

Introduction à l'optimisation théorique et numérique pour l'ingénieur

Mehmet Ersoy ^{*1}

¹ IMATH–Institut de Mathématiques de Toulon et du Var, Université du sud Toulon-Var, Bâtiment U,
BP 20132 - 83957 La Garde Cedex, France.

La Garde, Février 2012

*e-mail: Mehmet.Ersoy@univ-tln.fr

Avant-propos

Ce cours présente les bases de l'optimisation pour les sciences de l'ingénieur, illustrés par de nombreux exemples pratiques, industriels et exercices corrigés. La première partie présente le cadre générale et les premières définitions. La seconde partie est consacrée à l'aspect théorique de l'optimisation. Dans un premier chapitre des rappels sur la continuité et la différentiabilité des fonctions à plusieurs variables sont donnés. Le second chapitre est quant à lui destiné aux premiers résultats de base sur l'existence, les conditions nécessaires et/ou suffisantes d'existence d'optimum dans le cadre des problèmes sans contraintes. Ce chapitre est fondamentale pour mettre en place les notions utiles pour la dernière partie qui concernera quelques algorithmes numériques "classiques" d'optimisations sans contraintes.

En annexe, afin de mettre en pratique les notions vues précédemment, je propose une série de travaux pratiques. Un exemple complet, à savoir rédaction de l'algorithme à la main (sur papier), mise en oeuvre avec Fortran (par exemple), et présentation des résultats obtenus est à votre disposition dans cette annexe.

Ces notes de cours ont été rédigées dans le cadre du cours Optimisation Numérique à l'école I.S.I.T.V.

1. N'hésitez pas à me signaler les erreurs et les coquilles qui subsisteraient dans ces notes. De manière générale, vos suggestions sont les bienvenues, c'est grâce à elles que ces notes pourront être améliorées pour vos camarades des prochaines années. Vous en souhaitant bonne lecture.

Table des matières

I	Introduction, Motivations & exemples	4
1	Généralités	4
2	Formulation mathématique	4
3	Exemples	7
3.1	Problème de Didon & Bulles de savon	7
3.2	Le toboggan le plus rapide : la courbe brachistochrone	8
3.3	Identification de paramètres	9
3.4	En mécanique	9
3.5	Exemples industriels	11
II	Aspects théoriques	11
1	Rappels : Continuité & calcul différentiel	11
1.1	Ensembles ouverts et continuité	13
1.2	Calcul différentiel	15
2	Étude théorique des problèmes d'optimisations	19
2.1	Résultats généraux d'existence	20
2.2	Notions de convexité	23
2.2.1	Définitions	23
2.2.2	Propriétés des fonctions convexes	25
2.3	Résultats d'unicité	27
2.4	Conditions d'optimalité	30
2.4.1	Conditions d'optimalité du premier ordre	30
2.4.2	Conditions d'optimalité du second ordre	32
2.4.3	Quelques exercices supplémentaires	33
2.4.4	Applications à quelques exemples	37
III	Aspects numériques	39
1	Premières définitions	39
2	Méthode du gradient	40
3	Méthode de Newton	43
	Vous trouverez une liste non exhaustive de références bibliographiques pour approfondir certaines des notions utilisés dans ce cours.	

Références

- [1] A. Auslander. *Optimisation. Méthodes Numériques*. Masson 1976.

Bonnans

[2] J.F. Bonnans, J.C. Gilbert, C. Lemaréchal et C. Sagastizabal. *Optimisation numérique. Aspects théoriques et pratiques*,

Cartans

[3] H. Cartan. *Cours de calcul différentiel*, Hermann, Paris, 1977.

Ciarlet

[4] P.G. Ciarlet. *Introduction à l'analyse matricielle et à l'optimisation*. Masson 1990.

Demmel

[5] J. Demmel. *Applied Numerical Linear Analysis*. SIAM 1997. Collection Mathématiques et Applications, Vol. 27, Springer-verlag, 1998.

Temam

[6] I.Ekeland et R.Temam, *Analyse Convexe et Problèmes Variationnels*, Dunod, Gauthier-Villars, Paris, 1974.

Lions1

[7] B. Larrouturnou et P.L. Lions. *Méthodes mathématiques pour les sciences de l'ingénieur : optimisation et analyse numérique*, Cours de l'Ecole Polytechnique, 1994.

Première partie

Introduction, Motivations & exemples

1 Généralités

Optimisation est un mot qui vient du latin *optimum* qui se traduit par *meilleur*. L'optimisation est une branche des mathématiques qui, étant donné un critère, permet d'analyser et déterminer analytiquement ou numériquement le "meilleur" élément d'un ensemble donné. Cette première définition vague est à l'origine de nombreux problèmes historiques et joue un rôle fondamentale dans notre société actuelle. Cette théorie apparaît dans toutes les sciences de l'ingénieur et autres. Dès qu'il s'agit de rentabiliser, de faire mieux aux moindres efforts, il y a de l'optimisation : optimisation d'un trajet (contrôle aérien et sans vouloir faire de publicité, Mappy, google maps, ...), optimisation de forme (les voitures de courses, les avions notamment supersonique, ...). L'optimisation d'un problème donné, c'est trouver la/une configuration idéale en terme d'efforts, de temps (donc d'argent), de dimensionnement, ...

L'optimisation est donc un sujet omniprésent mais aussi très ancien (par exemple le voyageur de commerce). Il connaît un nouvel essor depuis l'apparition des ordinateurs et la recherche d'algorithme performant est toujours d'actualité. La recherche continue !

2 Formulation mathématique

Un problème d'optimisation peut s'écrire sous la forme générale suivante :

$$(\mathcal{P}) \left\{ \begin{array}{l} \min_{x \in \Omega} J(x) \\ \text{sous les contraintes} \\ g(x) \leq 0 \\ h(x) = 0, \end{array} \right. \quad (1) \quad \boxed{\text{po}}$$

où les fonctions $J : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $g : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^i$ et $h : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^e$ sont des fonctions (en générale non linéaires) et Ω est un sous ensemble ouvert (éventuellement $\mathbb{R}^n$ tout entier) de $\mathbb{R}^n$.

Il s'agit alors de déterminer x^* (si il existe) qui réalise le minimum de J (qu'on appelle aussi fonction *objectif* ou *cout*) sous les contraintes inégalités $g \leq 0$ et égalités $h = 0$. Si $g = (g_1, \dots, g_i)$ alors la contrainte $g \leq 0$ signifie $g_k(x) \leq 0, \forall x \in \Omega, \forall j = 1, \dots, i$. De même, la contrainte $h = 0$ signifie $h_k(x) = 0, \forall x \in \Omega, \forall j = 1, \dots, e$ si h_k dénote la k -ième composante de h .

Lorsqu'un tel problème admet une solution, on dit que le problème est réalisable. Si x^* est une de ces solutions alors on dit qu'un tel x^* réalise le minimum de J et on note

$$J(x^*) = \min \{J(x); x \in \Omega, g(x) \leq 0, h(x) = 0\}.$$

Nous aborderons donc systématiquement les points suivants :

- l'existence d'une solution x^* du problème $(\mathcal{P})$,
- les conditions nécessaires pour x^* ,
- les conditions suffisantes pour x^* ,
- le calcul effectif de x^* à l'aide de divers algorithmes dit de *descente*.

L'objectif de ce cours est donc la présentation d'outils théoriques et numériques pour la résolution du problème $(\mathcal{P})$. On adoptera la terminologie suivante pour un problème

- sans contraintes, un *problème libre* et
- avec contraintes égalités (seulement si $h(x) = 0$),
- avec contraintes inégalités (seulement si $g(x) \leq 0$),
- générale dans le cas où on considère les contraintes égalités et inégalités.

En générale, la plupart des problèmes réels et industriels ne sont pas donnés sous la forme (Po) . Le premier objectif sera donc de mettre le problème sous la forme (II) . Le cas d'un problème de maximisation $\max_{x \in \Omega} F(x)$ se ramène au cas du problème (II) en posant $f(x) = -F(x)$. En pratique, la mise en forme standard nécessite un peu plus de travail.

Précisons maintenant plus en détails le cadre de ce cours. Nous supposons $\Omega = \mathbb{R}^n$. De manière générale, on note X l'ensemble des contraintes sans contraintes, égalités et/ou inégalité. Plus précisément, on note l'ensemble

- sans contraintes lorsque

$$X = \mathbb{R}^n,$$

- avec contraintes égalités lorsque

$$X = \{x \in \mathbb{R}^n; h(x) = 0\},$$

- avec contraintes inégalités lorsque

$$X = \{x \in \mathbb{R}^n; g(x) \geq 0\},$$

- générale dans le cas où on considère les contraintes égalités et inégalités

$$X = \{x \in \mathbb{R}^n; h(x) = 0 \text{ et } g(x) \leq 0\}.$$

Nous noterons dans la suite $\mathcal{P}$ un des quatre problèmes précédent (lorsqu'il n'y a pas d'ambiguïté) et nous supposons que l'ensemble X est non vide. En l'état actuel, bien que ce cours concerne l'optimisation sans contraintes, certains résultats s'appliquent aussi aux autres cas.

Définition 2.1. Soient X un sous ensemble de $\mathbb{R}^n$ (éventuellement l'espace tout entier) et J une fonction de X dans $\mathbb{R}$. On dit que $x^* \in X$ réalise un minimum local de J sur X si il existe $r > 0$ tel que

$$\forall x \in B(x^*, r) \cap X \quad J(x^*) \leq J(x)$$

où $B(x, r) := \{y \in \mathbb{R}^n; \|y - x\| \geq r\}$ est la boule de rayon r centrée en x dans $\mathbb{R}^n$.

On dit que le problème $\mathcal{P}$ est réalisable si il admet au moins une solution x^* et on note

$$J(x^*) = \min\{J(x); x \in X\}.$$

Remarque 2.1.

- la notion de local s'étend à la notion de global lorsque le voisinage (ici la boule) correspond à l'espace tout entier.
- la notion de maximum s'obtient en remplaçant $\leq$ par $\geq$.
- Les minima et maxima strict s'obtiennent en remplaçant les inégalités par des inégalités strictes.
- Par abus de langage, nous dirons x^* est un minimum de J au lieu de x^* réalise un minimum de J .
- $\|\cdot\|$ est une norme quelconque de $\mathbb{R}^n$. On rappelle que toutes les normes sont équivalentes en dimension finie.

La notion de minimum et de maximum (local ou global) est illustré à la figure suivante

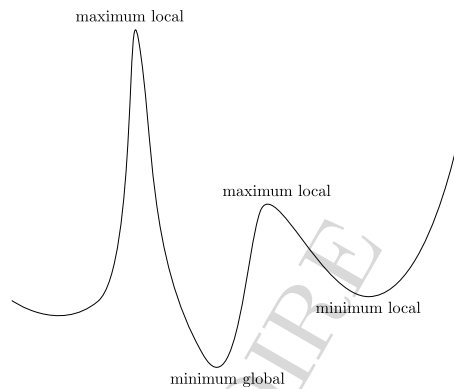


FIGURE 1 – Exemples de minima et maxima.

En principe, on s'intéresse à la recherche de minima car un problème de recherche de maxima se transforme en celui d'un minima via la

Proposition 2.1. Si x^* réalise un maximum (local ou global) de J sur X , alors x^* réalise un minimum (local ou global) de $-J$ sur X , i.e.

$$\max\{J(x), x \in X\} = -\min\{-J(x), x \in X\}.$$

Démonstration. Soit $x^* \in X$ tel que $J(x^*) = \max\{J(x); x \in X\}$. Alors

$$\forall x \in X \quad J(x) \leq J(x^*)$$

$$\iff$$

$$\forall x \in X \quad -J(x) \geq -J(x^*).$$

Par conséquent si x^* réalise le maximum de J sur X alors il réalise aussi le minimum de $-J$ sur X , i.e.

$$-J(x^*) = \min\{-J(x); x \in X\} \iff J(x^*) = -\min\{-J(x); x \in X\}.$$

□

Dans ce cours, nous supposons que le minimum est toujours atteint, ce qui justifie l'utilisation de min. Dans le cas contraire, on utilise la notation inf.

