

CALCUL DIFFÉRENTIEL ET OPTIMISATION

Yves Colin de Verdière
Institut Fourier (Grenoble)
Année 98-99

Table des matières

INTRODUCTION	7
1. Remarques générales	7
2. Introduction historique et motivations	7
3. Contenu	9
4. Notations	10
5. Annexe 1 : normes d'applications linéaires et bilinéaires	10
6. Annexe 2 : rappels sur les formes quadratiques	11
7. Annexe 3 : topologies faibles	12
8. Annexe 4 : notations o et O	12
Bibliographie	15
Chapitre 1. DIFFÉRENTIELLES	17
1. Introduction	17
2. Différentielle d'une application	18
3. Dérivées directionnelles	23
4. Champs de vecteurs	25
5. Matrices jacobiniennes	26
6. Notations dx_i	27
7. Accroissements finis	28
8. Convergence C^1 d'une suite d'applications	31
Chapitre 2. DÉRIVÉES D'ORDRE ≥ 2	33
1. Introduction	33
2. Différentielles secondes	33
3. Fonctions C^2	35
4. Forme quadratique associée à la différentielle seconde	37
5. Différentielles d'ordre arbitraire et dérivées partielles	37
6. Formules de Taylor à une variable	38
7. Interpolation	39
8. Formule de Taylor à plusieurs variables	40
9. Réciproque du lemme de Schwarz	40
Chapitre 3. SYSTÈMES D'ÉQUATIONS NON-LINÉAIRES	43
1. Rappel sur les systèmes linéaires en dimension finie (stabilité structurelle)	43

2.	Méthode de point fixe	44
3.	Inversion locale	46
4.	Théorèmes des fonctions implicites	48
5.	Stabilité structurelle des points fixes	49
6.	Sous-variétés de \mathbb{R}^n	49
7.	Calculs pratiques	52
8.	Résolution d'un système d'équations et méthode de Newton	54
Chapitre 4. EXTRÊMAS DES FONCTIONS NUMÉRIQUES		59
1.	Existence et topologie	59
2.	Conditions du 1er ordre	60
3.	Conditions du second ordre	61
4.	Extrêmas liés	62
5.	Catastrophes	64
6.	Minimax	64
Chapitre 5. CONVEXITÉ		67
1.	Définitions	67
2.	Continuité	68
3.	Opérations	69
4.	Convexité et dérivées secondes	69
5.	Minimas	71
6.	Convexité forte	71
7.	Théorème de projection	73
8.	Théorème de Hahn-Banach	74
9.	Algorithme de relaxation	75
10.	Algorithme de gradient à pas fixe	76
11.	Le cas des formes quadratiques	77
12.	Programmation linéaire	78
Chapitre 6. CALCUL DES VARIATIONS		83
1.	Exemples	83
2.	Equation d'Euler-Lagrange	85
3.	Théorème de Hilbert	86
4.	Brachistochrone	87
5.	Surface minimales de révolution	88
6.	Géodésiques	90
7.	Problème isopérimétrique en dimension 2	91
8.	Problème de Dirichlet et méthode des éléments finis	91
Chapitre 7. QUELQUES PROBLÈMES		93
1.	Exercices sur le chapitre 1	93
2.	Exercices sur le chapitre 2	96
3.	Exercices sur le chapitre 3	96

4. Exercices sur le chapitre 4	97
5. Exercices sur le chapitre 5	100
6. Problème	103
7. Problème	103
8. Examen de mai 1997	104
9. Examen de septembre 1997	105
10. Examen de juin 1998	107
11. Examen de septembre 1998	108
12. Partiel de mars 1999	109
13. Examen de juin 1999	111

INTRODUCTION

1. Remarques générales

Ce cours polycopié est un ensemble de notes du cours de la licence de mathématiques de l'Université de Grenoble 1, intitulé *Calcul différentiel et optimisation* (LO6) enseigné de 97 à 99.

Comme le lecteur le découvrira vite, les sujets sont traités avec une profondeur très variable et le temps consacré au cours oral n'a été en rien proportionnel à la longueur des notes.

Les étudiants ont été avertis que ces notes ne se substituent pas au cours oral, mais sont complémentaires, permettant en particulier de retrouver des énoncés précis, les notations utilisées ou des compléments et exemples non traités oralement.

Les prérequis sont essentiellement l'algèbre linéaire (et bilinéaire !) du premier cycle et la topologie des espaces métriques et des espaces vectoriels normés. Bien que le cadre naturel soit plutôt les espaces affines que les espaces vectoriels, j'ai ignoré ce détail, parlant suivant les cas de points ou de vecteurs d'un espace vectoriel.

J'espère développer ces notes pour en faire un texte plus utilisable et serai donc très reconnaissant pour toutes suggestions ou remarques qui pourraient m'être faites (yves.colin-de-verdiere@ujf-grenoble.fr).

2. Introduction historique et motivations

Le *calcul différentiel* (CD) ou *infinitésimal*, inventé par Newton (1642-1727) et Leibniz (1646-1716) à la fin du 17ème siècle, est souvent considéré comme une des grandes découvertes de l'humanité. Cette invention est en fait double : un concept mathématique (la différentielle d'une fonction), et une notation (dx).

Le CD permet en particulier de formuler un grand nombre de lois de la physique et de la mécanique en termes d'équations différentielles (EDO) (mécanique du point matériel et du solide, optique géométrique) et aux dérivées partielles (EDP) (propagation de la chaleur, des ondes, mécanique quantique et relativiste, mécanique des fluides). La raison en est que les EDO et les EDP ne font intervenir que le comportement à l'échelle infinitésimale qui est donné par les lois fondamentales de la physique. La résolution de ces équations permet de déduire

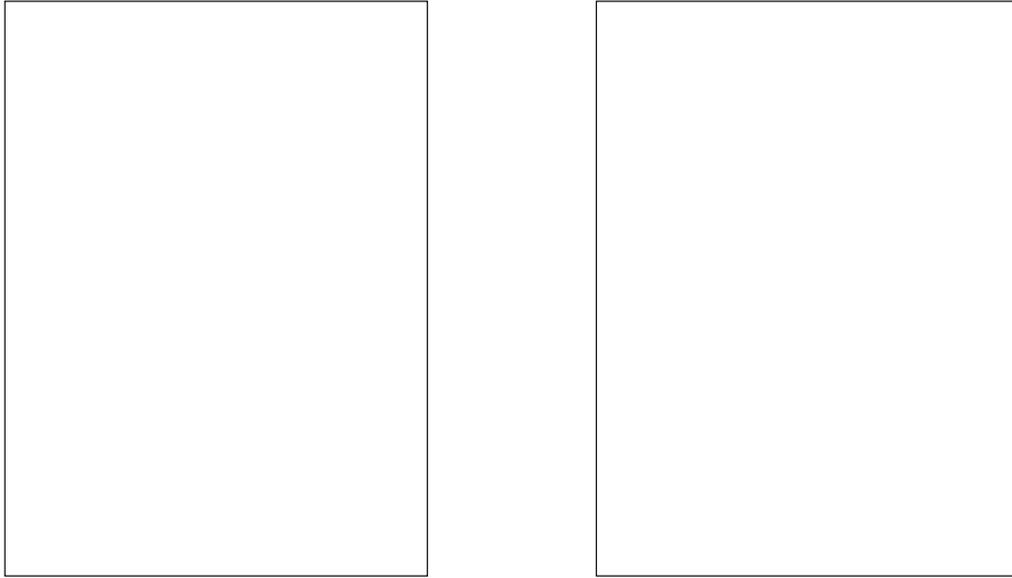


FIGURE 0.1. Gottfried Wilhelm von Leibniz (1646-1716) et Sir Isaac Newton (1643-1727)

des conséquences *macroscopiques* et quantitatives de ces lois physiques *microscopiques* (lois de Képler et attraction universelle, propagation de la chaleur dans les solides (travaux de Fourier), etc...).

Aujourd'hui le développement d'algorithmes numériques et l'accès aux ordinateurs permettent souvent une résolution numérique à grande échelle (prévisions météo, évolution à très long terme du système solaire).

Le CD permet aussi de traiter les problèmes d'optimisation : il s'agit de trouver le minimum (ou le maximum) d'une fonction numérique $f : X \rightarrow \mathbb{R}$ où X peut être une partie de \mathbb{R}^n ou un espace de fonctions ou de courbes. Par exemple trouver la forme d'une boîte cylindrique d'aire minimum contenant un volume donné (boîte de conserve !) ou la courbe fermée plane de longueur minimale entourant un domaine de surface donnée. On sait aujourd'hui qu'un grand nombre de lois de la physique se formulent en termes d'optimisation, l'exemple le plus classique étant la formulation variationnelle des équations de Newton

$$m \frac{d^2 x}{dt^2} = -\text{grad}V ,$$

c'est le principe dit de *moindre action* dû à Hamilton et Jacobi (un principe analogue pour l'optique avait déjà été découvert par Fermat). L'intérêt d'une formulation variationnelle d'une équation n'est pas purement académique : du point de vue théorique il permet par exemple souvent de montrer des théorèmes d'existence (perpendiculaire commune à 2 droites, géodésiques joignant 2 points d'une surface), de trouver les symétries (invariance Lorentzienne des équations

de Maxwell) et aussi de *quantifier la théorie* (intégrale de Feynman) ; il conduit aussi à des algorithmes numériques efficaces, par exemple méthodes de gradient et méthodes d'éléments finis.

Les outils de calcul différentiel utilisés pour résoudre ces problèmes (en particulier calcul différentiel d'ordre supérieur) donnent accès à d'autres branches des mathématiques, par exemple la topologie ou la géométrie différentielle.

3. Contenu

Le cours se compose de 7 chapitres

Chapitre 1 : différentielles.

Chapitre 2 : dérivées d'ordre ≥ 2 .

Chapitre 3 : systèmes d'équations non-linéaires.

Chapitre 4 : extrémums des fonctions numériques.

Chapitre 5 : convexité.

Chapitre 6 : calcul des variations.

Chapitre 7 : quelques problèmes.

Comme on le voit sur l'intitulé des chapitres, après les préliminaires sur les différentielles, le cours est centré sur les problèmes d'optimisation.

Le chapitre 1 s'inspire du livre [14] en utilisant les dx_i comme de vraies variables (linéaires) qui sont les coordonnées vues à travers un microscope de grossissement infini : il y a donc les coordonnées macroscopiques (x_i) et les microscopiques (dx_i).

Le chapitre 2 est consacré aux dérivées d'ordre supérieur. La définition de la différentielle seconde n'est pas la définition usuelle comme différentielle de la dérivée première ; elle a l'avantage de faire apparaître directement une forme bilinéaire symétrique.

Le chapitre 3 contient le théorème des fonctions implicites présenté sous forme d'un théorème de stabilité pour les systèmes d'équations non-linéaires (systèmes de Cramer non-linéaires). Il contient aussi une introduction aux sous-variétés de \mathbb{R}^n . La méthode de Newton y est présentée.

Le chapitre 4 contient les conditions classiques pour les extrémums. Elles permettent de caractériser les points où la fonctionnelle atteint son maximum ou son minimum. Il contient aussi la théorie des multiplicateurs de Lagrange.

Le chapitre 5 contient les fonctions convexes, les algorithmes de relaxation et de gradient à pas fixes pour la minimisation des fonctions convexes et un paragraphe sur les polyèdres convexes et polytopes et la programmation linéaire.

Le chapitre 6 contient les équations d'Euler-Lagrange, le théorème de Hilbert (condition suffisante de minimisation à l'aide de l'*invariant intégral*) et de nombreux exemples : brachistochrone, géodésiques, surfaces minimales de révolution, fil pesant, problème de Dirichlet pour les fonctions harmoniques. Sur cet exemple, il est facile de décrire une méthode d'éléments finis et de la comparer à une méthode de différences finies.

Le chapitre 7 contient des exercices proposés en TD et les problèmes d'examen des 2 années 96-97 et 97-98. *Je remercie Gérard Vinel de m'avoir fourni les fichiers de ses exercices que j'ai pu incorporer à l'ensemble.*

4. Notations

E, F, G sont des espaces vectoriels normés (evn) ou normés complets (eb). x_0 ou m un point courant de E, F .

La différentielle (voir la section 2 du chapitre 1):

$U \subset E$, $f : U \rightarrow F$: $x_0 \in U$, un point intérieur, $f'(x_0) \in L(E, F)$ la différentielle de f en x_0 . $y = f(x)$, $dy \in F$, $dx \in E$, $dy = f'(x_0)dx$.

$V \in E$, $\partial_V f(x_0)$ la dérivée directionnelle (voir la section 3 du chapitre 1).

$E = \mathbb{R}^p$, $F = \mathbb{R}$, $y = f(x_1, \dots, x_p)$, $\frac{\partial f}{\partial x_i}$ les dérivées partielles (voir la section 3.2 du chapitre 1).

$f(x_1, \dots, x_p) = (f_1(x), \dots, f_q(x)) : \mathbb{R}^p \rightarrow \mathbb{R}^q$, $J_f(x_0) = (\frac{\partial f_i}{\partial x_j}(x_0))$ la matrice jacobienne de f (p colonnes et q lignes) (voir la section 5 du chapitre 1).

$(E, \langle \cdot | \cdot \rangle)$ euclidien de dimension finie, $f : E \rightarrow \mathbb{R}$, $y = f(x)$,

$$dy = \langle \text{grad } f(x_0) | dx \rangle ,$$

(voir la section 2.2 du chapitre 1).

$f : I \rightarrow E$, $I \subset \mathbb{R}$, $y = f(x)$, $dy = f'(x)dx$ et $\frac{dy}{dx} = f'(x)$ est le vecteur vitesse.

$(x_1, x_2, \dots, x_n) \in U$, $f : U \rightarrow \mathbb{R}$, $a_i(x) = \frac{\partial f}{\partial x_i}(x)$, la linéarisée de la relation $f = 0$ au point x est

$$\sum_i a_i(x) dx_i = 0 .$$

Les dérivées partielles :

$$\frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} ,$$

avec $\alpha = (\alpha_1, \dots, \alpha_n)$, $|\alpha| = \alpha_1 + \dots + \alpha_n$.

La différentielle seconde (voir la section 2 du chapitre 2) :

$$x_0 \in U, V, W \in E, f''(x_0)(V, W) .$$

$$f''(x_0)(V, W) = \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j} V_i W_j .$$

La matrice hessienne de f au point x est notée $\text{Hess}_f(x)$.

5. Annexe 1 : normes d'applications linéaires et bilinéaires

Soient E, F, G des evn. On définit la norme de $A \in L(E, F)$ et de $B \in L_2(E \times F, G)$ (B est une application bilinéaire continue de $E \times F$ dans G) par :

$$\|A\| = \sup_{\|x\| \leq 1} \|Ax\|, \quad \|B\| = \sup_{\|x\| \leq 1, \|y\| \leq 1} \|B(x, y)\| .$$

Si $A \in L(E, L(F, G))$ et $B(x, y) = (A(x))(y)$ est l'application bilinéaire associée, on a :

$$\|A\| = \|B\| .$$

Cas euclidien : si E et F sont euclidiens, on a :

$$\|A\|^2 = \sup_{\|x\| \leq 1} \langle A^t A x | x \rangle ,$$

et la matrice symétrique positive $\mathcal{A} = A^t A$ est diagonalisable dans une base orthonormée de E , soit $0 \leq \lambda_1 \leq \dots \leq \lambda_{max}$ ses valeurs propres, on a alors :

$$\|A\| = \sqrt{\lambda_{max}} .$$

Si A est inversible, on a :

$$\|A^{-1}\| = \frac{1}{\sqrt{\lambda_1}} .$$

6. Annexe 2 : rappels sur les formes quadratiques

6.1. Formes bilinéaires symétriques. Soit E un espace vectoriel réel de dimension finie. Une forme bilinéaire symétrique B sur E est une application $(x, y) \rightarrow B(x, y)$ de $E \times E$ dans \mathbb{R} qui est linéaire par rapport à x et à y et telle que $\forall x, y \in E, B(x, y) = B(y, x)$. On note $S_2(E)$ l'espace vectoriel de ces applications.

Si $e_i, 1 \leq i \leq n$ est une base de E la matrice $(B) = (b_{i,j})$ de B est donnée par $b_{i,j} = B(e_i, e_j)$. Elle est symétrique et sa donnée détermine B .

6.2. Formes quadratiques. Si $B \in S_2(E)$, on note $q_B : E \rightarrow \mathbb{R}$ l'application $x \rightarrow q_B(x) = B(x, x)$. Dans les coordonnées associées à la base (e_i) , $q_B(x_1, \dots, x_n) = \sum_{i,j} b_{i,j} x_i x_j$ est un polynôme homogène de degré 2 sur \mathbb{R}^n . On peut retrouver B à partir de q_B soit par la formule précédente, soit par

$$B(x, y) = \frac{1}{2}(q_B(x+y) - q_B(x) - q_B(y)) ,$$

soit par dérivation

$$(q'_B(x))(y) = 2B(x, y) .$$

6.3. Indice, signature et positivité.

DÉFINITION 0.1. *L'indice $n_-(q)$ (de Morse) de la forme quadratique q est la dimension maximale d'un sev F de E tel que $\forall x \in F \setminus 0, q(x) < 0$.*

La nullité $n_0(q)$ est la dimension du noyau de B , i.e. $\ker B = \{x \mid \forall y \in E, B(x, y) = 0\}$.

On pose $n_+(q) = n_-(-q)$. Si $n_- = n_0 = 0, \forall x \in E \setminus 0, q(x) > 0$. On dit alors que q est définie positive.

La forme bilinéaire associée à une forme définie positive est un produit scalaire euclidien.

6.4. Diagonalisation et normes. Si E est euclidien et q une forme quadratique sur E , il existe une base orthonormée de E telle que

$$q(x) = \sum_{j=1}^n \lambda_j x_j^2 .$$

Les λ_j sont les valeurs propres de la matrice de q dans une b.o. quelconque de E .
Si les valeurs propres de q satisfont $\forall j, a \leq \lambda_j \leq b$, alors

$$a\|x\|^2 \leq q(x) \leq b\|x\|^2 .$$

7. Annexe 3 : topologies faibles

Tout le monde sait que les boules fermées des espaces de Hilbert de dimension infinie ne sont pas compactes pour la topologie associée à la norme hilbertienne. Il existe une autre topologie appelée *topologie faible* qui a cette propriété.

Soit E, F 2 ev sur \mathbb{R} et soit $b : E \times F \rightarrow \mathbb{R}$ une forme bilinéaire. On se restreint dans la suite au cas où $E = F$ est un Hilbert et $b(x, y) = \text{Re} \langle x|y \rangle$. On associe à cette situation la topologie sur E dont les suites convergentes sont les suites x_n telles qu'il existe $x_\infty \in E$ tel $\forall y \in F, b(x_n, y) \rightarrow b(x_\infty, y)$.

Il suffit que ce soit vrai pour $y \in Z$ où Z engendre F comme espace vectoriel (algébrique).

Donc, si F est un evn, on peut prendre $Z = B(0, 1)$.

Donc, si on envoie E par b dans $\mathcal{F}(Z, \mathbb{R})$ (l'espace vectoriel des fonctions de Z dans \mathbb{R}), il s'agit de la topologie de la convergence simple dans Z .

On a le :

THÉORÈME 0.1. *Si E est un Hilbert, $F = E$ et $b(x, y) = \text{Re} \langle x|y \rangle$, toute partie convexe fermée (pour la norme) bornée de E est compacte pour la topologie faible. Toute fonction continue convexe est sci pour la topologie faible.*

COROLLAIRE 0.1. *Si X est un convexe borné fermé d'un Hilbert et $f : X \rightarrow \mathbb{R}$ est convexe continue, elle atteint sa borne inférieure.*

8. Annexe 4 : notations o et O

DÉFINITION 0.2. *Soient f et g 2 fonctions définies au voisinage d'un point x_0 d'un espace topologique X et à valeurs dans des evn E et F .*

1) *On dit que $f = O(g)$ (sous-entendu en x_0) s'il existe un voisinage U de x_0 et un nombre $M > 0$ tels que*

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

2) *On dit que $f = o(g)$ (sous-entendu en x_0) si, pour tout $\varepsilon > 0$, il existe U voisinage de x_0 tel que*

$$\forall x \in U, \|f(x)\|_E \leq \varepsilon\|g(x)\|_F .$$

Ces notations sont dangereuses :

a) il faut toujours se rappeler que $f = O(g)$ avec g fixé ne définit pas f mais un ensemble : il serait plus logique mais moins pratique d'écrire $f \in O(g)$. Par exemple $f_1 = O(g)$ et $f_2 = O(g)$ n'impliquent pas $f_1 - f_2 = O(g) - O(g) = 0$ mais seulement $f_1 - f_2 = O(g)$.

b) Le point x_0 est en général sous-entendu.

$f = l + o(1)$ équivaut à $\lim_{x \rightarrow x_0} f(x) = l$. Les relations o et O obéissent à des règles simples du type :

$o(g)$ et $O(g)$ sont des espaces vectoriels.

$(f = o(g))$ implique $(f = O(g))$,

$(f = O(g) \text{ et } g = O(h))$ impliquent $(f = O(h))$,

$(f = o(g) \text{ et } g = O(h))$ impliquent $(f = o(h))$,

$(f = O(g) \text{ et } g = o(h))$ impliquent $(f = o(h))$,

$(f = o(g) \text{ et } g = o(h))$ impliquent $(f = o(h))$.

Si $f : X \rightarrow E$ est linéaire continue, $f = O(x)$ en $x = 0$.

Bibliographie

1. Avez, A., *Calcul différentiel*, Masson, 1983.
2. Bourguignon, J.P., *Calcul variationnel*, Ecole Polytechnique, 1988.
3. Cartan, H., *Cours de calcul différentiel*, Herman, 1990.
4. Cea, J., *Optimisation, théorie et algorithme*, Dunod, 1971.
5. Ciarlet, P.G., *Introduction à l'analyse numérique matricielle et à l'optimisation*, Masson, 1985.
6. Culioli, J.C., *Introduction à l'optimisation*, Ellipses, 1994.
7. Dantzig, G., *Linear programming and extensions*, Princeton U. P., 1963.
8. Ekeland, I. et Temam, R., *Analyse convexe et problèmes variationnels*, Dunod, 1974.
9. Hiriart-Urruty, J.B., *L'optimisation*, Que sais-je ?, 1996.
10. Larrouturou, B et Lions, P.L., *Optimisation et commande optimale*, Ecole Polytechnique, 1996.
11. Lascaux, P. et Theodor, R., *Analyse numérique matricielle appliquée à l'art de l'ingénieur*, Masson, 1987.
12. Laudenbach, F., —*Calcul différentiel et intégral*, Ecole Polytechnique, 1996.
13. Pham, F., *Géométrie et calcul différentiel sur les variétés*, Interéditions, 1992.
14. ———, *Les différentielles*, Masson, 1996.
15. Rockafellar, R.T., *Convex analysis*, Princeton , 1970.
16. Schwartz, L., *Analyse, II. Calcul différentiel et équations différentielles*, Hermann, 1983.

CHAPITRE 1

DIFFÉRENTIELLES

1. Introduction

Si I est un intervalle de \mathbb{R} et x_0 un point de I , on dit que $f : I \rightarrow \mathbb{R}$ est *dérivable* en x_0 si la limite

$$\lim_{x \rightarrow x_0, x \neq x_0} \frac{f(x_0 + h) - f(x_0)}{h}$$

existe et on la note $f'(x_0)$ appelée dérivée de f en x_0 . Si $l = f'(x_0)$ est cette limite, on peut réécrire ceci de façon équivalente :

$$(1.1) \quad f(x_0 + h) = f(x_0) + lh + o(h) .$$

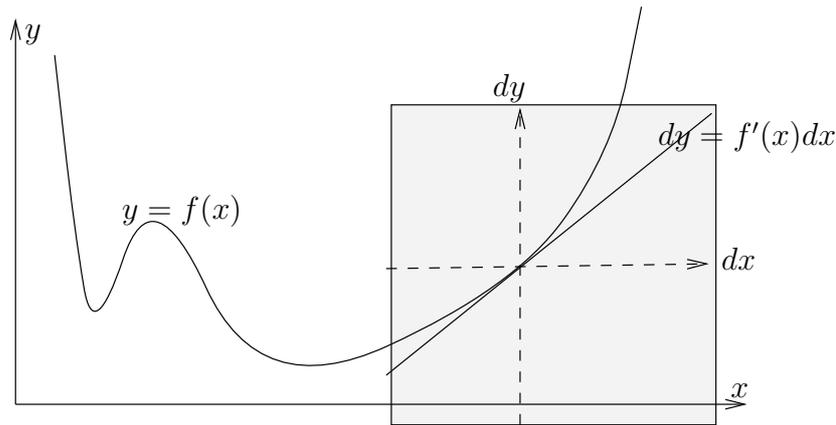


FIGURE 1.1. Une fonction $y = f(x)$ et sa différentielle $dy = f'(x)dx$

C'est sous cette forme qu'on va généraliser au cas des applications $f : U \rightarrow F$ où U est un domaine d'un evn E . On interprète l comme l'application linéaire (homothétie de rapport l de \mathbb{R}) $h_l : h \rightarrow lh$. L'équation 1.1 donne alors une approximation de $f(x_0 + h) - f(x_0)$ par h_l lorsque $h \rightarrow 0$, le sens précis étant donné par la notation $o(h)$.

On pourrait donner un nom à l'application linéaire h_l , il est plus efficace d'introduire des notations pour les variables qui entrent en jeu : on note dx (resp. dy) la variable réelle associée à x (resp. à y), doublant ainsi le nombre de variables

réelles du problème : au lieu de vivre dans un ouvert U du plan des (x, y) , on vit dans $U \times \mathbb{R}^2$. On écrit ainsi $dy = ldx$ ou plus exactement $dy = f'(x)dx$. La relation $y = f(x)$ se double donc d'une relation entre dy et dx qui s'écrit $dy = f'(x)dx$. Les variables dx, dy s'appellent alors *variables infinitésimales (ou microscopiques)* alors que les variables initiales (x, y) s'appellent *variables globales (ou macroscopiques)*.

En physique il est déjà plus important de donner des noms aux variables (énergie E , entropie S , potentiel V) sans préjuger de leurs relations. Bourbaki a mis l'accent en maths sur les relations plutôt que sur les variables et nous en sommes imprégnés... En informatique, on sait tous qu'il est fondamental d'identifier les variables par des noms.

Quel est le gain de cette notation (inventée par Leibniz) ?

D'abord, on fait une économie de lettres et on diminue le risque d'erreur une fois bien identifiées les noms des variables globales $(r, \theta, x, y, z, \dots)$. Ensuite, on a une règle très simple pour la dérivation des fonctions composées : si $z = g(y)$ et $y = f(x)$, on a $dz = g'(y)dy$ et $dy = f'(x)dx$ et la règle de dérivation des fonctions composées

$$(g \circ f)'(x) = g'(y) \cdot f'(x) \text{ avec } y = f(x)$$

dit qu'on peut substituer $dy = f'(x)dx$ dans $dz = g'(y)dy$ de façon que $dz = g'(y)f'(x)dx$ ce qui est cohérent avec la valeur de $(g \circ f)'$. On peut aussi écrire $\frac{dy}{dx} = f'(x)$ car on peut faire le quotient d'une variable réelle par une autre non nulle. Mais ceci cessera d'être vrai pour des variables vectorielles.

Le schéma formel général est donc le suivant : on se donne des variables globales x, y, z, \dots , puis les variables infinitésimales associées dx, dy, dz, \dots . A toute relation $f(x, y, z) = 0$ (f régulière), on associe une relation linéaire de la forme :

$$p(x, y, z)dx + q(x, y, z)dy + r(x, y, z)dz = 0$$

où p, q, r sont les dérivées partielles de f par rapport aux variables x, y, z . Cette procédure est consistante... De plus elle permet en la renversant de calculer les dérivées partielles. Un fois compris ce jeu, on se demande comment l'utiliser pour repasser de relations infinitésimales à des informations globales (équations aux dérivées partielles).

2. Différentielle d'une application

Dans ce chapitre, les espaces vectoriels normés (evn) ne sont pas forcément complets. Seule compte la topologie de l'evn, pas la norme choisie. En particulier, en dimension finie, il n'est pas nécessaire de préciser le choix d'une norme. Dans toute la suite, lorsqu'on parle de *domaine*, il s'agit d'une notion purement ensembliste : un domaine de E est un sous-ensemble de E .

DÉFINITION 1.1. Soit $f : D \rightarrow F$ une application d'un domaine D d'un evn E dans un evn F . Soit $x_0 \in D$ un point intérieur de D , on dit que f est

différentiable en x_0 , s'il existe $L \in L(E, F)$ (ie linéaire continue) telle que l'on ait :

$$f(x_0 + h) = f(x_0) + L(h) + o(h) ,$$

lorsque $h \rightarrow 0$.

On pourrait donner une définition plus forte :

DÉFINITION 1.2. f est fortement différentiable en x_0 si

$$f(x) - f(y) - L(x - y) = o(\|x - y\|)$$

lorsque x et y tendent vers x_0 .

On aurait pu prendre cela comme définition de la différentielle. C'est ce qu'il sera utile de faire pour les différentielles d'ordre ≥ 2 .

PROPOSITION 1.1. Si f est différentiable en x_0 , L est unique, on note $f'(x_0) = L$, on l'appelle la différentielle de f en x_0 .

Preuve

On a

$$f'(x_0)(V) = \lim_{h \rightarrow 0} \frac{f(x_0 + hV) - f(x_0)}{h}$$

et il suffit d'appliquer le théorème d'unicité des limites. □

Si $y = f(x)$ (noms des variables), on introduit les variables vectorielles (dx, dy) et on écrit $dy = L(dx)$, où $dx, (resp. dy) \in E (resp. F)$. On a donc $x \in U, y \in F, dx \in E, dy \in F$ et $y = f(x), dy = f'(x)dx$.

Par exemple, si $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ est donnée par $(x, y) = (r \cos \theta, r \sin \theta)$ la différentielle en (r_0, θ_0) peut s'écrire comme la matrice

$$(1.1) \quad f'(r_0, \theta_0) = \begin{pmatrix} \cos \theta_0 & -r_0 \sin \theta_0 \\ \sin \theta_0 & r_0 \cos \theta_0 \end{pmatrix}$$

ou sous la forme équivalente plus simple :

$$dx = \cos \theta_0 dr - r_0 \sin \theta_0 d\theta, \quad dy = \sin \theta_0 dr + r_0 \cos \theta_0 d\theta .$$

Remarque : si $E = F = \mathbb{R}$, $L(\mathbb{R}, \mathbb{R})$ est formé des homothéties qui sont caractérisées par un nombre qu'on appelle *dérivée* dans l'enseignement élémentaire. La différentielle est donc l'homothétie de rapport égal à la dérivée. En dimension finie, si $E = \mathbb{R}^p, F = \mathbb{R}^q$, $f'(x_0)$ admet une matrice appelée *matrice jacobienne*.

THÉORÈME 1.1. 1) Toute application différentiable en x_0 est continue en x_0 ;
2) la composition de 2 applications différentiables l'est et la différentielle est la composée des différentielles :

$$(g \circ f)'(x_0) = g'(f(x_0)) \circ f'(x_0) .$$

3) Si $f_i : U \rightarrow F_i, i = 1, \dots, n$ sont différentiables en $x_0, f = \bigoplus_{i=1}^n f_i$ l'est pour les normes naturelles sur $F = \bigoplus F_i$.

Preuve

Le 1) utilise la continuité de $f'(x_0)$.

