour $\mathbf{x}$ au voisinage de tout point $\mathbf{x}_0$ tel que $x_{0,1} x_{0,3} \neq 0$. Calculer les dérivées partielles au point $\mathbf{y}_0$ des fonctions obtenues lorsque $\mathbf{x}_0 = (1, 0, 1)$.

5. Soit $\mathbf{f}(\mathbf{x}) = (x_1, x_2^3, x_3^5)$. Vérifier que $\mathbf{f}$ est inversible au voisinage de $\mathbf{0}$ bien que $\mathbf{f}'(\mathbf{0})$ ne le soit pas. Explication ?

9 FONCTIONS IMPLICITES

Nous considérons le problème d'éliminer m inconnues d'un système de m équations différentiables en n inconnues : $f(\mathbf{x}) = \mathbf{0}$.

9.1 Le théorème

Dans le cas où $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ est linéaire, $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$, supposons que le rang de $\mathbf{A}$ soit égal à m ; alors (en vertu d'un théorème d'algèbre linéaire), il existe $1 \leq j_1 < j_2 < \cdots < j_m \leq n$ tels que

$$\det \begin{pmatrix} a_{1,j_1} & a_{1,j_2} & \cdots & a_{1,j_m} \\ a_{2,j_1} & a_{2,j_2} & \cdots & a_{2,j_m} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m,j_1} & a_{m,j_2} & \cdots & a_{m,j_m} \end{pmatrix} \neq 0$$

et les inconnues $x_{j_1}, x_{j_2}, \ldots, x_{j_m}$ peuvent être éliminées du système $\mathbf{A}\mathbf{x} = \mathbf{0}$, c'est-à-dire peuvent être exprimées comme fonctions linéaires des autres variables x_j : le système possède $p = n - m$ « degrés de liberté ».

Théorème 23 *Soient $E \subseteq \mathbb{R}^n$ un ensemble ouvert et $\mathbf{f} : E \to \mathbb{R}^m$ une fonction de classe $C^{(1)}$ dans E. Supposons que $\mathbf{f}(\mathbf{x}_0) = \mathbf{0}$ et que*

$$\frac{\partial(f_1, f_2, \ldots, f_m)}{\partial(x_1, x_2, \ldots, x_m)}(\mathbf{x}_0) \neq 0.$$

Alors il existe un voisinage ouvert U de $\mathbf{x}_0$ dans $\mathbb{R}^n$ et un voisinage ouvert W de $(x_{0,m+1}, x_{0,m+2}, \ldots, x_{0,n})$ dans $\mathbb{R}^{n-m}$ et il existe une fonction $\boldsymbol{\phi} : W \to \mathbb{R}^m$ de classe $C^{(1)}$ dans W telle que

$$\{\mathbf{x} \in U \mid \mathbf{f}(\mathbf{x}) = \mathbf{0}\} = \{\mathbf{x} \in U \mid x_i = \phi_i(x_{m+1}, x_{m+2}, \ldots, x_n) \ , \ 1 \leq i \leq m\}.$$

Démonstration.

Introduisons la fonction $\mathbf{g} : E \to \mathbb{R}^n$ en posant

$$\mathbf{g}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x}), x_{m+1}, x_{m+2}, \ldots, x_n).$$

Alors $\mathbf{g} \in C^{(1)}(E)$ et

$$\frac{\partial(g_1, g_2, \ldots, g_n)}{\partial(x_1, x_2, \ldots, x_n)}(\mathbf{x}) = \frac{\partial(f_1, f_2, \ldots, f_m)}{\partial(x_1, x_2, \ldots, x_m)}(\mathbf{x}).$$

Ainsi $\mathbf{g}(\mathbf{x}_0) = (0, 0, \ldots, 0, x_{0,m+1}, x_{0,m+2}, \ldots, x_{0,n})$ et $\mathbf{g}'(\mathbf{x}_0)$ est inversible.

En vertu du théorème (22), il existe un voisinage ouvert U de $\mathbf{x}_0$ dans $\mathbb{R}^n$ et un voisinage ouvert V de $(0, 0, \ldots, 0, x_{0,m+1}, x_{0,m+2}, \ldots, x_{0,n})$ dans $\mathbb{R}^n$ que nous pouvons choisir de la forme

$$V =]-\delta, \delta[^m \times W$$

où W est un pavé ouvert centré en $(x_{0,m+1}, x_{0,m+2}, \ldots, x_{0,n})$ dans $\mathbb{R}^{n-m}$ tels que la fonction $\mathbf{g}$ établisse une bijection entre U et V, la fonction inverse $\mathbf{g}^{-1}$ étant de classe $C^{(1)}$ dans V.