3 Exemples

On trouve dans la littérature de nombreux exemples de problèmes d'optimisation. Je propose la liste suivante qui permet de donner un aperçu des différents problèmes d'optimisations exposés dans le paragraphe précédent.

3.1 Problème de Didon & Bulles de savon

Virgile raconte dans l'Énéide que lorsque la reine Didon fonda la ville de Carthage, il ne lui fut alloué comme superficie que "ce que pourrait contenir une peau de boeuf". Elle découpa alors cette peau en fines lanières et encercla de cette cordelette sa future ville situé au bord de la mer. La question était donc de trouver la plus grande surface possible s'appuyant sur une droite (correspondant au rivage) et de frontière terrestre de longueur donnée. La réponse est bien sûr un demi disque. En termes mathématiques simplifiés, il faut déterminer la courbe plane de longueur fixée $l \geq 0$ qui enclose avec le segment reliant ses deux extrémités l'aire maximum, i.e. résoudre le problème suivant

$$\max \int_0^L y(x) dx$$

sous les contraintes

$$L \geq 0, \quad y(0) = 0, \quad l = \int_0^L \sqrt{1 + y'(x)^2} dx$$

où L est l'extrémité du segment et $y(x)$ la position de la courbe au dessus du point x du segment.

Cela signifie donc que parmi tous les domaines de périmètre fixé, le cercle minimise l'aire.

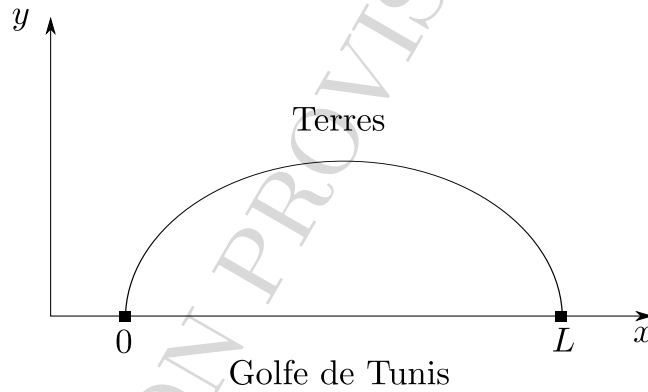


FIGURE 2 – Problème de Didon.

Un autre problème est celui de la bulle, à savoir pourquoi les bulles de savon prennent une forme sphérique? Par définition, une bulle est une fine pellicule d'eau et de savon qui renferme un certain volume d'air. A l'équilibre, une bulle de savon minimise sa surface sous la contrainte de volume fixé pour minimiser son énergie. On en déduit donc (a priori) que la sphère S est le domaine qui minimise la surface à volume fixé. C'est en fait l'inverse en dimension 3 du problème de Didon.



FIGURE 3 – La bulle de savon.

brachis

3.2 Le toboggan le plus rapide : la courbe brachistochrone

On se donne deux points dans l'espace A et B et on se pose la question sur la forme d'un toboggan qui permet d'atteindre le point B partant de A le plus rapidement possible. Mathématiquement, cela revient à déterminer la courbe qui relie A et B telle qu'un point matériel sous l'action de la gravité glissant le long de cette courbe arrive en B en un temps minimal.

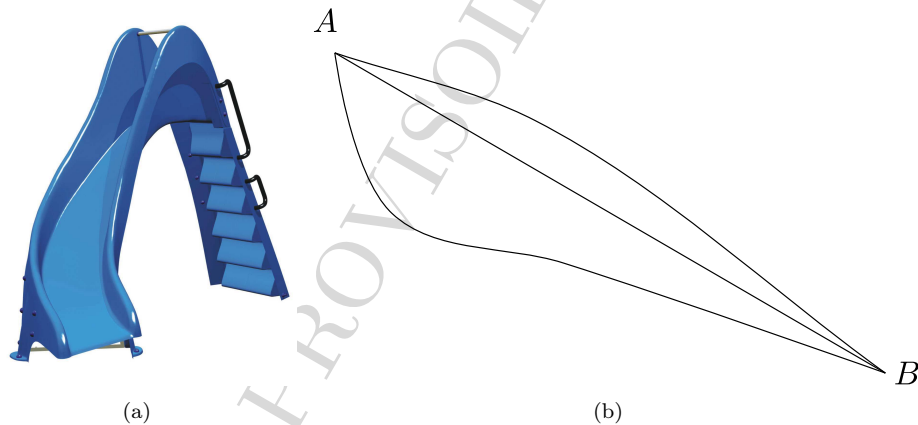


FIGURE 4 – Problème du toboggan .

On modélise ce problème, i.e. le toboggan par le graphe d'une fonction

$$f : [0, 1] \rightarrow \mathbb{R} \text{ avec } f(0) = h, \quad f(1) = 0$$

où h désigne la hauteur du toboggan (ici on prend $A = (0, h)$ et $B = (1, 0)$). La physique (conservation de l'énergie) impose alors que le temps T de descente du toboggan est donnée par la formule

$$T = \int_0^1 \frac{\sqrt{1 + f'(x)^2}}{2g(h - f(x))} dx$$

où g est la constante gravitationnelle. Minimiser ce temps T revient donc à trouver le “meilleur” f .

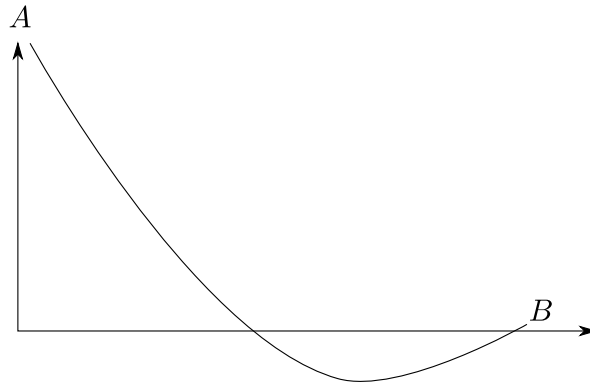


FIGURE 5 – La courbe brachistochrone : la courbe optimale d’un toboggan

La solution de ce problème est la courbe brachistochrone. Le problème fut clairement posé en 1696 par Jean Bernoulli et des solutions ont été apportées initialement par lui et son frère Jacques puis Newton, Leibniz, Hôpital, ...

Le problème avait été bien-sûr pas posé pour le toboggan mais dans le cadre de la trajectoire d’un rayon de lumière dans un milieu où la vitesse augmente selon la gravité terrestre, à savoir $c = \sqrt{2gh}$ où h représente la perte d’altitude par rapport au point de départ.

3.3 Identification de paramètres

Étant donnée un signal f , $f : \mathbb{R} \rightarrow \mathbb{R}$ dépendant de paramètres, disons (a, b, c, d) de la forme

$$f(t) = ae^{-bt} \cos(ct + d)$$

et un échantillons $[t_i, y_i]_{i=1, \dots, m}$ du signal $f(t)$, peut on identifier (a, b, c, d) ?

Oui, par exemple en minimisant la fonctionnelle

$$J(a, b, c, d) = \frac{1}{2} \sum_{i=1}^m (y_i - f(t_i))^2.$$

Le choix d’une fonctionnelle quadratique est arbitraire, la valeur absolue aurait également pu être considérée mais l’action du carré rend J différentiable (comme nous allons le voir dans la suite). C’est un exemple de base très important dans les sciences de l’ingénieur. En effet, vous serez souvent amenés à, étant donné un modèle, à identifier les paramètres de ce dernier, aussi appelé calibrage des paramètres.

3.4 En mécanique

On considère une corde horizontale de longueur 1 fixée à ses extrémités via une tension τ . On note $u(x)$, $x \in [0, 1]$ la déviation de la corde par rapport à l’origine. La corde fixée à ses extrémités ($u(0) = u(1) = 0$) est au repos si $u(x) = 0$, $x \in (0, 1)$. On peut montrer que l’énergie

explecorde

potentielle associée à une déformation donnée est

$$E(u) = \frac{1}{2} \int_0^1 \tau \left(\frac{d}{dx} u(x) \right)^2 dx.$$

En l'absence d'obstacle cette énergie est minimisée par $u(x) \equiv 0$. Si maintenant un obstacle est placé, intuitivement, seulement certains points de la corde seront en contact avec l'obstacle. Soit alors $v(x) \geq 0, x \in [0, 1]$ désignant cette obstacle.

En vu de connaître la réponse de la corde face à cette obstacle, il est donc naturel de résoudre le problème de minimisation suivant :

$$\begin{cases} \min_{u \in \Omega} E(u) \\ \text{sous les contraintes} \\ u(0) = u(1) = 0 \\ u(x) \geq v(x). \end{cases}$$

où Ω est un espace de fonctions. Ici l'inconnue est une fonction et un tel problème s'appelle *problème de calcul des variations*.

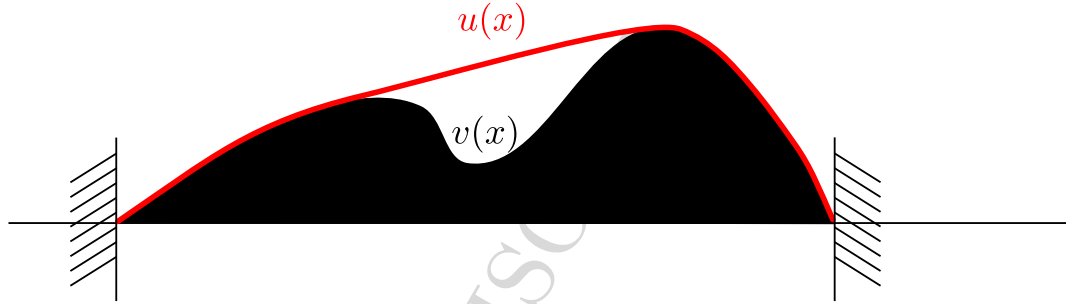


FIGURE 6 – Corde et obstacle.

En général, lorsqu'on ne sait pas apporter une solution analytique (ce qui est le cas en générale), on est souvent amené à discrétiser le problème pour ensuite résoudre le problème numériquement. Ici, on introduit

$$x_j = \frac{j}{N}, \quad j = 0, \dots, N$$

et le vecteur $U = (U_0, U_1, \dots, U_N, U_{N+1})$ avec

$$U_j \approx u(x_j) \quad j = 0, \dots, N.$$

On pose $U_0 = U_{N+1} = 0$ (ce sont nos conditions aux limites) et on procède à une discrétisation par différence finie de $\frac{d}{dx}u$. En d'autres termes, le problème discret s'écrit

$$\begin{cases} \min_U E_n(U) \\ \text{sous les contraintes} \\ G \leq 0 \quad j = 0 \dots N \end{cases}$$

où

$$E_n(U) = \frac{1}{2} U^t A U,$$

$$A = \frac{\tau}{N^2} \begin{pmatrix} 2 & -1 & & 0 \\ -1 & 2 & -1 & \\ 0 & \ddots & \ddots & \ddots \\ 0 & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix} \text{ et } G = (v(x_j) - U_j)_{j=1,\dots,N}.$$

On vient de transformer un problème d'optimisation non linéaire en un problème linéaire quadratique.