Pour le 2), on pose $y_0 = f(x_0)$ et on écrit :

$$g \circ f(x_0 + h) = g \circ f(x_0) + g'(y_0)(f(x_0 + h) - f(x_0)) + o(\|f(x_0 + h) - f(x_0)\|)$$

(différentiabilité de g en y_0), puis

$$f(x_0 + h) - f(x_0) = f'(x_0)(h) + o(\|h\|)$$

(différentiabilité de f en x_0). On a ainsi :

$$g \circ f(x_0 + h) - g \circ f(x_0) = g'(y_0)(f'(x_0)(h)) + g'(y_0)(o(\|h\|)) + o(\|h\|) ,$$

$$\text{car } f(x_0 + h) - f(x_0) = O(\|h\|).$$

□

PROPOSITION 1.2. Soit $r_\varepsilon(X) = \frac{f(x_0 + \varepsilon X) - f(x_0)}{\varepsilon}$, pour que f soit différentiable et de différentielle L en x_0 , il faut et il suffit que r_ε converge uniformément vers L sur une boule de centre 0 et de rayon $a > 0$.

2.1. Vitesse d'une fonction d'une variable. Si $f : I \rightarrow E$ ($x = f(t)$) où I est un intervalle de \mathbb{R} , f est différentiable en t_0 , intérieur à I , si et seulement si elle est dérivable en t_0 au sens où :

$$\lim_{t \rightarrow t_0, t \neq t_0} \frac{f(t) - f(t_0)}{t - t_0}$$

existe. Si cette limite vaut $V \in E$, on a : $dx = V dt$. On pourra noter la dérivée V sous la forme

$$V = \dot{f} = \frac{df}{dt} .$$

Il sera parfois commode de la noter comme la différentielle

$$V = f'(x) .$$

2.2. Gradient d'une fonction à valeurs réelles. Si E est un espace vectoriel euclidien de dimension finie muni du produit scalaire $\langle \cdot | \cdot \rangle$, on peut identifier E à $L(E, \mathbb{R})$ par l'isomorphisme j donné par :

$$j(x)(y) = \langle x | y \rangle .$$

DÉFINITION 1.3. Si $f : U \rightarrow \mathbb{R}$ où $U \subset E$ est différentiable en x_0 , on définit le gradient de f en x_0 comme le vecteur de E qui satisfait

$$f'(x_0)(V) = \langle \text{grad} f(x_0) | V \rangle$$

ou

$$dy = f'(x_0)(dx) = \langle \text{grad} f(x_0) | dx \rangle .$$

Dans le cas où E est muni d'une base orthonormée le vecteur gradient de f a pour coordonnées les dérivées partielles de f par rapport aux coordonnées dans cette base. Si $t \rightarrow \gamma(t)$ est une courbe différentiable tracée dans $U \subset E$, le théorème de dérivation des fonctions composées se réécrit :

$$\frac{d}{dt}f(\gamma(t)) = \langle \text{grad}f(\gamma(t)) | \dot{\gamma}(t) \rangle .$$

Propriétés du gradient :

lorsqu'il est non nul, le gradient est orthogonal aux surfaces de niveaux, dirigé du côté croissant de f et de longueur inversement proportionnelle à l'écartement des surfaces de niveaux.

EXEMPLE 1.1. *Lignes de niveau de $f(M) = d(A, M) + d(B, M)$ et tangentes aux ellipses. La fonction f est différentiable en tout point $M \neq A, B$ et son gradient est donné par :*

$$\text{grad}f(M) = \vec{u}_A + \vec{u}_B$$

où \vec{u}_A est le vecteur unitaire associé à \vec{AM} et de même pour \vec{u}_B . La tangente aux lignes de niveau de f qui sont des ellipses de foyers A et B est donc la perpendiculaire à $\vec{u}_A + \vec{u}_B$ qui est une bissectrice des droites (AM) , (BM) .

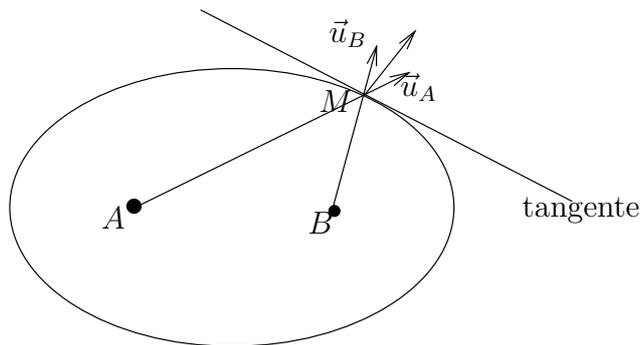


FIGURE 1.2. Tangente à l'ellipse

2.3. Exemples. La différentielle de l'application $A \rightarrow B = A^2 = A \circ A$ où $A \in L(E, E)$ est donnée par

$$dB = A \circ dA + dA \circ A .$$

La différentielle de $Z = z^5$ comme application de \mathbb{C} dans \mathbb{C} est donnée par $dZ = 5z^4 dz$ où $5z^4$ est la multiplication dans \mathbb{C} vue comme application \mathbb{R} -linéaire.

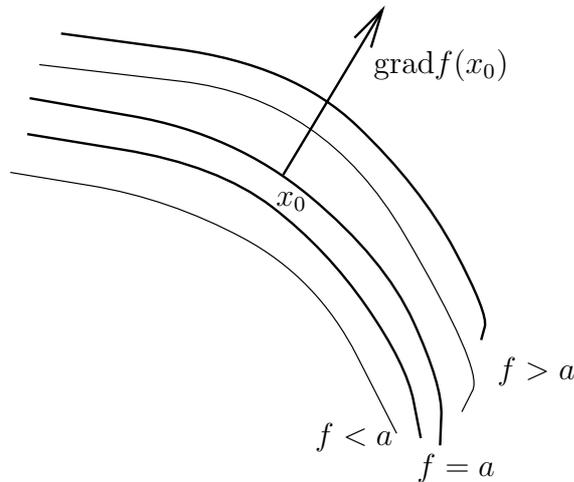


FIGURE 1.3. Gradient

2.4. Interprétation en terme de renormalisation (ou microscope). Si x_0 est donné et $f(x_0) = y_0$, on peut imaginer qu'on fait un zoom de grossissement $\frac{1}{\varepsilon}$ en ces deux points, si on note $x = x_0 + \varepsilon X_\varepsilon$ et $y = y_0 + \varepsilon Y_\varepsilon$, X_ε et Y_ε sont les coordonnées qu'on voit au microscope et f devient

$$f_\varepsilon(X_\varepsilon) = \frac{1}{\varepsilon}(f(x_0 + \varepsilon X_\varepsilon) - y_0) = Y_\varepsilon .$$

On dit qu'on a *renormalisé* f en f_ε . Le domaine de f_ε grossit avec ε comme $1/\varepsilon$. Cette idée de *renormalisation* est fondamentale dans beaucoup de domaines des maths ou de la physique.

On voit ainsi que, dans le champ du microscope, f_ε converge uniformément vers une application linéaire d'après la proposition 1.2.

Si on note $Y_\varepsilon = f_\varepsilon(X_\varepsilon)$, cette relation converge vers la relation infinitésimale $dy = f'(x)dx$.

2.5. Applications de classe C^1 . $f : U \rightarrow F$ (U ouvert de E) est dite C^0 si elle est continue.

DÉFINITION 1.4. On dit que f est C^1 (on dit aussi *continument dérivable*) si elle est différentiable en tout point de U et que $x \rightarrow f'(x)$ est une application continue de U dans $L(E, F)$.

Si $E = F = \mathbb{R}$, C^1 équivaut à dérivée continue. On définira plus loin la notion de classe C^k .

Il arrive très souvent qu'on voit instantanément qu'une application est C^1 en suivant la façon dont elle est construite à partir de polynômes, de fonctions différentiables d'une variable, etc.. Par exemple, l'application $z \rightarrow e^z = e^x(\cos y + i \sin y)$.

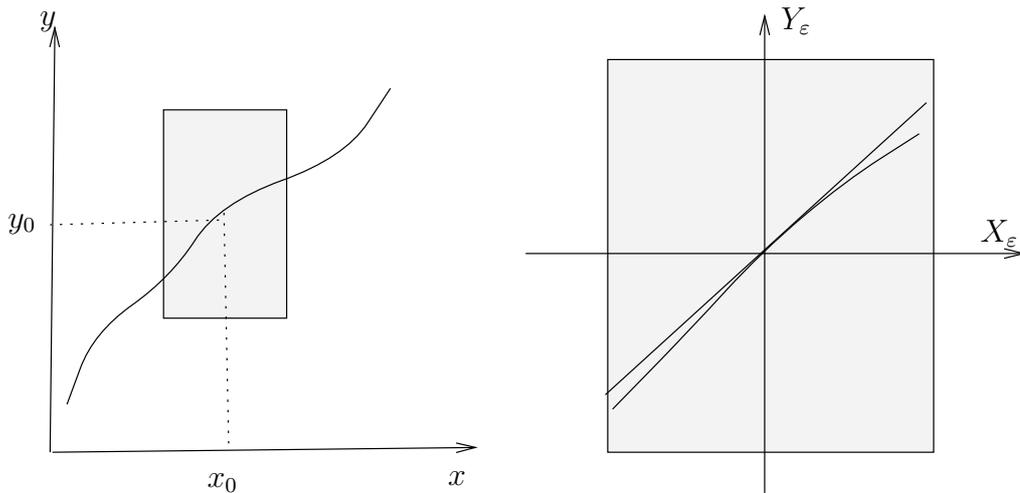


FIGURE 1.4. Zoom

2.6. Différentiabilité en un point non intérieur. Si $f : D \rightarrow F$ où $D \subset E$ et $x_0 \in D$, on dit que f est différentiable en x_0 si on a, **pour tout h tel que $x_0 + h \in D$**

$$f(x_0 + h) - f(x_0) = L(h) + o(h)$$

avec $L \in L(E, F)$.

Il revient au même de demander que f peut se prolonger au voisinage de x_0 en une application différentiable en x_0 (exercice). La seule difficulté est que L n'est pas toujours unique (par exemple si D est un sous-espace de E).

Exercice 1 : montrer que dans le cas où $D = [a, b] \subset \mathbb{R}$ la différentiabilité en a (resp. b) équivaut à la dérivabilité à droite en a (resp. à gauche en b).

Exercice 2 : soit $D = \{0, 1, \frac{1}{2}, \dots, \frac{1}{n}, \dots\}$, et $f(0) = a$, $f(\frac{1}{n}) = u_n$. Caractériser les suites (u_n) telles que f est différentiable en 0.

3. Dérivées directionnelles

DÉFINITION 1.5. Soit $f : D \rightarrow F$, $D \subset E$, et $x_0 \in D$, $f(x_0) = y_0$. Soit $V \in E$ tel qu'il existe $a > 0$ tel que si $0 \leq t \leq a$, $x_0 + tV \in D$, on dira que f admet une dérivée (directionnelle) suivant V , si $\varphi(t) = f(x_0 + tV)$ admet une dérivée à droite $W \in F$ en 0. On écrit alors

$$\partial_V f(x_0) = W .$$

Bien sûr si x_0 est intérieur à D et f différentiable en x_0 , on a :

$$\partial_V f(x_0) = f'(x_0)(V) .$$

Si une telle dérivée existe pour V , elle existe aussi pour λV , $\lambda > 0$ et $\partial_{\lambda V} f(x_0) = \lambda \partial_V f(x_0)$.

L'intérêt est par exemple pour les fonctions définies dans un domaine défini par des inéquations lorsqu'on est au bord. Certaines fonctions naturelles, comme les normes admettent des dérivées directionnelles à l'origine et pas de différentielles.

La fonction

$$f(x, y) = \frac{x^3}{x^2 + y^2}$$

admet des dérivées directionnelles suivant tous les vecteurs en O , mais n'y est pas différentiable, car f homogène de degré 1 donc insensible à la renormalisation donc on aurait l'identité $f = ax + by$ si f était différentiable.

3.1. Cas d'espaces fonctionnels (courbes, surfaces). Soit E un evn de fonctions de $[a, b]$ dans \mathbb{R} . Soit A, B donnés et $D = \{f \in E \mid f(a) = A, f(b) = B\}$. Soit $f \in D$ et $\varphi \in E$ telle que $\varphi(a) = \varphi(b) = 0$. Alors $f + t\varphi \in D$ et, si $F : D \rightarrow \mathbb{R}$, $\partial_\varphi F(f) = \frac{d}{dt}\big|_{t=0} F(f + t\varphi)$. Le calcul de ces dérivées fera l'objet du chapitre 6 (calcul des variations).

Exemple : $E = C^2([0, 1], \mathbb{R})$ et $F(f) = \int_0^1 \frac{1}{2} f'^2(t) - V(f(t)) dt$. Alors, on vérifie au moyen d'une intégration par parties que

$$(1.1) \quad \partial_\varphi F(f) = \int_0^1 -(f''(t) + \frac{\partial V}{\partial x}(f(t)))\varphi(t)dt + [f'(t)\varphi(t)]_0^1.$$

3.2. Dérivées partielles. Si $U \subset \mathbb{R}^n$ et $f : U \rightarrow F$, on considère les fonctions vectorielles d'une variable $f_i : x_i \rightarrow f(x_1, \dots, x_i, \dots, x_n)$ obtenues en fixant toutes les coordonnées sauf la i -ème. Si f_i est dérivable, on note

$$\frac{\partial f}{\partial x_i}$$

sa dérivée. Si e_i est le i -ème vecteur de base, on a :

$$\frac{\partial f}{\partial x_i} = \partial_{e_i} f,$$

que l'on note aussi $\partial_i f$.

L'existence des dérivées partielles n'assure pas la différentiabilité, mais on a le :

THÉORÈME 1.2. *Soit $f : U \rightarrow F$ où $U \subset \mathbb{R}^n$ est un ouvert. Si les dérivées partielles de f comme fonctions de U dans F existent et sont continues, f est de classe C^1 .*

Preuve

On va le prouver pour $E = \mathbb{R}^2$, $F = \mathbb{R}$. Soit $A = (x_0, y_0)$, $B = (x_0 + h, y_0)$ et $C = (x_0 + h, y_0 + k)$, on a en posant $p(x, y) = \frac{\partial f}{\partial x}$, $q(x, y) = \frac{\partial f}{\partial y}$:

$$(1.2) \quad f(C) - f(A) = \int_0^h p(x_0 + v, y_0)dv + \int_0^k q(x_0 + h, y_0 + u)du,$$

On pose $p_0 = p(A)$ et $q_0 = q(A)$. Pour tout $\varepsilon > 0$, il existe $a > 0$ tel que, si $|u| + |v| \leq a$, $|p(x_0 + v, y_0) - p_0| \leq \varepsilon$ et $|q(x_0 + v, y_0 + u) - q_0| \leq \varepsilon$. En reportant, il vient, pour $|h| + |k| \leq a$:

$$(1.3) \quad |f(C) - f(A) - p_0 h - q_0 k| \leq \varepsilon(|h| + |k|) .$$

Cela prouve que f est différentiable en A de différentielle $L(h, k) = p_0 h + q_0 k$.

Faire la démonstration pour $E = \mathbb{R}^n$ et F evn quelconque en exercice.

□

4. Champs de vecteurs

A l'aide des dérivées partielles, on fabrique les opérateurs différentiels. Pour cela, il est commode d'introduire les notations $\frac{\partial}{\partial x_i}$ qui sont des *opérateurs*, c'est à dire qu'ils opèrent sur les fonctions C^1 par $(\frac{\partial}{\partial x_i})f = \frac{\partial f}{\partial x_i}$. On peut aussi les voir comme des *champs de vecteurs* ; l'égalité $\frac{\partial f}{\partial x_i} = f'(x)(e_i)$ identifie $\frac{\partial}{\partial x_i}$ au champ de vecteurs e_i . De même on peut considérer l'opérateur $\frac{\partial}{\partial \theta}$ qui est la dérivée partielle par rapport à θ en polaires et on a l'égalité de champs de vecteurs :

$$(1.1) \quad \frac{\partial}{\partial \theta} = x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} .$$

que l'on peut voir comme un champ de vecteurs dans le plan.

Un champ de vecteurs X dans un ouvert U de E evn est une application (en général continue) de U dans E . Lorsque $E = \mathbb{R}^n$, on peut écrire $X = \sum_{i=1}^n a_i(x) e_i$ et en utilisant les notations de dérivées partielles :

$$X = \sum_{i=1}^n a_i(x) \frac{\partial}{\partial x_i} .$$

Un champ de vecteurs opère sur une fonction $f \in C^1(U, F)$ par $Xf(x) = f'(x)(X(x))$ ou encore dans \mathbb{R}^n :

$$Xf(x) = \sum_{i=1}^n a_i(x) \frac{\partial f}{\partial x_i} .$$

A un champ de vecteurs est aussi associé l'équation différentielle $dx/dt = X(x(t))$ mais c'est une autre histoire ...

Attention, $\frac{\partial}{\partial \theta}$ dépend des coordonnées (r, θ) et pas seulement de la fonction $\theta(x, y)$. Il en va différemment de $d\theta$.

5. Matrices jacobiennes

Si $f : U \rightarrow \mathbb{R}^q$ où U est un ouvert de \mathbb{R}^p , on peut écrire

$$f(x_1, \dots, x_i, \dots, x_p) = (f_1(x_1, \dots, x_p), \dots, f_j(x_1, \dots, x_p), \dots, f_q(x_1, \dots, x_p))$$

et si $x_0 = (a_1, \dots, a_i, \dots, a_p)$, la matrice A de $f'(x_0)$ (supposée exister) est la matrice à q lignes et p colonnes donnée par

$$A = \left(\frac{\partial f_j}{\partial x_i}(a_1, \dots, a_p) \right),$$

et appelée *matrice jacobienne* de f en x_0 et que l'on peut noter $J_f(x_0)$.

Interprétations géométriques :

1) Si $p = q$, la différentielle est un endomorphisme de \mathbb{R}^p . Son déterminant s'interprète de la façon suivante : il est > 0 si f préserve l'orientation, < 0 sinon.

La valeur absolue mesure la variation infinitésimale de volume ; on a la formule de changement de variables dans les intégrales multiples.

Plus précisément, donnons la :

DÉFINITION 1.6. *Si $U \subset \mathbb{R}^p$ et $V \subset \mathbb{R}^q$ sont deux ouverts, $F : U \rightarrow V$ est un difféomorphisme de U sur V si F est C^1 , bijective et F^{-1} est aussi C^1 .*

On a alors, pour tout $x_0 \in U$,

$$(F^{-1})'(f(x_0)) = (F'(x_0))^{-1},$$

ce qui implique $p = q$. C'est aussi vrai pour un homéomorphisme, mais plus difficile à prouver !

Si $p = q$, on lit sur la différentielle le fait que l'application est conforme, i.e. préserve les angles. Si $p = q = 2$, une application conforme qui préserve l'orientation vérifie les équations de Cauchy-Riemann et est dite *fonction holomorphe*. En effet la matrice jacobienne de $F = f + ig$ est alors celle d'une similitude directe et pour tout point z_0 il existe $a, b \in \mathbb{R}$ tels que :

$$(1.1) \quad \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{pmatrix} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

Cela équivaut à :

$$\frac{\partial f}{\partial y} + \frac{\partial g}{\partial x} = 0, \quad \frac{\partial f}{\partial x} - \frac{\partial g}{\partial y} = 0.$$

2) Si $p = 1$, l'application f s'interprète comme une courbe paramétrée et sa différentielle est $dt \rightarrow V dt$ où V le vecteur vitesse est donné par $V = \sum f'_i(t)e_i$.

3) Si $q = 1$, $f'(x_0)$ est une forme linéaire dont les composantes sont les $\partial_i f(x_0)$.

Une fois écrites les différentielles en termes de matrices jacobiennes, on peut relire le théorème de différentiation des fonctions composées (théorème 1.1) de façon à calculer les dérivées partielles des fonctions composées :

THÉORÈME 1.3. Si $f : U \rightarrow \mathbb{R}^p$ et $g : V \rightarrow \mathbb{R}^q$ où U est un ouvert de \mathbb{R}^n , f, g sont C^1 et $f(U) \subset V$, les matrices jacobiniennes de f au point x , $J_f(x)$, de g au point $y = f(x)$, $J_g(y)$ et de $g \circ f$ au point x , $J_{g \circ f}(x)$, satisfont :

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

où \circ est le produit des matrices. En particulier si $q = 1$, on a :

$$(1.2) \quad \frac{\partial g(f_1(x_1, \dots, x_n), \dots, f_q(x_1, \dots, x_n))}{\partial x_i} = \sum_{j=1}^q \frac{\partial g}{\partial y_j} \frac{\partial f_j}{\partial x_i} .$$

6. Notations dx_i

Soit U un ouvert de \mathbb{R}^N et $x = (x_1, \dots, x_i, \dots, x_N)$ le point courant. On note $(x_1, \dots, x_N; dx_1, \dots, dx_N)$ des coordonnées dans $U \times \mathbb{R}^N$. Lorsque $x = (x_1, \dots, x_N)$ est fixé, (dx_1, \dots, dx_N) sont les coordonnées *microscopiques* au point x . Si (x_1, \dots, x_N) satisfait une relation $F(x_1, \dots, x_N) = 0$ où $F : U \rightarrow \mathbb{R}$ est C^1 , on associe à la relation $F(x) = 0$, sa relation dérivée (ou microscopique)

$$\sum_i \frac{\partial F}{\partial x_i}(x) dx_i = 0 .$$

C'est la relation entre les coordonnées microscopiques. Elle exprime simplement la relation entre les (x_i) sous l'effet d'un zoom infini en x .

L'écriture avec les dx_i permet de calculer des dérivées ou des dérivées partielles de fonctions données *implicitement*.

Par exemple si $z = f(x, y)$ vérifie $F(x, y, z) = 0$ et $G(x, y, z) = 0$, les fonctions $\varphi(x, y) = F(x, y, f(x, y))$ et $\psi(x, y) = G(x, y, f(x, y))$ sont identiquement nulles. On suppose F, G, f de classe C^1 . On a

$$d\varphi = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy + \frac{\partial F}{\partial z} dz ,$$

avec $dz = p dx + q dy$ où p et q sont les dérivées partielles de f . Cette écriture est valide à cause du théorème de composition des différentielles. L'avantage de la notation est qu'on calcule dF sans se préoccuper du fait que $z = f(x, y)$ qui n'apparaît que dans la façon d'interpréter dz . L'écriture des différentielles au moyen des $d \dots$ est indépendante d'un choix *a priori* de variables indépendantes. Elle met l'accent sur des relations fonctionnelles entre des variables qui sont nommées par des lettres et non sur les applications entre ensembles. Par exemple, pour les coordonnées polaires, on ne se préoccupe pas de savoir si (r, θ) est fonction de (x, y) ou le contraire, on différencie $x = r \cos \theta$, $y = r \sin \theta$.

De même

$$d\psi = \frac{\partial G}{\partial x} dx + \frac{\partial G}{\partial y} dy + \frac{\partial G}{\partial z} dz .$$

De ces expressions de $d\varphi$ et $d\psi$, on peut en général tirer dz en fonction de dx et dy , ce qui détermine p et q par identification.

EXEMPLE 1.1. Soit $z = f(x^2 + y^2, e^{x \sin y})$. Posons $X = x^2 + y^2$, $Y = e^{x \sin y}$. On a :

$$dz = \frac{\partial f}{\partial X} dX + \frac{\partial f}{\partial Y} dY, \quad dX = 2(xdx + ydy), \quad dY = e^{x \sin y} (\sin y dx + x \cos y dy) .$$

Donc $dz = pdx + qdy$ avec $p = 2x \frac{\partial f}{\partial X} + e^{x \sin y} \sin y \frac{\partial f}{\partial Y}$ et $q = 2y \frac{\partial f}{\partial X} + e^{x \sin y} x \cos y \frac{\partial f}{\partial Y}$. p et q sont les dérivées partielles de z par rapport à x et y .

EXEMPLE 1.2. Soit $x^2 + 2y^2 = 3$. On en déduit la relation dérivée au point (x, y) : $2xdx + 4ydy = 0$. Considérons la fonction $y = g(x) = \sqrt{\frac{3-x^2}{2}}$ pour $|x| < 1$. On a $dy = g'(x)dx$ et par identification $g'(x) = -\frac{x}{2y}$.

7. Accroissements finis

7.1. Reconstruire une fonction à partir de sa différentielle. Le problème que nous allons traiter maintenant est celui de remonter d'une information infinitésimale à une information macroscopique. Ce sujet est très vaste puisqu'il contient les équations différentielles et aux dérivées partielles et donc une grande partie de la physique. Le problème de base est celui-ci : **une application $f : U \rightarrow F$ est-elle complètement déterminée par la connaissance de sa différentielle en tout point de U ?** La réponse est assez simple et bien connue :

THÉORÈME 1.4. *Supposons l'ouvert U connexe. Si $f, g \in C^1(U, F)$ ont mêmes différentielles en tout point de U , $f(x) - g(x)$ est constante.*

Preuve

Par différence, il suffit de montrer que, si $\forall x \in U, f'(x) = 0$, alors f est constante. Composant avec $L : F \rightarrow \mathbb{R}$ on obtient ainsi une fonction $\varphi = L \circ f : U \rightarrow \mathbb{R}$ telle que $\forall x \in U, \varphi'(x) = 0$ et on va montrer que φ est constante sur tout segment contenu dans U . par restriction au de φ au segment, on a une fonction $g : I = [a, b] \rightarrow \mathbb{R}$ dont la dérivée est identiquement nulle. Il est classique qu'une telle fonction est constante : on a $g(b) - g(a) = (b-a)g'(c)$ avec $c \in]a, b[$ (théorème de Rolle). Maintenant $L \circ f$ est constante sur chaque composante connexe de U et il suffit de savoir que $\forall x \neq y \in F$ il existe $L \in L(F, \mathbb{R})$ telle que $L(x - y) \neq 0$ (ce qui est vrai, et simple en dimension finie).

□

Remarquons que ce théorème laisse de côté la question suivante : *étant donné $F : U \rightarrow L(E, F)$ continue, existe-t-il $f : U \rightarrow F$ telle que $f'(x) = F(x)$?* Le théorème précédent montrent que 2 solutions différent d'une constante sur chaque composante connexe de U . Lorsque $U \subset \mathbb{R}^n$ et F est C^1 le lemme de Schwarz

(voir la section 3) donne une condition nécessaire (mais pas suffisante en général) : si $\omega = \sum_{i=1}^n F_i(x) dx_i$, pour qu'il existe f telle que $\omega = df$, on doit avoir

$$\forall i, j, \frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i} .$$

On peut maintenant restituer f à partir de sa différentielle au moyen d'une intégrale : si $\gamma : [0, 1] \rightarrow U$ est un chemin C^1 tel que $\gamma(0) = x$ et $\gamma(1) = y$, et si f est C^1 on a :

$$f(y) - f(x) = \int_0^1 f'(\gamma(t))(\gamma'(t)) dt .$$

Remarque : pour que l'intégrale précédente ait un sens, il faut a priori supposer F complet. En fait on s'en dispense, en définissant l'intégrale dans le complété \hat{F} de F et en considérant l'égalité précédente dans $F \subset \hat{F}$.

7.2. Inégalités des accroissements finis. On peut aller un peu plus loin et chercher des estimations (majorations) portant sur f à partir de contrôle de f' : si f' est partout petite, est ce que f est presque constante ?

Supposons toujours f de classe C^1 dans un ouvert U et soit $\gamma : [0, 1] \rightarrow U$ une application de classe C^1 (C^1 par morceaux suffit).

On a alors :

THÉORÈME 1.5. (*inégalité des accroissements finis*) Si $x_0 = \gamma(0)$ et $x_1 = \gamma(1)$,

$$\|f(x_1) - f(x_0)\| \leq \int_0^1 \|f'(\gamma(t))\| \|\gamma'(t)\| dt .$$

En particulier si $M = \sup_{t \in [0, 1]} \|f'(\gamma(t))\|$,

$$\|f(x_1) - f(x_0)\| \leq ML(\gamma) ,$$

où $L(\gamma)$ est la longueur de γ pour la norme de E . Si $a, b \in U$ sont tels que $[a, b] \subset U$ (toujours vrai si U est convexe) et $\|f'(x)\| \leq M$ sur U , $\|f(a) - f(b)\| \leq M\|a - b\|$.

Preuve

On a

$$f(x_1) - f(x_0) = \int_0^1 \frac{d}{dt} f(\gamma(t)) dt ,$$

et on utilise :

$$\left\| \int_0^1 g(t) dt \right\| \leq \int_0^1 \|g(t)\| dt .$$

□

Conséquences :

- a) C^1 implique localement Lipschitz.
- b) Si $M < 1$ et U convexe, f est contractante.

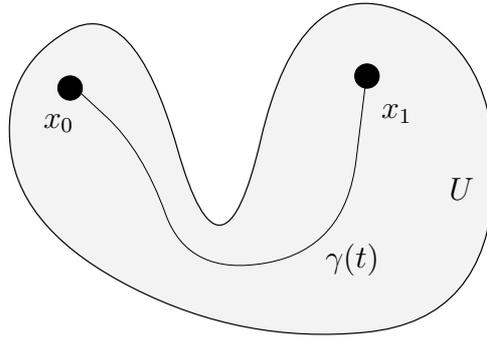


FIGURE 1.5. Accroissements finis

c) Si f est C^1 , f est fortement différentiable ; en effet, on pose $F(x) = f(x) - L(x)$, et donc $F'(x) = f'(x) - L$, où $L = f'(x_0)$ et on applique les accroissements finis dans une boule $B(x_0, r)$ telle que $\forall x \in B(x_0, r)$, $\|f'(x) - L\| \leq \varepsilon$.

d) Différentiabilité uniforme. On suppose que $K \subset U$ est un compact convexe et que $f : U \rightarrow F$ est de classe C^1 . Pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que, pour tout $x, y \in K$, on ait

$$\|f(x) - f(y) - f'(y)(x - y)\| \leq \varepsilon \|x - y\| .$$

On applique les accroissements finis à

$$x \rightarrow f(x) - f'(y)(x - y)$$

et la continuité uniforme de f' sur K .

7.3. Longueur d'un arc C^1 .

DÉFINITION 1.7. Soit $\gamma : [a, b] \rightarrow E$ une application C^1 (courbe tracée dans E). On suppose que E est muni d'une norme. Soit $\sigma = \{t_0 = a < t_1 < \dots < t_n = b\}$ une subdivision de σ . On pose $|\sigma| = \sup |t_{i+1} - t_i|$. On pose $L(\sigma) = \sum_{i=0}^{n-1} \|\gamma(t_{i+1}) - \gamma(t_i)\|$, $L(\sigma)$ est la longueur de la ligne polygonale associée à σ inscrite dans γ . La longueur de γ est $L(\gamma) = \sup_{\sigma} L(\sigma)$.

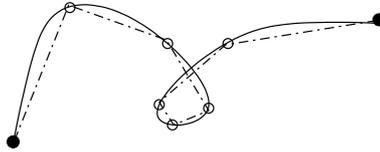


FIGURE 1.6. Longueur d'un arc de courbe

On a :

PROPOSITION 1.3.

$$L(\gamma) = \int_a^b \|\dot{\gamma}(t)\| dt = \lim_{\sigma \rightarrow 0} L(\sigma) .$$

Preuve

- Majoration : $\|\gamma(t_{i+1}) - \gamma(t_i)\| \leq \int_{t_i}^{t_{i+1}} \|\dot{\gamma}(t)\| dt$ par l'inégalité des accroissements finis.
- Minoration : ε étant donné, il existe $\alpha > 0$ tel que si $|t - t'| \leq \alpha$, on a :

$$\|\gamma(t) - \gamma(t') - \dot{\gamma}(t)(t - t')\| \leq \varepsilon |t - t'| ,$$

(différentiabilité uniforme sur $[a, b]$). On obtient ainsi que, si $|\sigma| \leq \alpha$,

$$|L(\sigma) - \sum_i \|\dot{\gamma}(t_i)\| (t_{i+1} - t_i)| \leq \varepsilon (b - a) .$$

On en déduit le résultat par convergence des sommes de Riemann vers l'intégrale.