Écrivons alors que les relations

$$\begin{gathered} y_i = f_i(x_1, x_2, \ldots, x_n)\ ,\ 1 \leq i \leq m \\ y_i = x_i\ ,\ m+1 \leq i \leq n \end{gathered}$$

sont équivalentes aux relations

$$\begin{gathered} x_i = g_i^{-1}(y_1, y_2, \ldots, y_n)\ ,\ 1 \leq i \leq m \\ x_i = y_i\ ,\ m+1 \leq i \leq n. \end{gathered}$$

Alors dire que $\mathbf{x} \in U$ revient à dire que

$$x_i = g_i^{-1}(y_1, y_2, \ldots, y_m, x_{m+1}, x_{m+2}, \ldots, x_n)\ ,\ 1 \leq i \leq m$$

et ajouter que $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ équivaut à écrire

$$\begin{aligned} x_i &= g_i^{-1}(0, 0, \ldots, 0, x_{m+1}, x_{m+2}, \ldots, x_n) \\ &= \phi_i(x_{m+1}, x_{m+2}, \ldots, x_n)\ ,\ 1 \leq i \leq m. \end{aligned}$$

C.Q.F.D.

Comme dans le théorème (22), il est illusoire, sauf dans les cas les plus simples, d'espérer trouver explicitement les fonctions

$$x_i = \phi_i(x_{m+1}, x_{m+2}, \ldots, x_n)\ ,\ 1 \leq i \leq m$$

à partir du système

$$f_i(x_1, x_2, \ldots, x_n) = 0\ ,\ 1 \leq i \leq n.$$

La règle de dérivation en chaîne permet quand même de calculer leurs dérivées partielles au point $(x_{0,m+1}, x_{0,m+2}, \ldots, x_{0,n})$. Dérivant les relations

$$f_i(\phi_1(x_{m+1}, x_{m+2}, \ldots, x_n), \ldots, \phi_m(x_{m+1}, x_{m+2}, \ldots, x_n), x_{m+1}, \ldots, x_n) = 0$$

$(1 \leq i \leq m)$ par rapport aux variables « indépendantes » x_j $(m+1 \leq j \leq n)$, on obtient un système de m équations linéaires en m inconnues

$$\frac{\partial \phi_1}{\partial x_j}, \frac{\partial \phi_2}{\partial x_j}, \ldots, \frac{\partial \phi_m}{\partial x_j}$$

qui s'écrit

$$\sum_{k=1}^{m} \frac{\partial f_i}{\partial x_k} \frac{\partial \phi_k}{\partial x_j} = -\frac{\partial f_i}{\partial x_j}\ ,\ 1 \leq i \leq m. \tag{1}$$

Comme son déterminant

$$\frac{\partial(f_1, f_2, \ldots, f_m)}{\partial(x_1, x_2, \ldots, x_m)}$$

est par hypothèse non nul, il admet une solution unique.

Le théorème précédent n'a, bien sûr, été énoncé en termes des variables $x_1, x_2, \ldots, x_m$ que pour des raisons de commodité d'écriture. L'hypothèse

$$\frac{\partial(f_1, f_2, \ldots, f_m)}{\partial(x_1, x_2, \ldots, x_m)}(\mathbf{x}_0) \neq 0$$

remplacée par l'hypothèse

$$\frac{\partial(f_1, f_2, \ldots, f_m)}{\partial(x_{j_1}, x_{j_2}, \ldots, x_{j_m})}(\mathbf{x}_0) \neq 0$$

permettra de résoudre pour les variables $x_{j_1}, x_{j_2}, \ldots, x_{j_m}$ en termes des autres variables x_j — il suffit de renuméroter les variables pour le voir.

9.2 Exemples

Lorsque $m = 1$, l'**hypersurface de niveau** $f(\mathbf{x}) = 0$ peut être considérée comme le graphe d'une fonction

$$x_n = \phi_n(x_1, x_2, \ldots, x_{n-1})$$

au voisinage de tout point $\mathbf{x}_0$ tel que

$$f(\mathbf{x}_0) = 0\ ,\ \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \neq 0.$$

On a alors

$$\frac{\partial \phi_n}{\partial x_j} = -\frac{\dfrac{\partial f}{\partial x_j}}{\dfrac{\partial f}{\partial x_n}}.$$

Exemple.

Considérons l'hyperbole $x_1^2 - x_2^2 = 1$. Dans le voisinage U du point $(1, 0)$ qu'est le demi-plan $x_1 > 0$, on peut écrire de façon équivalente

$$x_1 = \sqrt{1 + x_2^2}\ ,\ x_2 \in \mathbb{R}\ (= W).$$

Dans le voisinage U du point $(2, \sqrt{3})$ constitué par le premier quadrant $(x_1 > 0, x_2 > 0)$, on peut écrire

$$x_1 = \sqrt{1 + x_2^2}\ ,\ x_2 \in \mathbb{R}\ (= W)$$

ou

$$x_2 = \sqrt{x_1^2 - 1}\ ,\ x_1 \in]1, +\infty[\ (= W).$$

Exemple.