3.5 Exemples industriels

expleindus

Dans la plupart des problèmes industriels, un problème d'optimisation fait intervenir des équations aux dérivées partielles. La résolution de ces types de problèmes est souvent numérique car on ne sait pas les résoudre analytiquement. Il existe de nombreuses recherches actives dans ce domaine. Voici quelques exemples : optimisation de formes (ailes d'avion, voiture, ...) afin d'augmenter les performances aérodynamiques optimales, la chaise optimale (déterminer la

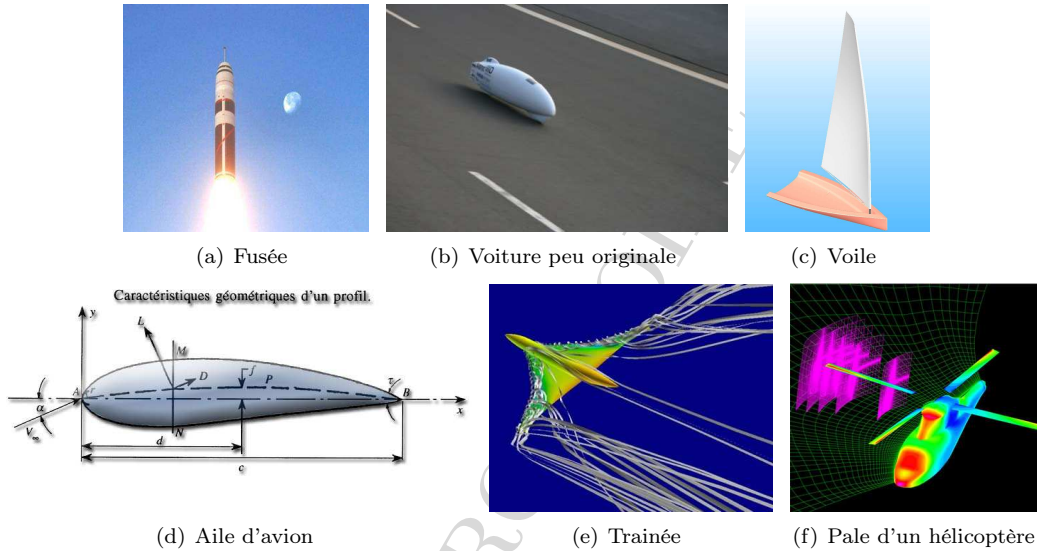


FIGURE 7 – Quelques exemples réels

Deuxième partie

Aspects théoriques

1 Rappels : Continuité & calcul différentiel

Dans la suite, des conditions de régularité seront demandées aux fonctions J, g et h . Ils est donc utile de faire les quelques rappels suivants.

Nous considérons ici des fonctions

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^m.$$

Nous allons nous focaliser aux cas où $n > 1$ et $m = 1$: c'est le cas d'une fonction de n variables réelles à valeurs dans $\mathbb{R}$. On appelle ces fonctions des fonctionnelles ou *champ scalaire*. Le cas $m > 1$ est une fonction vectorielle (ou champ vectoriel). Nous allons donc rappeler les notions de continuité et différentiabilité dans ce cadre présent. Lorsqu'aucune ambiguïté est possible nous noterons $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ la variable réelle et $f = (f_1, \dots, f_m)$ le champ vectoriel.

Définition 1.1. Une application $N : \mathbb{R}^n \rightarrow \mathbb{R}^+$ est dite une norme si et seulement si

- $\forall x \in \mathbb{R}^n \quad N(x) = 0 \longrightarrow x = 0,$
- $\forall x \in \mathbb{R}^n, \forall \lambda \in \mathbb{R}, \quad N(\lambda x) = \lambda N(x).$
- $\forall x \in \mathbb{R}^n, \forall y \in \mathbb{R}^n, \quad N(x + y) \leq N(x) + N(y).$

Les normes dans $\mathbb{R}^n$ sont les normes dites l_p , $1 \leq p < +\infty$:

$$N(x) := \|x\|_p = \left(\sum_{i=1}^n x_i^p \right)^{1/p}$$

et la norme infinie

$$N(x) := \|x\|_\infty = \max_{i=1, \dots, n} |x_i|.$$

Lorsque $p = 2$, on appelle la norme, norme euclidienne et elle est induite par le produit scalaire $\|x\|_2 = \sqrt{(x, x)}$.

Le produit scalaire de $x \in \mathbb{R}^n$ par $y \in \mathbb{R}^n$, qu'on note (x, y) (ou $x \cdot y$ ou $\langle x, y \rangle$) est définie par

$$x \cdot y = \sum_{i=1}^n x_i y_i.$$

On rappelle de nouveau qu'en dimension finie toutes les normes sont équivalents, à savoir si $\|\cdot\|_a$ et $\|\cdot\|_b$ sont deux normes dans $\mathbb{R}^n$ alors il existe $(\alpha, \beta) \in (\mathbb{R}_+^*)^2$ tel que

$$\forall x \in \mathbb{R}^n \quad \alpha \|x\|_b \leq \|x\|_a \leq \beta \|x\|_b.$$

Les normes les plus courantes en dimension finies sont les normes l^p où $p \in \mathbb{N}^*$. Elles sont définies par

$$\|x\|_p = \left(\sum_{i=1}^n x_i^p \right)^{\frac{1}{p}}.$$

Pour $p = 2$, c'est la norme euclidienne. Le cas $p = \infty$ correspond à la norme uniforme et elle est définie par

$$\|x\|_\infty = \max_{i=1, \dots, n} |x_i|$$

Exercice 1.1. Montrer que

$$\forall x \in \mathbb{R}^n \quad \|x\|_\infty \leq \|x\|_1 \leq n \|x\|_\infty \quad \text{et} \quad \|x\|_\infty \leq \|x\|_2 \leq \sqrt{n} \|x\|_\infty$$

Correction de l'exercice 1.1. Il suffit de remarquer que $\|x\|_2^2 \leq n \|x\|_\infty^2$

Exercice 1.2. Montrer que

$$\forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n, \quad |(x, y)| \leq \|x\|_2 \|y\|_2.$$

En déduire que l'application $\|\cdot\|_2 : \mathbb{R}^n \rightarrow \mathbb{R}^+$ est une norme.

Correction de l'exercice 1.2. Le produit scalaire est par définition géométrique donnée par $(x, y) = \cos((x, y)) \|x\|_2 \|y\|_2$. On a donc

$$|(x, y)| = |\cos((x, y))| \|x\|_2 \|y\|_2 \leq \|x\|_2 \|y\|_2.$$

Montrons que $\|\cdot\|_2$ est une norme. Les deux premières propriétés sont triviales. Pour montrer l'inégalité triangulaire, on procède comme suit :

$$\|x + y\|_2^2 = \|x\|_2^2 + 2(x, y) + \|y\|_2^2 \leq \|x\|_2^2 + 2\|x\|_2 \|y\|_2 + \|y\|_2^2 = (\|x\|_2 + \|y\|_2)^2$$

d'où le résultat.

Remarque 1.1. La difficulté pour montrer que l'application $\|\cdot\|_p : \mathbb{R}^n \rightarrow \mathbb{R}^+$ est une norme pour $p < \infty$ provient essentiellement de l'inégalité triangulaire. Cette inégalité porte aussi le nom d'inégalité de Minkowski (nous la démontrerons dans la partie sur les convexes [Section 2.2](#).) Cette inégalité est donc la suivante :

$$\forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n, \quad \left(\sum_{i=1}^n |x_i + y_i| \right)^{1/p} \leq \left(\sum_{i=1}^n |x_i| \right)^{1/p} + \left(\sum_{i=1}^n |y_i| \right)^{1/p}$$

Dans le cadre des sciences de l'ingénieur, les champs scalaires et vectoriels sont très fréquents en dimension 2 ou 3. Dans ce contexte, nous noterons $x = (x, y)$ au lieu de (x_1, x_2) dans $\mathbb{R}^2$ et $x = (x, y, z)$ au lieu de (x_1, x_2, x_3) dans $\mathbb{R}^3$.

1.1 Ensembles ouverts et continuité

Étant donné $x_0 \in \mathbb{R}^n$ et $r > 0$, l'ensemble des points tels que

$$\|x - x_0\| < r$$

est appelé une n -boule ouverte (car inégalité stricte). On la note $B(x_0, r)$. Ainsi en dimension 1 par exemple, le segment $(x_0 - r, x_0 + r)$ correspond à la 1-boule ouverte centrée en x_0 et de rayon r . En dimension 3, avec la norme euclidienne définie précédemment, c'est la boule au sens géométrique du terme. Lorsqu'aucune ambiguïté est possible, on parlera tout simplement de boule pour désigner une n -boule.

Les concepts de limite et continuité scalaire réel s'étendent facilement au cadre vectoriel. Pour ce faire rappelons ces notions.

Définition 1.2. Soit f une fonction de $\mathbb{R}$ dans $\mathbb{R}$. On dit que la fonction f admet une limite l en un point x_0 si

$$\forall \varepsilon > 0, \exists \eta > 0, |x - x_0| < \eta \Rightarrow |f(x) - l| < \varepsilon$$

et on note

$$l = \lim_{x \rightarrow x_0} f(x).$$

Définition 1.3. Soit f une fonction de $\mathbb{R}$ dans $\mathbb{R}$. On dit que la fonction f est continue en un point x_0 si

$$\forall \varepsilon > 0, \exists \eta > 0, |x - x_0| < \eta \Rightarrow |f(x) - f(x_0)| < \varepsilon.$$

Dans le cadre d'une fonction $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, on écrit

$$\lim_{x \rightarrow x_0} f(x) = l$$

pour désigner

$$\lim_{\|x-x_0\| \rightarrow 0} \|f(x) - l\| = 0$$

ce qui se réécrit aussi

$$\lim_{h \rightarrow 0} \|f(x_0 + h) - l\| = 0.$$

Par conséquent, une fonction f sera dite continue en x_0 si

$$\lim_{x \rightarrow x_0} f(x) = f(x_0).$$

La classe des fonctions continue, suivant les notations traditionnelles, sur Ω dans $\mathbb{R}^m$ est noté $C_0(\Omega, \mathbb{R}^m)$.

Comme il a été facile d'étendre ces notions aux cadre vectoriel, il est clair que toutes les propriétés de la dimension 1 sont encore vraies :

Théorème 1.1. *Si l et L sont les limites respectives de f et g en x_0 . Alors*

- $\lim_{x \rightarrow x_0} (f + g)(x) = l + L$
- $\lim_{x \rightarrow x_0} (fg)(x) = lL$
- $\lim_{x \rightarrow x_0} (\lambda f)(x) = \lambda l \quad \forall \lambda \in \mathbb{R}.$
- $\lim_{x \rightarrow x_0} \|f(x)\| = \|l\|$

Théorème 1.2. *Soit $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$. Notons $f(x) = (f_1(x), \dots, f_m(x))$ où $f_i(x)$ est la i -ème composante de f . Alors f est continue en $x_0 \in \Omega$ si et seulement si $\forall i = 1, \dots, m$ f_i est continue en x_0 .*

Théorème 1.3. *Soient $f : X \rightarrow Y$ et $g : Y \rightarrow Z$ deux fonctions vectorielles. Si g est continue en x_0 et si f l'est en $g(x_0)$ alors $f \circ g$ est continue en x_0 .*

Exemples 1.1.

- $(x, y) \rightarrow \sin(x^3 y)$ est continue dans $\mathbb{R}^2$ tout entier.
- $(x, y) \rightarrow \log(x^2 + y^2)$ est continue dans $\mathbb{R}^2 \setminus \{(0, 0)\}$.
- $(x, y) \rightarrow \begin{cases} \frac{xy}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon} \end{cases}$ est continue dans $\mathbb{R}^2$.

Exercice 1.3. *Les fonctions suivantes sont elles continues en $(0, 0)$?*

1. $f : (x, y) \rightarrow \frac{3x^2 + xy}{\sqrt{x^2 + y^2}}.$
2. $f : (x, y) \rightarrow \frac{xy^2}{x^2 + y^2}.$
3. $f : (x, y) \rightarrow \frac{xy}{x^2 + y^2}.$
4. $f : (x, y) \rightarrow \frac{\ln(1 + xy)}{x^2 + y^2}.$
5. $f : (x, y) \rightarrow \frac{\ln(1 + xy^2)}{x^2 + y^2}.$
6. $f : (x, y) \rightarrow \frac{x^2 + y}{x + y^2}.$

7. $f : (x, y) \rightarrow \frac{x^2 y^2}{x^2 + y^2} \exp(-x^2 - y^2).$

Correction de l'exercice 1.3.