□

8. Convergence C^1 d'une suite d'applications

On définit la topologie de la convergence C^1 locale de la façon suivante :

DÉFINITION 1.8. Une suite $f_n : U \rightarrow F$ d'applications de classe C^1 converge au sens C^1 vers f si f_n (resp. f'_n) converge localement uniformément vers f (resp. f').

On peut réécrire ceci en termes de quantificateurs : f_n converge au sens C^1 vers f si, pour tout x_0 de U , il existe un voisinage $V \subset U$ de x_0 tel que, pour tout $\varepsilon > 0$, il existe N tel que pour tout $n \geq N$ et tout $x \in V$, on ait :

$$\|f_n(x) - f(x)\| \leq \varepsilon \text{ et } \|f'_n(x) - f'(x)\| \leq \varepsilon .$$

Bien sûr, une suite de fonctions C^1 peut converger au sens C^0 vers une fonction C^1 : par exemple $f_n(x) = \frac{\sin nx}{n}$ converge au sens C^0 vers 0, mais pas au sens C^1 .
Un exemple :

PROPOSITION 1.4. Si f est C^1 , les applications renormalisées

$$f_\varepsilon(X) = \frac{f(x_0 + \varepsilon X) - f(x_0)}{\varepsilon}$$

convergent au sens C^1 sur toute boule vers $X \rightarrow f'(x_0)(X)$.

Preuve

On a :

$$f'_\varepsilon(X) = f'(x_0 + \varepsilon X) ,$$

d'où résulte la convergence C^1 .

□

CHAPITRE 2

DÉRIVÉES D'ORDRE ≥ 2

1. Introduction

On définit habituellement une fonction 2 fois différentiable en un point x_0 comme une fonction différentiable en tous les points x voisins de x_0 telle que $x \rightarrow f'(x)$ soit elle-même différentiable en x_0 ; cette différentielle est alors un élément de $L(E, L(E, F))$ que l'on identifie à une application bilinéaire dont on démontre qu'elle est symétrique (lemme de *Schwarz*). Cette démarche était déjà celle de Newton : on prend la *fluxion* de la *fluxion*. On la retrouve dans les manuels *modernes* de calcul différentiel (cf Cartan, Schwartz, etc...). Elle a l'inconvénient de rendre lointaine la nature (application bilinéaire symétrique) ainsi que la façon d'approcher numériquement la dérivée seconde. La démarche suivie dans ce texte est plus directe et évite ces inconvénients.

Considérons pour faire saisir notre propos le cas d'une fonction $f : [a, b] \rightarrow \mathbb{R}$ et plaçons nous en un point $x_0 \in]a, b[$. On considère le quotient suivant :

$$\delta = \frac{\frac{f(x_0+u+v)-f(x_0+v)}{u} - \frac{f(x_0+u)-f(x_0)}{u}}{v},$$

que l'on peut réécrire :

$$\delta = \frac{f(x_0+u+v) - f(x_0+v) - f(x_0+u) + f(x_0)}{uv}.$$

Si on prend d'abord la limite quand $u \rightarrow 0$ on obtient

$$\frac{\dot{f}(x_0+v) - \dot{f}(x_0)}{v}$$

qui converge vers la dérivée seconde \ddot{f} quand $v \rightarrow 0$. On peut aussi considérer directement la limite quand $u, v \rightarrow 0$. Si f est C^2 , on trouve la même limite.

2. Différentielles secondes

Soit $f : U \rightarrow F$ où $U \subset E$ est un ouvert et E, F sont des espaces vectoriels normés.

DÉFINITION 2.1. *On dit que f est 2-fois différentiable en x_0 s'il existe une application bilinéaire continue $B : E \times E \rightarrow F$ telle que*

$$(2.1) \quad \Delta_{x_0, u, v} f = f(x_0+u+v) - f(x_0+u) - f(x_0+v) + f(x_0) = B(u, v) + o(\|u\| \cdot \|v\|)$$

lorsque $u, v \rightarrow 0$. Cela signifie, $\forall \varepsilon > 0$, il existe $\alpha > 0$ tel que, pour tout u, v tels que $\|u\| \leq \alpha$ et $\|v\| \leq \alpha$, on ait :

$$\|f(x_0 + u + v) - f(x_0 + u) - f(x_0 + v) + f(x_0) - B(u, v)\| \leq \varepsilon \|u\| \cdot \|v\| .$$

On note $f''(x_0) = B$ la différentielle seconde en x_0 .

On voit immédiatement que B est symétrique, car le membre de gauche est symétrique en u, v .

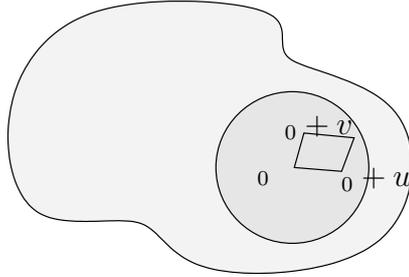


FIGURE 2.1. Le parallélogramme $P_{x_0, u, v}$

Géométriquement, les 4 points $x_0, x_0 + u, x_0 + v, x_0 + u + v$ sont les 4 sommets d'un parallélogramme $P_{x_0, u, v}$ et la condition $u, v \rightarrow 0$ signifie que les 4 sommets tendent vers x_0 .

Faisons quelques remarques :

a) même si $E = F = \mathbb{R}$, l'existence de la différentielle seconde n'implique pas à elle seule la différentiabilité ; il suffit de prendre pour f un application additive non linéaire : on aura $B = 0$.

b) La formule précédente donnera le schéma standard de différence finie pour le calcul numérique des dérivées secondes. En particulier, on écrira les discrétisations standard du laplacien comme laplacien combinatoire.

c) On voit immédiatement que la convexité de f implique la positivité de la différentielle seconde comme forme quadratique (i.e. $\forall u \in E, B(u, u) \geq 0$).

d) Si $E = F = \mathbb{R}$, la différentiabilité à l'ordre 2 en x_0 équivaut à l'existence de la limite

$$l = \lim_{u, v \rightarrow 0} \frac{f(x_0 + u + v) - f(x_0 + u) - f(x_0 + v) + f(x_0)}{uv} .$$

Lorsque f est C^1 et \dot{f} est aussi C^1 , le nombre l est égal à $\ddot{f}(x_0)$. Il est aussi égal à la limite de

$$\frac{f(x_0 + u) + f(x_0 - u) - 2f(x_0)}{u^2}$$

lorsque $u \rightarrow 0$ (faire $v = -u$). *Preuve*

On pose $\varphi(x) = f(x + v) - f(x)$ qui est C^1 . Par le théorème de Rolle :

$$\Delta_{x_0, u, v} f = \varphi(x + u) - \varphi(x) = u\dot{\varphi}(c)$$

avec $c \in]x, x + u[$. Puis $\dot{\varphi}(c)\dot{f}(c + v) - \dot{f}(c)$. Comme \dot{f} est C^1 , on peut appliquer Rolle à nouveau :

$$\Delta_{x_0, u, v} f = uv\ddot{f}(d)$$

avec $d \in]c, x_0 + c[$. Lorsque $u, v \rightarrow 0$, $c \rightarrow 0$ et par suite $d \rightarrow 0$.

Donc

$$\frac{\Delta_{x_0, u, v} f}{uv} \rightarrow \ddot{f}(x_0) .$$

□

DÉFINITION 2.2. Si $E = \mathbb{R}^n$ et $F = \mathbb{R}$, la différentielle seconde de f en x_0 est une application bilinéaire symétrique de $\mathbb{R}^n \times \mathbb{R}^n$ dans \mathbb{R} ; on appelle matrice hessienne ou simplement hessienne de f en x_0 la matrice (symétrique) de cette application. C'est la matrice notée $\text{Hess}_f(x_0)$ dont les éléments sont les $f''(x_0)(e_i, e_j)$.

3. Fonctions C^2

Si $f : U \rightarrow F$, $U \subset \mathbb{R}^n$, on peut définir les dérivées partielles secondes et plus de la façon suivante : si f admet une dérivée partielle par rapport à x_j dans U , notée $\frac{\partial f}{\partial x_j} : U \rightarrow F$, on peut considérer des dérivées partielles secondes

$$\frac{\partial \frac{\partial f}{\partial x_j}}{\partial x_i}$$

que l'on note

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

DÉFINITION 2.3. On dira que f est C^2 sur U si elle est C^1 et que $\Phi : x \rightarrow f'(x)$ de U dans $L(E, F)$ est aussi C^1 .

D'après le théorème 1.2 du chapitre 1, si $E = \mathbb{R}^p$ et $F = \mathbb{R}^q$, $f = (f_1, \dots, f_q)$ est C^2 si et seulement si les dérivées partielles

$$\frac{\partial^2 f_k}{\partial x_i \partial x_j}$$

existent et sont continues pour tous i, j, k avec $1 \leq k \leq q$, $1 \leq i, j \leq p$.

THÉORÈME 2.1. Si f est C^2 sur U , f est 2-fois différentiable en tout point de U , $x \rightarrow f''(x)$ est continue de U dans $B(E \times E, F)$ et on a :

$$(2.1) \quad \forall u, v \in E, f''(x_0)(u, v) = \frac{\partial^2}{\partial s \partial t} f(x_0 + su + tv)|_{s=t=0}$$

En particulier, si $E = \mathbb{R}^n$,

$$f''(x_0)(e_i, e_j) = \frac{\partial^2 f}{\partial x_i \partial x_j} .$$

COROLLAIRE 2.1. (lemme de Schwarz) Si $f : U \rightarrow \mathbb{R}$ avec $U \subset \mathbb{R}^n$ est C^2 , on a :

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

La hessienne de f en x_0 est alors la matrice symétrique des dérivées partielles secondes de f en x_0 .

Preuve

On va faire la démonstration dans le cas où $E = \mathbb{R}^n$ et on suppose qu'on a pris la norme $\|x\| = \sum |x_i|$. Calculons la dérivée partielle

$$\frac{\partial^2 f(x_0 + su + tv)}{\partial s \partial t} .$$

On trouve par un calcul sans malices :

$$\frac{\partial^2 f(x_0 + su + tv)}{\partial s \partial t} = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0 + su + tv) u_i v_j .$$

Donc, pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que, pour $\|u\| \leq \alpha$, $\|v\| \leq \alpha$, $|s| \leq 1$, $|t| \leq 1$, on ait :

$$\left\| \frac{\partial^2 f(x_0 + su + tv)}{\partial s \partial t} - \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) u_i v_j \right\| \leq \varepsilon \|u\| \|v\| .$$

On a besoin du :

LEMME 2.1. Si $\varphi : [0, 1] \times [0, 1] \rightarrow F$ admet une dérivée seconde partielle continue

$$\frac{\partial^2 \varphi}{\partial s \partial t} = A(s, t) ,$$

on a :

$$\varphi(1, 1) - \varphi(1, 0) - \varphi(0, 1) + \varphi(0, 0) = \int_0^1 \int_0^1 A(s, t) ds dt .$$

On applique le lemme à $\varphi(s, t) = f(x_0 + su + tv)$. On obtient ainsi

$$\Delta_{x_0, u, v} f = \int_0^1 \int_0^1 \frac{\partial^2}{\partial s \partial t} (f(x_0 + su + tv)) ds dt$$

Puis

$$\Delta_{x_0, u, v} f = \int_0^1 \int_0^1 \left(\sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0 + su + tv) u_i v_j \right) ds dt .$$

On en conclut :

$$\left\| \Delta_{x_0, u, v} f - \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) u_i v_j \right\| \leq \varepsilon \|u\| \|v\| .$$

□

4. Forme quadratique associée à la différentielle seconde

On va aussi considérer la forme quadratique (à valeurs vectorielles) associée notée encore

$$f''(x_0)(X, X) .$$

En termes des dérivées partielles :

$$f''(x_0)(X, X) = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) X_i X_j .$$

Rappelons que, si Q est une forme quadratique à valeurs scalaires sur un ev de dimension finie n , on définit son indice et sa nullité de la façon suivante : l'indice $n_-(Q)$ est le sup des dimensions des sev F de E où Q est strictement négative sur $F \setminus 0$. La nullité $n_0(Q)$ est la dimension du noyau de la matrice $A_Q = (a_{i,j})$ à ne pas confondre avec la dimension des espaces isotropes. On dit que la forme Q est *non dégénérée* si $n_0(Q) = 0$.

A changement de variables linéaire près, une forme quadratique est caractérisée par ces 2 entiers. On pose $n_+(Q) = n_-(-Q) = n - (n_- + n_0)$. Il existe une base de E telle que

$$Q(x_1, \dots, x_n) = (x_1^2 + \dots + x_{n_+}^2) - (x_{n_++1}^2 + \dots + x_{n_++n_-}^2) .$$

Invariance de la dérivée seconde en un point critique.

Si $f : U \rightarrow \mathbb{R}$ et $F : V \rightarrow U$, l'expression de la différentielle seconde de $f \circ F$ est compliquée, sauf en x_0 tel que, si $F(x_0) = y_0$, $f'(y_0) = 0$. On a alors :

$$(f \circ F)''(x_0)(X_1, X_2) = f''(y_0)(F'(x_0)X_1, F'(x_0)X_2) .$$

Cela permet de définir l'indice et la nullité d'un point critique x_0 de f comme ceux de $f''(x_0)$. On dit que le point critique est *non dégénéré* (ND en abrégé) si $f''(x_0)$ l'est.

5. Différentielles d'ordre arbitraire et dérivées partielles

On peut définir de façon analogue la différentielle d'ordre k de f en x_0 qui est par définition une application k -linéaire symétrique de $E \times \dots \times E$ dans F . On a aussi l'équivalence des notions de fonctions de classe C^k .

Si $f : U \rightarrow \mathbb{R}$, $U \subset \mathbb{R}^n$ est C^k , on peut définir des dérivées partielles d'ordre k :

$$\frac{\partial}{\partial x_{i_1}} \left(\frac{\partial}{\partial x_{i_2}} \left(\dots \left(\frac{\partial f}{\partial x_{i_k}} \right) \dots \right) \right) ,$$

pour toute suite d'indices i_1, \dots, i_k . A cause du lemme de Schwarz, l'ordre des dérivations importe peu pourvu que f soit C^k .

On est ainsi amené à noter les dérivées partielles de la façon suivante :

$$\frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}, \quad |\alpha| = \alpha_1 + \cdots + \alpha_n .$$

6. Formules de Taylor à une variable

6.1. Formules de Newton pour l'interpolation et formule de Taylor pour les polynômes. On cherche le polynôme P de degré N tel que, pour $i = 0, \dots, N$ $P(x_i) = a_i$ avec $x_i = x_0 + \varepsilon i$. On fabrique le tableau des différences successives des a_i dont on note Δ_k la premier terme de chaque ligne. Par exemple, voici ci-dessous le tableau correspondant à $a_i = x_i^3$ pour $x_0 = 4$, $\varepsilon = 1$ et $N = 3$:

$$(2.1) \quad \begin{array}{cccc} i & 0 & 1 & 2 & 3 \\ x_i & 4 & 5 & 6 & 7 \\ a_i & 64 & 125 & 216 & 343 \\ \Delta_1 & 61 & 91 & 127 & - \\ \Delta_2 & 30 & 36 & - & - \\ \Delta_3 & 6 & - & - & - \end{array}$$

On a alors la formule due à Newton :

$$(2.2) \quad P(x_0 + s\varepsilon) = a_0 + s\Delta_1 + \cdots + \frac{s(s-1)\cdots(s-(N-1))}{N!} \Delta_N .$$

Cette formule permet de calculer les interpolations de degré quelconque dans une table.

On en déduit la formule de Taylor pour un polynôme de degré N en posant $s\varepsilon = x$ et en faisant tendre ε vers 0. Il suffit de remarquer qu'on a alors :

$$\lim_{\varepsilon \rightarrow 0} \frac{\Delta_k}{\varepsilon^k} = P^{(k)}(x_0) .$$

On obtient :

$$P(x_0 + x) = \sum_{k=0}^N \frac{x^k}{k!} P^{(k)}(x_0) .$$

6.2. La formule de Taylor pour une fonction C^k . Soit f une fonction C^k d'une variable définie dans un intervalle ouvert I contenant 0 et 1. Soit $g(t) = f^{(k)}(t)$ la dérivée d'ordre k de f , il est clair que $f(t)$ est déterminée par g et les dérivées d'ordre $\leq k-1$ de f en 0 par la formule suivante :

$$f(t) = f(0) + tf'(0) + \cdots + \frac{t^{k-1}}{(k-1)!} f^{(k-1)}(0) + \int_0^t \int_0^{t_1} \cdots \int_0^{t_{k-1}} g(t_k) dt_1 \cdots dt_k .$$

On peut appliquer Fubini à cette intégrale $R(t)$ qui donne :

$$R(t) = \int_0^t \Delta(t, u) g(u) du ,$$

avec

$$\Delta(t, u) = \text{volume}(\{t = (t_1, \dots, t_{k-1}) \in \mathbb{R}^{n-1} | u \leq t_{k-1} \leq \cdots \leq t_1 \leq t\}) ,$$

qui vaut $\frac{1}{(k-1)!}(t-u)^{k-1}$.

Cela donne la formule de Taylor avec reste intégral et le développement limité à l'ordre k de f en 0.

THÉORÈME 2.2. *Si f est C^k sur un voisinage de 0 à valeurs dans un Banach F , on a :*

$$f(t) = f(0) + tf'(0) + \cdots + \frac{t^{k-1}}{(k-1)!}f^{(k-1)}(0) + \frac{t^k}{(k-1)!} \int_0^1 f^{(k)}(st)(1-s)^{k-1} ds .$$

COROLLAIRE 2.2. *Mêmes hypothèses que dans le théorème 2.2, mais $F = \mathbb{R}$, alors il existe $\tau \in]0, t[$ tel que*

$$f(t) = f(0) + tf'(0) + \cdots + \frac{t^{k-1}}{(k-1)!}f^{(k-1)}(0) + \frac{t^k}{k!}f^{(k)}(\tau) .$$

COROLLAIRE 2.3. *(formule de Taylor-Young) Mêmes hypothèses que le théorème 2.2*

$$f(t) = f(0) + tf'(0) + \cdots + \frac{t^{k-1}}{(k-1)!}f^{(k-1)}(0) + \frac{t^k}{k!}f^{(k)}(0) + o(t^k) .$$

7. Interpolation

La formule de Taylor décrit l'approximation polynomiale de la fonction f qui est asymptotiquement la meilleure quand $x \rightarrow x_0$.

On peut se poser le même problème de façon plus générale : si $I \subset \mathbb{R}$ est un intervalle et $a_0 < a_1 < \cdots < a_k$ sont des points de I et $f : I \rightarrow \mathbb{R}$, il est facile de voir qu'il existe un unique polynôme P de degré $\leq k$ tel que $\forall i, P(a_i) = f(a_i)$. Comment évaluer la différence $R(x) = f(x) - P(x)$? On a le résultat suivant :

PROPOSITION 2.1. *Supposons que f est C^{k+1} , alors pour tout $x \in I$, il existe ξ qui est dans le plus petit intervalle contenant x et les a_i tel que*

$$f(x) = P(x) + \frac{f^{(k+1)}(\xi)}{k!} \prod_{i=0}^k (x - a_i) .$$

Preuve

Soit $x = x_0$ fixé et ρ tel que

$$f(x_0) = P(x_0) + \frac{\rho}{k!} \prod_{i=0}^k (x_0 - a_i) .$$

Soit $F(x) = f(x) - P(x) - \frac{\rho}{k!} \prod_{i=0}^k (x - a_i)$, on a $F(a_i) = 0$ et $F(x_0) = 0$. Il existe donc ξ dans le plus petit intervalle contenant les a_i et x_0 tel que $F^{(k+1)}(\xi) = 0$ (appliquer le théorème de Rolle k fois). Et on a aussi $F^{(k+1)}(x) = f^{(k+1)}(x) - \rho$.

□

On utilisera en particulier ceci lorsque $k = 1$ et f est C^2 ; le polynôme P est alors l'interpolation affine.

8. Formule de Taylor à plusieurs variables

Soit $f : U \rightarrow \mathbb{R}$ de classe C^k et $m_0 \in U$, il existe un unique polynôme $P(X) = \sum_{l=0}^k P_l(X)$ de degré $\leq k$ tel que les P_l sont homogènes de degré l tel que :

$$f(x_0 + X) = P(X) + o(\|X\|^k) .$$

$P_l(X)$ s'écrit à l'aide des dérivées partielles de f :

$$P_l(X) = \frac{1}{l!} \sum_{|\alpha|=l} \frac{\partial^\alpha f}{\partial x^\alpha}(x_0) X^\alpha .$$

Ecriture pour $k = 1$: définition de la différentielle. Ecriture pour $k = 2$:

$$f(x_0 + X) = f(x_0) + f'(x_0)(X) + \frac{1}{2} f''(x_0)(X, X) + o(\|X\|^2) .$$

On peut obtenir une expression intégrale du reste.

Comme on a l'unicité, on peut souvent calculer les polynômes de Taylor directement par un calcul direct de développement limité.

Exercice : calculer le développement de Taylor à l'ordre 4 en 0 de

$$e^{xy^2} \sin x .$$

9. Réciproque du lemme de Schwarz

Soit U un ouvert de \mathbb{R}^n et $\omega = \sum_i a_i(x) dx_i$ une fonction sur $U \times \mathbb{R}^n$ à valeurs réelles. On se placera dans le cas où les a_i sont C^1 . On dira que ω est une forme différentielle (de degré 1) sur U .

Si il existe f telle que $df = \sum a_i dx_i$, on dira que ω est *exacte* ou que c'est une *différentielle totale*.

Si f est C^2 , le lemme de Schwarz implique alors que

$$\forall i, j, \quad \frac{\partial a_i}{\partial x_j} = \frac{\partial a_j}{\partial x_i} .$$

Si les a_i sont C^1 et vérifie la condition précédente, on dira que ω est *fermée*.

Toute forme exacte est fermée, la réciproque est fautive, mais localement vraie, on a le

LEMME 2.2. *Si U est un rectangle et ω fermée dans U , il existe $f : U \rightarrow \mathbb{R}$ telle que $df = \omega$.*

Si $U = \mathbb{R}^2 \setminus \{0\}$, la forme

$$"d\theta" = \frac{xdy - ydx}{x^2 + y^2}$$

est fermée, mais non exacte. La fonction θ n'est pas définie globalement sur U (ligne de changement d'heure), mais elle est définie localement à un multiple près de 2π et donc sa différentielle $d\theta$ est bien définie, fermée et non exacte.

Si $\gamma : [a, b] \rightarrow U$ est une courbe C^1 et ω est une 1-forme sur U , on peut définir $\int_{\gamma} \omega$ et ce nombre ne dépend pas du paramétrage de γ , mais seulement de son orientation.

Si $\omega = df$, $\int_{\gamma} \omega = f(\gamma(b)) - f(\gamma(a))$; si $d\omega = 0$, l'intégrale sur 2 chemins homotopes à extrémités fixées dans U est la même.

SYSTÈMES D'ÉQUATIONS NON-LINÉAIRES

Dans tout ce chapitre, les evn sont complets.

1. Rappel sur les systèmes linéaires en dimension finie (stabilité structurelle)

Soit $\sum_j a_{i,j}x_j = b_i$ un système de n équations linéaires ($1 \leq i \leq n$) à n inconnues x_j , $1 \leq j \leq n$. On dit que ce système est de Cramer si la matrice $A = (a_{i,j})$ est inversible. Dans ce cas, on a existence et unicité de la solution. On remarque que cette condition est ouverte : l'ensemble des A inversibles est un ouvert de l'espace vectoriel de toutes les matrices $n \times n$ (par exemple, écrire l'équation $\det(A) \neq 0$ qui fait apparaître l'image réciproque d'un ouvert de \mathbb{R} par une application continue).

Un meilleur argument (valable dans les Banach et n'utilisant pas le déterminant) consiste à écrire la série de Neumann suivante qui converge pour $\|B\|\|A^{-1}\| < 1$:

$$\sum_{i=0}^{\infty} (-1)^i (A^{-1}B)^i A^{-1} = A^{-1} - A^{-1}BA^{-1} + \dots ,$$

et donne l'inverse de $A + B$.

Il faut remarquer que cette série peut s'obtenir naturellement de la façon suivante : on réécrit l'équation $(A + B)x = y$ sous la forme équivalente

$$x = A^{-1}y - A^{-1}Bx = T(x) ,$$

où l'application T est contractante si $\|A^{-1}B\| < 1$. Les sommes partielles de la série coïncident alors avec les images de x par les itérés successifs de T à $(-1)^n(A^{-1}B)^n x$ près.

Cette approche n'a pas grand chose à voir avec la linéarité.

Si on note $GL(E)$ (E Banach) l'ensemble des homéomorphismes linéaires (groupe) de E dans E . On voit ainsi que $GL(E)$ est un ouvert de $L(E, E)$ (quelle topologie ?).

On peut aussi à partir de la série calculer la différentielle de $C = A^{-1}$:

$$dC = -A^{-1}dAA^{-1} .$$

En particulier, on voit que l'incertitude sur A^{-1} est contrôlée en terme de la norme de A^{-1} , cela est lié à ce que les numériciens appellent *conditionnement* du système linéaire.

2. Méthode de point fixe

Dans ce §, je note $B'(x, r)$ la boule *fermée* de centre x et de rayon r et $B(x, r)$ la boule ouverte. Soit $F : U \rightarrow E$ où $U \subset E$ est un ouvert et E est un Banach. On généralise ainsi la notion de système de Cramer :

DÉFINITION 3.1. *Un point $x_0 \in U$ tel que $F(x_0) = 0$ est dit stable si la différentielle $F'(x_0)$ est inversible.*

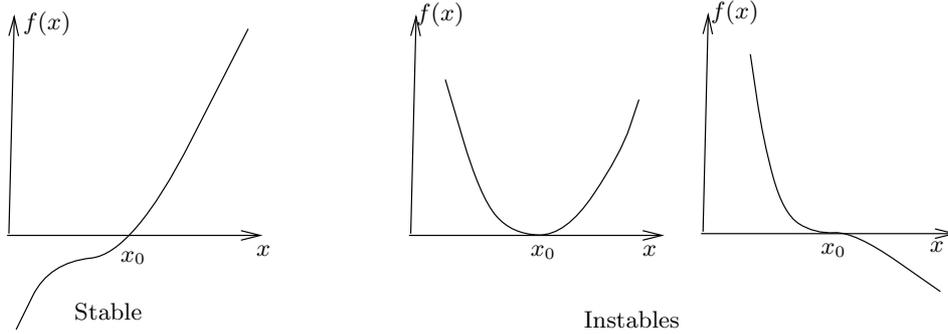


FIGURE 3.1. Zéros stables et instables

THÉORÈME 3.1. *Soit $F_\varepsilon = F + \varepsilon g : U \rightarrow E$ où $F, g \in C^1(U, E)$. On suppose que la norme C^1 de g sur U est $\leq M$, que $F(x_0) = 0$ et x_0 est un zéro stable de F . Alors il existe des constantes > 0 , r_0 , ε_0 et C ne dépendant que de F et de M telles que :*

- (1) *Pour tout ε avec $|\varepsilon| \leq \varepsilon_0$, il existe un unique $x(\varepsilon) \in B(x_0, r_0)$ tel que $F_\varepsilon(x(\varepsilon)) = 0$*
- (2) *On a :*

$$\|x(\varepsilon) - x_0\| \leq C\varepsilon .$$

COROLLAIRE 3.1. *Sous les mêmes hypothèses et si E est de dimension finie, on peut remplacer $B(x_0, r_0)$ par tout compact $K \subset U$ tel que $x_0 \in \text{Intérieur}(K)$ si x_0 est le seul zéro de F dans K .*

Preuve

La stratégie est de réécrire l'équation

$$F(x) + \varepsilon g(x) = 0$$

sous une forme équivalente

$$T_\varepsilon(x) = x$$

où T_ε est une application contractante de rapport $\frac{1}{2}$ de la boule $B'(x_0, r_0)$ qui est complète, car fermée dans l'espace de Banach E .

On pose $L = F'(x_0)$. On réécrit l'équation $F_\varepsilon(x) = 0$ sous la forme équivalente suivante

$$x = L^{-1}(L(x) - F_\varepsilon(x)) = T_\varepsilon(x) ,$$

avec

$$T_\varepsilon(x) = x - L^{-1} \circ F(x) - \varepsilon L^{-1} \circ g(x) .$$

Il est clair que $T'_0(x_0) = 0$ et donc $T'_\varepsilon(x)$ est petit pour ε petit et x proche de x_0 .

On choisit d'abord r_0 (ne dépendant que de F) tel que, si $\|x - x_0\| \leq r_0$, on ait

$$\|Id - L^{-1} \circ F'(x)\| \leq \frac{1}{4} ,$$

grâce à la continuité de $x \rightarrow F'(x)$ en x_0 .

Puis ε_0 , ne dépendant que de F et M , tel que, si $|\varepsilon| \leq \varepsilon_0$, on ait :

- $\|T_\varepsilon(x_0) - x_0\| \leq \frac{r_0}{3}$
- $\|\varepsilon L^{-1} \circ g'(x)\| \leq \frac{1}{4}$.

Alors, pour $|\varepsilon| \leq \varepsilon_0$, T_ε est une application contractante de rapport $1/2$ de $B'(x_0, r_0)$ dans lui-même et même dans $B(x_0, r_0)$: en effet, $\|T_\varepsilon(x_0) - x_0\| \leq r_0/3$ et, par les accroissements finis

$$\|T_\varepsilon(x) - T_\varepsilon(x')\| \leq \frac{1}{2}\|x - x'\|.$$

Donc

$$\|T_\varepsilon(x) - x_0\| \leq \|T_\varepsilon(x) - T_\varepsilon(x_0)\| + \|T_\varepsilon(x_0) - x_0\| \leq \frac{r_0}{2} + \frac{r_0}{3} < r_0 .$$

On en déduit, car la boule $B'(x_0, r_0)$ est complète, que la suite $T_\varepsilon^n(x_0)$ converge vers l'unique point fixe $x(\varepsilon)$ de T_ε et qu'on a la majoration

$$\|T_\varepsilon^n(x_0) - x_0\| \leq \left(1 + \frac{1}{2} + \dots + \frac{1}{2^{n-1}}\right) \|T_\varepsilon(x_0) - x_0\| ,$$

et en passant à la limite :

$$\|x(\varepsilon) - x_0\| \leq 2\|T_\varepsilon(x_0) - x_0\| ,$$

Puis $\|T_\varepsilon(x_0) - x_0\| = |\varepsilon| \|L^{-1} \circ g'(x)\| \leq C|\varepsilon|$, ce qui conclut la preuve.

□

On voit ainsi qu'une solution stable de $F(x) = 0$ est stable par petite perturbation au sens C^1 : il y a persistance d'une unique solution proche de la solution initiale.

Exemple : $F_\varepsilon(x) = 2x + \varepsilon(1 + x^5)$, $K = [-1, +1]$. Déterminer ε_0 .

2.1. Perturbations C^0 . Dans le théorème 3.1, on peut se contenter de g continue, mais alors on n'a plus unicité :

THÉORÈME 3.2. *Soit $F_\varepsilon = F + \varepsilon g : U \rightarrow E$ où $F \in C^1(U, E)$ et g continue. On suppose que $F(x_0) = 0$ et x_0 est un zéro stable de F . Alors il existe des constantes > 0 , r_0 et ε_0 telles que : Pour tout ε avec $|\varepsilon| \leq \varepsilon_0$, il existe un $x(\varepsilon) \in B(x_0, r_0)$ tel que $F_\varepsilon(x(\varepsilon)) = 0$*

Il n'est pas question de prouver ce théorème dans le cadre de ce cours sauf avec $E = \mathbb{R}$, cas où il résulte du théorème des valeurs intermédiaires.