Considérons l'ellipsoïde

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} + \frac{x_3^2}{c^2} = 1.$$

Dans le demi-espace $x_3 < 0$, voisinage U d'un de ses points $\mathbf{x}_0$ tel que $x_{0,3} < 0$, l'équation qui la définit est équivalente à

$$x_3 = -\,c\sqrt{1 - \frac{x_1^2}{a^2} - \frac{x_2^2}{b^2}} \quad \text{pour} \quad \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} < 1.$$

(ici W est l'ellipse $x_1^2/a^2 + x_2^2/b^2 < 1$).

9.3 Exercices

Justifier ses réponses.

1. On considère l'équation

$$x_1x_2 - x_3 \log x_2 + e^{x_1x_2} - 1 = 0$$

au voisinage du point $(0, 1, 2)$. Peut-on l'y résoudre pour x_1 ? pour x_2 ? pour x_3 ?

2. On considère le système d'équations

$$x_1 + x_2 + x_3 = 1\ ,\ x_1^2 + x_2^2 + x_3^2 = 1$$

au voisinage du point $(1, 0, 0)$. Quelles variables peut-on y éliminer ?

3. On considère le système d'équations

$$x_1^2 - 2x_2x_4 + x_3^2 = 0 \ , \ x_1^3 - x_2^3 + x_3^3 + x_4^3 = 0$$

au voisinage du point $(-1, 1, 1, 1)$. Vérifier que l'on peut l'y résoudre pour x_1 et x_3 et calculer les dérivées partielles

$$\frac{\partial x_1}{\partial x_2} \ , \ \frac{\partial x_1}{\partial x_4} \ , \ \frac{\partial x_3}{\partial x_2} \ , \ \frac{\partial x_3}{\partial x_4}.$$

4. On considère le système d'équations

$$\begin{aligned} x_1 + 2x_2 + 3x_3 + 10x_4 &= 0 \\ 4x_1 + 5x_2 + 6x_3 + x_4^2 &= 0 \\ 7x_1 + 8x_2 + 9x_3 + x_4^3 &= 0 \end{aligned}$$

au voisinage de l'origine. Vérifier qu'elles déterminent x_1, x_2 et x_4 comme fonctions de x_3 et calculer la dérivée à l'origine de chacune de ces fonction.

5. Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction de classe $C^{(2)}$ telle que $f(\mathbf{x}_0) = 0$ et que

$$\frac{\partial f}{\partial x_2}(\mathbf{x}_0) \neq 0.$$

Calculer

$$\frac{d^2 x_2}{dx_1^2}$$

le long de la **courbe de niveau** $f(\mathbf{x}) = 0$ au voisinage de $\mathbf{x}_0$.

6. Soit $f : \mathbb{R}^3 \to \mathbb{R}$ une fonction de classe $C^{(1)}$ telle que $f(\mathbf{x}_0) = 0$ et que

$$\frac{\partial f}{\partial x_1}(\mathbf{x}_0)\frac{\partial f}{\partial x_2}(\mathbf{x}_0)\frac{\partial f}{\partial x_3}(\mathbf{x}_0) \neq 0.$$

Alors l'équation $f(\mathbf{x}) = 0$ détermine chaque variable x_j comme fonction des deux autres au voisinage du point $\mathbf{x}_0$. Montrer que

$$\frac{\partial x_1}{\partial x_2}\frac{\partial x_2}{\partial x_3}\frac{\partial x_3}{\partial x_1} = -1.$$

7. Déduire le théorème des fonctions inverse du théorème des fonctions implicites.

10 OPTIMISATION SOUS CONTRAINTES

Nous considérons le problème d'optimiser une fonction de n variables $f(\mathbf{x})$ sous m contraintes différentiables $\mathbf{g}(\mathbf{x}) = \mathbf{0}$.

10.1 Variétés différentiables

Un ensemble $M \subseteq \mathbb{R}^n$ est une **variété différentiable de dimension** p si l'on peut, au moins localement, le paramétrer par p variables indépendantes. De façon précise, on suppose qu'à chaque $\mathbf{x}_0 \in M$ correspondent un voisinage ouvert $U \subseteq \mathbb{R}^n$ de $\mathbf{x}_0$, un ensemble ouvert $T \subseteq \mathbf{R}^p$ et une fonction $\mathbf{h} : T \to \mathbb{R}^n$ de classe $C^{(1)}$ tels que

$$M \cap U = \{\mathbf{x} \mid \mathbf{x} = \mathbf{h}(\mathbf{t}) \ , \ \mathbf{t} \in T\}$$

et que le rang de $\mathbf{h}'(\mathbf{t})$ est égal à p en tout point $\mathbf{t} \in T$. Cette dernière condition entraîne que l'on peut localement obtenir les paramètres $t_1, t_2, \ldots, t_p$ en termes de coordonnées $x_{j_1}, x_{j_2}, \ldots, x_{j_p}$ et signifie que M est vraiment localement « à p dimensions ».