1. Il suffit de remarquer que $|f(x, y)| \leq 4 \|(x, y)\|_2^2$ (puisque $2|xy| \leq x^2 + y^2$) et $\lim_{\|(x, y)\| \rightarrow 0} \|(x, y)\| = 0$, donc f est continue en $(0, 0)$.
2. Il suffit de remarquer que $|f(x, y)| \leq x$ (puisque $y^2 \leq x^2 + y^2$), donc f est continue en $(0, 0)$.
3. Si on considère les suites $(1/n, 1/n)$ et $(1/n, 0)$ qui convergent toutes les deux vers $(0, 0)$ mais $\lim_{n \rightarrow \infty} f(1/n, 1/n) = 1/2 \neq 0 = \lim_{n \rightarrow \infty} f(1/n, 0)$. Donc f n'est pas continue en $(0, 0)$.
4. On montre que f n'est pas continue en $(0, 0)$ en considérant les suites précédentes ou écrire le développement en $(0, 0)$

$$\frac{\ln(1 + xy)}{x^2 + y^2} \approx_{(0,0)} \frac{xy}{x^2 + y^2}$$

qui n'est pas continue en $(0, 0)$.

5. On a

$$\frac{\ln(1 + xy^2)}{x^2 + y^2} \approx_{(0,0)} \frac{xy^2}{x^2 + y^2}$$

donc f est continue en $(0, 0)$.

6. On a $f(x, x) = f(y, y) = 1$ et $f(0, y) = 1/y \rightarrow \infty$ quand $y \rightarrow 0$. Donc f n'est pas continue en $(0, 0)$.
7. On a $|f(x, y)| \leq \frac{\|(x, y)\|_2^4}{\|(x, y)\|_2^2} \exp(-\|(x, y)\|_2^2) = re^{-r} \rightarrow_{r \rightarrow 0} 0$ avec $r = \|(x, y)\|_2^2$ donc f est continue en $(0, 0)$.

1.2 Calcul différentiel

Nous rappelons ici quelques notions de base pour le calcul différentiel utile dans la suite dans le cadre d'un champ scalaire, i.e. on considère ici le cas d'une fonction de plusieurs variables réelles $x = (x_1, \dots, x_n)$ à valeur dans $\mathbb{R}$ (i.e. $m = 1$).

Définition 1.4. On dit que f est différentiable (ou dérivable au sens de Fréchet) en $x_0 \in \mathbb{R}^n$ si il existe une application linéaire $d_{x_0}f : \mathbb{R}^n \rightarrow \mathbb{R}$ (noté aussi $f'(x_0)$) telle que

$$\lim_{\|h\| \rightarrow 0} \frac{f(x_0 + h) - f(x_0) - d_{x_0}f(h)}{\|h\|} = 0.$$

On appelle $d_{x_0}f$ la différentiel de f au point x_0 . On note aussi

$$f(x_0 + h) = f(x_0) + d_{x_0}f(h) + \|h\| \varepsilon(h)$$

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

Dans la définition ci-dessus il faut lire et comprendre $d_{x_0}f(h)$ comme la différentiel de f au point x_0 appliquée en h .

Exercice 1.4. Montrer que la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \rightarrow 2x + y - x^2 - y^2$ est différentiable en tout point de $\mathbb{R}^2$.

Correction de l'exercice 1.4. Prenons $h = (u, v)^t$ et $(x_0, y_0) \in \mathbb{R}^2$. On a donc $f(x_0 + u, y_0 + v) = 2x_0 + y_0 - x_0^2 - y_0^2 + (2 - 2x_0)u + (1 - 2y_0)v + \sqrt{u^2 + v^2}\varepsilon(u, v)$ avec $\varepsilon(u, v) = -\sqrt{u^2 + v^2}$. La différentielle de cette fonction existe et est donnée par $(2 - 2x_0)u + (1 - 2y_0)v$. C'est bien une application linéaire continue de $\mathbb{R}^2$ dans $\mathbb{R}$. De plus $\varepsilon(u, v) \rightarrow 0$ lorsque $\sqrt{u^2 + v^2} \rightarrow 0$. Donc f est bien différentiable en tout point de $\mathbb{R}^2$.

Rappelons que :

Définition 1.5. Une application f d'un espace vectoriel E à valeurs dans F sur un corps $\mathbb{K}$ est une application linéaire ou morphisme d'espaces vectoriels si

1. $\forall (x, y) \in E \times F, f(x + y) = f(x) + f(y)$ et
2. $\forall x \in E, \lambda \in \mathbb{K}, f(\lambda x) = \lambda f(x)$.

Dans le cas qui nous préoccupe, une application linéaire de $\mathbb{R}^n \rightarrow \mathbb{R}$ (ou forme) est de la forme

$$\forall x \in \mathbb{R}^n, f(x_1 e_1 + \dots + x_n e_n) = x_1 f(e_1) + \dots + x_n f(e_n)$$

où e_i est le i ème vecteur de la base canonique de $\mathbb{R}^n$ et $f(e_i) = a_i$. On note $\mathcal{L}(\mathbb{R}^n)$ l'ensemble des formes linéaires de $\mathbb{R}^n$ dans $\mathbb{R}$ et plus généralement $\mathcal{L}(E, F)$ l'ensemble des formes linéaires de E dans F .

De manière générale nous avons le résultat suivant

Théorème 1.4. Soit $f \in \mathcal{L}(E, F)$ alors les propositions suivantes sont équivalentes :

1. f est continu en 0.
2. f est continu sur E .
3. f est uniformément continu i.e.

$$\forall \varepsilon > 0, \exists \eta > 0; \forall (x, y) \in E \times E, \|x - y\| \leq \eta \Rightarrow \|f(x) - f(y)\| < \varepsilon.$$

4. f est Lipschitz
5. f est borné sur la boule unité
6. f est borné sur la sphère unité

Démonstration.

1. Montrons que 1 implique 3 (par définition 3 implique 2). Supposons f continu en 0 alors

$$\forall \varepsilon > 0, \exists \eta > 0; \forall y \in E, \|y\| \leq \eta \Rightarrow \|f(y)\| < \varepsilon.$$

Donc pour ε et η donné, on a

$$\forall (x, y) \in E \times E, \|x - y\| \leq \eta \Rightarrow \|f(x) - f(y)\| = \|f(x - y)\| < \varepsilon.$$

Donc f est uniformément continu sur E donc continu.

2. Montrons que 3 implique 4. Supposons f uniformément continu en 0. Prenons $\varphi = 1$ il existe $\eta > 0$ tel que $\|x\| \leq \eta \Rightarrow \|f(x)\| \leq 1$ (on a donc que f est continue sur la boule unité). En particulier, pour tout x non nul, on a aussi $\left\| f\left(\eta \frac{x}{\|x\|}\right) \right\| \leq 1 \Rightarrow \|f(x)\| \leq \frac{\|x\|}{\eta}$ (vrai également si $x = 0$). Donc $\forall (x, y) \in E^2$,

$$\|f(x) - f(y)\| = \|f(x - y)\| \leq k \|x - y\| \text{ avec } k = \frac{1}{\eta}.$$

3. L'implication $4 \rightarrow 5$, $5 \rightarrow 6$ et $6 \rightarrow 1$ sont évidentes. □

Définition 1.6. On dit que f est dérivable dans la direction $y \in \mathbb{R}^n$ au point $x_0 \in \mathbb{R}^n$ si

$$\lim_{t \rightarrow 0^+} \frac{f(x_0 + ty) - f(x_0)}{t} = 0$$

existe et est linéaire par rapport à y . Lorsque $y = e_i$ le i -ème vecteur de la base canonique de $\mathbb{R}^n$, on dit que f admet une dérivée partielle par rapport à x_i et on la note $\frac{\partial}{\partial x_i} f$ (ou $\partial_{x_i} f$ ou encore lorsqu'aucune ambiguïté est possible $\partial_i f$). On a donc

$$\lim_{t \rightarrow 0} \frac{f(x_0 + te_i) - f(x_0)}{t} = \partial_{x_i} f(x_0).$$

On peut donc maintenant présenter le résultat important suivant :

Théorème 1.5. Si f est différentiable en $x_0 \in \mathbb{R}^n$ alors toutes les dérivées partielles existent et on a

$$d_{x_0} f(h) = \nabla f(x_0) \cdot h = \sum_{i=1}^n \partial_{x_i} f(x_0) h_i \in \mathbb{R}$$

où $h \in \mathbb{R}^n$ et

$$\nabla f(x_0) = \begin{pmatrix} \partial_{x_1} f(x_0) \\ \vdots \\ \partial_{x_n} f(x_0) \end{pmatrix}$$

est le vecteur gradient (simplement appelé gradient de f)

En vue des définitions et du théorème précédent, on peut donc écrire :

$$f(x_0 + h) = f(x_0) + \nabla f(x) \cdot h + \|h\| \varepsilon(h)$$

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

Exercice 1.5. On introduit la célèbre fonction de Rosenbrock (1960). C'est une fonction non convexe qui est souvent utilisé en optimisation numérique (comme "cas test"). Cette fonction est aussi connue sous le de fonction banane. La fonction réalise un minimum global à l'intérieur d'une longue vallée étroite de forme parabolique mais la convergence numérique vers ce minimum est extrêmement difficile. Ce qui explique pourquoi une telle fonction est souvent utilisé comme cas test en optimisation. Elle est donnée par

$$f(x, y) = (1 - x)^2 + \alpha(y - x^2)^2.$$

Le minimum global (indépendant du coefficient α) vaut $(1, 1)$.

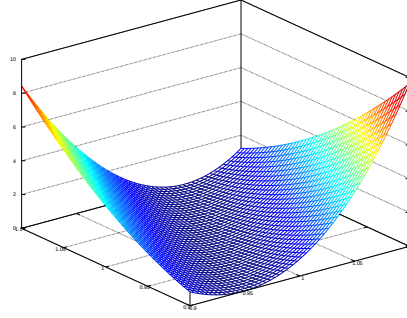


FIGURE 8 – Fonction de Rosenbrock $\alpha = 1$

Calculer son gradient, donner l'expression de sa différentiel appliquée en h et en déduire son développement de Taylor à l'ordre 1.

Dans le cas d'une fonction vectorielle, par exemple à valeur dans $\mathbb{R}^n$, la notion de dérivabilité reste la même. Cependant, cette fois ci la dérivée (ou différentielle) en x n'est pas un vecteur mais une matrice qu'on appelle matrice Jacobienne. Par définition elle correspond à la matrice

$$d_x f = \text{Jac}(f(x)) = (\partial_{x_j} f_i(x))_{1 \leq i, j \leq n}.$$

Par suite on peut alors définir la dérivée d'ordre supérieur de $f : \mathbb{R}^n \rightarrow \mathbb{R}$ comme la dérivée de df , i.e.

Définition 1.7. On dit que $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est dérivable sur un sous ensemble X de $\mathbb{R}^n$ si f est dérivable en tout point de X .

L'application $df : X \rightarrow \mathbb{R}^n$ est une application vectorielle tel que $df(x) = \nabla f(x)$. Dans ce cas, pour calculer la dérivée seconde de f , il suffit d'appliquer la notion hessienne sur df . On peut donc identifier cette dérivée seconde, qu'on note $D^2 f(x)$ à la $\text{Jac}(df(x))$. Il vient que

$$D^2 f(x) = (\partial_{x_i x_j} f(x))_{1 \leq i, j \leq n}.$$

Théorème 1.6. Si f est de classe C^2 sur un ouvert de $\mathbb{R}^n$ en tout point alors on a

$$\partial_{x_i x_j} f(x) = \partial_{x_j x_i} f(x)$$

Dans le cadre d'une fonction régulière de classe C^2 , la matrice Hessienne est donc symétrique.