Preuve

Soit $a < x_0 < b$ tels que $F(a)F(b) < 0$. Il existe alors ε_0 tel que, pour $|\varepsilon| \leq \varepsilon_0$, on ait :

$$(F(a) + \varepsilon g(a))(F(b) + \varepsilon g(b)) < 0$$

et le théorème des valeurs intermédiaires donne un zéro $x(\varepsilon)$ de $F + \varepsilon g$ dans $]a, b[$.

□

Remarquons que l'unicité est définitivement perdue : prendre $f(x) = x$ avec $x_0 = 0$ et $g(x) = x \sin \frac{1}{x}$. Alors $F + \varepsilon g$ admet une infinité de zéros dans tout voisinage de 0.

3. Inversion locale

On peut appliquer ce qui précède à l'équation

$$f(x) = y$$

avec y proche de $y_0 = f(x_0)$. Plus précisément, on pose $y = y_0 + \varepsilon a$ avec $\|a\| = 1$. Et on obtient l'équation : $f(x) - y_0 - \varepsilon a = 0$. La condition de stabilité s'écrit $f'(x_0)$ inversible.

DÉFINITION 3.2. $f : U \rightarrow V$ ($U \subset E$, $V \subset F$ ouverts) est un difféomorphisme C^k (avec $k \geq 1$) de U sur V si f est bijective, de classe C^k ainsi que son inverse.

PROPOSITION 3.1. *Si $f : U \rightarrow V$ est un difféomorphisme et que $U \subset E$ et $V \subset F$ sont non vides, on a : $\dim E = \dim F$.*

Soit $x_0 \in U$, $y_0 = f(x_0)$, $A = f'(x_0)$ et $B = g'(y_0)$, on a : $B \circ A = Id_E$ et $A \circ B = Id_F$. Donc E et F sont des espaces vectoriels isomorphes.

Remarque : la même proposition est vraie avec $k = 0$, i.e. si f est un homéomorphisme. C'est un théorème dur sauf si $\dim E = 1$.

On a le théorème suivant dit d'*inversion locale* :

THÉORÈME 3.3. *Soit $f : U \rightarrow F$ de classe C^k , avec $k \geq 1$, soit $x_0 \in U$ tel que $f'(x_0)$ est inversible. Alors il existe des voisinages ouverts V_1 de x_0 et V_2 de*

$y_0 = f(x_0)$ tels que f est un difféomorphisme C^k de V_1 sur V_2 . La différentielle de $g = f^{-1}$ est donnée par :

$$g'(y) = f'(g(y))^{-1} .$$

Preuve

- (1) f est une bijection de V_1 sur V_2 . On applique le théorème 3.1 à l'équation $f(x) - y_0 - \varepsilon a$ avec $\|a\| = 1$. On en déduit l'existence d'un $r_0 > 0$ et d'un $\varepsilon_0 > 0$ tels que cette équation ait une unique solution dans $B(x_0, r_0)$ pour $|\varepsilon| \leq \varepsilon_0$. Autrement dit l'équation $f(x) = y$ admet une unique solution dans $B(x_0, r_0)$ si $y \in B(y_0, \varepsilon_0)$. On pose $V_1 = B(x_0, r_0) \cap f^{-1}(B(x_0, \varepsilon_0))$ et $V_2 = B(x_0, \varepsilon_0)$.
- (2) f^{-1} est différentiable. Il suffit de montrer la différentiabilité en y_0 car on peut appliquer le théorème en tous les points de V_2 si on a choisit r_0 pour que $f'(x)$ soit inversible lorsque $x \in B(x_0, r_0)$. On a, en notant $L = f'(x_0)$:

$$y - y_0 = L(x - x_0) + o(\|x - x_0\|) ,$$

puis à cause de la majoration 2 du théorème 3.1

$$\|x - x_0\| = O(\|y - y_0\|) ,$$

d'où

$$x - x_0 = L^{-1}(y - y_0) + o(\|y - y_0\|) .$$

- (3) f^{-1} est C^k . On part de l'expression $g'(y) = f'(g(y))^{-1}$. Si f est C^2 cette expression est C^1 , car on sait déjà que g l'est. On amorce ainsi une preuve par récurrence sur k .

□

On a ainsi les :

COROLLAIRE 3.2. *Si $f : U \rightarrow E$ est C^1 de différentielle inversible en tout point, f est ouverte (l'image de tout ouvert de U est un ouvert de E) et localement injective (pour tout $x_0 \in U$, il existe un voisinage W de x_0 tel que la restriction de f à W est injective).*

COROLLAIRE 3.3. *Si $f : U \rightarrow F$ est C^k , injective et que $f'(x)$ est partout inversible, alors :*

i) $f(U) = V$ est un ouvert,

ii) f est un difféomorphisme C^k de U sur V . En particulier f est un homéomorphisme.

EXEMPLE 3.1. 1) Si $f :]a, b[\rightarrow \mathbb{R}$ est C^k à dérivée strictement positive en tout point, f est un difféo. de $]a, b[$ sur son image. C'est ainsi qu'on construit les fonctions comme arcsinus, arctangente, racine carrée, etc..

2) Soit $f : \{x > 0, y > 0\} \rightarrow \mathbb{R}^2 = \mathbb{C}$ définie par $z \rightarrow z^2$ c'est un difféo. sur le demi-plan $\Im z > 0$.

3) Les coordonnées non-linéaires (polaires, sphériques).

4) L'application $A \rightarrow A^2$ sur les matrices $n \times n$ symétriques définies positives, application $z \rightarrow \frac{z}{(1+z)^2}$ sur le disque unité de \mathbb{C} .

5) L'application $(x, y) \rightarrow (X = \cosh x \cos y, Y = \sinh X \sin y)$ est un difféo. local hors de l'origine. Montrer que c'est un difféo. du rectangle $]a, b[\times]c, d[$ sur un quart d'ellipse.

4. Théorèmes des fonctions implicites

On va déduire du théorème d'inversion locale le théorème des fonctions implicites.

Soit E, F, G 3 Banach, U un ouvert de $E \times F$, $f : U \rightarrow G$ une application C^k avec $k \geq 1$, $(x_0, y_0) \in U$ tels que $f(x_0, y_0) = 0$. On pose $L = \frac{\partial f}{\partial y}(x_0, y_0) \in L(F, G)$, $K = \frac{\partial f}{\partial x}(x_0, y_0) \in L(E, G)$.

THÉORÈME 3.4. (des fonctions implicites) Si on suppose que L est inversible, il existe A , (resp. B) des boules de centres x_0 , (resp. y_0) telles que, pour tout $x \in A$, l'équation $f(x, y) = 0$ admette une unique solution $y = g(x)$ dans B . De plus $g : A \rightarrow B$ est C^k et

$$g'(x_0) = -L^{-1} \circ K .$$

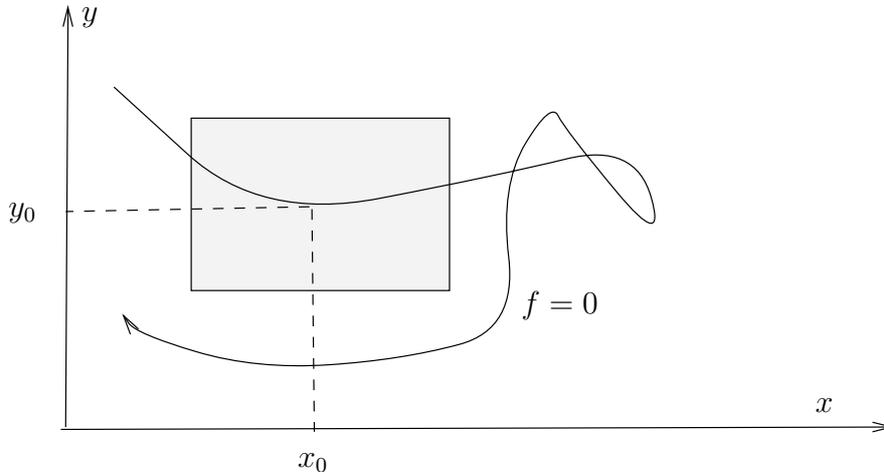


FIGURE 3.2. fonctions implicites

Preuve

On introduit $F(x, y) = (x, f(x, y))$ et on montre que l'hypothèse L inversible implique que $F'(x_0, y_0)$ est inversible ; en effet

$$dF = (dx, df = Kdx + Ldy)$$

que l'on inverse facilement en

$$(dx = dx, dy = L^{-1}(df - Kdx)) .$$

On en déduit qu'il existe un voisinage A de x_0 et un voisinage B de y_0 tel que si on applique le théorème 3.3 avec $V_1 = A \times B$, et $\Omega = V_2 \cap \{f = 0\}$, on a $\Omega = A$ et pour tout $x \in A$, il existe un unique $y = g(x) \in B$ tel que $f(x, g(x)) = 0$ et g est de classe C^k . On a

$$Kdx + L \circ g'(x)dx = 0$$

d'où le calcul de $g'(x)$.

□

Remarque : calcul de la différentielle de la fonction implicite

$Kdx + Ldy = 0$ où K, L sont les dérivées partielles de f par rapport à x, y . L'hypothèse faite que L est inversible donne :

$$dy = -L^{-1} \circ Kdx .$$

Le fait que l'équation donnant dy soit régulière justifie l'existence de la fonction implicite !

EXEMPLE 3.1. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ une application C^1 . Soit $X = \{f(x, y) = 0\} \subset \mathbb{R}^2$. Au voisinage de tout point de X tel que les 2 dérivées partielles ne s'annulent pas simultanément, l'ensemble X est le graphe d'une application C^1 $y = \varphi(x)$ ou $x = \psi(y)$.

5. Stabilité structurelle des points fixes

Soit $f : U \rightarrow U$ et $x_0 \in U$ tel que $f(x_0) = x_0$, on suppose que $L = f'(x_0)$ n'a pas 1 comme valeur propre. Si f_ε converge au sens C^1 vers f , f_ε admet pour ε petit un unique point fixe dans un voisinage V de x_0 qui tend vers x_0 avec ε : on dit que le point x_0 est *structurellement stable*.

6. Sous-variétés de \mathbb{R}^n

6.1. Motivations. Il s'agit de définir une classe de sous-ensembles réguliers de \mathbb{R}^n évitant par exemples les points de type cusp, points coniques, intersection de nappes, etc ...

Ces sous-ensembles X sont isotropes, ont une dimension $d \leq n$ et tout point de X admet un voisinage homéorphe (et même) difféomorphe à une boule ouverte de \mathbb{R}^d .

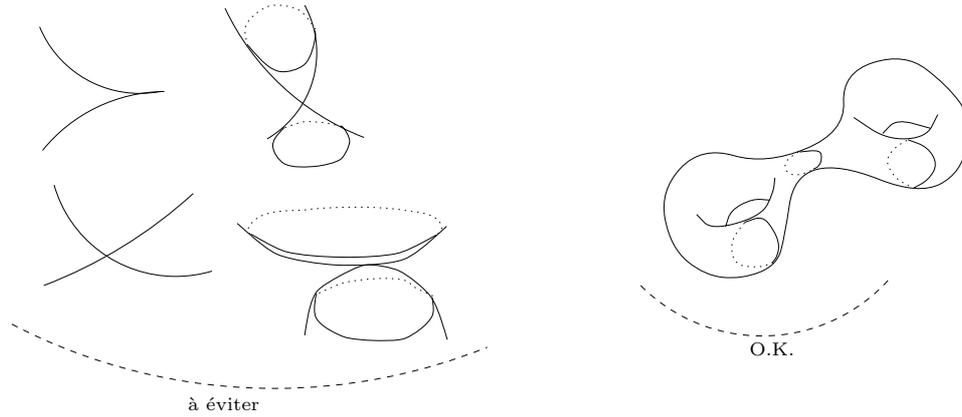


FIGURE 3.3. sous-variétés

6.2. Graphes. Soit $\mathbb{R}^n = \mathbb{R}^d \oplus \mathbb{R}^q$ avec $q = n - d$, où cette décomposition résulte d'une partition de l'ensemble $\{1, \dots, n\}$ en une partie à d éléments $\{i_1, \dots, i_d\}$ et son complémentaire.

Soit $U \subset \mathbb{R}^d$ un ouvert et $f : U \rightarrow \mathbb{R}^q$ de classe C^1 au moins. Le graphe X_f de f est le sous-ensemble de \mathbb{R}^n défini par :

$$X_f = \{(x, f(x)) | x \in U\} .$$

6.3. Contingents et espaces tangents. Soit $X \subset \mathbb{R}^n$ et $z_0 \in X$; définissons le cône tangent (ou contingent) $C_{z_0}X$ d'un sous-ensemble X en z_0 comme l'ensemble des vecteurs vitesses à l'instant $t = 0$ des courbes C^1

$$\gamma : [0, 1[\rightarrow X \subset \mathbb{R}^n$$

telles que $\gamma(0) = z_0$.

$C_{z_0}X$ est un cône : il est invariant par homothétie positive.

PROPOSITION 3.2. *Si f est différentiable en $x = x_0$, le cône tangent $C_{z_0}X_f$ est le sous-espace vectoriel de \mathbb{R}^n d'équation $Y = f'(x_0)X$; on l'appelle espace tangent (vectoriel) à X_f en z_0 et le note $T_{z_0}X_f$. L'espace tangent (affine) a pour équation :*

$$y - f(x_0) = f'(x_0)(x - x_0) .$$

6.4. Sous-variétés de \mathbb{R}^n .

DÉFINITION 3.3. *Une sous-variété de classe C^k ($k \geq 1$) et de dimension d de \mathbb{R}^n est un sous-ensemble X de \mathbb{R}^n tel que, pour tout $z_0 \in X$, il existe une partition de $\{1, \dots, n\}$ en un sous-ensemble à d éléments et un sous-ensemble à $n - d = q$ éléments et une application $f : U \rightarrow \mathbb{R}^q$, U ouvert de \mathbb{R}^d , de classe C^k , telle que X_f soit un voisinage de z_0 dans X , i.e. il existe un ouvert $\Omega \subset \mathbb{R}^n$ tel que $z_0 \in \Omega$ et $X \cap \Omega = X_f$.*

Cette définition est un peu difficile à mettre en oeuvre ; on peut vérifier que les sphères de $\{\sum_{i=1}^n x_i^2 = 1\}$ sont des sous-variétés de classe C^∞ et de dimension $n-1$ de \mathbb{R}^n . Localement, à permutation près des coordonnées, $U = \{\sum_{i=1}^{n-1} x_i^2 < 1\}$ et $f(x_1, \dots, x_{n-1}) = \pm \sqrt{1 - (x_1^2 + \dots + x_{n-1}^2)}$.

Compte-tenu de ce qui précède, on voit qu'une sous-variété de \mathbb{R}^n de dimension d admet en tout point un espace tangent, sous-espace vectoriel de \mathbb{R}^n de dimension d . On est aussi amené à considérer l'espace affine tangent qui est l'espace affine de direction $T_{z_0}X$ passant par z_0 .

6.5. Sous-variétés définies par des équations. Soit $\Omega \subset \mathbb{R}^n$ et $f = (f_1, \dots, f_q) : \Omega \rightarrow \mathbb{R}^q$ une application de classe C^k . Soit $Z_f = \{z \in \Omega | f(z) = 0\}$, on aimerait expliciter des conditions pour que Z_f soit une sous-variété de \mathbb{R}^n .

THÉORÈME 3.5. *Si en tout $z_0 \in Z_f$ la différentielle $f'(z_0) : \mathbb{R}^n \rightarrow \mathbb{R}^q$ est surjective, Z_f est une sous-variété de dimension $d = n - q$ de \mathbb{R}^n ; son espace tangent en z_0 est le noyau de $f'(z_0)$.*

Preuve

On choisit $\{i_1, \dots, i_q\}$ tels que la sous-matrice associée à ce choix de q colonnes de $f'(z_0)$ soit inversible. On décompose ainsi $\mathbb{R}^n = \mathbb{R}^d \oplus \mathbb{R}^q$ et on écrit

$$Z_f = \{(x, y) \in \mathbb{R}^d \oplus \mathbb{R}^q | f(x, y) = 0\} .$$

L'hypothèse faite sur le choix des coordonnées permet d'appliquer le théorème des fonctions implicites au voisinage de z_0 .

□

Gradients.

Si $(E, \langle \cdot | \cdot \rangle)$ est un espace euclidien et $f : U \rightarrow \mathbb{R}$ de classe C^1 (U ouvert de E), on a défini le gradient de f en un point comme le vecteur $\text{grad}f(x_0)$ qui satisfait :

$$\forall W \in E, \langle \text{grad}f(x_0) | W \rangle = f'(x_0)(W) .$$

Soit $X_a = f^{-1}(a)$. Soit $x_0 \in X_a$ tel que $f'(x_0) \neq 0$, alors X_a admet un plan tangent T_{x_0} en x_0 .

PROPOSITION 3.3. *Le vecteur $\text{grad}f(x_0)$ est normal à T_{x_0} et orienté vers la région $f > a$, i.e.*

$$f(x_0 + \varepsilon \text{grad}f(x_0)) > a$$

pour $\varepsilon > 0$ petit.

6.6. Intersections de surfaces. Soient $f, g : \mathbb{R}^3 \rightarrow \mathbb{R}$ 2 applications de classe C^1 et soient $X = \{f = 0\}$, $Y = \{g = 0\}$ et supposons que $f'(z) \neq 0$ en tout point z de X et $g'(z) \neq 0$ en tout point de Y . X et Y sont alors 2 surfaces (variétés de dimension 2) de \mathbb{R}^3 . Soit $Z = X \cap Y$. On dit que X et Y sont transversales en $z_0 \in X \cap Y$ si les espaces tangents $C_{z_0}X$ et $C_{z_0}Y$ se coupent suivant une droite D_{z_0} .

Alors si c'est le cas en tout $z_0 \in X \cap Y$, $X \cap Y$ est une courbe (sous-variété de dimension 1 de \mathbb{R}^3) d'espace tangent D_{z_0} .

EXEMPLE 3.1. *L'ensemble des quadrilatères convexes d'un plan euclidien à isométrie près peut être vu comme un ouvert de \mathbb{R}^5 : on considère le quadrilatère $Q = \{O, A, B, C\}$ avec $O = (0, 0)$, $A = (a, 0)$, $B = (b, c)$, $C = (d, e)$ et les inéquations $a > 0$, $c > 0$, $e > 0$, $be - cd > 0$, $e(b - a) - c(d - a) > 0$.*

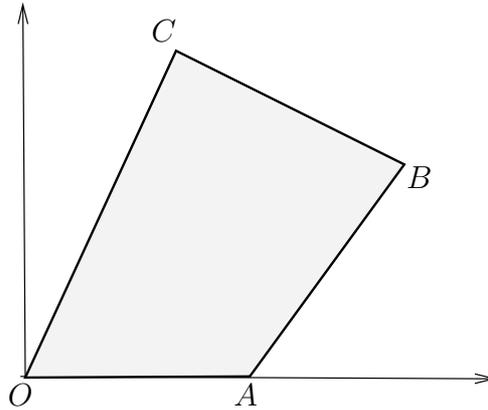


FIGURE 3.4. Un quadrilatère convexe

Ceux d'aire fixée égale à 1 forment une sous-variété A de dimension 4, les losanges une sous-variété L de dimension 2. L'intersection $A \cap L$ est une sous-variété de dimension 1. On considère les vecteurs tangents à A obtenus en ne bougeant qu'un des sommets. Est ce qu'ils engendrent l'espace tangent au point correspondant à un carré ? Ecrire une relation linéaire entre eux (penser à une rotation globale du carré autour de son centre).

7. Calculs pratiques

Linéarisation : soit $f : U \rightarrow \mathbb{R}$, $U \subset \mathbb{R}^n$, une application de classe C^1 . Considérons l'équation $f(x_1, \dots, x_n) = 0$. On associe à cette relation (non-linéaire) entre les variables x_i une relation linéaire entre les dx_i en chaque point $x \in U$ qui s'écrit :

$$\sum a_i(x) dx_i = 0, \quad a_i(x) = \frac{\partial f}{\partial x_i}(x).$$

Si $f_j(x_0) = b_j$, $1 \leq j \leq n$ et que le système linéarisé associé est de Cramer, on peut appliquer le théorème d'inversion locale au système $f_j(x) = y_j$ pour y proche de (b_j) .

EXEMPLE 3.1. : position à partir de 2 angles.

Supposons qu'on repère un point $m = (x, y)$ du demi-plan $y > 0$ par les angles $\alpha, \beta \in]0, \pi[$ de l'axe orienté Ox avec les demi-droites orientées Am et Bm avec $A = (-a, 0)$, $B = (a, 0)$.

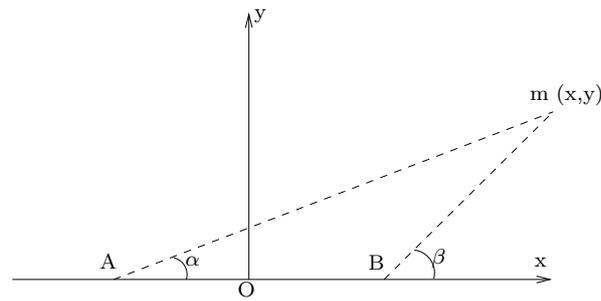


FIGURE 3.5. position à partir de 2 angles

On a les relations

$$(x + a) \sin \alpha - y \cos \alpha = 0, \quad (x - a) \sin \beta - y \cos \beta = 0,$$

qu'on pourrait résoudre explicitement en (x, y) . On peut calculer directement dx et dy en fonction de $d\alpha$ et $d\beta$ de la façon suivante qui marcherait même si on ne pouvait pas calculer explicitement x et y .

$$\sin \alpha \, dx - \cos \alpha \, dy = -(\cos \alpha (x + a) + \sin \alpha \, y) d\alpha,$$

$$\sin \beta \, dx - \cos \beta \, dy = -(\cos \beta (x - a) + \sin \beta \, y) d\beta.$$

qui est un système de Cramer facile à résoudre :

$$dx = \frac{1}{\sin(\beta - \alpha)} (r_1 \cos \beta \, d\alpha - r_2 \cos \alpha \, d\beta), \quad dy = \dots,$$

où $r_1 = \|Am\|$, $r_2 = \|Bm\|$.

EXEMPLE 3.2. : dérivées d'ordre supérieur des f.i. (ex: courbure d'une courbe implicite).

Soit $X = \{f(x, y) = 0\} \subset \mathbb{R}^2$ et (x_0, y_0) un point régulier de X . Supposons $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$ et que $y = \varphi(x)$ soit la fonction implicite correspondante. Supposons $f \in C^2$ et soit

$$p_0, q_0, r_0, s_0, t_0$$

les dérivées partielles premières et secondes de f en (x_0, y_0) . Alors φ est C^2 , on souhaite calculer $y''_0 = \varphi''(x_0)$. On écrit

$$p + qy' = 0$$

et on dérive par rapport à x :

$$r + 2sy' + ty'^2 + qy'' = 0.$$

Utilisation de développements limités.

8. Résolution d'un système d'équations et méthode de Newton

8.1. Existence. En général, les méthodes topologiques garantissent l'existence de solutions.

EXEMPLE 3.1. : *THÉORÈME DES VALEURS INTERMÉDIAIRES.* Si $f : [a, b] \rightarrow \mathbb{R}$ est continue et $f(a)f(b) < 0$, l'équation $f(x) = 0$ admet au moins une solution dans $[a, b]$.

EXEMPLE 3.2. : *DEGRÉ.* Soit $f : D \rightarrow \mathbb{R}^2$ où D est le disque unité et $F : S^1 \rightarrow \mathbb{R}^2$ sa restriction au bord. Si l'indice de F par rapport à O est non nul, on est sûr que l'équation $f = 0$ admet au moins une solution.

EXEMPLE 3.3. : *ÉCRITURE VARIATIONNELLE.* On suppose que l'équation $f = 0$ équivaut à une équation $F' = 0$ où $F : X \rightarrow \mathbb{R}$. C'est souvent le cas pour les problèmes qui ont une origine physique. Il suffit alors de savoir que F a un extremum.

EXEMPLE 3.4. : *NORMALES À UNE SURFACE S ISSUES D'UN POINT x_0 .* Les pieds des normales sont les points critiques de la fonction numérique sur S définie par $x \rightarrow d^2(x, x_0)$.

On verra plus loin des méthodes de type gradient pour traiter ces cas.

8.2. Solutions approchées. Soit $f(a) = 0$. Une solution approchée de $f(x) = 0$ est un point x tel que $\|x - a\|$ est petit.

Bien sûr, ce n'est pas un bon critère puisque a n'est pas connu.

En général, on se contente de dire que $f(x)$ est petit. Rappelons la :

DÉFINITION 3.4. Si $f : U \rightarrow F$ est C^1 , une solution a de $f(x) = 0$ est dite stable si $f'(a)$ est inversible.

Supposons que dans le compact D , $f(x) = 0$ n'ait que des solutions stables et soit $Z = f^{-1}(0)$. Alors, on a :

$$d(x, Z) \leq C\|f(x)\| .$$

8.3. Méthode de Newton. Si F est seulement continue, seules des méthodes de type dichotomie permettent de localiser les solutions. Elles sont très lentes.

Si F est différentiable, on commence par trouver des approximations grossières par dichotomie ou méthode aléatoire et on utilise ensuite une méthode de point fixe.

Soit à résoudre $F(x) = 0$.

DÉFINITION 3.5. Soit x_0 , l'équation linéarisée en x_0 est l'équation :

$$F(x_0) + F'(x_0)(x - x_0) = 0 .$$

x est solution de cette équation signifie que $(x, 0)$ est un point d'intersection du plan tangent affine au graphe de F et de l'espace $E \times 0$.

Si $F'(x_0)$ est inversible, on obtient un x_1 donné par :

$$x_1 = x_0 - F'(x_0)^{-1}(F(x_0)) .$$

On espère que x_1 sera plus proche de la solution que ne l'est x_0 . On pose $T_{x_0}(x) = x - F'(x_0)^{-1}(F(x))$. Toute solution a de $F(a) = 0$ est point fixe de T_{x_0} .

Lorsque l'inversion de $F'(x_0)$ est couteuse, on itère

$$T_{x_0}(x) = x - F'(x_0)^{-1}(F(x)) .$$

On a $T'_{x_0}(x) = Id - F'(x_0)^{-1} \circ F'(x)$. Donc si x_0 est assez proche de la solution a , supposée stable, T_{x_0} est contractante et la suite $x_{n+1} = T_{x_0}x_n$ converge vers a .

Si cette inversion n'est pas couteuse, on linéarise au point x_1 et on obtient :

$$T(x) = x - F'(x)^{-1}(F(x)) ,$$

c'est la fameuse *méthode de Newton*.

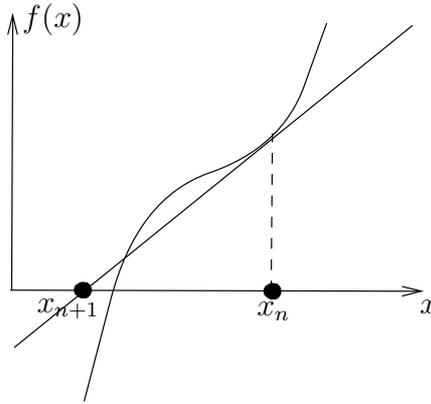


FIGURE 3.6. La méthode de Newton

DÉFINITION 3.6. • Une suite $x_n \in E$ converge géométriquement vers a avec vitesse k si on a une inégalité

$$\|x_n - a\| \leq Ck^n$$

avec $k < 1$.

• Une suite $x_n \in E$ converge rapidement vers a si on a une inégalité

$$\|x_n - a\| \leq \frac{1}{A}(A\|x_0 - a\|)^{2^n}$$

avec $A\|x_0 - a\| < 1$.

On remarque que la convergence rapide implique la convergence géométrique avec vitesse k arbitraire.

THÉORÈME 3.6. Soit a un point fixe de T supposé de classe C^1 , si $\|T'(a)\| < 1$, la suite $x_{n+1} = T(x_n)$ converge géométriquement vers a . On peut prendre pour k tout réel compris entre $\|T'(a)\|$ et 1. Ceci s'applique à $T_{x_0} = x - F'(x_0)^{-1}(F(x))$ pour x_0 assez proche de a solution stable de $F(x) = 0$.

Dans le cas où $T'(a) = 0$, on a le :

THÉORÈME 3.7. *Si T est C^2 , $T(a) = a$ et $T'(a) = 0$, la suite $x_{n+1} = T(x_n)$ converge rapidement vers a pour x_0 assez proche de a .*

On a besoin du :

LEMME 3.1. *Si $\|x_{n+1} - a\| \leq A\|x_n - a\|^2$ et $\|x_0 - a\| < 1/A$, la suite (x_n) converge rapidement vers a .*

Preuve

On écrit

$$A\|x_{n+1} - a\| \leq (A\|x_n - a\|)^2 .$$

Donc

$$A\|x_n - a\| \leq (A\|x_0 - a\|)^{2^n} ,$$

qui est la majoration cherchée. □

On peut alors prouver le théorème 3.7 : *Preuve*

Par Taylor, on a :

$$T(x) = T(a) + O(\|x - a\|^2)$$

et on applique le lemme 3.1. □

On a alors le

THÉORÈME 3.8. *Si F est C^3 , si $F(a) = 0$ et $F'(a)$ inversible, pour x_1 assez proche de a , la suite de Newton*

$$x_{n+1} = x_n - F'(x_n)^{-1}(F(x_n))$$

converge rapidement vers a .

Preuve

Si $T(x) = x - F'(x)^{-1}(F(x))$, on vérifie facilement que T est C^2 et vérifie $T'(a) = 0$. Il suffit donc d'appliquer le théorème 3.7. □

Ex : pour $F(x) = x^2 - 2$ la suite x_n converge rapidement vers $\sqrt{2}$ pour $x_1 > 0$.

La méthode de Newton a l'avantage d'une grande vitesse de convergence, mais l'évaluation de $F'(x)^{-1}$ peut être longue en grande dimension. D'autre part il faut partir d'un x_0 qui soit déjà une solution approchée de $F(x) = 0$.

Ex : appliquer la méthode de Newton pour trouver les solutions complexes de $z^3 + az + b = 0$.

Ex : appliquer la méthode de Newton pour trouver la racine carrée d'une matrice symétrique définie positive.

Un exemple : trouver les cercles tangents à 3 cercles donnés.

8.4. Méthode d'homotopie. Supposons qu'on veuille résoudre $F(x) = 0$ et qu'on sache résoudre $F_0(x) = 0$. Si on peut trouver une famille F_t , $0 \leq t \leq 1$ d'équations telles que $(t, x) \rightarrow F_t(x)$ est C^1 , $F_1 = F$ et toute solution $x(t)$ de $F_t(x(t)) = 0$ est stable, alors, sous certaines hypothèses supplémentaires, on peut trouver $x(t)$ solution de $F_t(x(t)) = 0$ dépendant de façon C^1 de t .

Hypothèse possible : il existe un compact $K \subset U$ tel que $F_t^{-1}(0) \subset K$.

Numériquement, on discrétise l'intervalle $(0, 1)$ et on résout par Newton en prenant comme donnée initiale, pour résoudre $F_{t_k}(x) = 0$, $x(t_{k-1})$.

Applications : cercles tangents à 3 cercles.

EXTRÊMAS DES FONCTIONS NUMÉRIQUES

1. Existence et topologie

Soit $f : D \rightarrow \mathbb{R}$ une fonction. Supposons f minorée et soit $m = \inf_{x \in D} f(x)$. m s'appelle le *minimum* (ou *valeur minimum* ou *borne inférieure*) de f . On ne demande pas que cette valeur soit atteinte. Si f n'est pas minorée, on pose $\inf f = -\infty$.

Une suite x_n de points de X telle que $f(x_n)$ tend vers $m = \inf f$ est appelée *suite minimisante pour f* . Une telle suite existe toujours !

Tout $x \in D$ tel que $f(x) = m$ s'appelle un *point de minimum de f sur D* .

Pour les maximas, changer f en $-f$ pour se ramener aux minimas.