Si $\mathbf{h} : T \to \mathbb{R}^n$ est un paramétrage d'une portion de M et $\psi : S \to T$ est un difféomorphisme de classe $C^{(1)}$ entre S et T, alors $\mathbf{h}_1 = \mathbf{h} \circ \psi$ en est un autre paramétrage possible.

Soit $\mathbf{x}_0 = \mathbf{h}(\mathbf{t}_0)$ un point d'une variété M de dimension p. L'**espace tangent** à M en $\mathbf{x}_0$ est le translaté par $\mathbf{x}_0$ de l'espace

$$T_M(\mathbf{x}_0) = \left\{ \mathbf{u} \mid \mathbf{u} = \sum_{k=1}^{p} \mu_k \, \frac{\partial \mathbf{h}}{\partial t_k}(\mathbf{t}_0) \right\} .$$

C'est un espace de dimension p qui ne dépend pas du paramétrage de M au voisinage de $\mathbf{x}_0$ choisi.

L'**espace normal** à M en $\mathbf{x}_0$ est le translaté par $\mathbf{x}_0$ du complémentaire orthogonal de $T_M(\mathbf{x}_0)$:

$$N_M(\mathbf{x}_0) = \left\{ \mathbf{v} \mid \mathbf{v} = \sum_{i=1}^{m} \lambda_i \, \mathbf{n}_i \right\} ,$$

où les vecteurs $\mathbf{n}_1, \mathbf{n}_2, \ldots, \mathbf{n}_m$ sont des vecteurs linéairement indépendants tels que

$$\mathbf{n}_i \cdot \frac{\partial \mathbf{h}}{\partial t_k}(\mathbf{t}_0) = 0 \quad \text{pour tout} \quad 1 \leq i \leq m \ , \ 1 \leq k \leq p.$$

Une variété M peut être définie de façon implicite.

Soient $E \subseteq \mathbb{R}^n$ un ensemble ouvert et $\mathbf{g} : E \to \mathbb{R}^m$ une fonction de classe $C^{(1)}$ telle que le rang de $\mathbf{g}'(\mathbf{x})$ soit égal à m en tout point. Alors

$$M = \{\mathbf{x} \mid \mathbf{g}(\mathbf{x}) = \mathbf{0}\}$$

est une variété de dimension $p = n - m$. Cela découle du théorème des fonctions implicites. À chaque $\mathbf{x}_0 \in M$ correspond un voisinage ouvert $U \subseteq \mathbb{R}^n$ de $\mathbf{x}_0$ et des indices $j_1, j_2, \ldots, j_m$ tels que

$$\frac{\partial(g_1, g_2, \ldots, g_m)}{\partial(x_{j_1}, x_{j_2}, \ldots, x_{j_m})}(\mathbf{x}) \neq 0 \ \text{ pour tout } \ \mathbf{x} \in U$$

et que, par conséquent, $x_{j_1}, x_{j_2}, \ldots, x_{j_m}$ peuvent s'écrire comme fonctions continûment dérivables des autres coordonnées x_j. Lorsque $j_1 = 1, j_2 = 2, \ldots, j_m = m$ par exemple, on a

$$x_j = \phi_j(x_{m+1}, x_{m+2}, \ldots, x_n) \ , \ (x_{m+1}, x_{m+2}, \ldots, x_n) \in W$$

W étant un voisinage ouvert de $(x_{0,m+1}, x_{0,m+2}, \ldots, x_{0,n})$ dans $\mathbb{R}^{n-m}$. Un paramétrage local de M est alors $(t_1 = x_{m+1}, t_2 = x_{m+2}, \ldots, t_p = x_n)$

$$h_j(x_{m+1}, x_{m+2}, \ldots, x_n) = \begin{cases} \phi_j(x_{m+1}, x_{m+2}, \ldots, x_n) & \text{pour } \ 1 \leq j \leq m, \\ x_j & \text{pour } \ m+1 \leq j \leq n. \end{cases}$$