Exercice 1.6. Vérifier ce résultat avec la fonction de Rosenbrock.

Enfin, pour terminer, il est commode de rappeler la formule de Taylor avec reste intégrale

Théorème 1.7. Soit $f : X \rightarrow Y$ une application de classe C^{n+1} (X et Y deux espaces de Banach). Soit a et h tel que $[a, a+h] \in X$, alors

$$f(a+h) = f(a) + f'(a) \cdot h + \dots + \frac{1}{n!} f^{(n)}(a) \cdot h^n + \int_0^1 \frac{(1-t)^n}{n!} f^{(n+1)}(a+th) \cdot h^{n+1} dt.$$

Exercice 1.7. Donner le développement de Taylor à l'ordre deux de la fonction de Rosenbrock.

Exercice 1.8. La fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \rightarrow \frac{x^3 y}{x^4 + y^2}$ prolongée par 0 à l'origine est elle continue ? différentiable ? C^1 ?

Correction de l'exercice 1.8. La fonction est bien-sûr continue à l'origine puisque

$$\left| \frac{x^3 y}{x^4 + y^2} \right| \leq \left| \frac{x^3 y}{2x^2 y} \right| \leq \left| \frac{x}{2} \right|.$$

Notons $u : \mathbb{R} \rightarrow \mathbb{R}^2$, $x \rightarrow (x, x^2)$. On a $(f \circ u)(x) = x/2$, donc $(f \circ u)' = 1/2$. D'autres part, puisque $f(x, 0) = f(0, y) = 0$. Si f était différentiable, on aurait $f'(0, 0) = 0$ ce qui ne peut être le cas car $(f \circ u)' = 1/2$. Donc f n'est pas différentiable.

Exercice 1.9. La fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \rightarrow \frac{xy^3}{x^4 + y^2}$ prolongée par 0 à l'origine est elle continue ? différentiable ? C^1 ?

Correction de l'exercice 1.9. Calculons les dérivées partielles. On a

$$\partial_x f = -\frac{y^3(3x^4 - y^2)}{(x^4 + y^2)^2}$$

et

$$\partial_y f = \frac{xy^2(3x^4 + y^2)}{(x^4 + y^2)^2}.$$

Elles sont clairement continue sauf peut être à l'origine. Calculons alors

$$|\partial_x f| \leq \left| \frac{3x^4 y^3}{(x^4 + y^2)^2} \right| + \left| \frac{y^5}{(x^4 + y^2)^2} \right| \leq \left| \frac{3x^4 y^3}{(2x^2 y)^2} \right| + \left| \frac{y^5}{(y^2)^2} \right| \leq |7y/4|.$$

De même, on montre que

$$|\partial_y f| \leq |7x/4|.$$

Donc par passage à la limite à l'origine, on établit que cette fonction est C^1 .

2 Étude théorique des problèmes d'optimisations

On s'intéresse dans ce chapitre à la résolution du problème suivant :

Trouver le minimum d'une fonctionnelle avec ou sans contraintes

ce qui se traduit mathématiquement par la résolution du problème de minimisation

– sans contrainte :

$$\min_{x \in \mathbb{R}^n} J(x),$$

– avec contrainte :

$$\min_{x \in X} J(x).$$

Nous nous limiterons aux cas de la dimension finie pour la simple raison que tout problème continu (posé en dimension infinie, par exemple un espace hilbertien) peut être réduit à un problème de dimension finie par discrétisation (c.f. exemple 3.4).

Nous donnerons quelques résultats d'existences, conditions nécessaires et suffisantes pour savoir si le problème d'optimisation est réalisable, à savoir l'existence d'un x^* tel que

$$J(x^*) = \min \{J(x); x \in \mathbb{R}^n\}$$

pour le problème sans contrainte et

$$J(x^*) = \min \{J(x); x \in X\}$$

pour le problème avec contraintes.

En générale, l'ensemble X contient des contraintes dites

– d'égalités lorsque

$$X = \{x \in \mathbb{R}^n; h_i(x) = 0, i = 1 \dots e\},$$

– d'inégalités lorsque

$$X = \{x \in \mathbb{R}^n; g_i(x) \leq 0, i = 1 \dots i\},$$

– mixtes lorsque

$$X = \{x \in \mathbb{R}^n; g_i(x) \leq 0, i = 1 \dots i \text{ et } h_i(x) = 0, i = 1 \dots e\}.$$

Dans la littérature, on rencontre aussi la terminologie et plus précisément la classification suivante :

– de programmation linéaire lorsque J est linéaire et X un polyèdre convexe de la forme

$$X = \{x \in \mathbb{R}^n, Bx \leq b\}$$

où B est une matrice de taille $m \times n$ et $b \in \mathbb{R}^m$,

- de programmation quadratique lorsque $J(x) = \frac{1}{2}Ax \cdot x - b \cdot x$ est quadratique et X encore un polyèdre convexe,
- de programmation convexe lorsque J et X sont convexes
- de programmation différentiable lorsque J et les contraintes sont une ou deux fois différentiable.

2.1 Résultats généraux d'existence

La plupart des théorèmes d'existence de minimum sont des variantes du théorème classique due à Weierstrass :

theo1

Théorème 2.1. Soit J une fonction continue de l'intervalle $[a, b]$ dans $\mathbb{R}$. Alors J est borné et atteint ses bornes. En particulier, le problème de minimisation

$$\min \{J(x); x \in [a, b]\}$$

admet au moins une solution.

On a en particulier le résultat similaire suivant :

theo2

Théorème 2.2. Soit X un sous ensemble fermé, borné et non vide de $\mathbb{R}^n$ et J une fonction continue de $\mathbb{R}^n$ dans $\mathbb{R}$. Alors le problème

$$\min\{J(x); x \in X\}$$

admet au moins une solution.

Démonstration. X étant fermé et borné dans $\mathbb{R}^n$, de toute suite dans X , on peut extraire une sous suite convergente. Autrement dit, soit $(x_n)_n \subset X$, il existe une sous suite d'indice croissant n_k telle que la suite $(x_{n_k})_k$ converge dans X . Par conséquent pour toute suite minimisante $(x_n) \subset X$, i.e. telle que

$$\lim_{n \rightarrow +\infty} J(x_n) = \min\{J(x); x \in X\}$$

on peut extraire une suite x_{n_k} convergente, soit x^* sa limite. Par continuité de J , il s'en suit que

$$\lim_{k \rightarrow \infty} J(x_{n_k}) = J(x^*) = \min\{J(x); x \in X\}.$$

□

Dans le cadre de la dimension finie, un ensemble compact est un ensemble fermé et borné. En d'autres termes le théorème 2.2 est l'analogue du théorème 2.1 pour un problème en dimension n . Si l'ensemble X est non borné, par exemple l'espace tout entier $\mathbb{R}^n$ (le cas d'un problème libre typiquement) alors il faut une condition nécessaire sur la fonctionnelle J . Un théorème dans ce contexte permet d'assurer l'existence d'un tel minimum :

theo3

Théorème 2.3. Soit X un sous ensemble fermé de $\mathbb{R}^n$ et J une fonctionnelle continue. Si X est non borné et $\lim_{\|x\| \rightarrow +\infty} J(x) = +\infty$ (coercivité) alors le problème

$$\min\{J(x); x \in X\}$$

admet au moins une solution.

Démonstration. Soit x_n une suite minimisante. J étant coercitive, on vérifie que la suite $(x_n)_n$ est bornée et on conclue comme dans la démonstration du théorème 2.2. Montrons que J coercitive implique que la suite $(x_n)_n$ est bornée. Supposons que ce ne soit pas le cas, alors cette suite on peut extraire une sous suite encore noté x_n telle que $\|x_n\| \rightarrow +\infty$. Par coercivité de J on aurait alors $J(x_n) \rightarrow +\infty$ ce qui contredit le fait $\lim_{n \rightarrow +\infty} J(x_n) < +\infty$. □

Remarque 2.1. La notion de fonction coercitive est aussi appelée fonction infini à l'infini, ce qui exprime :

$$\forall X \in \mathbb{R}, \exists R > 0, \forall x \in X \subset \mathbb{R}^n, \|x\| \geq R \Rightarrow J(x) \geq X.$$

Pour montrer qu'une fonction $J : X \rightarrow \mathbb{R}$ est coercitive, on utilise en pratique la proposition suivante

Proposition 2.1. Si il existe $g : \mathbb{R} \rightarrow \mathbb{R}$ tel que $\lim_{t \rightarrow +\infty} g(t) = +\infty$ et

$$J(x) \geq g(\|x\|)$$

alors J est coercitive.

Exemples 2.1. Les fonctions suivantes sont coercitives :

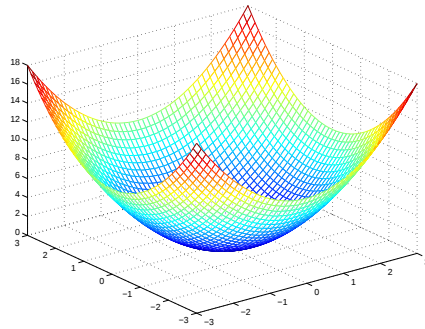
- $J(x) = \|x\|, x \in \mathbb{R}^n$.
 - $J(x) = Ax \cdot x - bx, A \in M_n(\mathbb{R})$ symétrique définie positive
- et $J(x) = b \cdot x + c, x \in \mathbb{R}^n, b \in \mathbb{R}^n, c \in \mathbb{R}$ n'est jamais coercitive.

Exercice 2.1. Les fonctions suivantes sont elles coercitives ?

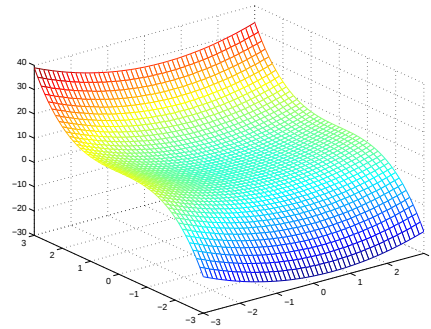
- $J(x) = x^3 + x + 1, x \in \mathbb{R}$
- $n = 2, J(x) = x^2 + y^2 - ax - by - c$ avec a, b et $c \in \mathbb{R}$.
- $n = 2, J(x) = x^2 - y^2$

Correction de l'exercice 2.1.

- $J(x) = x^3 + x + 1, x \in \mathbb{R}$ n'est pas coercitive car $\lim_{x \rightarrow \pm\infty} J(x) = \pm\infty$.
- $J(x) = \|x\|$ est clairement coercitive.
- $n = 2, J(x) = x^2 + y^2 - ax - by - c$, est coercitive quelque soit les valeurs de a, b et c (c.f Théorème ^{theo4} 2.6)
- $n = 2, J(x) = x^2 - y^2$ n'est pas coercitive. En effet, la suite $(0, n)$ est telle que $\|x_n\| \rightarrow \infty$ lorsque $n \rightarrow +\infty$ alors que $J(x_n) = -n^2$ ne converge pas vers $+\infty$.
- $n = 2, J(x) = x^2$ n'est pas coercitive. De même, en considérant la suite précédente x_n , on $J(x_n) = 0$.
- Une fonctionnelle affine $J(x) = b \cdot x + c, b \in \mathbb{R}^n, c \in \mathbb{R}$ n'est jamais coercitive. En effet, la suite $x_n = n \frac{b_1}{|b_1|} e_1$ est telle que $\|x_n\| = n \rightarrow +\infty$ et $J(x_n) = b_n n \rightarrow -\infty$ lorsque $n \rightarrow +\infty$
- $J(x) = Ax \cdot x + bx + c, A \in M_n(\mathbb{R})$ définie positive, $b \in \mathbb{R}^n$ (c.f Théorème ^{theo4} 2.6)



(a)



(b)

FIGURE 9 – Exemple d'une fonction coercitive (gauche) et non coercitive (droite).