Les théorèmes d'existence sont souvent donnés par la topologie, alors que la caractérisation est liée au calcul infinitésimal : si on a un point de minima, la fonction renormalisée en ce point en a aussi un.

Le théorème de base est :

THÉORÈME 4.1. *Si $f : D \rightarrow \mathbb{R}$ est continue et D compact, f a des points de minimum.*

Ce théorème ne s'applique en général pas directement (domaines de définition ouverts, non continuité (dim ∞)).

Preuve

On extrait d'une suite minimisante une sous-suite convergente (compacité) et la limite est un point de minimum (continuité).

□

DÉFINITION 4.1. *$f : D \rightarrow \mathbb{R}$ est dite semi-continue inférieurement (sci) en x_0 si $\forall \varepsilon > 0, \exists V$ voisinage de x_0 tel que*

$$\forall x \in V, f(x) \geq f(x_0) - \varepsilon .$$

On peut aussi dire que les $f^{-1}(]-\infty, a])$ sont tous fermés. Si $f(x) = \sup f_\alpha(x)$ où les f_α sont continues, f est sci.

Si on a une suite minimisante pour f qui converge vers a et que f est sci en a , a est un point de minimum de f .

On a le :

THÉORÈME 4.2. Soit $f : D \rightarrow \mathbb{R}$, f semi-continue inférieurement, soit $m = \inf f$ et supposons qu'il existe $K \subset D$ compact et $a > 0$ tels que :

$$(x \notin K) \text{ implique } (f(x) \geq m + a) ,$$

alors f a au moins un point de minimum dans K .

Si $D \subset \mathbb{R}^n$ est un ouvert, une condition suffisante pour appliquer le théorème précédent est que f soit continue et $f(x)$ tend vers $+\infty$ lorsque x tend vers l'infini de D .

Rappelons que $x_n \rightarrow \infty_D$ si $\forall K$ compact de D , il existe N tel que $\forall n \geq N$, $x_n \notin K$.

Dans un Hilbert, on utilisera le fait que les boules fermées sont compactes pour la topologie faible et des propriétés de semi-continuité.

EXEMPLE 4.1. Soit $X = \{A_1, A_2, \dots, A_N\}$ un sous-ensemble fini d'un espace euclidien E de dimension finie dont on note d la distance. La fonction $f : E \rightarrow \mathbb{R}$ définie par :

$$f(M) = \sum_i d(M, A_i)$$

est continue et on a, lorsque $M \rightarrow \infty$, $f(M) \sim Nd(O, M) \rightarrow \infty$.

EXEMPLE 4.2. Soit C une courbe fermée simple d'un espace euclidien. Soit $f : C^3 \rightarrow \mathbb{R}$ la fonction définie par :

$$f(P, Q, R) = d(P, Q) + d(Q, R) + d(R, P) .$$

Elle a un maximum sur C^3 (compacité et continuité). Ce maximum est obtenu pour un triplet (P, Q, R) de points 2 à 2 distincts. En effet, si $P' \neq P$,

$$f(P, P, R) < f(P', P, R) .$$

EXEMPLE 4.3. Soit

$$X = \{f \in C^1([-1, +1], \mathbb{R}) \mid f(-1) = -1, f(1) = 1\}$$

et $J : X \rightarrow \mathbb{R}$ définie par :

$$J(f) = \int_{-1}^{+1} x^2 f'(x)^2 dx .$$

Montrer que $\inf_X J(f) = 0$ et que ce inf n'est pas atteint.

2. Conditions du 1er ordre

THÉORÈME 4.3. Si x_0 est un point de minimum de $f : X \rightarrow \mathbb{R}$, et $V \in C_{x_0} X$ est tel que $\partial_V f(x_0)(V)$ existe, alors $\partial_V f(x_0)(V) \geq 0$.

En effet, si la dérivée directionnelle $\partial_V f(x_0)$ est < 0 et si $\gamma : [0, a[\rightarrow X$ est un arc C^1 issu de x_0 et de vitesse initiale V , pour $\varepsilon > 0$ assez petit, $f(\gamma(t)) < f(x_0)$.

COROLLAIRE 4.1. *Si x_0 est un point de minimum intérieur au domaine $D \subset E$ et f est différentiable en x_0 , $f'(x_0) = 0$.*

En effet, alors les dérivées directionnelles vérifient $\partial_V f(x_0) = -\partial_{-V} f(x_0)$.

DÉFINITION 4.2. *Si $f : U \rightarrow \mathbb{R}$, l'équation $f'(a) = 0$ s'appelle équation d'Euler du problème variationnel associé à f . Les solutions de l'équation d'Euler s'appellent points critiques de f ou extrémales du problème variationnel. Les valeurs de f en ces points s'appellent valeurs critiques de f .*

THÉORÈME 4.4. *Si $D \subset E$ est défini au voisinage de x_0 par $g(x) \geq 0$ avec $g'(x_0) \neq 0$ (domaine à bord) et $g(x_0) = 0$, et si f est différentiable et atteint son minimum en x_0 , on a $f'(x_0) = \lambda g'(x_0)$, $\lambda \geq 0$.*

L'hypothèse implique qu'il existe un voisinage U de x_0 dans E tel que $D \cap U = \{x \in U \mid g(x) \geq 0\}$. Si $g'(x_0)(V) > 0$, $x_0 + \varepsilon V \in D$ pour $\varepsilon > 0$ assez petit, et donc $f'(x_0)(V) \geq 0$. Si L, M sont 2 formes linéaires sur E , $L \neq 0$, telles que $L(V) > 0$ implique $M(V) \geq 0$, il existe $\lambda \geq 0$ tel que $M = \lambda L$ (choisir des coordonnées dans E telles que $L = x_1$).

Par exemple, pour une fonction d'une variable sur un intervalle fermé $[a, b]$, a point de minimum de f implique $f'_+(a) \geq 0$.

3. Conditions du second ordre

Soit q une forme quadratique sur un evn E .

On dit que q est définie positive s'il existe $C > 0$ telle que $q(x) \geq c\|x\|^2$. Cela veut dire que $n_-(Q) = n_0(Q) = 0$.

Soit maintenant f de classe C^2 et m_0 un point critique de f . La formule de Taylor à l'ordre 2 en m_0 s'écrit alors :

$$f(m_0 + X) = f(m_0) + \frac{1}{2}f''(m_0)(X, X) + o(\|X\|^2).$$

La forme quadratique $f''(m_0)$ s'appelle alors la *hessienne* de f en m_0 .

THÉORÈME 4.5. *Dans la situation précédente, si m_0 est un point de minimum de f , $f''(m_0) \geq 0$ ($n_- = 0$).*

Réciproquement, si $f''(m_0)$ est définie positive, m_0 est un point de minimum local strict de f .

Si le point critique m_0 est ND, il est structurellement stable au sens suivant : si f_ε converge au sens C^2 vers f , f_ε admet pour ε petit un unique point critique m_ε ND proche de m_0 et de même indice.

LEMME 4.1. (de Morse). *Si m_0 est un point critique de f non dégénéré au sens que $f''(m_0)$ est une forme quadratique non dégénérée, il existe un difféomorphisme F d'un voisinage de m_0 sur un voisinage de 0 tel que $f \circ F^{-1}(y) = f(m_0) + \frac{1}{2}f''(m_0)(y)$.*

La formule de Taylor au second ordre devient exacte dans ces coordonnées.

Application au tracé des lignes de niveaux d'une fonction réelle de 2 variables réelles : on peut utiliser le lemme de Morse pour avoir l'allure près des points critiques ND.

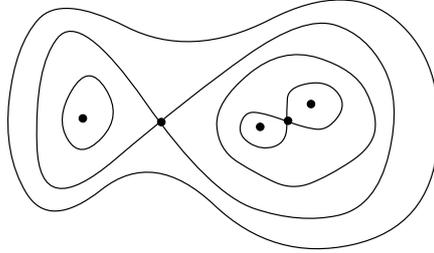


FIGURE 4.1. Lignes de niveau et points critiques

Si on a un point critique d'indice 0 ou 2 les lignes de niveaux proches de ces points ressemblent aux ellipses données comme ligne de niveaux de la différentielle seconde.

Dans le cas d'un point d'indice 0, la ligne de niveau a un point double transversal dont les directions des tangentes sont les directions isotropes de la différentielle seconde.

4. Extrémas liés

Soit à chercher les points extrémas de $y = f(x)$ sur une sous-variété X de E définie par des équations implicites $f_1(x) = \dots = f_p(x) = 0$.

THÉORÈME 4.6. *Si $m_0 \in X$ est un point de minimum local de f , il existe des λ_i tels que*

$$f'(m_0) = \sum \lambda_i f'_i(m_0) .$$

Les λ_i s'appellent les multiplicateurs de Lagrange. Ils ont souvent une interprétation physique.

On a besoin du

LEMME 4.2. *Soit L_j , $j = 1, \dots, q$ formes linéaires sur un espace vectoriel E , $F = \{x \in E \mid \forall j, L_j(x) = 0\}$, alors une forme linéaire L sur E est nulle sur F si et seulement si il existe λ_j tels que $L = \sum \lambda_j L_j$.*

Preuve

On applique le lemme précédent avec F l'espace tangent en m_0 à X , $L_j = f'_j(m_0)$ et $L = f'(m_0)$.

□

Application 1 : fonction distance à une hypersurface.

Soit X une hypersurface de \mathbb{R}^n muni de la métrique euclidienne canonique. Soit $x_0 \notin X$ et $f : X \rightarrow \mathbb{R}$ définie par $f(x) = d(x_0, x)$.

Les points critiques de f sont les points tels que le vecteur x_0x est proportionnel au gradient de $f : x_0x$ est normal à X .

L'indice de la différentielle seconde de f en x dépend de la position de x sur la normale à X en x : plus précisément l'indice change lorsqu'on franchit un des centres de courbure de la surface en x .

Application 2 : mesures de Gibbs.

Soit $X = \{1, \dots, N\}$ l'ensemble (fini) des états d'un système physique. Soit $E_i \in \mathbb{R}$ l'énergie de l'état i .

Un *état statistique* du système est une probabilité $p = (p_i)$ sur X (ie. $\sum p_i = 1$, $p_i \geq 0$). L'énergie d'un état statistique $\mathcal{E}(p)$ est l'espérance de $i \rightarrow E_i$, soit

$$\mathcal{E}(p) = \sum p_i E_i .$$

Soit $E_- = \inf E_i$, $E_+ = \sup E_i$ et $I = [E_-, E_+]$. L'énergie d'un état statistique est situé dans I .

L'entropie d'un état statistique p est :

$$S(p) = \sum p_i \text{Log } p_i .$$

Un état statistique est dit d'équilibre à l'énergie $\mathcal{E} \in I$ s'il maximise l'entropie parmi les états statistiques d'énergie \mathcal{E} . On note $P_{\mathcal{E}} = \{p \mid \sum p_i = 1, p_i \geq 0, \sum p_i E_i = \mathcal{E}\}$. Un tel état existe pour chaque $\mathcal{E} \in I$ par compacité de $P_{\mathcal{E}}$ et continuité de S . Il est unique si $E_- < E_+$ par convexité et situé dans l'intérieur du polyèdre $P_{\mathcal{E}}$.

L'écriture des multiplicateurs de Lagrange donne l'existence de 2 nombres λ et μ tels que :

$$\forall i, \text{Log } p_i = \lambda - \mu E_i ,$$

et donc

$$p_i = \frac{1}{Z} e^{-\mu E_i} ,$$

avec $Z = \sum e^{-\mu E_i}$.

Le multiplicateur de Lagrange μ s'écrit en thermodynamique $1/kT$ où k est la constante de Boltzmann et T la température. On vérifie que S est une fonction différentiable de \mathcal{E} et on a la relation :

$$\frac{\partial S}{\partial \mathcal{E}} = \frac{1}{kT} .$$

Application 3 :

extrêmes d'une forme quadratique $q(X) = (AX|X)$ sur la sphère unité d'un espace euclidien. Indices des points critiques. Minimax.

5. Catastrophes

Soit $f_\alpha : X \rightarrow \mathbb{R}$ une famille de fonctions continues sur un compact X , on suppose que $\alpha \in A$ espace topologique et que

$$\alpha \rightarrow f_\alpha$$

est continue de A dans $C(X, \mathbb{R})$, muni de la convergence uniforme. Il n'est pas difficile de voir que $\alpha \rightarrow m(\alpha) = \inf_{x \in X} f_\alpha(x)$ dépend continument de α .

Soit $Z_\alpha = \{x \in X \mid f_\alpha(x) = m(\alpha)\}$ l'ensemble compact Z_α ne dépend en général pas continument de α . C'est le cas cependant aux points α_0 tels que Z_{α_0} est réduit à un point.

L'exemple type est donné par $f_a(x) = x^4 - x^2 + ax$ avec $X = [-1, +1]$. On a : $Z_0 = \{\pm 1/\sqrt{2}\}$, alors que Z_a ne comporte qu'un point pour $a \neq 0$ voisin de 0.

Considérons la question de la différentiabilité de $m(\alpha)$. Supposons maintenant que $A \subset E$ est un ouvert d'un evn et que $\alpha \rightarrow f_\alpha(x)$ est continue pour la topologie C^2 uniforme.

On dira que α est régulier si Z_α a un seul élément $x(\alpha)$ et que f_α admet un $x(\alpha)$ un minimum non dégénéré. Alors c'est encore le cas pour les α' voisins de α et $\alpha \rightarrow x(\alpha)$ ainsi que $\alpha \rightarrow m(\alpha)$ sont différentiables. De plus

$$dm = \partial_\alpha f_\alpha(x(\alpha)) d\alpha .$$

Les α non réguliers sont appelés points de bifurcations. Ex : trouver les points de bifurcation pour

$$f_{u,v}(x) = \frac{x^4}{4} + u \frac{x^2}{2} + vx, \quad x \in \mathbb{R} .$$

6. Minimax

Une fonction $f : X \rightarrow \mathbb{R}$ ($X \subset \mathbb{R}^n$) différentiable a en général des points critiques qui ne sont ni des maximas, ni des minimas locaux. De tels points s'appellent *points cols* pour une raison évidente liée à l'exemple où f est la fonction altitude d'un point en fonction de la latitude et de la longitude de ce point. Un point critique non dégénéré est un point col si son indice $\in \{1, \dots, n-1\}$.

C'est en général la topologie de X qui force l'existence de points cols.

Ex : soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction de classe C^1 telle que $f(x, y) = x^2 - y^2$ en dehors du disque D de centre O et de rayon 1. Alors f admet un point critique qui ne soit pas un extremum local dans le disque D .

La hauteur h du col est définie par $h = \inf_\gamma (\sup_t f(\gamma(t)))$ où γ décrit les chemins $\gamma : \mathbb{R} \rightarrow \mathbb{R}^2$ tels que $\gamma(t) = (0, t)$ pour $|t| > 1$.

Montrons que f admet un point critique à la hauteur h qui est un col. S'il n'y a pas de points critiques tels que $f = h$, $Z = \{f = h\}$ est une sous-variété de \mathbb{R}^2 qui en dehors de D est formée de 4 arcs d'hyperbole.

Il faut regarder les différents cas de figures pour se convaincre qu'il est impossible que Z ne contienne pas de point de croisement.

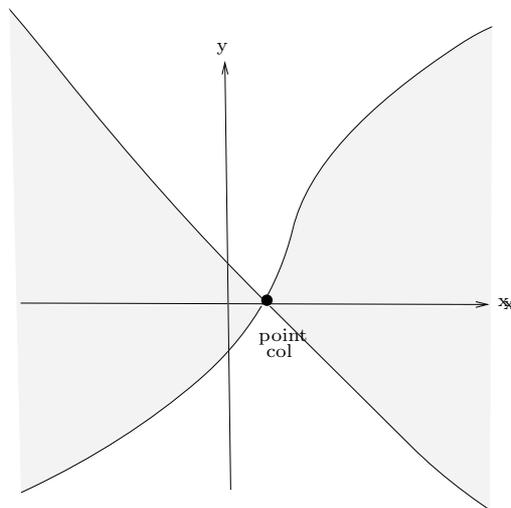
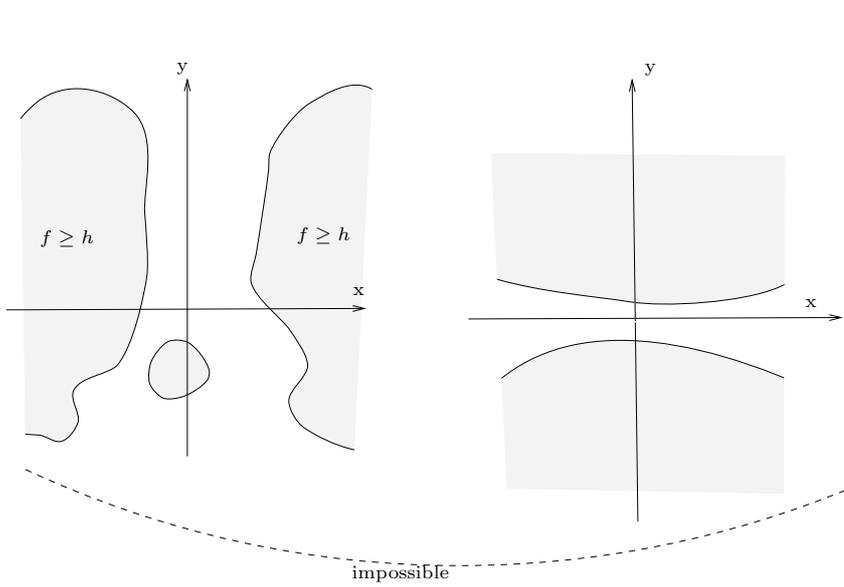


FIGURE 4.2. existence de cols

L'exemple précédent a été considérablement généralisé par Morse et ses successeurs. Cette théorie (de Morse) est maintenant un outil de base aussi bien en topologie, qu'en calcul des variations.

CHAPITRE 5

CONVEXITÉ

1. Définitions

E est un evn.

DÉFINITION 5.1. *Un convexe $X \subset E$ est un sous-ensemble de E qui contient tout segment dont il contient les extrémités :*

$$\forall x, y \in X, \forall t \text{ tel que } 0 \leq t \leq 1, tx + (1 - t)y \in X .$$

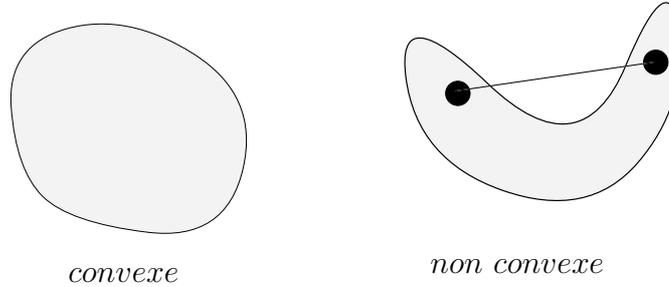


FIGURE 5.1. Convexité

Les intersections de convexes sont convexes. Si $L \in E'$, les demi-espaces $L(x) \leq a$ sont convexes.

DÉFINITION 5.2. *Un polyèdre convexe est une intersection finie de demi-espaces fermés. Un polytope est un polyèdre convexe compact.*

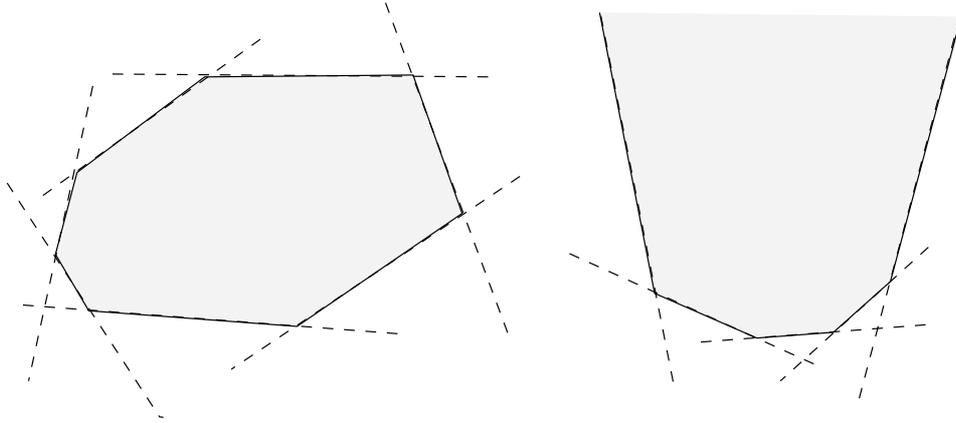
DÉFINITION 5.3. *Un domaine D est dit strictement convexe s'il est convexe et que, pour tout couple $x, y \in \partial D$ avec $x \neq y$, le segment $]x, y[$ ne rencontre pas ∂D .*

DÉFINITION 5.4. *Une fonction $f : X \rightarrow \mathbb{R}$ est dite convexe si X l'est et*

$$\forall x, y \in X, \forall 0 \leq t \leq 1, f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y) .$$

Une définition équivalente est de dire que l'épigraphe de f , $\mathcal{E}_f = \{(x, z) \mid x \in X, z \geq f(x)\}$ est convexe.

On définit de même des ensembles et des fonctions *strictement convexes* :

FIGURE 5.2. Un polytope et un polyèdre de \mathbb{R}^2

DÉFINITION 5.5. $f : D \rightarrow \mathbb{R}$ est strictement convexe si $\forall x, y \in D$ tels que $x \neq y$, pour tout t avec $0 < t < 1$, on a :

$$f(tx + (1-t)y) < tf(x) + (1-t)f(y) .$$

DÉFINITION 5.6. Soit a_1, \dots, a_n n points de E et $t_j \geq 0$ avec $\sum_{j=1}^n t_j = 1$. Le point $b = \sum t_j a_j$ est appelé barycentre des points a_j avec les coefficients t_j .

L'ensemble des barycentres de n points donnés est convexe, c'est le plus petit convexe qui les contient, appelé aussi *enveloppe convexe*.

On peut montrer que tout polytope est enveloppe convexe d'un nombre fini de points (ses sommets) et réciproquement.

On peut construire les barycentres à partir de barycentres de 2 points : on considère d'abord le barycentre a de a_1, a_2 avec les coefficients $t_1/S, t_2/S$ et $S = t_1 + t_2$, puis le barycentre cherché est celui de a, a_3, \dots, a_n avec les coefficients S, t_3, \dots, t_n . On montre ainsi l'inégalité :

$$f\left(\sum t_j a_j\right) \leq \sum t_j f(a_j)$$

dès que f est convexe.

2. Continuité

PROPOSITION 5.1. Si x_0 est intérieur à D et que f est convexe et majorée dans un voisinage de x_0 , elle est continue en x_0 et, plus précisément, si $f(x) - f(x_0) \leq M$ sur la boule fermée B de centre x_0 et de rayon a , on a :

$$(5.1) \quad \forall x \in B, |f(x) - f(x_0)| \leq \frac{M}{a} \|x - x_0\|$$

COROLLAIRE 5.1. Si $f : E \rightarrow \mathbb{R}$ est convexe et majorée elle est constante.

COROLLAIRE 5.2. Si E est de dimension finie N et $f : X \rightarrow \mathbb{R}$ convexe, elle est continue en tout point intérieur de X

Preuve

Un tel point admet en effet une base de voisinages formée de cubes centrés en ce point qui sont les enveloppes convexes de leurs 2^N sommets et donc f est majorée sur ces cubes s'ils sont contenus dans X .

□

Preuve

(de la proposition 5.1) Il suffit de montrer l'estimation (5.1) dans le cas d'une fonction convexe d'une variable avec $x_0 = f(x_0) = 0$. Elle repose sur la figure 5.3.

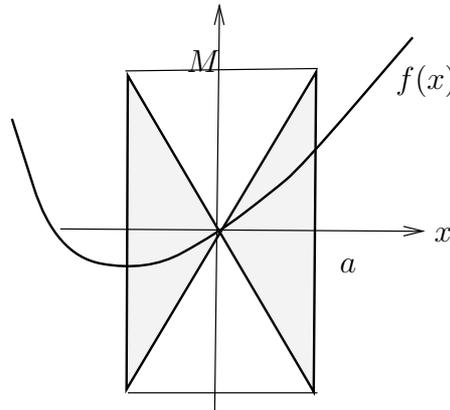


FIGURE 5.3. Continuité d'une fonction convexe

□

3. Opérations

PROPOSITION 5.2. *Si $f, g : X \rightarrow \mathbb{R}$ sont convexes et $\lambda \geq 0$, $\mu \geq 0$, $\lambda f + \mu g$ est convexe.*

Si $f_\alpha : X \rightarrow \mathbb{R}$ est une famille de fonctions convexes, leur sup est fini sur un sous convexe $Y \subset X$ et y est convexe.

Si $F : \mathbb{R} \rightarrow \mathbb{R}$ est convexe et croissante et si $f : D \rightarrow \mathbb{R}$ est convexe, $F \circ f : D \rightarrow \mathbb{R}$ est convexe.

4. Convexité et dérivées secondes

THÉORÈME 5.1. *Une fonction de classe C^2 sur un ouvert convexe X est convexe si et seulement si la dérivée seconde est une forme quadratique positive en tout point.*

Si la dérivée seconde est définie positive en tout point la fonction est strictement convexe, la réciproque est fautive (ex: $x \rightarrow x^4$ est strictement convexe).

Preuve

1) Si f est C^2 , on a :

$$f''(x_0)(X, X) = \lim_{\varepsilon \rightarrow 0} \frac{f(x_0 + \varepsilon X) + f(x_0 - \varepsilon X) - 2f(x_0)}{\varepsilon^2},$$

et ce rapport est ≥ 0 à cause de la convexité de f .

2) Montrons le :

LEMME 5.1. Si $f''(x_0)(X, X) \geq 0$ pour tout $x_0 \in D$ et $X \in E$, on a :

$$\forall x \in D, f(x_0 + X) \geq f(x_0) + f'(x_0)(X).$$

Cela résulte de la formule de Taylor pour la fonction $\varphi(t) = f(x_0 + tX)$.

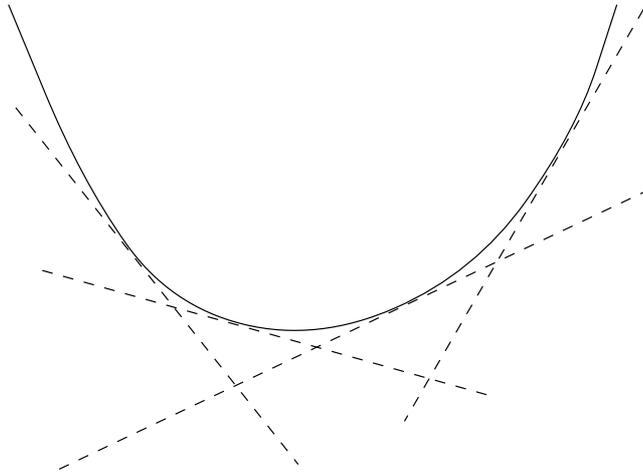


FIGURE 5.4. Une fonction convexe C^2

On en déduit que

$$f(x) = \sup_{x_0 \in D} L_{x_0}(x)$$

où $L_{x_0}(x) = f(x_0) + f'(x_0)(x - x_0)$ (car $f(x_0) = L_{x_0}(x_0)$ et le lemme). f étant un sup de fonctions affines, donc convexes est elle-même convexe (voir la proposition 5.2).

Autre démonstration : il suffit de traiter le cas où $f : [a, b] \rightarrow \mathbb{R}$ est C^2 et $\forall x, f''(x) \geq 0$. Soustrayant $L : [a, b] \rightarrow \mathbb{R}$ affine telle que $f(a) - L(a) = f(b) - L(b) = 0$, on suppose par l'absurde qu'il existe $x \in]a, b[$ tel que $f(x) > 0$. Soit x_0 tel que $f(x_0) = \sup f(x)$; il est clair que $f(x_0) > 0$ et comme $f'(x_0) = 0$ et $\forall x \in [x_0, b]$, $f''(x) > 0$, on en déduit que f est croissante sur $[x_0, b]$ et donc $f(b) > 0$ et une contradiction.

□

5. Minimas

THÉORÈME 5.2. (a) L'ensemble des points de minimum d'une fonction convexe est convexe, il est réduit à au plus un point si f est strictement convexe.

(b) Tout minimum local est global.

(c) Si D est strictement convexe et f linéaire non nulle, si f atteint son minimum c'est en un point unique de ∂D .

Preuve

Le (a) est trivial, le (b) résulte du fait que si $f(x) \geq f(x_0)$, on a $f(y) \geq f(x_0)$ pour tout y de la demi-droite d'origine x_0 orientée par $x_0 \vec{x}$. Le (c) est clair : le minimum ne peut pas être atteint à l'intérieur car $df \neq 0$ et à cause de la stricte convexité, il ne peut y avoir qu'un point de minimum au bord. □

Dérivées directionnelles :

THÉORÈME 5.3. Soit $f : X \rightarrow \mathbb{R}$ une fonction convexe et $x_0 \in X$ tel que $x_0 + tV \in X$ pour $t \geq 0$ assez petit. Alors la dérivée directionnelle $\partial_V f(x_0) \in \mathbb{R} \cup -\infty$ existe. De plus

$$f(x_0 + tV) \geq f(x_0) + t\partial_V f(x_0).$$

Tout point x_0 tel que $\forall V, \partial_V f(x_0) \geq 0$ est un minimum global de f ; si f est C^1 c'est le cas de tout point critique de f .

Preuve

La suite des

$$\frac{f(x_0 + tV) - f(x_0)}{t}$$

est décroissante quand $t \rightarrow 0_+$. □

Si $x_0 + tV \in X$ pour tout t voisin de 0 ces dérivées sont finies et $\partial_V f(x_0) \geq -\partial_{-V} f(x_0)$. En particulier, il existe une fonction affine de t qui minore $f(x_0 + tV)$ tout en valant $f(x_0)$ en $t = 0$.

COROLLAIRE 5.3. Si f est convexe sur X et x_0 est intérieur à X , il est équivalent que toutes les dérivées directionnelles en x_0 soient ≥ 0 et que x_0 soit un point de minimum (global) de f sur X .

6. Convexité forte

Dans la suite E est un espace euclidien de dimension finie ou non.

DÉFINITION 5.7. Une fonction convexe $f : X \rightarrow \mathbb{R}$ est dite fortement convexe si f est continue et s'il existe $\alpha > 0$ telle que

$$(5.1) \quad f(x) - \frac{\alpha}{2} \|x\|^2$$

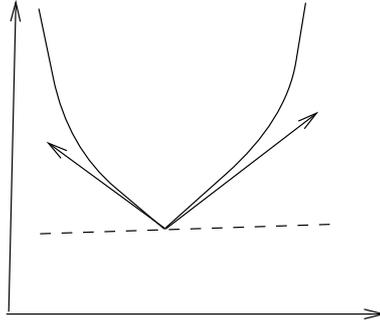


FIGURE 5.5. Dérivées à droite et à gauche d'une fonction convexe

est convexe. On dit aussi que f est α -convexe.

Une fonction α -convexe vérifie :

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2} - \frac{\alpha}{8}\|x-y\|^2.$$

Par exemple, le carré de la norme d'un espace euclidien est 2-convexe.

Une fonction fortement convexe est strictement convexe. La réciproque est fautive comme le montre l'exemple $f(x) = x^4$.

PROPOSITION 5.3. Si f est C^2 , la α -convexité équivaut à

$$\forall x \in X, \forall V \in E, f''(x)(V, V) \geq \alpha\|V\|^2.$$

Preuve

C'est clair car la différentielle seconde de $f - \frac{\alpha}{2}\|\cdot\|^2$ en x_0 est $f''(x_0) - \alpha < \cdot | \cdot >$.

□

LEMME 5.2. Si $f : X \rightarrow \mathbb{R}$ est α -convexe, elle admet une minoration

$$f(x) \geq Q(x) - a,$$

où Q est une forme quadratique définie positive. En particulier une fonction α -convexe est minorée.