On a

$$\mathbf{h}'(x_{m+1}, x_{m+2}, \ldots, x_n) = \begin{pmatrix} \dfrac{\partial \phi_1}{\partial x_{m+1}} & \dfrac{\partial \phi_1}{\partial x_{m+2}} & \cdots & \dfrac{\partial \phi_1}{\partial x_n} \\ \dfrac{\partial \phi_2}{\partial x_{m+1}} & \dfrac{\partial \phi_2}{\partial x_{m+2}} & \cdots & \dfrac{\partial \phi_2}{\partial x_n} \\ \vdots & \vdots & \cdots & \vdots \\ \dfrac{\partial \phi_m}{\partial x_{m+1}} & \dfrac{\partial \phi_m}{\partial x_{m+2}} & \cdots & \dfrac{\partial \phi_m}{\partial x_n} \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Les équations définissant les vecteurs normaux $\mathbf{n}_i$ s'écrivent ici

$$\sum_{j=1}^{m} n_{i,j} \frac{\partial \phi_j}{\partial x_{m+k}} + n_{i,m+k} = 0.$$

Comme (équation (1))

$$\sum_{j=1}^{m} \frac{\partial g_i}{\partial x_j} \frac{\partial \phi_j}{\partial x_{m+k}} = -\frac{\partial g_i}{\partial x_{m+k}},$$

on doit avoir

$$\mathbf{n}_i = \mathbf{grad}\, g_i(\mathbf{x}_0)\ ,\ 1 \leq i \leq m$$

et donc

$$N_M(\mathbf{x}_0) = \left\{ \mathbf{v} \mid \mathbf{v} = \sum_{i=1}^{m} \lambda_i\, \mathbf{grad}\, g_i(\mathbf{x}_0) \right\}.$$

Cette description, établie pour le cas où $j_1 = 1, j_2 = 2, \ldots, j_m = m$, reste bien entendu valable pour $j_1, j_2, \ldots, j_m$ quelconque.

10.2 Exemples

Une hypersurface de niveau

$$M = \{\mathbf{x} \mid g(\mathbf{x}) = c\}$$

est une variété de dimension $n-1$ si $\mathbf{grad}\, g(\mathbf{x})$ ne s'annule pas sur M. L'espace normal en $\mathbf{x}_0$ est la droite passant par $\mathbf{x}_0$ et portée par le vecteur $\mathbf{grad}\, g(\mathbf{x}_0)$:

$$\mathbf{x}_0 + \lambda\, \mathbf{grad}\, g(\mathbf{x}_0)\ ,\ -\infty < \lambda < \infty.$$

L'espace tangent est l'hyperplan orthogonal à cette droite :

$$\{\mathbf{x} \mid \mathbf{grad}\, g(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) = 0\}.$$

Exemple.

Une sphère

$$S = \{\mathbf{x} \mid (\mathbf{x} - \mathbf{a})^T (\mathbf{x} - \mathbf{a}) = r^2\}$$

est une variété de dimension $n-1$. Si $\|\mathbf{x}_0 - \mathbf{a}\| = r$, soit $x_{0,k} - a_{0,k} \neq 0$, par exemple, $x_{0,k} - a_{0,k} > 0$. Alors, au voisinage de $\mathbf{x}_0$, S peut être paramétrée par les coordonnées x_j telles que $j \neq k$:

$$x_k = a_k + \sqrt{r^2 - \sum_{j \neq k} (x_j - a_j)^2} \quad \text{si} \quad \sum_{j \neq k} (x_j - a_j)^2 < r^2.$$

L'espace normal en $\mathbf{x}_0$ est la droite

$$\mathbf{x}_0 + \lambda(\mathbf{x}_0 - \mathbf{a})\ ,\ -\infty < \lambda < \infty.$$

L'espace tangent est l'hyperplan d'équation

$$(\mathbf{x}_0 - \mathbf{a}) \cdot (\mathbf{x} - \mathbf{x}_0) = 0.$$

Une **courbe paramétrée** C est définie par une fonction $\mathbf{h} :]a, b[\to \mathbb{R}^n$ de classe $C^{(1)}$ dont la dérivée $\mathbf{h}'(t)$ ne s'annule pas :

$$C = \{\mathbf{x} \mid \mathbf{x} = \mathbf{h}(t)\}.$$

C'est une variété de dimension 1. L'espace tangent en $\mathbf{x}_0 = \mathbf{h}(t_0)$ est la droite

$$\mathbf{x}_0 + \mu\, \mathbf{h}'(t_0).$$

L'espace normal est l'hyperplan d'équation

$$\mathbf{h}'(t_0) \cdot (\mathbf{x} - \mathbf{x}_0) = 0.$$

Aux **points de rebroussement** où $\mathbf{h}'(t_0) = \mathbf{0}$, la « courbe » n'admettrait pas de tangente.

Exemple.
La fonction $\mathbf{h} :] - 1, 1[\to \mathbb{R}^2$,

$$\mathbf{h}(t) = (1 + t^3, 1 - t^{4/3}),$$

ne définit pas une variété car il y a un point de rebroussement lorsque $t = 0$.