Les deux théorèmes précédent permettent de montrer l'existence d'au moins un minimum mais donne aucune information sur l'unicité.

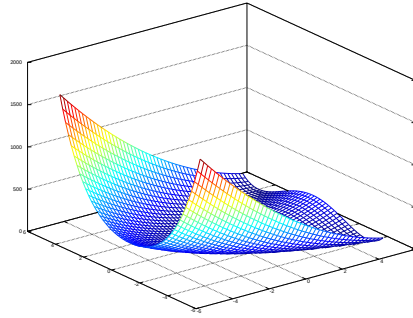


FIGURE 10 – Exemple d'un minima (globaux) d'une fonction coercitive (non convexe).

Le critère de convexité joue un rôle fondamentale dans la quête de l'unicité. Commençons par rappeler quelques notions liées à la convexité.

2.2 Notions de convexité

La convexité dans le cadre de la théorie de l'optimisation est une notion très importante aussi bien d'un point de vue théorique (existence comme nous allons le voir) mais aussi numérique. Les problèmes quadratiques sont à la base de nombreux algorithmes non linéaires.

2.2.1 Définitions

Définition 2.1 (Ensemble et fonction convexe).

- Un ensemble C est dit convexe si

$$\forall (x, y) \in C^2, \forall \lambda \in [0, 1], \lambda x + (1 - \lambda)y \in C$$

- Une fonction J est dite convexe si

$$\forall (x, y) \in C^2, \forall \lambda \in [0, 1], J(\lambda x + (1 - \lambda)y) \leq \lambda J(x) + (1 - \lambda)J(y)$$

- Une fonction J est dite strictement convexe si

$$\forall (x, y) \in C^2, x \neq y, \forall \lambda \in [0, 1], J(\lambda x + (1 - \lambda)y) < \lambda J(x) + (1 - \lambda)J(y)$$

La définition d'un ensemble convexe signifie que tout segment est inclus dans l'ensemble. Dans le cas d'une fonction de $\mathbb{R}$ dans lui-même, la convexité correspond au fait que la corde est au dessus du graphe.

Définition 2.2. *Domaine de J* Soit $J : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction convexe. On appelle domaine de J et on note $\text{Dom}(J)$ l'ensemble

$$\text{Dom}(J) = \{x \in \mathbb{R}^n; J(x) < +\infty\}.$$

Ce domaine est convexe et lorsqu'il est non vide on dit que J est propre

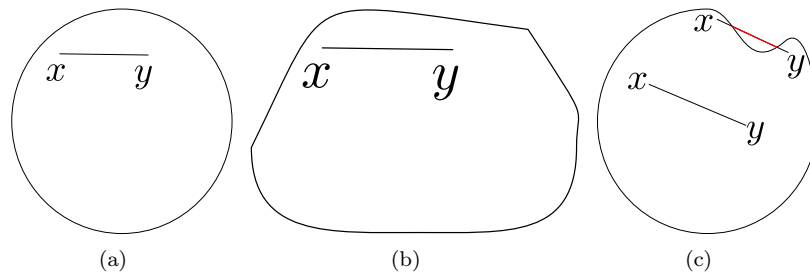


FIGURE 11 – Exemple d'ensemble strictement convexe (à gauche), convexe (milieu) et non convexe (à droite).

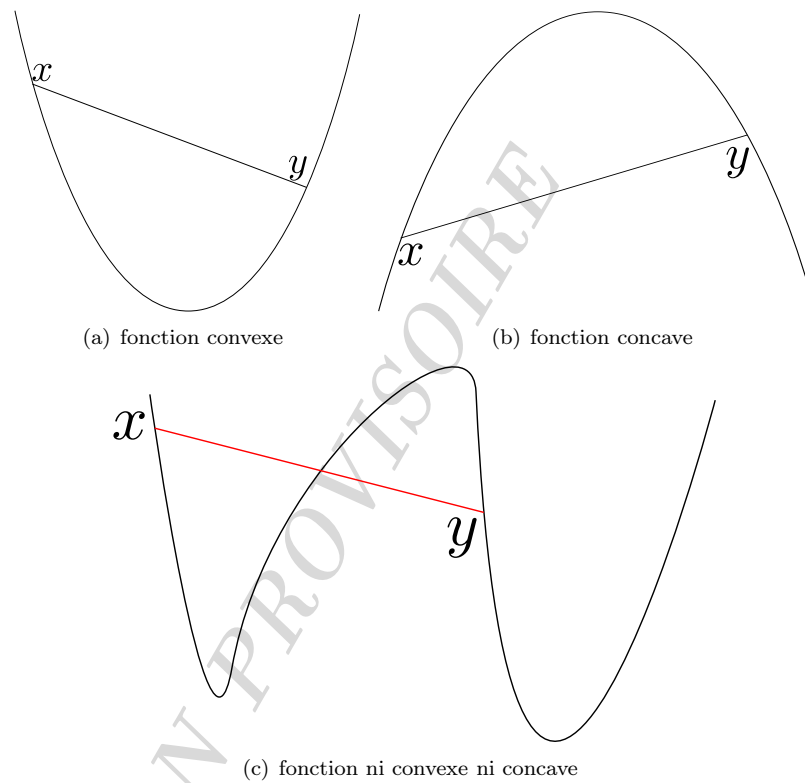


FIGURE 12 – Exemple de fonction convexe et non convexe.

Remarque 2.2. La fonction dite concave est telle que $-f$ soit convexe. Dans les problèmes de maximisation, c'est la notion de convexe est remplacée par concave.

Exemples 2.2. Montrer que

- un intervalle $[a, b]$ est convexe.
- $\mathbb{R}^*$ n'est pas convexe.
- $J(x) = \|x\|$ est strictement convexe

- $J(x) = b \cdot x + c$, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$ est seulement convexe.
- Soit A une matrice symétrique semi-définie positive de taille n , $J(x) = Ax \cdot x - b \cdot x$. Alors J est convexe. Si de plus A est définie positive alors la forme quadratique J est strictement convexe.

Exercice 2.2.

- Montrer qu'une norme est convexe.
- Montrer que l'indicatrice d'un ensemble K est convexe si et seulement si K est convexe.
- Montrer que $|xy|$ n'est pas convexe autour de l'origine.
- Montrer que pour tout p et q conjugué, i.e. $\frac{1}{p} + \frac{1}{q} = 1$, on a

$$\forall x, y > 0 \quad xy \leq \frac{x^p}{p} + \frac{y^q}{q}.$$

C'est l'inégalité de Young. Retrouver ce résultat simplement pour $p = 2$.

Correction de l'exercice 2.2. la fonction $x \rightarrow \ln(x)$ est concave (puisque sa dérivée seconde est $-1/x^2 < 0$). Par ailleurs, si $a + b = 1$, on a

$$a \ln(x) + b \ln(y) \leq \ln(ax + by)$$

$$\ln(xy) = \ln(x) + \ln(y) = \frac{1}{p} \ln(x^p) + \frac{1}{q} \ln(y^q) \leq \ln\left(\frac{x^p}{p} + \frac{y^q}{q}\right)$$

d'où le résultat.

Ou encore, on peut montrer ce résultat en utilisant la convexité de la fonction $x \rightarrow \exp(x)$. En effet, on a :

$$xy = \exp(\ln(xy)) = \exp\left(\frac{1}{p} \ln(x^p) + \frac{1}{q} \ln(y^q)\right) \leq \frac{1}{p} \exp(\ln x^p) + \frac{1}{q} \exp(\ln y^q).$$

2.2.2 Propriétés des fonctions convexes

Dans le cadre de ce cours, à l'aide du résultat suivant, toute fonctions convexes est continue. Plus précisément, on a

Théorème 2.4 (Continuité). Toute fonction convexe propre sur un espace de dimension finie est continue sur l'intérieure de son domaine.

Voici maintenant la différentiabilité des fonctions convexes au sens de Gâteaux. Nous supposons que J est propre.

Définition 2.3. On dit que J est Gâteaux-différentiable en $u \in \mathbb{R}^n$ si la dérivée directionnelle $J'(u; v) = \lim_{t \rightarrow 0} \frac{J(u + tv) - J(u)}{t}$ existe dans toutes les directions $v \in \mathbb{R}^n$ et l'application $v \mapsto J'(u; v)$ est linéaire continue.

Il s'en suit que si J est différentiable au sens classique (i.e. Fréchet-différentiable) alors J est Gâteaux-différentiable. Attention, la réciproque est fautive et voici un contre exemple : la fonction

$$f(x, y) = \begin{cases} y & \text{if } x = y^2 \\ 0 & \text{sinon} \end{cases}$$

est Gâteaux-différentiable en $(0, 0)$ mais pas différentiable au sens classique.

Dans la suite on présente quelques critères pratiques pour déterminer si une fonctionnelle est convexe ou non.

Proposition 2.2 (Caractérisation du premier ordre). *Soit J une fonction Gâteau différentiable de C (convexe) dans $\mathbb{R}^n$. Alors J est convexe si et seulement si*

$$\forall (u, v) \in C^2, \quad J(v) \geq J(u) + (\nabla J(u), v - u).$$

C'est l'analogie de la courbe au dessus du graphe de J en dimension 1.

Démonstration. Supposons J convexe. Soient $(u, v) \in C^2$. J étant convexe on a :

$$\forall t \in (0, 1) J(tu + t(v - u)) \leq t(J(v) - J(u)).$$

En divisant par $t > 0$ et par passage à la limite en 0^+ , on trouve le résultat escompté.

Réciproquement, en appliquant successivement $u + t(v - u)$ pour $t \in [0, 1]$ et u , puis $u + t(v - u)$ et v , on obtient :

$$\begin{aligned} J(u) &\geq J(u + t(v - u)) - t(\nabla J(u + t(v - u)), v - u), \\ J(v) &\geq J(u + t(v - u)) + (1 - t)(\nabla J(u + t(v - u)), v - u). \end{aligned}$$

Enfin en sommant ces deux inégalités, on en déduit

$$(1 - t)J(u) + tJ(v) \geq J(u + t(v - u)).$$

□

propo1

Proposition 2.3 (Caractérisation du premier ordre). *Soit $J : C \rightarrow \mathbb{R}^n$ Gâteau différentiable et C convexe. Alors J est convexe si et seulement si*

$$\forall (u, v) \in C^2, (\nabla J(u) - \nabla J(v), u - v) \geq 0.$$

On dit que ∇J est monotone.

Le résultat précédent reste vrai si on remplace convexe par strictement convexe et l'inégalité par une inégalité stricte.

Pour terminer ce paragraphe, on définit aussi la dérivée (Gâteau) seconde de J en u comme étant la dérivée de la fonction (vectorielle) $u \rightarrow \nabla J(u)$ qu'on note $D^2 J(u)$. Par abus de langage, on l'appelle encore Hessien. En particulier, nous avons la caractérisation suivante des fonctions convexes à l'aide du Hessien.

propo2

Proposition 2.4 (Caractérisation du second ordre). *Soit $J \in C^2(\mathbb{R}^n, \mathbb{R})$. Alors*

- *J est convexe si et seulement si $D^2 J$ est symétrique et semi définie positive en tout point de $\mathbb{R}^n$.*
- *Si $D^2 J$ est symétrique et définie positive en tout point de $\mathbb{R}^n$ alors J est strictement convexe*

Remarque 2.3. *Attention, la réciproque du deuxième point dans la proposition est fausse. En particulier, la fonction $x \in \mathbb{R} \rightarrow x^4$ est strictement convexe mais $f''(x) \geq 0$ et $f''(0) = 0$.*

En fait à l'aide de ces précieux rappels, on va pouvoir obtenir des résultats d'unicité aux problèmes d'optimisation, donner des caractérisations de ces solutions sous formes de conditions nécessaires et/ou suffisantes, enfin donner des algorithmes.

quad

Exercice 2.3. *On considère la fonctionnelle quadratique $J(x) = \frac{1}{2}Ax \cdot x - b \cdot x$ avec A une matrice de taille n et b un vecteur de taille n .*

- Montrer que $J'(x) = \frac{1}{2}(A + A^t)x - b$. En déduire les conditions sur A pour que J soit convexe et strictement convexe.
- Calculer la dérivée seconde de J . En déduire des conditions sur A pour que J soit convexe et strictement convexe.