Preuve

On minore f par une forme affine L : c'est toujours possible ; par exemple si f est différentiable en x_0 , $f(x) \geq f(x_0) + f'(x_0)(x - x_0)$. On se ramène ainsi au cas où $f \geq 0$, car la forme affine est minorée par $-\varepsilon Q(x) - A$.

Ensuite, on écrit

$$f(x) + f(x_0) \geq 0 + \frac{\alpha}{4}\|x - x_0\|^2.$$

□

PROPOSITION 5.4. Une fonction fortement convexe définie sur un convexe fermé de E de dimension finie (ou plus généralement sur un convexe complet d'un evn) atteint son minimum en un unique point x_0 . De plus toute suite minimisante, ie telle que $f(x_n) \rightarrow \inf f$ converge vers le point de minimum x_0 . Et on a :

$$\|x - x_0\|^2 \leq \frac{4}{\alpha}(f(x) - f(x_0)) .$$

Preuve

On suppose $m = \inf f$. Soit x et y tels que $m \leq f(x) \leq m + \varepsilon$ et $m \leq f(y) \leq m + \varepsilon$. On a donc :

$$\frac{\alpha}{8}\|x - y\|^2 \leq \varepsilon .$$

Donc toute suite x_n telle que $f(x_n) \rightarrow m$ (suite minimisante) est de Cauchy. Si x_0 est la limite de la suite (x_n) , $f(x_0) = m$.

□

Soit $f : E \rightarrow \mathbb{R}$ de classe C^2 . Supposons $\forall x, V, f''(x)(V, V) \geq \alpha\|V\|^2$. On a alors :

$$(5.2) \quad f\left(\frac{x+y}{2}\right) \leq \frac{1}{2}(f(x) + f(y)) - \frac{\alpha}{8}\|x - y\|^2 ,$$

$$(5.3) \quad f(x) \geq f(y) + f'(y)(x - y) + \frac{\alpha}{2}\|x - y\|^2 ,$$

$$(5.4) \quad (f'(y) - f'(x))(y - x) \geq \alpha\|x - y\|^2 .$$

Ces estimations sont encore vraies sous la seule hypothèse de la α -convexité.

7. Théorème de projection

Soit $X \subset (E, d)$ un fermé d'un espace métrique. Si $x_0 \in E$, on appelle projection de x_0 sur X tout point de minimum de la fonction $f : X \rightarrow \mathbb{R}$ définie par $f(x) = d(x_0, x)$. Un tel point existe dès que X est compact ou que E est un evn de dimension fini et X fermé.

THÉORÈME 5.4. Si E est un espace euclidien de dimension finie et $X \subset E$ est un convexe fermé (ou plus généralement X complet dans un espace préhilbertien), tout point $x_0 \notin X$ admet une projection unique sur X , noté $p_X(x_0)$.

Cette projection p_X contracte les distances et le point $y_0 = p_X(x_0)$ est caractérisé par :

$$(5.1) \quad \forall z \in X, z \neq y_0 : \langle y_0 - x_0 | z - y_0 \rangle \geq 0 .$$

X est donc contenu dans le demi-espace passant par y_0 et orthogonal à $x_0 y_0$. Toute suite minimisante est de Cauchy et donc converge vers la projection.

Preuve

Cela résulte de la proposition 5.4 appliquée au carré de la distance à x_0 comme fonction sur X ($F(y) = \|y - x_0\|^2$) et pour la caractérisation 5.1 du calcul de la dérivée directionnelle de F : si $V = z - y_0$ avec $z \in X$, on a :

$$\partial_V F(y_0) = 2 \langle y_0 - x_0 | z - y_0 \rangle .$$

On peut aussi donner une démonstration directe de résultat important : soit $(x_n) \in X$ une suite minimisante alors (x_n) est une suite de Cauchy. Si on pose $m = \inf_{x \in X} d(x_0, x)$ et qu'on a :

$$m \leq d(x_n, x_0) \leq m + \varepsilon \text{ et } m \leq d(x_m, x_0) \leq m + \varepsilon$$

on a aussi

$$m \leq d\left(\frac{x_n + x_m}{2}, x_0\right)$$

car le point $z = \frac{x_n + x_m}{2}$ est dans X . On en déduit : $\|x_n - x_m\|^2 \leq 4((m + \varepsilon)^2 - m^2)$.

□

8. Théorème de Hahn-Banach

Ici E est supposé de dimension finie.

THÉORÈME 5.5. *Soit $f : X \rightarrow \mathbb{R}$ convexe continue où X est supposé convexe fermé, x_0 un point de X et $\lambda_0 < f(x_0)$, il existe alors une fonction affine $L : E \rightarrow \mathbb{R}$ telle que $L(x_0) = \lambda_0$ et $\forall x \in X, L(x) \leq f(x)$.*

COROLLAIRE 5.4. *Une fonction convexe continue définie sur un convexe fermé de E est sup de fonctions affines (et réciproquement).*

Preuve

On a besoin du

LEMME 5.3. *Si $Y \subset F$ est un convexe fermé et $y_0 \notin Y$, il existe une forme linéaire M sur F qui sépare strictement y_0 et Y : il existe $\alpha \in \mathbb{R}$ tel que*

$$M(y_0) < \alpha < M(Y) .$$

Si z_0 est la projection de y_0 sur Y , on prend pour M la forme produit scalaire avec $z_0 - y_0$.

On applique le lemme avec Y l'épigraphe de f et $y_0 = (x_0, \lambda_0)$. $M(x, \lambda) = K(x) + \beta\lambda$. On voit que $\beta \geq 0$ (faire tendre λ vers $+\infty$). Puis $\beta > 0$, appliquer l'inégalité sur M à $(x_0, f(x_0))$. On a donc

$$\forall x \in X, \beta(f(x) - f(x_0)) > K(x) - K(x_0) ,$$

et donc

$$f(x) > L(x) = \frac{1}{\beta}(K(x_0) - K(x)) + \lambda_0 .$$

□

9. Algorithme de relaxation

Soit à minimiser la fonction C^2 et α -convexe $f : \mathbb{R}^N \rightarrow \mathbb{R}$, où \mathbb{R}^N est muni de la structure euclidienne canonique. On construit une suite $X^n = (x_1^n, \dots, x_N^n)$ de la façon suivante. On pose $X_1^n = X^n$ et soit $1 \leq i \leq N$ et

$$X_i^n = (x_1^{n+1}, \dots, x_{i-1}^{n+1}, x_i^n, \dots, x_N^n) ,$$

on calcule X_{i+1}^n en minimisant par rapport à la variable $t = x_i$. Lorsqu'on arrive à $X_{N+1}^n = X^{n+1}$, on recommence et on calcule X_1^{n+1} en minimisant par rapport à x_1 . En fait ce qui compte est l'ordre cyclique sur les i , $1 \leq i \leq N$.

On passe ainsi en N étapes de

$$X^n = X_N^n = (x_1^n, \dots, x_N^n)$$

à

$$X^{n+1} = (x_1^{n+1}, \dots, x_N^{n+1}) .$$

On remarque aussi que l'algorithme dépend du choix d'une base de E .

PROPOSITION 5.5. *Si f est C^2 et α -convexe, l'algorithme de relaxation converge vers le point de minimum de f .*

Exemple : le problème de Dirichlet discret.

Preuve

La suite des $f(X_i^n)$ est décroissante et minorée, donc converge vers une limite l .

On applique (5.3) à f comme fonction de x_i :

$$f(X_i^n) - f(X_{i+1}^n) \geq \frac{\alpha}{2}|x_i^n - x_i^{n+1}|^2 = \frac{\alpha}{2}\|X_i^n - X_{i+1}^n\|^2 .$$

On en déduit $X_i^n - X_{i+1}^n \rightarrow 0$ et donc, comme l'ensemble de ces points est borné, que les $\partial_i f(X_j^n)$ tendent vers 0 lorsque $n \rightarrow \infty$ (car $\partial_i f(X_{i+1}^n) = 0$).

De (5.4), on déduit ainsi, si X_0 est le point de minimum de f :

$$\alpha\|X_i^n - X_0\|^2 \leq f'(X_i^n)(X_i^n - X_0) ,$$

et donc

$$\|X_i^n - X_0\| \leq \frac{1}{\alpha}\|f'(X_i^n)\| ,$$

ce qui prouve la convergence.

□

10. Algorithme de gradient à pas fixe

Les hypothèses sur f sont les mêmes, mais on ne suppose plus $E = \mathbb{R}^N$: E est un espace euclidien. On suppose en outre que

$$\|f'(x) - f'(y)\| \leq C\|x - y\| .$$

Soit $\mu > 0$, on considère l'algorithme

$$X_{n+1} = X_n - \mu \text{grad}f(X_n) ,$$

appelé algorithme de gradient à pas μ .

PROPOSITION 5.6. *L'algorithme de gradient à pas μ converge vers le point de minimum dès que*

$$0 < \mu < \frac{2\alpha}{C^2} .$$

Preuve

Soit a le point de minimum de f , on a :

$$X_{n+1} - a = X_n - a - \mu(\text{grad}f(X_n) - \text{grad}f(a)) ,$$

Elevant au carré, il vient :

$$\|X_{n+1} - a\|^2 = \|X_n - a\|^2 + \mu^2 \|f'(X_n) - f'(a)\|^2 - 2\mu(f'(X_n) - f'(a))(X_n - a) ,$$

et utilisant (5.4) :

$$\|X_{n+1} - a\|^2 \leq \|X_n - a\|^2(1 + C^2\mu^2 - 2\mu\alpha) ,$$

et si $0 < \mu < \frac{2\alpha}{C^2}$,

$$0 < 1 + C^2\mu^2 - 2\mu\alpha < 1 ,$$

ce qui prouve la convergence. □

Remarque : si on fait $\mu \rightarrow 0$, on trouve l'équation différentielle

$$X'(t) = -\text{grad}f(X(t)) ,$$

dont les solutions convergent toutes vers le point de minimum a dès que f est C^2 (théorème de Cauchy) et que a est le seul point de minimum.

En effet, $\varphi(t) = f(X(t))$ vérifie

$$\varphi'(t) = -\|\text{grad}f(X(t))\|^2$$

et donc tend vers une limite l en décroissant. Si $l > f(a)$ le gradient est minoré en norme sur la trajectoire et donc $\varphi(t) \rightarrow -\infty$, d'où contradiction.

Exercice : que donne la méthode du gradient à pas fixe sur la fonction x^{2n} ?
 $x_{N+1} = x_N - \mu x_N^{2n-1}$.

11. Le cas des formes quadratiques

Si $f(x) = \frac{1}{2} \langle Ax|x \rangle - \langle b|x \rangle$, où A est définie positive, on voit que, si x_0 est le point de minimum, et $y_n = x_n - x_0$:

$$y_{n+1} = (1 - \mu A)y_n .$$

Et donc la méthode converge si les valeurs propres $0 < \lambda_1 \leq \dots \leq \lambda_N$ de A vérifient

$$1 > 1 - \mu\lambda_i > -1 .$$

La convergence est optimale pour

$$\mu = \frac{2}{\lambda_1 + \lambda_N}$$

et alors elle est géométrique de raison

$$\frac{\lambda_N - \lambda_1}{\lambda_N + \lambda_1} .$$

Soit maintenant $B : E \rightarrow E$ une application linéaire inversible, où E est un evn de dimension finie, et considérons le système d'équations linéaires

$$Bx = c .$$

On appelle *conditionnement* de B et on note $\delta(B)$ le nombre ≥ 1

$$\delta(B) = \|B\| \cdot \|B^{-1}\| .$$

Si on considère les équations :

$$Bx_0 = c_0, \quad B(x_0 + y) = c_0 + d ,$$

on a

$$\frac{\|y\|}{\|x_0\|} \leq \delta(B) \frac{\|d\|}{\|c_0\|} ,$$

qui contrôle l'erreur relative sur la solution en termes de l'erreur relative sur le second membre.

Si la norme sur E est euclidienne et qu'on pose

$$A = B^t B$$

on a :

$$\|B\| = \sqrt{\lambda_N} , \\ \|B^{-1}\| = \frac{1}{\sqrt{\lambda_1}} ,$$

et donc

$$\delta(B) = \sqrt{\frac{\lambda_N}{\lambda_1}} .$$

Soit $f(x) = \frac{1}{2} \|Bx - c\|^2$, le point de minimum de f est la solution de $Bx = c$.

On a :

$$f(x) = \frac{1}{2} \langle Ax|x \rangle - \langle B^t|c \rangle + \frac{1}{2} \|c\|^2 .$$

La méthode de gradient à pas fixe évoquée plus haut converge de façon optimale avec une raison :

$$\frac{\delta(B)^2 - 1}{\delta(B)^2 + 1} ,$$

liée au conditionnement de B .

12. Programmation linéaire

Un problème de programmation linéaire est un problème de minimisation dans \mathbb{R}^n d'une fonctionnelle linéaire sur un sous-ensemble défini par un ensemble fini d'inéquations linéaires, i.e. une intersection finie de demi-espaces affines.

La méthode la plus célèbre, méthode dite *du simplexe*, repose sur une étude des polyèdres et polytopes de \mathbb{R}^n . Nous nous contenterons d'une étude théorique, la mise en place pratique étant extrêmement technique !!

DÉFINITION 5.8. *Un polyèdre (convexe) P est un sous-ensemble fermé de \mathbb{R}^n défini par un nombre fini d'inéquations :*

$$\forall i, 1 \leq i \leq N, L_i(x) \leq a_i ,$$

où $L_i \in (\mathbb{R}^n)'$ et $a_i \in \mathbb{R}$. Un polytope est un polyèdre (convexe) compact.

Dans la suite de ce chapitre : P est toujours un polyèdre convexe, polyèdre signifie polyèdre convexe, et K un polytope.

La dimension d'un polyèdre est la dimension du plus petit sous-espace affine qui le contient.

Les polytopes de dimension 1 sont les intervalles compacts, ceux de dimension 2 les polygones bornés convexes.

DÉFINITION 5.9. *Un sommet de P est un point $x_0 \in P$ tel qu'il existe un hyperplan H affine de \mathbb{R}^n avec*

$$H \cap P = \{x_0\} .$$

Il existe donc une forme linéaire constante sur H qui atteint son min en l'unique point x_0 .

Une facette est l'ensemble des points de minimum d'une forme linéaire. C'est un polyèdre.

Les sommets sont donc exactement les facettes réduites à un point. Une arête est une facette de dimension 1.

PROPOSITION 5.7. *Si $A : E \rightarrow F$ est linéaire et P un polyèdre (resp. polytope) de E , $A(P)$ est un polyèdre (resp. polytope).*

COROLLAIRE 5.5. *Si $f : P \rightarrow \mathbb{R}$ est linéaire et minorée, elle atteint sa borne inférieure.*

Un polyèdre de \mathbb{R} minoré est un intervalle fermé à gauche.

Remarque : un polyèdre peut ne pas avoir de sommets, par exemple si P est un sous-espace affine de $\dim \geq 1$.

Preuve

(de la proposition 5.7).

On se ramène au cas où A est surjective, puis au cas où $A : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ est la projection oubliant la première coordonnée :

$$A(x_0, x') = x' .$$

Si P est défini par

$$L_i(x') \leq x_0, \quad L_j(x') \geq x_0, \quad L_k(x') \leq 0 ,$$

son image est définie par :

$$\forall i, j, k, \quad L_j(x') - L_i(x') \geq 0, \quad L_k(x') \leq 0 .$$

□

PROPOSITION 5.8. *Si $x_0 \in P$, et $Z(x_0) = \{i | L_i(x_0) = a_i\}$, alors il existe un voisinage U de x_0 dans E tel que*

$$U \cap K = \{L_i(x) \leq a_i \mid i \in Z(x_0)\} .$$

C'est l'intersection d'un cône, le cône tangent en x_0 , avec U .

On élimine les inéquations qui donnent $L_i(x_0) \neq a_i$. Cela sélectionne le voisinage.

PROPOSITION 5.9. *L'ensemble des sommets d'un polyèdre P est fini. Ce nombre est majoré par C_N^n où N est le nombre d'équations.*

Si x_0 est un sommet, il existe n hyperplans $L_i(x) = a_i$ dont x_0 soit exactement l'intersection. Si on prend les L_i de la proposition précédente les $\{L_i(x) = a_i\}$ ont x_0 comme intersection, sinon x_0 serait point milieu de 2 points distincts de P et x_0 ne serait point de minimum unique d'aucune forme linéaire. On prend un système libre maximal (de rang n) dans les L_i .

PROPOSITION 5.10. *Si $f : K \rightarrow \mathbb{R}$ est une fonction affine sur le polytope K , elle atteint son inf en au moins un sommet $x_0 \in S(K)$. En particulier, tout polytope non vide a un ensemble fini non vide de sommets et est l'enveloppe convexe de ses sommets.*

Preuve

On le prouve par récurrence sur la dimension n de K en supposant la forme linéaire f non triviale : l'ensemble des points de min est alors un polytope de dimension $< n$.

On a ainsi besoin du

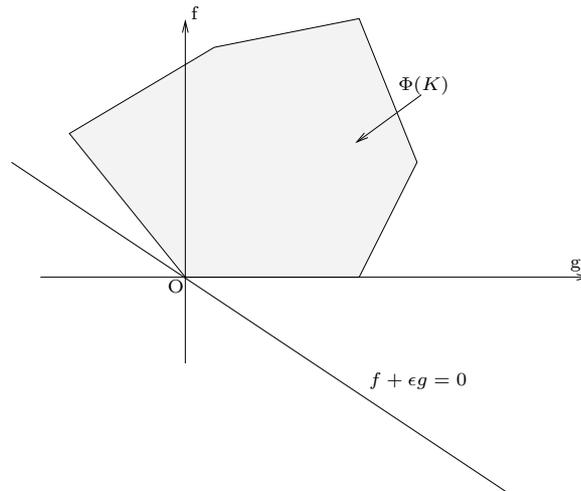


FIGURE 5.6. Existence de sommets

LEMME 5.4. *Si x_0 est sommet d'une facette L de K , c'est un sommet de K . Le résultat reste vrai pour un polyèdre.*

Soit f affine, ≥ 0 sur K et telle que $f^{-1}(0) \cap K = L$ et g affine, ≥ 0 sur L et $g^{-1}(0) \cap L = x_0$. On considère l'image de K par $\Phi = (g, f)$; c'est un polytope M de \mathbb{R}^2 contenu dans $y \geq 0$ et tel que $M \cap y = 0 = [0, A]$ (c'est l'image par $(g, 0)$ de L). On considère alors $f + \epsilon g$ pour $\epsilon > 0$ petit ; cette forme affine atteint son minimum uniquement au point $\Phi^{-1}(0) = x_0$.

□

Il reste à prouver l'affirmation sur l'enveloppe convexe $C = C(S(K))$ qui est contenue dans K . Si $x_0 \in K \setminus C$, on peut trouver par Hahn-Banach une forme affine f nulle en x_0 et $\geq a > 0$ sur C . Tout sommet où f atteint son minimum n'est pas dans C , c.q.f.d.

COROLLAIRE 5.6. *Toute arête compacte d'un polyèdre P est l'enveloppe convexe de 2 sommets de P .*

Une telle arête est un polytope de dimension 1 donc un intervalle compact $[a, b]$. Les extrémités de cette facette sont des sommets du polytope K .

PROPOSITION 5.11. *Soit $f : K \rightarrow \mathbb{R}$ et $m = \inf f$. Pour tout sommet x_0 de K tel que $f(x_0) > m$, il existe une arête (x_0, x_1) issue de x_0 telle que $f(x_1) < f(x_0)$.*

Soit t une coordonnée affine sur \mathbb{R}^n tel que $t(x_0) = 0$ et $t > 0$ sur $K \setminus x_0$. Alors $(t = \epsilon > 0) \cap K = K_1$ est un polytope tel que l'enveloppe convexe de K_1 et de x_0 est un voisinage de x_0 dans K .

Il y a une bijection entre les sommets de K_1 et les arêtes de K issues de x_0 . f a un min local en x_0 ssi $f \geq f(x_0)$ sur K_1 . Sinon, on peut trouver une arête

descendante : si x_1 est un sommet de K_1 tel que $f(x_1) < f(x_0)$, il existe une arête de K dont l'un des sommets est x_0 et contenant x_1 : en effet si f est définie sur $t = \varepsilon$ comme atteignant son minimum = 0 en x_1 uniquement et étendue en une forme affine F prenant la valeur 0 en x_0 , l'arête cherchée est l'ensemble des points de minimum de F sur K .

Il reste à observer que, par convexité, les points de min local et global sont les mêmes.

Méthode du simplexe :

soit à minimiser $f : K \rightarrow \mathbb{R}$ où K est un polytope (il suffit que les $K \cap \{f \leq m\}$ soient des polytopes), on part d'un sommet x_0 et si il n'y a pas d'arêtes issue de x_0 sur laquelle f est strictement décroissante, on a fini, sinon, on descend et on recommence avec l'autre extrémité.

CHAPITRE 6

CALCUL DES VARIATIONS

1. Exemples

Dans ce chapitre, on aborde des méthodes pour traiter des problèmes d'optimisation où l'inconnue est une fonction $f : [a, b] \rightarrow \mathbb{R}$ vérifiant des conditions aux limites

$$f(a) = A, f(b) = B,$$

où A et B sont donnés, et la fonctionnelle $\Phi(f)$ est de la forme suivante : on se donne une fonction $L : [a, b] \times \mathbb{R} \times \mathbb{R}$ de 3 variables, le *lagrangien* $L(x, y, v)$, et on pose

$$\Phi(f) = \int_a^b L(x, f(x), f'(x)) dx.$$

Bien sûr on peut généraliser à des fonctions $f : D \rightarrow E$ où D est un domaine de \mathbb{R}^p !!

Beaucoup de problèmes de géométrie ou de physique prennent cette forme.

1.1. Le brachistochrone. Il s'agit de construire un toboggan entre le point $O = (0, 0)$ et le point $P = (b, B)$, $b > 0, B > 0$ du plan des (x, y) avec l'axe des y dirigé vers le bas et on considère $g = 1, m = 2$.

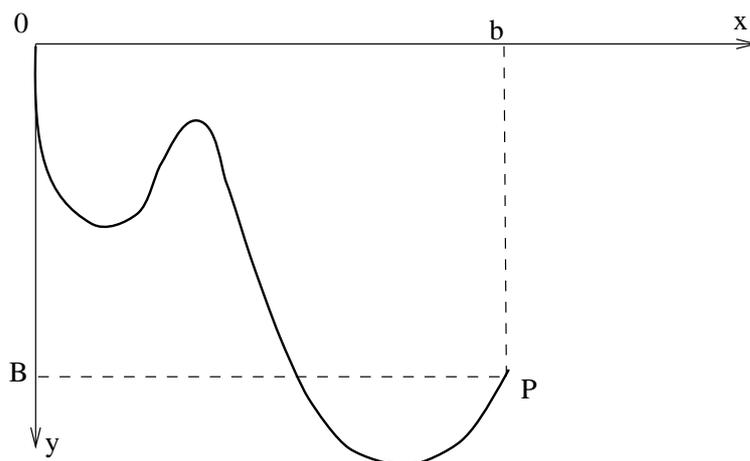


FIGURE 6.1. Le brachistochrone

La conservation de l'énergie totale s'écrit :

$$(\dot{x}^2 + \dot{y}^2) - y = 0 ,$$

ou encore :

$$(1 + y'^2)\dot{x}^2 = y ,$$

et donc $T = \Phi(f)$ est donné par :

$$T = \int_0^b \sqrt{\frac{1 + y'^2}{y}} dx .$$

On est ainsi amené à minimiser l'intégrale T parmi les fonctions $y = f(x)$ telles que $f(0) = 0$ et $f(b) = a$. Le lagrangien est $L(x, y, v) = \sqrt{\frac{1+v^2}{y}}$.

1.2. Les surfaces minimas de révolution et le fil pesant. Soit une surface de révolution d'axe Ox située entre les cercles (d'axes Ox) de centres $(a, 0, 0)$ et $(b, 0, 0)$ avec $a < b$ et de rayons $A, B > 0$.

Si on suppose que la méridienne est donnée par $y = f(x)$, $f(x) > 0$ et $f(a) = A, f(b) = B$, l'aire vaut :

$$A = 2\pi \int_a^b y \sqrt{1 + y'^2} dx .$$

Le problème de la position d'équilibre d'un fil pesant entre les points (a, A) et (b, B) est voisin : l'énergie totale est donnée par

$$E = \int_a^b y \sqrt{1 + y'^2} dx ,$$

mais on a une contrainte sur la longueur totale :

$$l = \int_a^b \sqrt{1 + y'^2} dx .$$

Le lagrangien du problème est $L = y\sqrt{1 + v^2}$.

1.3. Les géodésiques d'une surface. Soit $z = K(x, y)$ une surface de \mathbb{R}^3 et $y = f(x)$, $z = K(x, f(x))$ une courbe tracée sur celle-ci entre les points $(a, A, K(a, A))$ et $(b, B, K(b, B))$.

La longueur de cette courbe est :

$$L = \int_a^b \sqrt{E(x, y)y'^2 + 2F(x, y)y' + G(x, y)} dx ,$$

avec $p = \partial_x K$, $q = \partial_y K$ et $E = 1 + q^2$, $F = pq$, $G = 1 + p^2$. On cherche une courbe de longueur minimale d'origine A et d'extrémité B .

On va d'abord déterminer des conditions que doit vérifier le minimum : l'équation d'Euler-Lagrange.

Attention à 2 choses :

1) un problème du type précédent peut très bien n'avoir aucune solution.

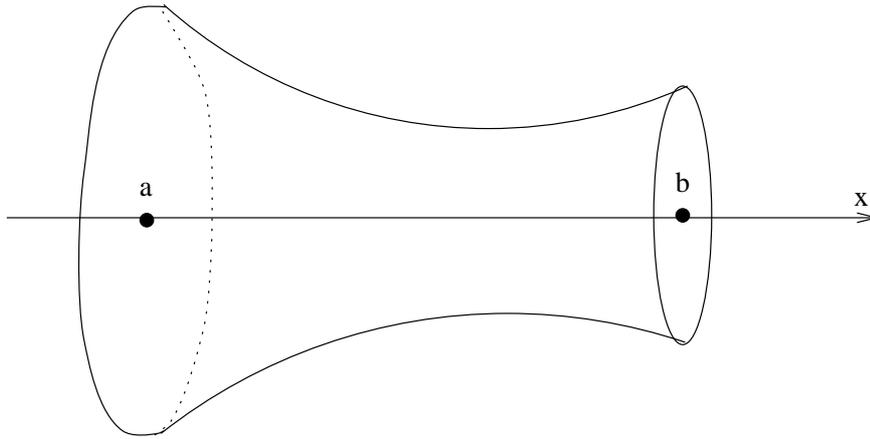


FIGURE 6.2. Surface de révolution

2) Mmême si l'équation d'Euler-Lagrange a des solutions, elles peuvent ne pas donner le minimum.

2. Equation d'Euler-Lagrange

On suppose que $L : U \rightarrow \mathbb{R}$ (U ouvert de \mathbb{R}^3) est C^2 , l'équation d'Euler-Lagrange associée à $L(x, y, v)$ est l'équation différentielle d'inconnue $y = f(x)$, $y' = f'(x)$ donnée par

$$\Omega(x, y(x), y'(x)) = -\frac{d}{dx}\left(\frac{\partial L}{\partial v}(x, y(x), y'(x))\right) + \frac{\partial L}{\partial y}(x, y(x), y'(x)) = 0.$$

C'est une équation différentielle du second ordre en y . Cela veut dire qu'il faut en principe 2 conditions pour avoir une solution bien déterminée ; dans notre problème les 2 conditions sont en général données par les conditions aux extrémités de l'intervalle $f(a) = A$, $f(b) = B$.

DÉFINITION 6.1. Toute solution de l'équation d'Euler-Lagrange du problème s'appelle extrémale du lagrangien L .

THÉORÈME 6.1. Soit $y_\varepsilon = y_0(x) + \varepsilon z(x)$ avec y_0, z de classe C^2 , et $z(a) = z(b) = 0$, on a

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \Phi(y_\varepsilon) = \int_a^b \Omega(x, y_0(x), y_0'(x)) z(x) dx .$$

Preuve

On a :

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \Phi(y_\varepsilon) = \int_a^b \left(\frac{\partial L}{\partial y}(x, y_0, y_0') z + \frac{\partial L}{\partial v}(x, y_0, y_0') z' \right) dx ,$$

et on intègre par parties le 2ème terme :

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \Phi(y_\varepsilon) = \int_a^b \left(\frac{\partial L}{\partial y}(x, y_0, y'_0)z - \frac{d}{dx} \left(\frac{\partial L}{\partial v}(x, y_0, y'_0) \right) z \right) dx ,$$

car le terme tout intégré est nul aux 2 bornes.

□

COROLLAIRE 6.1. *Si $y_0(x)$ est C^2 et vérifie*

$$\Phi(y_0) = \inf_{y(a)=A, y(b)=B} \Phi(y) ,$$

où le inf porte sur les fonctions C^1 vérifiant les conditions aux limites, $y_0(x)$ vérifie l'équation d'E-L :

$$\Omega(x, y_0, y'_0) = 0 .$$

Preuve

On a alors

$$\int_a^b \Omega(x, y_0(x), y'_0(x))z(x)dx = 0$$

pour tout z , C^2 et nulle aux extrémités de l'intervalle. Cela implique le résultat.

□

EXEMPLE 6.1. *Principe de moindre action de Hamilton-Jacobi et équations de Newton. Si on prend $L(x, y, v) = \frac{1}{2}v^2 - V(y)$, l'équation d'Euler-Lagrange est :*

$$\frac{d^2y}{dx^2} = -\frac{dV}{dy} ,$$

c'est l'équation de Newton du mouvement d'une particule dans la potentiel V .

3. Théorème de Hilbert

Soit $L(x, y, v)$ un lagrangien et Ω un ouvert de \mathbb{R}^2 .

DÉFINITION 6.2. *Un champ d'extrémales est une application*

$$(x, y) \rightarrow v(x, y)$$

de classe C^1 sur Ω telle que toute fonction $f : I \rightarrow \mathbb{R}$, C^2 telle que $\forall x \in I$, $(x, f(x)) \in \Omega$ et que $f'(x) = v(x, f(x))$ est une extrémale pour le lagrangien L .

THÉORÈME 6.2. *Supposons que L est convexe en v et Ω simplement connexe et admettant un champ d'extrémales $v(x, y)$, alors toute extrémale f_0 dont le graphe est contenu dans Ω est un minimum de $\Phi(f)$ parmi les fonctions dont le graphe a mêmes extrémités et est contenu dans Ω .*

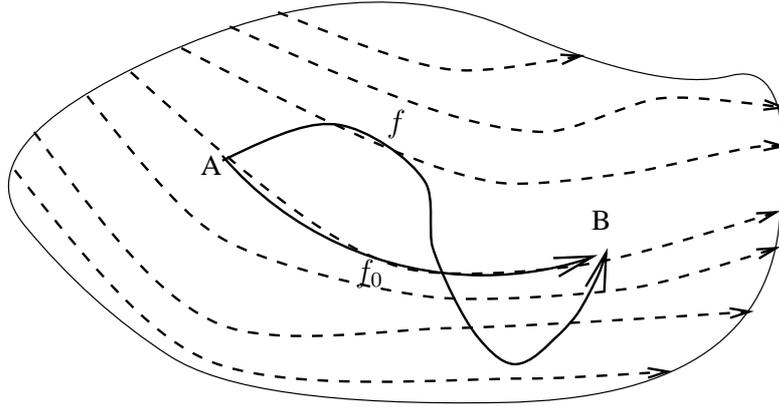


FIGURE 6.3. Champ d'extrémales

Preuve

On suppose pour simplifier la preuve que L est indépendant de x . Si $f'(x) = v(x, f(x))$, $f''(x) = p(x, f(x)) + q(x, f(x))v(x, f(x))$, où p, q sont les dérivées partielles de v . On peut écrire l'équation d'E-L pour f en remplaçant f' et f'' comme plus haut :

$$\frac{\partial^2 L}{\partial v \partial y} v + \frac{\partial^2 L}{\partial v^2} (p + qv) = \frac{\partial L}{\partial y},$$

où cette identité est satisfaite avec $v = v(x, y)$.