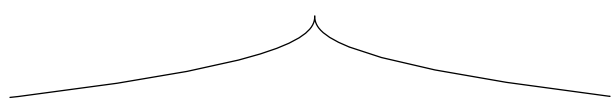

FIG. 9 – Un point de rebroussement

Une **surface paramétrée** S dans $\mathbb{R}^3$ est définie par une fonction $\mathbf{h} : T \to \mathbf{R}^3$ de classe $C^{(1)}$ dans l'ensemble ouvert $T \subseteq \mathbb{R}^2$ telle que les vecteurs

$$\frac{\partial \mathbf{h}}{\partial t_1}(\mathbf{t}) \text{ et } \frac{\partial \mathbf{h}}{\partial t_2}(\mathbf{t})$$

sont linéairement indépendants en tout point $\mathbf{t} \in T$:

$$S = \{\mathbf{x} \mid \mathbf{x} = \mathbf{h}(\mathbf{t})\}.$$

C'est une variété de dimension 2 dont l'espace $T_M(\mathbf{x}_0)$ est engendré par les vecteurs précédents évalués en $\mathbf{t}_0$. L'espace $N_M(\mathbf{x}_0)$ consiste alors des multiples de

$$\mathbf{n}(\mathbf{t}_0) = \frac{\partial \mathbf{h}}{\partial t_1}(\mathbf{t}_0) \times \frac{\partial \mathbf{h}}{\partial t_2}(\mathbf{t}_0).$$

L'intersection de deux telles surfaces S_1 et S_2 est une variété de dimension 1 si leurs normales

$$\mathbf{n}_1(\mathbf{x}_0) \text{ et } \mathbf{n}_2(\mathbf{x}_0)$$

sont linéairement indépendantes en tout point $\mathbf{x}_0 \in S_1 \cap S_2$. L'espace tangent à $S_1 \cap S_2$ en $\mathbf{x}_0$ est la droite

$$\mathbf{x}_0 + \mu\, \mathbf{n}_1(\mathbf{x}_0) \times \mathbf{n}_1(\mathbf{x}_0)\ ,\ -\infty < \mu < \infty.$$

Exemple.

La normale au point $\mathbf{x}_0$ au plan d'équation

$$ax_1 + bx_2 + cx_3 = 1\ ,\ a > b > c > 0,$$

est

$$\mathbf{n}_P(\mathbf{x}_0) = \mathbf{n}_P = (a, b, c)$$

et celle à la sphère

$$x_1^2 + x_2^2 + x_3^2 = r^2$$

est

$$\mathbf{n}_S(\mathbf{x}_0) = \mathbf{x}_0.$$

On a

$$\mathbf{n}_P \times \mathbf{n}_S(\mathbf{x}_0) \neq \mathbf{0}$$

sauf dans le cas où plan et sphère sont tangents en $\mathbf{x}_0$ — le plan est alors l'espace tangent à la sphère. Sauf pour ce cas exceptionnel, l'intersection du plan et de la sphère est un cercle. Son centre est le point du plan qui est le plus près de l'origine, c'est-à-dire

$$\mathbf{c} = \frac{\mathbf{n}_P}{\|\mathbf{n}_P\|^2}$$

et son rayon est

$$r = \sqrt{r^2 - \frac{1}{\|\mathbf{n}_P\|^2}}.$$

10.3 Extremums liés

Théorème 24 *Soient* $E \subseteq \mathbb{R}^n$ *un ensemble ouvert,* $f : E \to \mathbb{R}$ *une fonction numérique dérivable définie sur* E *et* $M \subseteq E$ *une variété différentiable. Une condition nécessaire pour que la restriction* f/M *de* f *à* M *admette un extremum local en* $\mathbf{x}_0 \in M$ *est que* $\mathbf{grad} f(\mathbf{x}_0) \in N_M(\mathbf{x}_0)$*.*

Démonstration.

Soit $\mathbf{h} :]-\delta, \delta[^p \to \mathbb{R}^n$ un paramétrage de M au voisinage de $\mathbf{x}_0$ tel que $\mathbf{x}_0 = \mathbf{h}(\mathbf{0})$. Soit $\mathbf{u} \in T_M(\mathbf{x}_0)$ quelconque. Si

$$\mathbf{u} = \sum_{k=1}^{p} \mu_k \frac{\partial \mathbf{h}}{\partial t_k}(\mathbf{0}),$$

considérons sur M la courbe paramétrée $\mathbf{c} :]-\eta, \eta[\to \mathbb{R}^n$ définie par

$$\mathbf{c}(s) = \mathbf{h}(s\,\mu_1, s\,\mu_2, \ldots, s\,\mu_p).$$

Pour cette courbe, $\mathbf{c}(0) = \mathbf{x}_0$ et $\mathbf{c}'(0) = \mathbf{u}$. La fonction $s \mapsto f(\mathbf{c}(s))$ admettant en 0 un extremum local, sa dérivée doit s'annuler, ce qui donne

$$\sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\mathbf{x}_0) \frac{dc_j}{ds}(0) = \mathbf{grad} f(\mathbf{x}_0) \cdot \mathbf{u} = 0.$$

Ainsi $\mathbf{grad} f(\mathbf{x}_0) \in N_M(\mathbf{x}_0)$. C.Q.F.D.