Correction de l'exercice 2.3. Soit $y \in \mathbb{R}^n$, le calcul de la quantité $J(x + ty) - J(x)$ donne

$$J(x + ty) - J(x) = \frac{t}{2}(Ay \cdot x + Ax \cdot y - 2b \cdot y) + t^2 Ay \cdot y = \frac{t}{2}((A + A^t)x \cdot y + Ax \cdot y - 2b \cdot y).$$

Donc en divisant par t et par passage à la limite, on trouve

$$J'(x) = \frac{1}{2}(A + A^t)x - b.$$

Utilisons maintenant la seconde caractérisation du premier ordre et montrons que la dérivée première est monotone, à savoir

$$(J'(x) - J'(y), x - y) \geq 0$$

On a

$$J'(x) - J'(y) = \frac{A + A^t}{2}(x - y).$$

Par conséquent $(J'(x) - J'(y), x - y) \geq 0$ si A est symétrique et semi-définie positive. Qui plus est si A est définie positive alors, par exemple soit $\lambda > 0$ la plus petite des valeurs propres. Posons $\alpha = \lambda$ et on a

$$(Ax, x) \geq \alpha \|x\|^2.$$

En appliquant ceci au résultat précédent, on en déduit la stricte convexité. Pour cela nous avons du utiliser quelques résultats de l'analyse matricielle.

Rappels 2.1.

- Une matrice A est dite semi définie positive si la forme quadratique $(Ax, x) \geq 0, \quad \forall x \in \mathbb{R}^n$.
- Une matrice A est dite définie positive si de plus elle satisfait $(Ax, x) = 0 \Rightarrow x = 0$ pour $x \in \mathbb{R}^n$.
- Les valeurs propres des matrices définies positive sont strictement positive (puisque si x est un vecteur propre associé à la valeur propre λ alors $(Ax, x) = \lambda(x, x) \Rightarrow \lambda > 0$).

Calculons maintenant la dérivée seconde. Comme précédemment il suffit de calculer $J'(x + ty) - J'(x)$. On trouve

$$J'(x + ty) - J'(x) = tAy$$

d'où

$$J''(x) = A.$$

Pour conclure, il suffit donc d'appliquer la caractérisation du second ordre.

2.3 Résultats d'unicité

Grâce à l'aide de la notion de fonctions convexes et de la notion d'ellipticité ci dessous, on va donner un critère d'unicité.

Définition 2.4. On dit que $J : \mathbb{R}^n \rightarrow \mathbb{R}$ est elliptique si il existe $\alpha > 0$ tel que

$$\exists \alpha > 0; \quad \forall (x, y) \in \mathbb{R}^{2n} \quad (\nabla J(x) - \nabla J(y), x - y) \geq \alpha \|x - y\|^2 \quad (2) \quad \boxed{\text{ellip}}$$

Le premier résultat est le suivant :

theo5

Théorème 2.5. Soit $J : \mathbb{R}^n \rightarrow \mathbb{R}$ strictement convexe. Alors le problème $\mathcal{P}$ admet une unique solution.

Démonstration. Soient $x_1 \neq x_2$ réalisant le minimum de J , i.e. $J(x_1) = J(x_2) = m$. J étant strictement convexe, on a alors

$$J\left(\frac{x_1 + x_2}{2}\right) < \frac{J(x_1) + J(x_2)}{2} = m.$$

Donc par contradiction $x_1 = x_2$. □

La stricte convexité peut être obtenue via la notion d'ellipticité : Plus précisément, on a :

theo4

Théorème 2.6. Soit J une fonction de classe $C^1(\mathbb{R}^n, \mathbb{R})$ α -elliptique. Alors J est strictement convexe et coercitive.

Il s'en suit que dans ce cadre là, le problème $\mathcal{P}$ admet aussi une unique solution.

Démonstration. D'après ce qui précède, la condition ^{ellip}(2) implique que ∇J est monotone et donc que J est strictement convexe. Pour obtenir la coercivité, on procède comme suit. En écrivant la formule de Taylor avec reste intégrale, il vient que

$$J(y) = J(x) + \int_0^1 (\nabla J(x + t(y-x)), y-x) dt$$

se qui s'écrit aussi

$$J(y) = J(x) + (\nabla J(x), y-x) + \int_0^1 (\nabla J(x + t(y-x)) - \nabla J(x), y-x) dt.$$

J étant α -elliptique, on obtient :

$$J(y) \geq J(x) + (\nabla J(x), y-x) + \alpha \|x-y\|^2 \int_0^1 t dt = J(x) + (\nabla J(x), y-x) + \frac{\alpha \|x-y\|^2}{2}$$

d'où

$$J(y) \geq J(x) - \|\nabla J(x)\| \|y-x\| + \frac{\alpha \|x-y\|^2}{2}.$$

Donc J est coercitive (en fixant $x = 0$ par exemple). Par suite, J admet un unique minimum x^* sur $\mathbb{R}^n$. De plus si on applique l'inégalité précédente avec $x = x^*$ et $y \neq x^*$ quelconque, comme

$J(y) - J(x^*) \geq -\|\nabla J(x^*)\| \|y-x^*\| + \frac{\alpha \|x^*-y\|^2}{2}$. α étant strictement positive, il s'en suit que le minimum x^* est caractérisé par $\nabla J(x^*) = 0$. □

regress

Exercice 2.4. On considère le problème de la régression polynomiale au sens des moindres carrés. Soit n un entier strictement positif et $(x_i, y_i)_{1 \leq i \leq N}$ un nuage de points. On cherche un polynôme p_n de degré inférieur ou égal à n qui l'approche au sens suivant :

$$\text{on cherche } a \text{ tel que le problème suivant soit } \sum_{i=1}^N |y_i - p_n(x_i)|^2 \text{ minimale}$$

où $a \in \mathbb{R}^{n+1}$ est le coefficient du polynôme p_n .

- Écrire ce problème sous la forme $\mathcal{P}$ (on peut aussi raisonner avec $N = n$ dans toute la suite).
- Montrer que J est continue.
- Montrer qu'il existe une fonction $g : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto g(x) = \alpha x^2 - 2\beta x + \gamma$ où α, β, γ dépendent de $V^t V$. En déduire que J est coercive.
- Montrer que J est strictement convexe.
- Conclure.

Correction de l'exercice 2.4.

- Comme on cherche un polynôme p_n de degré inférieur ou égal à n qui approche le nuage de point, on écrit :

$$p_n(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n.$$

Le problème devient alors

$$\min_{a \in \mathbb{R}^{n+1}} J(a) \text{ avec } J(a) = \|Va - y\|_2^2$$

où

$$V = \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^n \\ 1 & x_2 & x_2^2 & \dots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_N & x_N^2 & \dots & x_N^n \end{pmatrix}$$

est la matrice de Vandermonde de taille $N \times n$,

$$a = (a_0, \dots, a_n)^t \text{ et } y = (y_1, \dots, y_N)^t.$$

- J est clairement continue comme composé de fonctions continues. Montrons que J est coercitive. On a

$$J(a) = \|Va\|_2^2 - 2(Va, y) + \|y\|_2^2 \geq \alpha \|a\|_2^2 - 2\beta \|a\|_2^2 + \|y\|_2^2$$

où α est la plus petite valeur propre de la matrice définie positive $V^t V$ (il suffit qu'au moins $n+1$ valeurs de x_i soient distinctes) et β la plus grande. Notons $g(x) = \alpha x^2 - \beta x + \gamma$ avec $\gamma = \|y\|_2^2$, on a $J(a) \geq g(\|a\|)$ avec $\lim_{x \rightarrow +\infty} g(x) = +\infty$. Donc J est coercitive. On peut d'ores et déjà affirmer que le problème

$$\min\{J(a); a \in \mathbb{R}^{n+1}\}$$

admet au moins une solution (c.f. théorème [theo3](#)).

- Montrons que J est strictement convexe. Pour cela on utilise la proposition [propo1](#), i.e. la caractérisation du second ordre. Calculons donc la dérivée seconde. La fonction J est clairement de classe $C^2(\mathbb{R}^n, \mathbb{R})$ et sa dérivée première est donnée par

$$J'(a) = 2V^t(Va - y).$$

En effet, étant donné $b \in \mathbb{R}^{n+1}$, on a

$$J(a + tb) - J(a) = 2t(Va - y, Vb) + t^2 \|Vb\|_2^2.$$

Donc, en divisant par t et par passage à la limite, on trouve bien ce résultat. De même, en écrivant $J'(a + tb) - J'(a)$, on trouve

$$\lim_{t \rightarrow 0^+} \frac{J'(a + tb) - J'(a)}{t} = 2V^t V.$$

Donc dès que $n + 1$ x_i sont distincts, la matrice $V^t V$ est définie positive. On en déduit donc que J est strictement convexe.

– Par le théorème 2.5, le problème

$$\min\{J(a); a \in \mathbb{R}^n\}$$

admet une unique solution $a^* \in \mathbb{R}^{n+1}$.

Nous savons désormais dans quelle cadre le problème d'optimisation est réalisable. Il s'agit dans la suite de donner des caractérisations en terme de conditions nécessaires et/ou suffisantes pour pouvoir calculer la (ou les) solution(s).

Dans une première partie nous allons nous consacrer à l'étude où $X = \mathbb{R}^n$, i.e. aux problèmes d'optimisations sans contraintes générale.

2.4 Conditions d'optimalité

Ces conditions seront liées à la différentiabilité de la fonctionnelle J . On considère donc que dans toute cette partie J est suffisamment régulière (au sens de Gâteaux). Il existe principalement deux type de conditions nécessaires, celles du premier et second ordre, i.e. qui concerne la première et seconde dérivée de J .

2.4.1 Conditions d'optimalité du premier ordre

Cette condition dite d'Euler s'énonce sous la forme suivante (et a déjà été rencontré dans la preuve du théorème 2.6).

Théorème 2.7 (CN du premier ordre). Soit $J : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonctionnelle Gâteaux différentiable sur $\mathbb{R}^n$. Si x^* réalise un minimum (local ou global) de J sur $\mathbb{R}^n$ alors

$$\nabla J(x^*) = 0.$$

Démonstration. Soit x^* un tel minimum. Alors pour tout $x \in B(x^*, r)$ (dans la boule ouverte de centre x^* et de rayon $r > 0$), on a

$$J(x^*) \leq J(x).$$

Prenons x de la forme $x = x^* + tu \in B(x^*, r)$ avec $0 < t < \frac{r}{\|u\|}$. Puisque J est Gâteaux différentiable,

$$\lim_{t \rightarrow 0^+} \frac{J(x^* + tu) - J(x^*)}{t} = (\nabla J(x^*), u) \geq 0 \quad \forall u \in \mathbb{R}^n.$$

u étant arbitraire, on a donc $\nabla J(x^*) = 0$. □

Exercice 2.5. Écrire la condition du premier ordre pour $J(x) = \frac{1}{2}(Ax, x) - (b, x)$ où A est une matrice symétrique définie positive.

Définition 2.5. Un point $x^* \in \mathbb{R}^n$ vérifiant l'équation dite d'Euler $\nabla J(x^*) = 0$ est appelé point d'équilibre ou stationnaire.

Cependant, attention la réciproque du théorème est en générale fausse. Ce dernier permet seulement de sélectionner les candidats opportuns qui peuvent être un minimum, un maximum ou ni l'un ni l'autre (comme dans le cas de la dimension 1, un point d'inflexion, par exemple $x \mapsto x^3$).