Soit $y = g(x)$ une courbe contenue dans Ω , on a :

$$\Phi(g) = \int_a^b L(x, g(x), g'(x)) dx \geq \int_a^b (g'(x) - v) \frac{\partial L}{\partial v}(x, y, v) dx + \int_a^b L(x, y, v) dx,$$

grâce à la convexité. On intègre donc la forme différentielle $\alpha = (L(x, y, v) - v \frac{\partial L}{\partial v}) dx + L(x, y, v) dy$ sur le graphe de g . On vérifie que $d\alpha = 0$ et donc cette intégrale a la même valeur que sur le graphe de f_0 , c'est-à-dire $\Phi(f_0) = \int_a^b L(x, f_0(x), f_0'(x)) dx$.

□

4. Brachistochrone

L'équation d'Euler-Lagrange s'écrit :

$$\frac{d}{dx} \left(\frac{1}{\sqrt{y}} \frac{y'}{\sqrt{1+y'^2}} \right) = -\frac{1}{2} y^{-\frac{3}{2}} \sqrt{1+y'^2}.$$

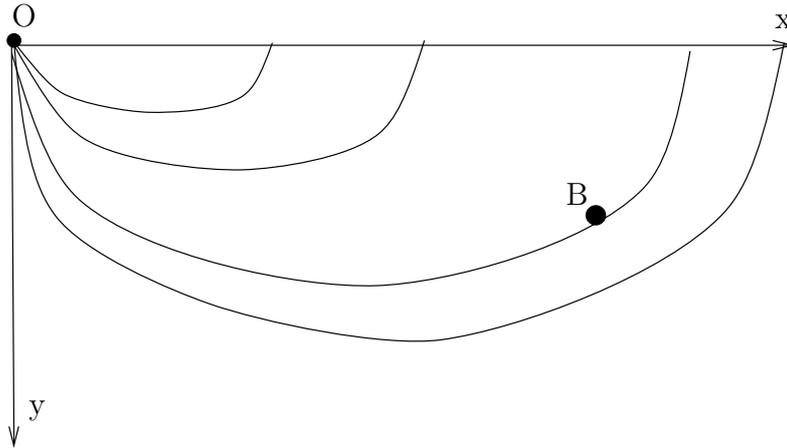


FIGURE 6.4. Champ d'extrémales pour le brachistochrone

On cherche une famille d'extrémales paramétrées par φ avec $y' = 1/\tan \varphi$.
On obtient :

$$\frac{d}{dx}\left(\frac{\cos \varphi}{y}\right) = -\frac{1}{2} \frac{1}{y^{3/2} \sin \varphi},$$

puis :

$$y = R(1 - \cos 2\varphi),$$

et :

$$dx = 2R(1 - \cos 2\varphi)d\varphi,$$

$$x = x_0 + 2R\varphi - R \sin 2\varphi.$$

Si $x_0 = 0$, on obtient une famille de cycloïdes homothétiques dont les premières arches C_R , $R > 0$ couvrent le quart de plan $x > 0, y > 0$.

On peut appliquer le théorème de Hilbert.

Si on pose $\alpha = b/B$, on voit que le toboggan idéal descend, puis remonte si $\alpha > \pi/4$.

5. Surface minimales de révolution

5.1. Surfaces de révolution. L'équation d'Euler-Lagrange s'écrit :

$$\frac{d}{dx}\left(\frac{yy'}{\sqrt{1+y'^2}}\right) = \sqrt{1+y'^2}.$$

On cherche des courbes paramétrées par φ avec $y' = \sinh \varphi$. On a donc :

$$\frac{d}{dx}(y \tanh \varphi) = \cosh \varphi,$$

$$\frac{dy}{y} = \tanh \varphi d\varphi,$$

$$y = C \cosh \varphi, \quad x = \frac{x-a}{C},$$

qui donne :

$$y = C \cosh \frac{x-a}{C}.$$

Si regarde les courbes passant par $(0, 1)$, elles remplissent un région Ω d'allure parabolique.

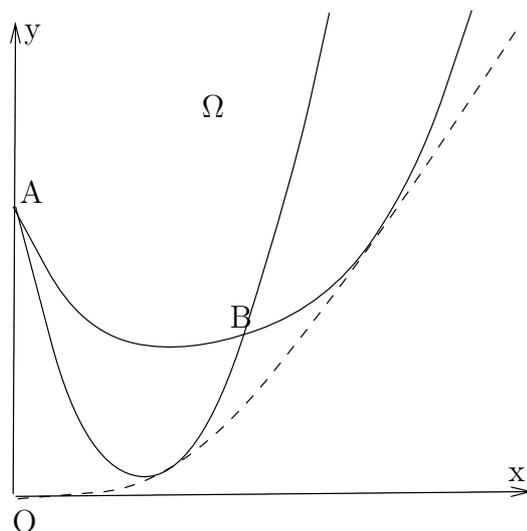


FIGURE 6.5. Surfaces minimales

Tout point de l'intérieur de Ω est joint à $(0, 1)$ par 2 courbes. D'après le théorème de Hilbert, la courbe la plus haute est minimisante. Si on considère une bulle de savon s'appuyant sur deux cercles co-axiaux de rayon 1, la bulle cylindrique va éclater pour un éloignement l donné par :

$$l = 2C \operatorname{Arsinh}\left(\frac{1}{C}\right), \quad 1 = C \cosh \frac{1}{\sqrt{1-C^2}}.$$

5.2. Chainettes. La théorie des multiplicateurs de Lagrange s'applique au même problème avec la contrainte $L = \int_a^b \sqrt{1+y'^2} dx$ et donc

$$\frac{d}{dx} \frac{yy'}{\sqrt{1+y'^2}} - \frac{y'}{\sqrt{1+y'^2}} = \lambda \frac{d}{dx} \frac{y'}{\sqrt{1+y'^2}}.$$

Cela donne :

$$y = A + B \cosh \frac{x-a}{B}.$$

6. Géodésiques

L'équation d'Euler-Lagrange s'écrit :

$$\frac{d}{dx} \left(\frac{Ey' + F}{\sqrt{Ey'^2 + 2Fy' + G}} \right) = \partial_y \sqrt{Ey'^2 + 2Fy' + G} .$$

6.1. Métrique hyperbolique. Il s'agit du cas

$$y > 0, \quad ds^2 = \frac{dx^2 + dy^2}{y^2} ,$$

et donc des extrémales de

$$\int \frac{\sqrt{1 + y'^2}}{y} dx ,$$

qui sont les demi-cercles de centre $(a, 0)$ situés dans le demi-plan $y > 0$. D'après le théorème de Hilbert, ils minimisent tous.

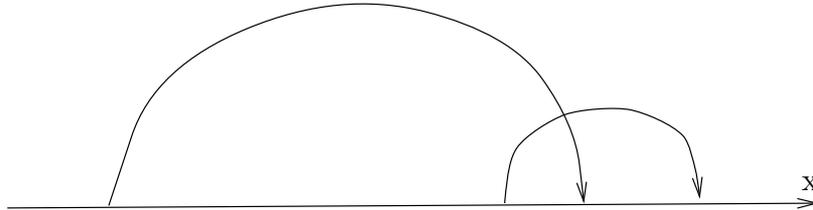


FIGURE 6.6. Géodésiques en géométrie hyperbolique

6.2. Optique. On considère les extrémales de $\int n(x, y) ds$ où $n(x, y)$ est l'indice de réfraction au point de coordonnées (x, y) . Si on suppose $n(x, y) = n(y)$ donné par $0 < y \leq 1, n(y) = y$ et $y \geq 1, n(y) = 1$ on obtient le phénomène des mirages :

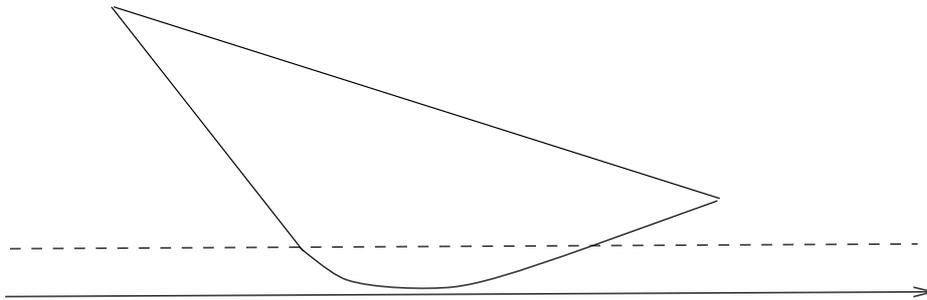


FIGURE 6.7. Mirages

7. Problème isopérimétrique en dimension 2

Il s'agit de trouver la courbe fermées de longueur minimale qui entoure une région d'aire π . On s'attend à trouver les cercles de rayon 1 !

Soit donc $\gamma : \mathbb{R}/2\pi\mathbb{Z} \rightarrow \mathbb{R}^2$ une courbe C^2 fermée simple définie par $\gamma(t) = (x(t), y(t))$ avec $x'^2 + y'^2 = 1$ et donc la longueur est 2π . Si A est l'aire limitée par la courbe, on a :

$$2A = 2 \int_{\gamma} xdy = 2 \int_0^{2\pi} xy' dt \leq \int_0^{2\pi} (x^2 + y'^2) dt ,$$

et on peut supposer que le centre de gravité de la courbe est à l'origine. On a alors :

$$\int_0^{2\pi} x^2 dt \leq \int_0^{2\pi} x'^2 dx ,$$

ainsi qu'on peut le montrer par les séries de Fourier :

$$x(t) = \sum_{n \neq 0} a_n e^{int}$$

et donc

$$\int_0^{2\pi} x^2(t) dt = 2\pi \sum_{n \neq 0} |a_n|^2$$

et

$$\int_0^{2\pi} x'^2(t) dt = 2\pi \sum_{n \neq 0} n^2 |a_n|^2 ,$$

ce qui permet de conclure que

$$2A \leq 2\pi .$$

On peut même traiter le cas d'égalité et montrer que c'est un cercle.

8. Problème de Dirichlet et méthode des éléments finis

Soit D un domaine compact à bord C^1 de \mathbb{R}^2 et considérons le problème de minimiser parmi les fonctions $f : D \rightarrow \mathbb{R}$ de classe C^1 et de valeur donnée $\varphi : \partial D \rightarrow \mathbb{R}$ sur le bord l'intégrale de Dirichlet :

$$J(f) = \frac{1}{2} \int_D \|\text{grad} f(x, y)\|^2 |dxdy| .$$

Il s'agit de minimiser une forme quadratique sur un espace affine et on s'attend à ce que l'équation d'Euler soit linéaire, c'est le cas. On trouve :

$$\delta J = - \int_D \Delta f \delta f |dxdy| ,$$

où $\delta f : D \rightarrow \mathbb{R}$ est nulle au bord. L'équation d'Euler-Lagrange est donc

$$\Delta f = 0 .$$

On a en fait un théorème général d'existence que je donne sans preuve :

THÉORÈME 6.3. *Sous les hypothèses précédentes, il existe une unique fonction $f : D \rightarrow \mathbb{R}$ de classe C^2 telle que $\Delta f = 0$ dans D et $f|_{\partial D} = \varphi$.*

L'unicité est facile, car elle résulte du principe du maximum. L'existence est plus délicate.

Si on veut résoudre numériquement ce problème dans le cas où D est le carré $[0, 1] \times [0, 1]$, on doit le remplacer par un problème avec un nombre fini d'inconnues. Décrivons un tel problème.

Soit $C_N = \{0, \dots, N\} \times \{0, \dots, N\}$ le carré discret à $(N + 1)^2$ points. L'inconnue est une fonction $f : C_N \rightarrow \mathbb{R}$. On pose $B = \{0, N\} \times \{0, \dots, N\} \cup \{0, \dots, N\} \times \{0, N\}$, B est le bord de C . On se donne $\varphi : B \rightarrow \mathbb{R}$. Pour $0 < i < N, 0 < j < N$ posons

$$\Delta_N f(i, j) = f(i, j) - \frac{1}{4} \sum_{\alpha=\pm 1, \beta=\pm 1} f(i + \alpha, j + \beta) .$$

Si on identifie C_N avec le réseau de pas $1/N$ du carré unité, $\frac{1}{N^2} \Delta$ donne une bonne approximation de Δ (voir définition des dérivées secondes). On veut résoudre $\Delta f = 0$ avec $f|_B = \varphi$. On a un système linéaire de $(N - 1)^2$ équations à $(N - 1)^2$ inconnues. Il suffit de montrer l'unicité qui résulte du principe du maximum :

LEMME 6.1. *Soit f telle que $\Delta f = 0$, si f atteint un max local en (i_0, j_0) dans $C_N \setminus B$, f est constante aux points voisins de (i_0, j_0) .*

Remarquons qu'on peut obtenir f en minimisant

$$q(f) = \sum_{i=0}^{N-1} \sum_{j=1}^{N-1} (f(i+1, j) - f(i, j))^2 + \sum_{i=1}^{N-1} \sum_{j=0}^{N-1} (f(i, j+1) - f(i, j))^2 ,$$

sur l'espace affine des f qui valent φ sur le B . On peut ainsi résoudre ce problème par une méthode de gradient.

QUELQUES PROBLÈMES

1. Exercices sur le chapitre 1

1.1. Exercice. Étudier, suivant les valeurs de α la continuité, l'existence de dérivées suivant tout vecteur, la différentiabilité de

$$f(x, y) = \begin{cases} (x^2 + y^2)^\alpha \sin\left(\frac{1}{x^2 + y^2}\right) & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

1.2. Exercice. Soit $f(x, y)$ la fonction de \mathbb{R}^2 dans \mathbb{R} définie par $f(0, 0) = 0$ et pour $(x, y) \neq (0, 0)$:

$$f(x, y) = \frac{x^6}{(y - x^2)^2 + x^8} .$$

- 1) f est-elle continue ?
- 2) Prouver que f admet des dérivées directionnelles en $(0, 0)$ et les calculer.
- 3) f est-elle différentiable en $(0, 0)$?

1.3. Exercice. Soit M le point de coordonnées (x, y) du plan euclidien \mathbb{R}^2 (on note d la distance), $A = (1, 0)$, $B = (-1, 0)$, $O = (0, 0)$. On considère la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par

$$f(x, y) = d(A, M).d(B, M) .$$

- 1) En quels points la fonction f est-elle différentiable ?
- 2) Soit $\Omega = \{\mathbb{R}^2 \setminus \{A, B, O\}\}$. Montrer que les ensembles de niveaux $f(x, y) = a$ rencontrent Ω suivants des sous-variétés que l'on notera X_a . Pour quelles valeurs de a sont-elles connexes ?
- 3) Déterminer l'ensemble Z des points M de Ω où l'espace tangent à la sous-variété $f(x, y) = f(M)$ est parallèle à l'axe $x'Ox$.
- 4) Près du point $(0, 1)$, la variété X_2 est le graphe d'une fonction $y = g(x)$. Trouver le développement limité à l'ordre 4, $g(x) = \sum_{i=0}^4 a_i x^i + O(x^5)$, de g en $x = 0$.
- 5) Représenter sur une même figure l'ensemble Z et les ensembles X_a pour

$$a = 0, 0.5, 1, 1.5, 2, 2.5$$

1.4. Exercice. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la fonction définie par

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

Pour quelles valeurs de α cette fonction est-elle continue? admet-elle des dérivées suivant tout vecteur? est-elle différentiable?

1.5. Exercice. Montrer que la fonction de \mathbb{R}^2 dans \mathbb{R} définie par $f(x, y) = 2x^2y/(x^4 + y^2)$ si $(x, y) \neq (0, 0)$ et $f(0, 0) = 0$ admet des dérivées suivant tout vecteur en $(0, 0)$. Cette fonction est-elle différentiable en ce point?

1.6. Exercice. L'application $X \mapsto^t X \cdot X$ de $M_n(\mathbb{R})$ dans $M_n(\mathbb{R})$ est-elle différentiable? Dans l'affirmative déterminer sa différentielle en X_0 . En quel(s) point(s) cette différentielle est-elle nulle?

1.7. Exercice. Soient U un ouvert de \mathbb{R}^n muni de sa structure euclidienne, f une application de U dans \mathbb{R} et x_0 un point de U . On suppose qu'il existe deux applications h_1 et h_2 différentiables en x_0 telles que $h_1 \leq f \leq h_2$. Montrer que si $h_1(x_0) = h_2(x_0)$, alors f est différentiable en x_0 .

Soit f une application 1-lipschitzienne sur U (*i.e.* pour tout x et tout y de U on a $|f(x) - f(y)| \leq \|x - y\|$). On suppose qu'il existe un segment $[a, b]$ inclus dans U tel que $f(b) - f(a) = \|b - a\|$. Montrer que f est différentiable en tout point de $]a, b[$. (On pourra établir que $f(b) - f(a) - \|x - b\| \leq f(x) - f(a) \leq \|x - a\|$)

Application: Si C est une partie fermée de \mathbb{R}^n muni de sa structure euclidienne on note $U = \mathbb{R}^n \setminus C$ et $d_C(x) = \inf_{z \in C} \|x - z\|$.

Montrer, en utilisant l'inégalité triangulaire, que d_C est 1-lipschitzienne.

On suppose maintenant que C est convexe. Pour tout $x \in U$ il existe alors un point unique $p(x) \in C$ tel que $d_C(x) = \|x - p(x)\|$ et $p(x)$ est le seul point u de C tel que $\langle x - u, z - u \rangle \leq 0$ pour tout $z \in C$. Dédurre de ce qui précède que d_C est différentiable sur U .

1.8. Exercice. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ la fonction définie par

$$f(x, y) = (e^x \cos y, e^x \sin y).$$

a) Déterminer l'image de \mathbb{R}^2 par f puis l'image de la bande

$$\mathcal{B} = \{(x, y); \quad 0 < y < 2\pi\} .$$

b) Déterminer l'image par f des droites d'équations $x = x_0$ et $y = y_0$. En déduire l'image du rectangle de sommets $(0, 0)$, $(0, h)$, (k, h) et $(k, 0)$. Calculer l'aire de l'image.

c) Montrer que f est différentiable en tout point et calculer sa différentielle. Reprendre la question b) avec f' et calculer l'aire de l'image par $f'(0, 0)$ du rectangle considéré.

1.9. Exercice. Reprendre l'exercice ci-dessus avec $f : (x, y) \mapsto (x^2 - y^2, 2xy)$.

1.10. Exercice. En utilisant les coordonnées polaires, déterminer toutes les fonctions de classe C^1 de $(\mathbb{R}^+)^2 \rightarrow \mathbb{R}$ vérifiant

$$x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} = a \sqrt{x^2 + y^2}.$$

1.11. Exercice. Soit g une application continûment différentiable de \mathbb{R}^n dans \mathbb{R}^n et soit $I = [0, 1]$.

Étudier la continuité puis la différentiabilité de $\Phi : C^0(I, \mathbb{R}^n) \rightarrow C^0(I, \mathbb{R}^n)$ définie par $\Phi(f) = g \circ f$ (pour démontrer la différentiabilité de Φ on utilisera le théorème des accroissements finis avec reste intégral).

1.12. Exercice. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction de classe C^1 telle que $f'(0) = 0$. Soit $C_f = \{(x, y) \in \mathbb{R}^2 \mid y = f(x)\}$. Montrer qu'il existe un difféomorphisme C^1 , F , de \mathbb{R}^2 dans \mathbb{R}^2 tel que $F(C_f) = \{y = 0\}$.

1.13. Exercice. Construire un difféo. C^1 du de plan sur un demi-plan. Même question avec un quart de plan et un plan.

1.14. Exercice. On identifie \mathbb{R}^2 à \mathbb{C} par $z = x + iy$. Prouver que l'application $f(z) = \frac{z-i}{z+i}$ est un difféo. du demi-plan $\{y > 0\}$ sur le disque unité de centre 0. Prouver que ce difféo. préserve les angles. Dessiner l'image des demi-droites verticales et des droites horizontales.

1.15. Exercice. Soit f une fonction continue sur $] - 1, 1[$ telle qu'il existe un réel λ , $0 < |\lambda| < 1$ de sorte que

$$f(x) = \lambda x + o(x)$$

lorsque x tend vers 0.

- 1) Montrer que $f(0) = 0$.
- 2) Montrer que pour tout $\varepsilon > 0$ vérifiant $|\lambda| + \varepsilon < 1$, il existe $\eta > 0$ tel que

$$\forall x \in] - \eta, \eta[, |f(x)| \leq (|\lambda| + \varepsilon)|x|.$$

Montrer alors que

$$f(] - \eta, \eta]) \subset] - \eta, \eta[.$$

Dans la suite, on suppose de plus que

$$f(x) = \lambda x + O(x^2)$$

lorsque x tend vers 0.

- 3) Montrer qu'il existe $\eta > 0$ et un réel positif M tel que

$$\forall x \in] - \eta, \eta[, \forall n \in \mathbb{N}^*, |f(f^{\circ n}(x)) - \lambda f^{\circ n}(x)| \leq M |f^{\circ n}(x)|^2.$$

- 4) En déduire que pour η suffisamment petit, la suite de fonctions $h_n(x) := f^{\circ n}(x)/\lambda^n$ converge uniformément sur $] -\eta, \eta[$ vers une fonction continue h nulle en 0 et vérifiant $h \circ f = \lambda h$ sur $] -\eta, \eta[$. (Indication : estimer les différences $h_{n+1}(x) - h_n(x)$.)
- 5) Montrer que

$$h(x) = x + O(x^2)$$

lorsque x tend vers 0.

1.16. Exercice. Limites en $(0, 0)$ des fonctions

$$\frac{x^3 + y^3}{x^2 + y^2}, \frac{xy}{x + y}, \frac{x \sin y - y \sin x}{x^2 + y^2}.$$

1.17. Exercice. Si f est une fonction de classe C^1 sur \mathbb{R} , montrer que la fonction définie par $\frac{f(x) - f(y)}{x - y}$ si $x \neq y$ et par $f'(x)$ sinon est continue sur \mathbb{R}^2 .

2. Exercices sur le chapitre 2

2.1. Exercice. Soit f une fonction de classe C^2 sur $\mathbb{R}^2 \setminus \{0\}$. On pose $x = r \cos \theta$, $y = r \sin \theta$ et $g(r, \theta) = f(x, y)$. Calculer

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

en fonction de r , θ et des dérivées partielles de g .

2.2. Exercice. Calculer les dérivées partielles premières et secondes de la fonction implicite $z = f(x, y)$ définie par $z^3 + ze^x + \cos y = 0$.

2.3. Exercice. 1) Prouver que toutes les solutions $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ de classe C^2 de

$$\frac{\partial^2 f}{\partial X \partial Y} = 0$$

sont de la forme $f(X, Y) = F(X) + G(Y)$.

2) En déduire les solutions de l'équation des ondes

$$\frac{\partial^2 \varphi}{\partial t^2} - v^2 \frac{\partial^2 \varphi}{\partial x^2} = 0$$

au moyen du changement de variables $X = x + vt$, $Y = x - vt$.

3. Exercices sur le chapitre 3

3.1. Exercice. On considère l'équation $y = 1 + xy^{1/5}$ avec $y > 0$. Prouver qu'il existe pour x voisin de 0 une unique solution proche de 1 notée $y = f(x)$. Ecrire le développement limité à l'ordre 3 de φ en $x = 0$.

3.2. Exercice. Montrer que le système

$$x^2 + y^2 + z^2 = 3, \quad x^2 + 3xy - 2z = 0$$

admet au voisinage de $x = 1$ une unique solution proche de $(-1, 1)$ notée $(\varphi(x), \psi(x))$.
Ecrire les d.l. à l'ordre 2 en 1 de ces fonctions.

3.3. Exercice. Prouver que pour $|t| < 1/\sqrt{2}$, l'équation

$$\sin tx + \cos tx = x$$

admet une unique solution $\varphi(t)$ et calculer son d.l. à l'ordre 2 au voisinage de 0.

3.4. Exercice. On fixe un entier $n \geq 1$. On note par P l'ensemble des matrices symétriques définies positives $n \times n$.

1) Montrer que P peut-être considéré naturellement comme un ouvert de $\mathbb{R}^{n(n+1)/2}$.

2) Montrer que, pour tout $A \in P$, il existe $B \in P$ unique tel que $B^2 = A$.

3) Calculer la différentielle de $f : A \rightarrow A^2$ ($f : P \rightarrow P$) et montrer qu'elle est inversible.

4) Montrer que f est un difféomorphisme C^∞ de P sur P .

3.5. Exercice. Montrer que $P_n(x) = x^n - x - 1$ admet une unique racine x_n dans l'intervalle $]1, 2[$. Montrer que x_n admet un développement à tout ordre par rapport à $1/n$. Déterminer ce développement à l'ordre 3.

3.6. Exercice. On se donne 3 points non alignés A, B, C du plan. On se donne aussi 2 réels a, b . On considère le système d'équations où M est un point du plan :

$$d(A, M) - d(B, M) = a, \quad d(A, M) - d(C, M) = b.$$

Prouver qu'une solution M de ce système est stable si et seulement si M n'est sur aucune des droites AB, BC, CA .

3.7. Exercice. Etudier les points critiques et tracer les lignes de niveau dans $[0, 2\pi] \times [0, 2\pi]$ de

$$f(x, y) = \cos x + \cos y + \cos(x + y).$$

4. Exercices sur le chapitre 4

4.1. Exercice. Soient D_1 et D_2 2 droites affines non parallèles et non concurrentes de \mathbb{R}^3 . Prouver que la fonction $f : D_1 \times D_2 \rightarrow \mathbb{R}$ atteint son minimum en 1 point unique (M_1, M_2) . Prouver que la droite M_1M_2 est l'unique,perpendiculaire commune à D_1 et D_2 .

4.2. Exercice. Soit $X = \{z = f(x, y)\}$ une surface C^2 de \mathbb{R}^3 . On suppose $f(0, 0) = 0$ et $f'(0, 0) = 0$. Soit $A = (0, 0, a)$ avec $a \neq 0$. Prouver que la fonction $f_a : X \rightarrow \mathbb{R}$ définie par $f_a(M) = d(A, M)$ a un point critique en $(0, 0, 0)$ et déterminer suivant la valeur de a la nature de ce point critique.

4.3. Exercice. On considère $f(x, y) = 2x^3 + 6xy^2 - 3x^2 + 3y^2$. Déterminer les points critiques de f ainsi que ses extrema. Si $\Gamma = \{(x, y); f(x, y) = 0\}$, décrire Γ près du ou des points critiques de f appartenant à Γ .

4.4. Exercice. Déterminer les triangles d'aire donnée et de périmètre minimal, puis ceux de périmètre donné et d'aire maximale. Est-il prévisible que ce sont les mêmes ?

4.5. Exercice. Dans la figure suivante dans le plan euclidien, les points A et B sont fixes de distance euclidienne $a > 0$.

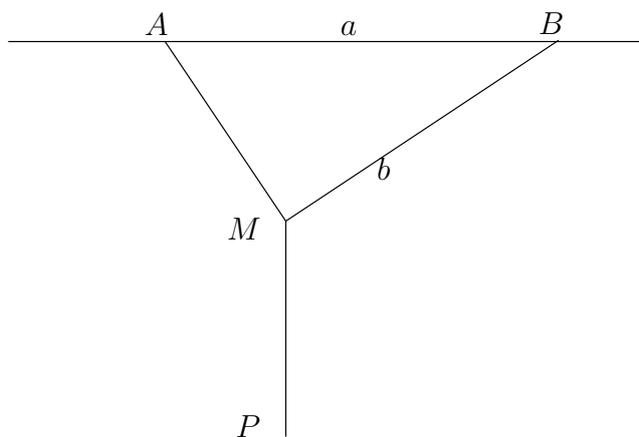


FIGURE 7.1

Les points P et M sont variables, mais les distances $b = d(B, M)$ et $l = d(A, M) + d(M, P)$ sont fixées. Déterminer le point M qui maximise la distance de P à la droite AB .

4.6. Exercice. Calculer le volume maximal du parallépipède rectangle contenu dans $\{x \geq 0, y \geq 0, z \geq 0\}$ ayant un sommet en $(0, 0, 0)$ et le sommet opposé sur le paraboloid d'équation $-z - x^2 - 4y^2 + 4 = 0$.

4.7. Exercice. L'ensemble des matrices 2×2 étant identifié à \mathbb{R}^4 on considère l'application f :

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto a^2 + b^2 + c^2 + d^2.$$

Déterminer les extrema de f sur $\mathbf{SL}(2, \mathbb{R})$ (déterminant 1). Montrer que ces extrema sont atteints sur $\mathbf{SO}(2, \mathbb{R})$.

4.8. Exercice. Soient x_1, \dots, x_n des réels strictement positifs. Montrer que

$$\text{Min}\left\{\frac{1}{n} \sum u_i x_i; u_i > 0, \prod u_i = 1\right\} = \left(\prod x_i\right)^{\frac{1}{n}}.$$

En déduire que

$$\left(\prod x_i\right)^{\frac{1}{n}} + \left(\prod y_i\right)^{\frac{1}{n}} \leq \left(\prod (x_i + y_i)\right)^{\frac{1}{n}}$$

4.9. Exercice. 1) On considère 2 points A et B du plan \mathbb{R}^2 euclidien situés de part et d'autre de la droite $y = 0$. Soit, pour $M = (x, 0)$, $f(M) = d(A, M) + \nu d(B, M)$, avec $\nu > 0$. Montrer que f atteint son minimum en un point M_0 unique de $\{y = 0\}$ et donner la relation entre les angles α et β qui caractérise ce point (loi de Snell-Descartes).

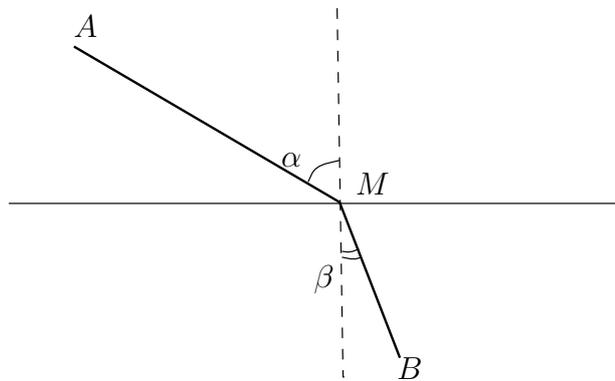


FIGURE 7.2

2) On considère maintenant 2 droites parallèles et 2 points A et B situés dans les 2 demi-plans séparés par la bande limitée par ces 2 droites. On se donne 2 réels > 0 ν_A et ν_B . Prouver en utilisant la méthode précédente qu'il existe un unique rayon lumineux de A à B qui satisfait la loi de Snell-Descartes en supposant le demi-plan de A d'indice ν_A , celui de B d'indice ν_B et la bande centrale d'indice 1.

4.10. Exercice. Prouver que la fonction $f(x, y) = \cos x + 2 \cos y$ atteint son maximum en un point unique sur le domaine

$$D = \{(x, y) | 0 \leq x \leq \frac{\pi}{2}, 0 \leq y \leq \frac{\pi}{2}, \sin x + \sin y = 1\}.$$

Calculer numériquement la valeur de ce maximum.

Interprétation géométrique : position d'équilibre d'un système de 2 masses ponctuelles P et Q de même masse dans un champ de pesanteur uniforme vertical assujetties, par des barres rigides, par rapport à un point fixe A situé à distance horizontale 1 d'un axe vertical D aux conditions $d(A, P) = 1$ et $d(P, Q) = 1$ et $Q \in D$.