Chercher les extremums liés d'une fonction $f(\mathbf{x})$ sous m contraintes

$$g_1(\mathbf{x}) = g_2(\mathbf{x}) = \cdots = g_m(\mathbf{x}) = 0$$

revient donc en pratique à chercher les extremums libres de la **fonction de Lagrange**

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \, g_i(\mathbf{x}).$$

Les conditions nécessaires

$$\frac{\partial L}{\partial x_1} = \frac{\partial L}{\partial x_2} = \cdots = \frac{\partial L}{\partial x_n} = 0$$

correspondent à la condition d'orthogonalité

$$\mathbf{grad} f = \sum_{i=1}^{m} \lambda_i \, \mathbf{grad}\, g_i$$

et les conditions nécessaires

$$\frac{\partial L}{\partial \lambda_1} = \frac{\partial L}{\partial \lambda_2} = \cdots = \frac{\partial L}{\partial \lambda_m} = 0$$

correspondent aux contraintes

$$g_1 = g_2 = \cdots = g_m = 0.$$

Exemple.

Déterminer la distance d'un point donné $\mathbf{x}_0$ à une variété M donnée de façon implicite,

$$M = \{\mathbf{x} \mid \mathbf{g}(\mathbf{x}) = 0\},$$

revient à minimiser la fonction

$$f(\mathbf{x}) = \|\mathbf{x} - \mathbf{x}_0\|^2$$

sous les contraintes

$$g_1(\mathbf{x}) = g_2(\mathbf{x}) = \cdots = g_m(\mathbf{x}) = 0$$

et conduit aux équations de Lagrange

$$2(\mathbf{x} - \mathbf{x}_0) = \sum_{i=1}^{m} \lambda_i \,\mathbf{grad}\, g_i(\mathbf{x})$$
$$\mathbf{g}(\mathbf{x}) = \mathbf{0}.$$

Dans le cas le plus simple, $m = 1$ et $g(\mathbf{x}) = \mathbf{a} \cdot (\mathbf{x} - \mathbf{b})$, l'unique solution de ces équations est

$$\mathbf{x} = \mathbf{x}_0 + \frac{\mathbf{a} \cdot (\mathbf{b} - \mathbf{x}_0)}{\|\mathbf{a}\|^2}\, \mathbf{a}$$

ce qui conduit au minimum global de la distance

$$\frac{|\mathbf{a} \cdot (\mathbf{b} - \mathbf{x}_0)|}{\|\mathbf{a}\|}.$$

Exemple.

Pour déterminer les valeurs extrêmes de la fonction quadratique

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$$

sur la sphère unité

$$S = \{\mathbf{x} \mid \|\mathbf{x}\| = 1\},$$

les équations de Lagrange sont

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \quad , \quad \mathbf{x}^T\mathbf{x} = 1.$$

Les multiplicateurs de Lagrange sont donc les valeurs propres

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$$

de la matrice $\mathbf{A}$ et les points critiques sont les vecteurs propres normalisés $\mathbf{x}_j$ associés. Comme

$$f(\mathbf{x}_j) = \lambda_j,$$

on a

$$\inf\{f(\mathbf{x}) \mid \|\mathbf{x}\| = 1\} = \lambda_1,$$
$$\sup\{f(\mathbf{x}) \mid \|\mathbf{x}\| = 1\} = \lambda_n$$

et

$$\lambda_1 \|\mathbf{x}\|^2 \leq f(\mathbf{x}) \leq \lambda_n \|\mathbf{x}\|^2.$$

10.4 Exercices

Justifier ses réponses.

1. Vérifier que l'espace $T_M(\mathbf{x}_0)$ est indépendant du choix des paramètres au voisinage de $\mathbf{x}_0$.
2. Déterminer espace tangent et espace normal au cercle défini paramétriquement par
$$x_1 = r\cos\theta_1 \sin t\ ,\ x_2 = r\sin\theta_1 \sin t\ ,\ x_3 = r\cos t$$
(r et θ_1 sont fixés).
3. Déterminer espace tangent et espace normal au cercle défini implicitement par
$$x_1 + x_2 + x_3 = 1\ ,\ x_1^2 + x_2^2 + x_3^2 = 1.$$
4. Montrer que le graphe d'une fonction dérivable $f :]a, b[\to \mathbb{R}$,
$$\{(x_1, x_2) \mid x_2 = f(x_1)\ ,\ a < x_1 < b\},$$
est une courbe différentiable sans point de rebroussement dans $\mathbb{R}^2$.