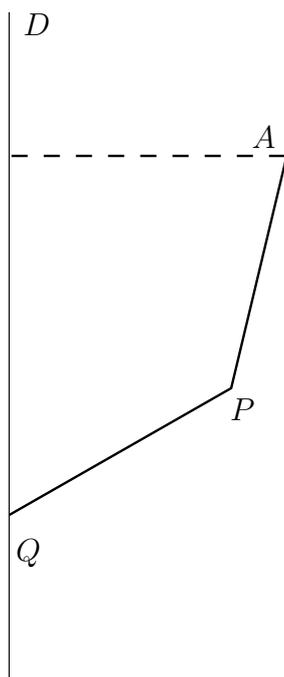


FIGURE 7.3

4.11. Exercice. 1) Trouver les courbes de longueur minimale tracée sur un cube de \mathbb{R}^3 qui joint 2 sommets opposés du cube.

2) Soit P un plan perpendiculaire à l'axe qui joint deux sommets opposés du cube. On pose $l(P)$ = la longueur de l'intersection de P avec le cube. Trouver le maximum de $l(P)$ et les points de maximum.

5. Exercices sur le chapitre 5

5.1. Exercice. Soit $f : U \rightarrow \mathbb{R}$ où U est un domaine convexe de E evn. On suppose f continue et $\forall x, y \in U$:

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2} .$$

Prouver que f est convexe. La continuité est-elle nécessaire ?

5.2. Exercice. Soit

$$f(x, y) = x^4 + y^4 + x^2 + xy + y^2 - 3x - 5y .$$

- 1) Montrer que f est α -convexe et préciser α (pour la norme euclidienne).
- 2) Tracer les courbes $\partial_x f = 0$ et $\partial_y f = 0$. En déduire que le point de min de f , z_0 , est situé dans le carré

$$K = [0, 1] \times [0, 1] .$$

3) Soit $F_\mu : (x, y) \rightarrow (x, y) - \mu \text{grad } f(x, y)$. Montrer que pour $\mu_0 = 0, 1$, $F_{\mu_0}(K) \subset K$.

4) Est ce que la méthode de gradient à pas μ_0 converge pour toute donnée initiale ?

5) Montrer que $F_{\mu_0}|_K$ est contractante de rapport < 1 en norme sup.

6) En déduire une méthode de recherche de z_0 .

5.3. Exercice. Résoudre graphiquement le problème suivant :

si $P = \{2x + y \leq 8, x + 2y \leq 7, y \leq 3\}$, minimiser $f(x, y) = -x - y$ sur P .

5.4. Exercice. 1) Soit K le polytope défini par les équations :

$$K = \{(x, y, z) \mid 0 \leq x \leq 1, 0 \leq y \leq 2, 0 \leq z \leq 3, x + y + z \leq 5\}.$$

Faire la liste des facettes de K .

2) Trouver le sup K de $x + y + z$ et de $x + y + z + a(x + y)$ (on pourra déterminer l'image par $(x, y, z) \rightarrow (x + y + z, x + y)$ de K).

5.5. Exercice. Soit

$$f(x, y) = x^4 + y^4 + x^2 + xy + y^2 - 2x - 3y.$$

1) Prouver que f est α -convexe sur \mathbb{R}^2 en précisant la valeur de α .

2) Montrer que les courbes $\frac{\partial f}{\partial x}(x, y) = 0$ et $\frac{\partial f}{\partial y}(x, y) = 0$ se coupent dans le carré

$$K = [0, 1] \times [0, 1].$$

3) En déduire que le point de minimum de f est dans K .

4) Soit $\mu_0 = 10^{-1}$; prouver que, si $(x_0, y_0) \in K$, et qu'on applique à f la méthode de gradient de pas μ_0 avec comme initialisation le point (x_0, y_0) les itérés successifs restent dans K .

5) Sous les mêmes hypothèses qu'en 4), montrer que la méthode de gradient à pas fixe μ_0 converge.

6) Retrouver ainsi le résultat de 3).

5.6. Exercice. Soit f une fonction différentiable sur un convexe X . Montrer que f est convexe si et seulement si pour tout x, y dans X on a

$$f(y) - f(x) \geq f'(x) \cdot (y - x).$$

5.7. Exercice. On considère $f_a(x, y) = x^2 + y^2 + axy - 2x - 2y$. Pour quelles valeurs de a cette fonction est-elle convexe, strictement convexe, α -convexe? (Dans ce dernier cas préciser la valeur de α)

Déterminer les points critiques de f , ses minima locaux.

On note $\varphi(a) = \inf_{\mathbb{R}^2} f_a(x, y)$. Justifier l'existence de $\varphi(a)$ et le déterminer lorsque f_a est convexe.

Soit $X = \{0 \leq x \leq 1, y \geq 0\}$. On note $m(a) = \inf_X f_a$, montrer que m_a est atteint. Pour quelles valeurs de a $m(a) = \varphi(a)$?

5.8. Exercice. Soit $X = \{(a, b) ; a + tb \geq 0 \forall t \in [-1, 1]\}$. Montrer que X est convexe fermé dans le plan. Représenter X .

Trouver l'ensemble des polynômes $Q : t \mapsto a + tb$ dont la projection orthogonale sur X est le polynôme nul.

5.9. Exercice. Soit ABC un triangle euclidien, prouver qu'il existe un unique point M qui minimise la fonction $f(M) = d(A, M) + d(B, M) + d(C, M)$. On pourra montrer que f est α -convexe dans les boules. Caractériser géométriquement M (distinguer suivant que ABC a ou non un angle $\geq 2\pi/3$).

5.10. Exercice. On considère le problème de minimisation suivant :

minimiser $f(z_1, \dots, z_{N-1}) = \sum_{j=1}^{N-1} y_j$ où $z_j = x_j + iy_j$ sous les contraintes

$$z_0 = a, \quad z_N = b, \quad \text{et } \forall j = 0, \dots, N-1, \quad \|z_{j+1} - z_j\| \leq 1.$$

Il s'agit de trouver les positions d'équilibre d'une chaîne constituée de $(N-1)$ masses égales reliées par des barres de longueur 1 dans un champ de gravitation constant les extrémités étant fixées (les points a et b).

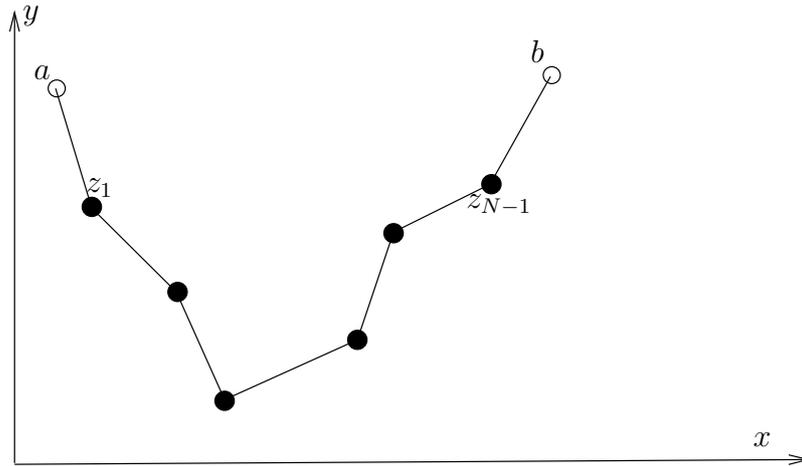


FIGURE 7.4. Un système de barres articulées

1) Montrer que toute solution du problème vérifie

$$\forall i = 0, \dots, N-1, \quad \|z_{i+1} - z_i\| = 1.$$

2) Montrer que le domaine de définition de f est strictement convexe et en déduire l'existence et l'unicité d'une solution.

3) Considérant le problème avec contraintes $\forall i = 0, \dots, N-1, \|z_{i+1} - z_i\| = 1$ écrire les multiplicateurs de Lagrange et les interpréter physiquement en termes des forces de tension qui s'exercent sur les barres.

6. Problème

Soient U un voisinage ouvert convexe de 0 dans \mathbb{R}^2 et f une fonction de classe C^k (i.e. qui admet des dérivées partielles continues jusqu'à l'ordre k) à valeurs réelles définie sur U .

a) Montrer que si $f(0, 0) = 0$ il existe des fonctions u et v définies sur U de classe C^{k-1} telles que

$$f(x, y) = xf_1(x, y) + yf_2(x, y).$$

(Posant $\varphi(u) = f(ux, uy)$, on pourra intégrer $\varphi'(t)$ sur $[0, 1]$)

b) Si en outre $f'(0, 0) = 0$ il existe des fonctions a , b et c de classe C^{k-2} telles que $f(x, y) = x^2a(x, y) + 2xyb(x, y) + y^2c(x, y)$ et

$$r = a(0, 0) = \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(0, 0), \quad s = b(0, 0) = \frac{1}{2} \frac{\partial^2 f}{\partial x \partial y}(0, 0), \quad t = c(0, 0) = \frac{1}{2} \frac{\partial^2 f}{\partial y^2}(0, 0).$$

c) On note Σ la surface d'équation $z = f(x, y)$ et l'on suppose que $f(0, 0) = 0$ et $f'(0, 0) = 0$. Donner l'équation du plan tangent à Σ en $(0, 0, 0)$.

d) On conserve les notations précédentes et on suppose $k \geq 3$. Montrer que si $s^2 - rt \neq 0$ il existe des voisinages U_1 et U_2 de O dans \mathbb{R}^2 et une bijection de classe C^{k-2} : $F : (x, y) \mapsto (X(x, y), Y(x, y))$ telles que

$$f \circ F^{-1}(X, Y) = X^2 + \epsilon Y^2$$

avec $\epsilon = \pm 1$.

(S'inspirer de la décomposition d'une forme quadratique puis en appliquant le théorème des fonctions implicites à une fonction bien choisie, montrer que l'on obtient une bijection de classe C^{k-2} sur des ouverts convenables)

e) Si $s^2 - rt > 0$ décrire l'intersection de Σ et de son plan tangent en 0 .

7. Problème

1) On note E_n l'espace \mathbb{R}^n muni de sa structure euclidienne et d la distance associée. Montrer que pour tout A in E_n la fonction $f_A : M \mapsto d(M, A)$ est convexe et que si $d(\lambda M + (1 - \lambda)M', A) = \lambda d(M, A) + (1 - \lambda)d(M', A)$ avec $\lambda \in]0, 1[$, alors M , M' et A sont alignés.

Montrer que f_A est C^∞ sur $E_n \setminus \{A\}$, calculer son gradient et sa différentielle seconde.

2) En déduire que si A_1 , A_2 et A_3 sont trois points non alignés de E_2 la fonction $f : M \in E \mapsto d(A_1, M) + d(A_2, M) + d(A_3, M)$ est strictement convexe.

Montrer que f atteint un minimum en un point unique et que ce point ne peut être à l'extérieur du triangle de sommets A_1 , A_2 et A_3 .

3) On note f_i la fonction $M \mapsto \sum_{j \neq i} d(M, A_j)$ et α_i le gradient de cette fonction au point A_i . En appliquant une formule de Taylor à f_i , montrer que si $\|\alpha_i\| \leq 1$, alors A_i est le minimum global de f . Que peut-on alors dire de l'angle de sommet A_i ?

Par contre, si $\|\alpha_i\| > 1$, le minimum est atteint à l'intérieur du triangle de sommets A_1, A_2 et A_3 ("point de Fermat" du triangle).

4) On considère maintenant k points A_i non alignés dans E_n . On pose $f(M) = \sum_i d(A_i, M)$ et, pour tout $i \in \{1, \dots, k\}$ on note $f_i(M) = \sum_{j \neq i} d(M, A_j)$.

Montrer que f est strictement convexe et atteint un minimum dans le sous espace affine engendré par les A_i . Montrer, toujours avec les notations du 3), que si $\|\alpha_i\| \leq 1$, alors A_i est le minimum global de f (même méthode). Si pour tout i , $\|\alpha_i\| > 1$ la fonction f admet un unique point critique et atteint en ce point un minimum global.

8. Examen de mai 1997

Exercice 1

Soit A une matrice symétrique définie positive sur un espace euclidien E de dimension finie N . On note $0 < \lambda_1 \leq \dots \leq \lambda_N$ les valeurs propres de A . Soit $f(x) = \frac{1}{2} \langle Ax|x \rangle - \langle b|x \rangle$, où $\langle \cdot | \cdot \rangle$ est le produit scalaire de E .

1) Écrire l'algorithme de gradient à pas μ pour la recherche du point de minimum a de f .

2) Pour quelles valeurs de μ l'algorithme converge-t-il vers a pour toute donnée initiale ?

3) Quelle est la valeur optimale de μ pour la vitesse de convergence ?

Exercice 2

Soit $E = C([-1, +1], \mathbb{R})$. On considère le problème de minimisation suivant :

$$(\mathcal{P}) \text{ Min}_{p \in \mathcal{C}} \int_{-1}^1 (\cos t - p(t))^2 dt ,$$

où

$$\mathcal{C} = \{p \text{ polynôme de degré } \leq 1 \text{ réel tel que } p'(0) = -1 \} .$$

1) Justifier l'existence et l'unicité d'une solution de (\mathcal{P}) .

2) Résoudre (\mathcal{P}) , i.e. trouver le point de minimum de la fonctionnelle considérée sur \mathcal{C} .

Exercice 3

Soient $X = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$ et, pour $a > 0$, $Y_a = \{(x, y, z) \in \mathbb{R}^3 \mid (x - \frac{1}{2})^2 + y^2 = a^2\}$.

1) Montrer que X et Y_a sont des sous-variétés de \mathbb{R}^3 .

2) Pour quelles valeurs de a , l'intersection $C_a = X \cap Y_a$ est-elle une sous-variété de dimension 1 de \mathbb{R}^3 ?

3) Si $M_o = (x_o, y_o, z_o) \in X$ et $a^2 = (x_o - \frac{1}{2})^2 + y_o^2$, déterminer lorsqu'elle existe la direction de la tangente en M_o à C_a .

Exercice 4

Soit, pour $a \in \mathbb{R}$,

$$f_a(x, y) = x^2 + y^2 + axy - 2x - 2y .$$

1) Pour quelles valeurs de a , f_a est-elle convexe, strictement convexe, α -convexe ?

Dans ce dernier cas, préciser la valeur de α .

2) Déterminer les points critiques des f_a et, parmi ceux-ci, ceux qui sont des points de minimum local (resp. points de minimum local stricts).

3) Soit $\varphi(a) = \inf_{(x,y) \in \mathbb{R}^2} f_a(x, y) \in \mathbb{R} \cup \{-\infty\}$. Déterminer explicitement $\varphi(a)$ (on pourra commencer par les a pour lesquels f_a est convexe).

4) Soit $X = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq 1, y \geq 0\}$.

On pose $m(a) = \inf_{(x,y) \in X} f_a(x, y)$.

Montrer que, $\forall a \in \mathbb{R}$, il existe $(x(a), y(a)) \in X$ tel que

$$f_a(x(a), y(a)) = m(a) .$$

5) Pour quelles valeurs de a , a-t-on $m(a) = \varphi(a)$?

On note I l'ensemble de ces valeurs de a .

6) On suppose dans ce 6) que $a \notin I$. Montrer que $(x(a), y(a))$ est un point de la frontière ∂X de X .

On pose :

$$Y = \{(0, y) \mid y \geq 0\}, Z = \{(x, 0) \mid 0 \leq x \leq 1\}, W = \{(1, y) \mid y \geq 0\} .$$

On a ainsi :

$$\partial X = Y \cup Z \cup W .$$

Déterminer

$$R(a) = \inf_{(x,y) \in Y} f_a(x, y), S(a) = \inf_{(x,y) \in Z} f_a(x, y), T(a) = \inf_{(x,y) \in W} f_a(x, y) .$$

7) Calculer $m(a)$ pour tout a et tracer sur un même dessin $\varphi(a)$ et $m(a)$.

8) Pouvait-on montrer a priori (i.e. sans calculs) la continuité de $a \rightarrow m(a)$ et que $a \rightarrow -m(a)$ est convexe ?

9. Examen de septembre 1997*Exercice 1*

Soit $\mathbb{R}^2 = \mathbb{C}$ muni de la norme euclidienne

$$\|(x, y)\|^2 = x^2 + y^2 .$$

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ donnée par

$$f(x, y) = \|1 + e^{ix} + e^{iy}\|^2 .$$

1) Soit $A = (\pi, 0)$, prouver que A est un point critique de f .

2) Ecrire la formule de Taylor-Young à l'ordre 2 pour f en A .

3) Le point A est-il un point de maximum local, de minimum local ou un col pour f ?

4) Représenter l'allure des courbes de niveaux de f au voisinage de A en indiquant les régions $f(x, y) > f(A)$ et $f(x, y) < f(A)$.

Exercice 2

Soit $f_a(x, y) = x^4 + y^4 + 4axy + (\sin \pi a)(x - 2y)$.

1) Pour quelles valeurs du paramètre réel a , la fonction $f_a : \mathbb{R}^2 \rightarrow \mathbb{R}$ est-elle convexe ?

2) Déterminer les points critiques de f_1 et parmi ceux-ci ceux qui sont des minimas locaux.

3) On pose $\varphi(a) = \inf_{(x,y) \in \mathbb{R}^2} f_a(x, y)$. Calculer $m = \varphi(1)$.

4) On considère le système de 2 équations (\star_a) , à 2 inconnues (x, y) , dépendant du paramètre a donné par :

$$(\star_a) \quad 4x^3 + 4ay + \sin(\pi a) = 0, \quad 4y^3 + 4ax - 2\sin(\pi a) = 0.$$

Soit

$$K = \left[\frac{1}{2}, \frac{3}{2}\right] \times \left[-\frac{3}{2}, -\frac{1}{2}\right],$$

et $(1, -1) \in K$ l'unique solution de (\star_1) contenue dans K .

Prouver qu'il existe un intervalle $I =]-\alpha, \alpha[$, $\alpha > 0$ tel que, si $a \in I$, (\star_a) admet une unique solution dans K , notée $(x(a), y(a))$, et que l'application $a \rightarrow (x(a), y(a))$ est C^∞ de I dans \mathbb{R}^2 .

Calculer

$$l = \frac{d}{da} \Big|_{a=1} f_a(x(a), y(a)).$$

Exercice 3

Soit E l'espace vectoriel de dimension 2 formé des polynômes de degré ≤ 1 $P(t) = a + bt$. On munit E du produit scalaire euclidien

$$(P|Q) = \int_{-1}^{+1} P(t)Q(t)dt.$$

1) Soit X l'ensemble des $P \in E$ qui vérifient

$$\forall t \in [-1, +1], P(t) \geq 0.$$

Prouver que X est un polyèdre convexe fermé de E et le représenter dans le plan des (a, b) .

2) Soit $P_0(t) = t$, déterminer la projection (pour le produit scalaire euclidien précédent) de P_0 sur X .

3) Quels sont les $P \in E$ dont la projection est le polynôme $P = 0$?

Représenter cet ensemble dans le plan des (a, b) .

10. Examen de juin 1998

Exercice 1

Soit

$$f(x, y) = y^3 + e^x y + \sin x .$$

- 1) Prouver que, pour tout $x \in \mathbb{R}$, l'équation $f(x, y) = 0$ a une unique solution qu'on note $y = \varphi(x)$.
- 2) Prouver que φ est C^∞ . Calculer $\varphi(0)$, $\varphi'(0)$ et $\varphi''(0)$.

Exercice 2

Soit

$$f(x, y) = 2x + y + x^2 + xy + y^3 .$$

- 1) Chercher les points critiques de f .
- 2) La fonction f est-elle convexe ?
- 3) Déterminer la nature du point critique de f , noté A , qui est situé sur l'axe $y = 0$.
- 4) Tracer sommairement les lignes de niveaux de f près de A en indiquant les tangentes à la ligne de niveau et les régions où $f > f(A)$

Exercice 3

Soit $D_n \subset \mathbb{R}^n$ (avec $n \geq 2$) défini par :

$$D_n = \{(x_1, \dots, x_n) \mid \forall i, x_i \geq 0, \sum_{i=1}^n x_i = 1\} ,$$

et $f : \mathbb{R}^n \rightarrow \mathbb{R}$ définie par :

$$f(x_1, \dots, x_n) = \sum_{1 \leq i < j \leq n} x_i x_j .$$

Le but de l'exercice est de déterminer les nombres m_n et M_n définis par

$$m_n = \inf_{x \in D_n} f(x)$$

et

$$M_n = \sup_{x \in D_n} f(x) .$$

- 1) Calculer m_n sans faire de calculs.
- 2) Calculer M_2 .
- 3) Soit $Z = \{x \in \mathbb{R}^n \mid \sum_{i=1}^n x_i = 1\}$. Chercher le point critique C_n de $F = f|_Z$ à l'aide des multiplicateurs de Lagrange. On pose $K_n = f(C_n)$. Prouver que la suite K_n est croissante.
- 4) En utilisant ce qui précède et une récurrence sur n , prouver que f atteint son maximum sur D_n uniquement au point C_n et calculer M_n .

5) Prouver que la restriction de f à D_n est concave (i.e. $-f$ est convexe) et retrouver ainsi une preuve du résultat précédent.

Exercice 4

Soit $D \subset \mathbb{R}^2$ défini par :

$$D = \{(x, y) \in \mathbb{R}^2 \mid 2x \geq y^2 \text{ et } y \geq 1\} .$$

1) Soit $g : \mathbb{R} \rightarrow \mathbb{R}$ une fonction. Soit $E_g = \{(x, y) \mid y \geq g(x)\}$. Prouver que E_g est un sous-ensemble convexe de \mathbb{R}^2 si et seulement si g est une fonction convexe.

2) Dessiner D et prouver que D est *convexe*.

3) Soit $A = (\frac{1}{2}, 1)$, déterminer le cône tangent à D au point A . On note Y ce cône.

4) Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ de classe C^1 . Dans quelle région Z de \mathbb{R}^2 doit se trouver nécessairement le gradient W de f au point A si A est un point de minimum de f . Dessiner Y et Z sur une même figure.

5) Soit, pour $\alpha \in I = [-\pi, \pi]$, f_α définie par :

$$f_\alpha(x, y) = (x - \frac{1}{2}) \cos \alpha + (y - 1) \sin \alpha .$$

On pose :

$$F(\alpha) = \inf_{M \in D} f_\alpha(M)$$

où $F(\alpha) \in \mathbb{R} \cup \{-\infty\}$. Pour quelles valeurs de α a-t-on $F(\alpha) = -\infty$? On note J cet ensemble de valeurs de α .

6) Calculer $F(\alpha)$ pour $\alpha \in I$ et tracer le graphe de F pour $\alpha \in I \setminus J$.

11. Examen de septembre 1998

Exercice 1

Soit $f(x, y) = 2x^3 + 6xy^2 - 3x^2 - 3y^2$.

1) Déterminer les points critiques de f et leur nature.

2) Soit $A = (a, b)$ celui des points cols tel que $b > 0$. Déterminer les tangentes aux lignes de niveaux de f passant par ce point et dessiner l'allure de la ligne de niveau de A au voisinage du point A . Indiquer les régions où $f > f(A)$ et $f < f(A)$.

Exercice 2

Calculer les bornes supérieures et inférieures de la fonction $f(x, y, z) = xyz$ lorsque (x, y, z) varie sur le domaine

$$D = \{(x, y, z) \mid x^2 + 4y^2 + 9z^2 \leq 3\} .$$

Exercice 3

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ et $a \in \mathbb{R}$.

1) Montrer que si f est convexe, le domaine

$$D_a = \{(x, y) \mid f(x, y) \leq a\}$$

est convexe.

2) Si D_a est convexe pour tout a , cela implique-t-il que f est convexe ?

3) Prouver que le domaine D défini par $x^4 + x^2 + y^2 + xy + 2x + 3y \leq 9$ est convexe.

4) Prouver que la courbe C d'équation

$$x^4 + x^2 + y^2 + xy + 2x + 3y = 9$$

admet une tangente au point $(1, 1)$ et la déterminer.

5) A l'aide d'un dessin, montrer le position de D par rapport à C .

Exercice 4

Soit $f(x) = x + \sqrt{1 + x^2}$.

1) Prouver que $f : \mathbb{R} \rightarrow \mathbb{R}$ est strictement convexe, C^2 et vérifie

$$0 < f'(x) < 2, \quad 0 < f''(x) \leq 1.$$

2) Soit $F_a : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par :

$$F_a(x, y) = f(x - y) + x^2 + axy + y^2.$$

Pour quelles valeurs de a , F_a est-elle convexe, strictement convexe ?

3) Prouver que pour $|a| < 2$, F_a est α -convexe en précisant une valeur de α .

4) Prouver que F_1 admet un unique point de minimum et écrire la méthode de gradient à pas μ pour cette fonction. Donner une valeur de μ pour laquelle la méthode converge et dire à quelle vitesse.

12. Partiel de mars 1999

12.1. Exercice. Soit f la fonction de \mathbb{R}^2 dans \mathbb{R} définie par

$$f(x, y) = \frac{xy^3}{x^2 + 2y^2} \text{ si } (x, y) \neq (0, 0) \text{ et } f(0, 0) = 0.$$

Montrer que f est de classe \mathcal{C}^1 sur \mathbb{R}^2 .

12.2. Exercice. On note $U = \{(x, y) \in \mathbb{R}^2 \mid x + y \neq 0\}$.

Trouver toutes les fonctions $\varphi : \mathbb{R}^* \rightarrow \mathbb{R}$ de classe \mathcal{C}^2 telles que la fonction $f : U \rightarrow \mathbb{R}$ définie par $f(x, y) = \varphi(x + y)$ vérifie

$$\forall (x, y) \in U, \quad \frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{2f(x, y)}{(x + y)^2}.$$

Parmi ces fonctions, donner celles se prolongeant de façon \mathcal{C}^2 en 0.

12.3. Problème. *Les questions 1 et 2 sont indépendantes.*

Dans tout le problème, \mathbb{R}^2 et \mathbb{R}^3 sont munis de la norme euclidienne, et si A est un compact de \mathbb{R}^2 , on rappelle que A est négligeable si et seulement si pour tout $\varepsilon > 0$, il existe une famille finie de boules $(B_k)_{1 \leq k \leq n}$ telle que

$$A \subset \bigcup_{k=1}^n B_k \text{ et } \sum_{k=1}^n \text{aire}(B_k) \leq \varepsilon.$$

1. Dans cette question, $\gamma : [0, 1] \rightarrow \mathbb{R}^2$ est une fonction de classe \mathcal{C}^1 .
 - a. Montrer qu'il existe une constante $M > 0$ telle que pour tout $n \geq 1$ et pour tout $k \in \{0, \dots, n-1\}$, on a :

$$\gamma([k/n, (k+1)/n]) \subset B(\gamma(k/n), M/n),$$

où $B(x, r)$ désigne la boule fermée de centre x et de rayon r .

- b. En déduire que $\gamma([0, 1])$ est un compact négligeable de \mathbb{R}^2 .
2. Dans cette question, $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ est une fonction de classe \mathcal{C}^2 . On pose

$$\Sigma = \{x \in \mathbb{R}^3 \mid f'(x) = 0\}.$$

- a. Montrer que Σ est fermé.
- b. Soit $R > 0$ fixé et $K_R = [-R, R]^3 \subset \mathbb{R}^3$.
Montrer qu'il existe $\lambda > 0$ tel que pour tout $(x, y) \in K_R \times K_R$, on a

$$\|f'(x) - f'(y)\| \leq \lambda \|x - y\|.$$

En déduire (ou montrer directement), que pour tout $(x, y) \in K_R \times K_R$, on a

$$\|f(x) - f(y) - f'(y)(x - y)\| \leq \frac{\lambda}{2} \|x - y\|^2.$$

- c. Pour tout n entier fixé, on découpe le cube K_R en n^3 cubes de cotés $2R/n$. Soit C l'un de ces n^3 cubes.
Montrer que si $C \cap \Sigma \neq \emptyset$, alors $f(C)$ est inclus dans une boule de rayon $6\lambda R^2/n^2$.
- d. En déduire que $f(\Sigma \cap K_R)$ est un compact négligeable de \mathbb{R}^2 .

Quelques commentaires :

- (i) il existe des applications *continues* de $[0, 1]$ dans \mathbb{R}^2 dont l'image est le carré $[0, 1] \times [0, 1]$, elles sont appelées **courbes de Peano**. La question 1 du problème montre que ces courbes ne sont pas de classe \mathcal{C}^1 .
- (ii) le problème montre deux cas particuliers du théorème de Sard : si $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ est une fonction de classe \mathcal{C}^k avec $k > \max(0, m - n)$, alors l'image par f des points $x \in \mathbb{R}^m$ dont la différentielle $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ n'est pas surjective est négligeable dans \mathbb{R}^n muni de la mesure de Lebesgue.

13. Examen de juin 1999

Exercice 1. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la fonction

$$f(x, y) = \sup(x^2, y^2) .$$

- Tracer la courbe de niveau $f = 1$.
- Calculer les dérivées $\partial_V f(m_0)$ avec $V = (1, 0)$, $V = (0, 1)$ et $V = (1, 1)$, $m_0 = (1, 1)$.
- En quels points de \mathbb{R}^2 la fonction f est-elle différentiable ?
- La fonction f est-elle convexe ?

Exercice 2. Prouver que la fonction $f(x, y, z) = xy^2z^3$ est majorée et atteint sa borne supérieure M sur la sphère S de \mathbb{R}^3 d'équation $x^2 + y^2 + z^2 = 1$. Déterminer M et les points $m \in S$ où ce sup est atteint.

Problème. \mathbb{R}^2 est muni de la structure euclidienne $\langle \cdot | \cdot \rangle$ et de la distance d canoniques.

a) Soient \vec{u}, \vec{v} et \vec{w} 3 vecteurs de \mathbb{R}^2 de norme euclidienne 1. Prouver que les vecteurs $\vec{V} = \vec{u} - \vec{v}$ et $\vec{W} = \vec{v} - \vec{w}$ sont liés si et seulement si 2 ou 3 des vecteurs \vec{u}, \vec{v} et \vec{w} sont égaux.

b) Soient $A = (1, 0)$, $B = (-\frac{1}{2}, \frac{\sqrt{3}}{2})$ et $C = (-\frac{1}{2}, -\frac{\sqrt{3}}{2})$ les 3 sommets d'un triangle équilatéral de centre $O = (0, 0)$. Pour $M = (x, y) \in \mathbb{R}^2 \setminus \{A, B, C\}$, on pose $X(M) = d(A, M) - d(B, M)$ et $Y(M) = d(B, M) - d(C, M)$. Calculer $\text{grad}X$ et $\text{grad}Y$ en termes des vecteurs

$$\vec{u} = \frac{A\vec{M}}{\|AM\|}, \quad \vec{v} = \frac{B\vec{M}}{\|BM\|}, \quad \vec{w} = \frac{C\vec{M}}{\|CM\|} .$$

c) Soit $F(x, y) = (X, Y)$. Prouver que, pour tout point M_0 intérieur au triangle ABC , F est un difféomorphisme C^∞ d'une boule ouverte de centre M_0 et de rayon $r(M_0) > 0$ sur un ouvert de \mathbb{R}^2 .

d) Calculer la matrice jacobienne J de F au point O ainsi que celle de son inverse local G au point $O = F(O)$.

e) Prouver que, si ϵ est donné assez petit, il existe un unique point $M(\epsilon) \in B(O, r(O))$ tel que $X(M(\epsilon)) = \epsilon$ et $Y(M(\epsilon)) = 0$.

f) Calculer la dérivée

$$\frac{dM}{d\epsilon}$$

en $\epsilon = 0$.

g) Prouver que, si $r > 0$ est assez petit, il existe un unique cercle passant par B et C et tangent intérieurement (resp. extérieurement) au cercle de centre A et de rayon r . Calculer le développement limité à l'ordre 1 en $r = 0$ du rayon de ces cercles. Faire une figure.