5. Les **coordonnées cylindriques** ρ, ϕ et z sur $\mathbb{R}^3$ sont définies par

$$x_1 = \rho \cos\phi \ , \ x_2 = \rho \sin\phi \ , \ x_3 = z$$

$(\rho > 0 \ , \ -\pi < \phi \leq \pi \ , \ -\infty < z < \infty)$.

– On considère le **cylindre**

$$S = \{\mathbf{x} = (\cos\phi, \sin\phi, z)\}.$$

Calculer les vecteurs tangents

$$\frac{\partial \mathbf{x}}{\partial \phi} \ \text{ et } \ \frac{\partial \mathbf{x}}{\partial z}$$

et vérifier que

$$\frac{\partial \mathbf{x}}{\partial \phi} \times \frac{\partial \mathbf{x}}{\partial z} \neq \mathbf{0}.$$

– Considérons ensuite l'**hélice**

$$C = \{\mathbf{x} = (\cos t, \sin t, h\,t)\}$$

($h > 0$ est le **pas** de l'hélice). Exprimer le vecteur tangent

$$\frac{\partial \mathbf{x}}{\partial t}$$

en un point $\mathbf{x}_0$ de la courbe C en terme des vecteurs

$$\frac{\partial \mathbf{x}}{\partial \phi} \ \text{ et } \ \frac{\partial \mathbf{x}}{\partial z}$$

au même point.

6. Considérons la sphère

$$S = \{\mathbf{x} = (\cos\theta_1 \sin\theta_2, \sin\theta_1 \sin\theta_2, \cos\theta_2)\}.$$

– Vérifier que

$$\frac{\partial \mathbf{x}}{\partial \theta_1} \times \frac{\partial \mathbf{x}}{\partial \theta_2} = -\sin\theta_2\,\mathbf{x}.$$

– Vérifier que les **méridiens** (les courbes $\theta_1 = c_1$) et **les parallèles** (les courbes $\theta_2 = c_2$) se coupent bien à angle droit.

7. Soient $m, n, p \in \mathbb{N}$. Déterminer le maximum de l'expression

$$x_1^{2m+1} x_2^{2n+1} x_3^{2p+1}$$

sous la contrainte $x_1 + x_2 + x_3 = 1$.

8. Déterminer le minimum de l'expression
$$\int_{-1}^{1}(\sin\pi\, t - a - b\, t - c\, t^2)^2 dt$$
sous la contrainte $a + b + c = 0$.
9. Soient $0 < a_1 < a_2 < \cdots < a_n$, $0 < b_1 < b_2 < \cdots < b_n$,
$$\mathbf{A} = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix} \quad \text{et} \quad \mathbf{B} = \begin{pmatrix} b_1 & 0 & \cdots & 0 \\ 0 & b_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & b_n \end{pmatrix}.$$
Déterminer
$$\sup\{\mathbf{x}^T\mathbf{A}\mathbf{x} \mid \mathbf{x}^T\mathbf{B}\mathbf{x} = 1\}$$
et
$$\inf\{\mathbf{x}^T\mathbf{A}\mathbf{x} \mid \mathbf{x}^T\mathbf{B}\mathbf{x} = 1\}.$$
10. La formule de Héron pour l'aire du triangle de côtés a, b et c est
$$\sqrt{s(s-a)(s-b)(s-c)}$$
où
$$s = \frac{a+b+c}{2}.$$
Pour quel triangle cette aire est-elle maximale ?
11. L'**entropie** d'une distribution de probabilité est
$$-\sum_{j=1}^{n} p_j \log p_j.$$
(C'est une mesure du degré d'aléatoire de la distribution). Pour quelle distribution de probabilité est-elle maximale ?
12. La **moyenne arithmétique** de n nombres positifs est
$$A = \frac{1}{n}\sum_{j=1}^{n} x_j$$
et leur **moyenne géométrique** est
$$G = \sqrt[n]{x_1 x_2 \cdots x_n}.$$
Montrer que
$$G \leq A$$
en précisant le cas d'égalité.

Références

[1] Srishti D. Chatterji. *Cours d'Analyse 1.* Presses polytechniques et univrsitaires romandes, Lausane, 1997.
Manuel de premier cycle
Très complet, exercices
Une deuxième partie porte sur les intégrales multiples.

[2] Walter Rudin. *Principes d'analyse mathématique.* Ediscience, Paris, 1995.
Manuel de premier cycle, exercices
Un chapitre résume le cours
Math-Info QA 300 R 8212 1995.

Index