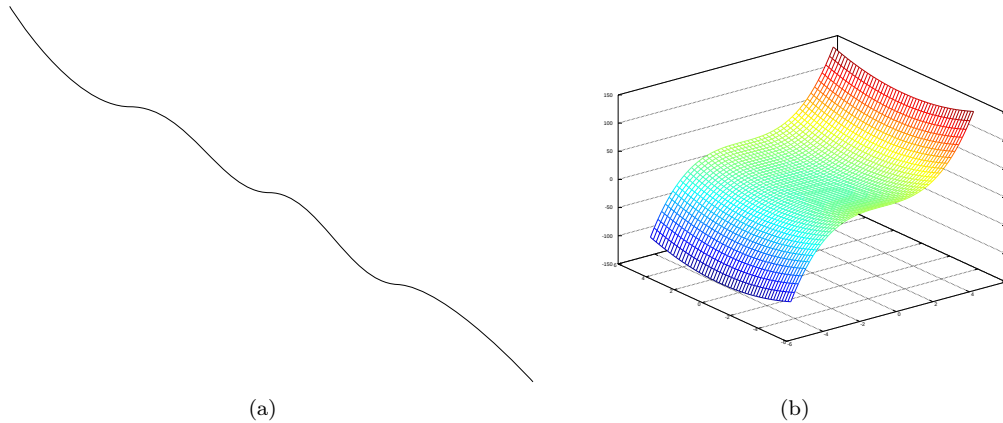


FIGURE 13 – Points critiques non nécessairement optimaux.

On peut néanmoins obtenir une condition nécessaire et suffisante dans le cadre de la convexité. En effet,

CNS1a

Théorème 2.8 (CNS du premier ordre cas convexe). *Soit $J : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonctionnelle Gâteau différentiable sur $\mathbb{R}^n$. On suppose J convexe. Alors x^* réalise le minimum de J sur $\mathbb{R}^n$ si et seulement si x^* est un point d'équilibre.*

Démonstration. D'après le théorème précédent, cette condition est nécessaire. Montrons que si de plus J est convexe, cette condition est aussi suffisante. Soit $x^* \in \mathbb{R}^n$ un point d'équilibre. Par convexité de J (c.f. proposition 2.3), il vient que

$$\forall x \in \mathbb{R}^n, \quad J(x) \geq J(x^*) + (\nabla J(x^*), x - x^*) = J(x^*)$$

x^* réalise un minimum de J sur $\mathbb{R}^n$. □

Exercice 2.6. On considère la fonctionnelle $J(x) = \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2$ dans $\mathbb{R}^2$. Étudier ce problème de minimisation.

Par conséquent, en allégeant les hypothèses, à savoir en demandant la convexité local (autour de x^*), on obtient un minimum local.

Corollaire 2.1. *Soit $J : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonctionnelle Gâteau différentiable sur $\mathbb{R}^n$. Soit $x^* \in \mathbb{R}^n$ et on suppose J convexe dans un voisinage $\mathcal{V}(x^*)$ de x^* . Alors x^* réalise le minimum local de J sur $\mathbb{R}^n$ si et seulement si x^* est un point d'équilibre.*

Les conditions du premier ordre autour sont obtenues principalement grâce à la convexité de J . Cependant, les problèmes dans la nature ne sont pas toujours décrit par une fonctionnelle convexe. En demandant un peu plus de régularité à la fonctionnelle, on peut aussi obtenir des conditions nécessaires dites du second ordre.

2.4.2 Conditions d'optimalité du second ordre

Commençons par énoncer le théorème suivant :

Théorème 2.9. *Supposons que J est deux fois différentiable sur $\mathbb{R}^n$ et x^* un minimum local de J . Alors*

1. x^* est un point d'équilibre
2. $D^2J(x^*)$ est semi-définie positive, i.e.

$$\forall x \in \mathbb{R}^n, \quad (D^2J(x^*)x, x) \geq 0.$$

Démonstration. Le premier point a déjà été vérifié. Vérifions le point deux. Considérons le développement de Taylor à l'ordre deux de la fonction $\varphi : t \mapsto J(x^* + tx)$, $x \in \mathbb{R}^n$. Compte tenu des hypothèses : x^* est un minimum local et $\nabla J(x^*) = 0$, il vient que :

$$0 \leq J(x^* + tx) - J(x^*) = \frac{t^2}{2}(D^2J(x^*)x, x) + o(t^2).$$

Donc en divisant par t et par passage à la limite, on obtient le résultat escompté. $\square$

Remarque 2.1.

- Rappelons qu'une matrice est semi-définie positive si et seulement si toutes ses valeurs propres sont positives ou nulles.
- La réciproque du théorème est fautive (contre exemple $x \mapsto x^3$) mais comme dans le cas des conditions du premier ordre, il permet de sélectionner les candidats potentiels.

En supposant des conditions supplémentaires, on peut obtenir une condition suffisante :

theo6

Théorème 2.10. *Supposons que J est deux fois différentiable sur $\mathbb{R}^n$ et $x^* \in \mathbb{R}^n$ tel que :*

1. $\nabla J(x^*) = 0$,
2. $\exists \alpha > 0, \forall x \in \mathbb{R}^n \quad (D^2J(x^*)x, x) \geq \alpha \|x\|^2$.

Alors x^ réalise un minimum local strict de J .*

Démonstration. Il suffit de considérer le développement de Taylor à l'ordre deux de la fonction φ vu précédemment. $\square$

Remarque 2.2. *La condition d'ellipticité est obtenue en pratique en prenant α comme étant la plus petite des valeurs propres de la matrice Hessienne.*

Exercice 2.7. Soit $J(x, y) = 3x^4 - 4x^2y + y^2$. Étudier le problème de minimisation sur $\mathbb{R}^2$.

Correction de l'exercice 2.7. On voit immédiatement que $(0, 0)$ est un point critique de J . Pour vérifier si il réalise le minimum, on utilise le théorème précédent. Notamment en calculant la matrice Hessienne :

$$D^2J(x, y) = \begin{pmatrix} 12x^2 - 4y & -4x \\ -4x & 2 \end{pmatrix}$$

Donc, en $(0, 0)$, on a

$$D^2J(0, 0) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}.$$

Les valeurs propres sont donc 0 et 2. La matrice est donc semi définie positive et on ne peut donc pas conclure de cette manière. Remarquez que $J(0, 1) = 1 > J(0, 0) = 0 > J(1/2, 1/2) = -0.0625$. Donc ce point critique est ni un minimum ni un maximum. Par ailleurs, on peut montrer que $(0, 0)$ est le seul point critique de J .

Exercice 2.8. Considérons de nouveau l'exemple de la fonctionnelle quadratique

$$J(x) = \frac{1}{2}(Ax, x) - (b, x)$$

dans $\mathbb{R}^n$. Montrer que le problème $\mathcal{P}$ est bien posé.

Correction de l'exercice 2.8. Les points critiques de J sont les solutions du système $Ax = b$. Si de plus A (i.e D^2J) est définie positif alors il existe un unique x^* (globale).

2.4.3 Quelques exercices supplémentaires

A présent, voici une liste d'exercices pour mettre en pratique cette partie de cours.

Exercice 2.9. Soit $J : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par

$$J(x, y) = x^3 - 3x^2 + 3x + y^3 + 3xy - 3y.$$

- Trouver les points critiques de la fonction J .
- Déterminer la nature des points critiques de J .
- La fonction admet elle un minimum ou maximum global sur $\mathbb{R}^2$?

Correction de l'exercice 2.9.

- On doit résoudre le système $\nabla J(x, y) = 0$, i.e.

$$\begin{cases} \partial_x J(x, y) = 3x^2 - 6x + 3 + 3y = 0 \\ \partial_y J(x, y) = 3y^2 + 3x - 3 = 0 \end{cases} \implies \begin{cases} (x-1)^2 + y = 0 \\ y^2 + x - 1 = 0 \end{cases}$$

De la première équation, on en déduit que $x = 1 \pm y$. En injectant dans la seconde on trouve,

$$(x, y) \in \{(0, -1), (1, 0)\}.$$

- Pour déterminer la nature de ces points, on calcule la Hessienne, à savoir,

$$D^2J(x, y) = \begin{pmatrix} 6x - 6 & 3 \\ 3 & 6y \end{pmatrix}.$$

En particulier, on a

$$D^2J(1, 0) = \begin{pmatrix} 0 & 3 \\ 3 & 0 \end{pmatrix}.$$

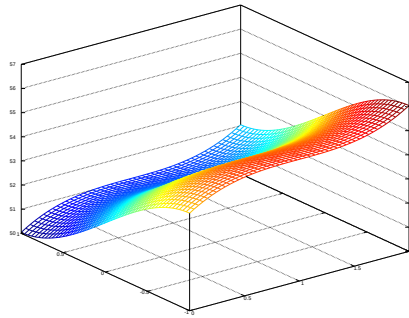
Les valeurs propres étant $\lambda = \pm 3$, la matrice n'est ni définie positive ni définie négative. C'est en fait un point selle (ou point col).

Au point $(0, -1)$, la matrice

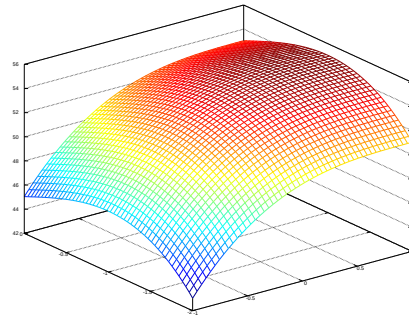
$$D^2J(0, -1) = \begin{pmatrix} -6 & 3 \\ 3 & -6 \end{pmatrix}.$$

admet comme valeurs propres $\lambda = -1$ et $\lambda = -3$. On a donc un maximum local strict (c.f. théorème 2.10) en ce point.

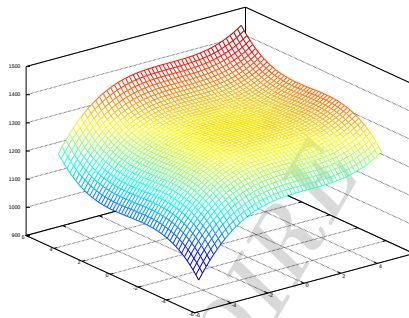
- Sur la droite $y = 0$, par exemple, $J(x, 0) = x^3 - 3x^2 + 3x$ est tel que $\lim_{x \rightarrow \pm\infty} J(x, 0) = \pm\infty$. Donc il ne peut exister de minimum ou maximum globale.



(a) En $(1, 0)$



(b) En $(0, -1)$



(c) Aperçu globale

FIGURE 14 – Points critiques de la fonction $J(x, y) = x^3 - 3x^2 + 3x + y^3 + 3xy - 3y$.

Exercice 2.10. Soit $J : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par

$$J(x, y) = x^4 + y^4 + 2x^2y^2 - x^2 - y^2.$$

- Trouver les points critiques de la fonction J et représenter les graphiquement.
- Déterminer la nature des points critiques de J (indication : on peut éventuellement poser $X = x^2 + y^2$).

Correction de l'exercice 2.10.

- L'équation $\nabla J(x, y) = 0$ fournit le système suivant

$$\begin{cases} \partial_x J(x, y) = 4x^3 + 4xy^2 - 2x = 0 \\ \partial_y J(x, y) = 4y^3 + 4yx^2 - 2y = 0 \end{cases} \implies (0, 0) \text{ ou } \begin{cases} 2x^2 + 2y^2 - 1 = 0 \\ 2y^3 + 2x^2 - 1 = 0 \end{cases}$$

Soient les points critiques suivants :

$$\nabla J(x, y) = 0 \implies (x, y) \in \left\{ (0, 0) \text{ ou } x^2 + y^2 = \frac{1}{2} \right\}.$$

La matrice Hessienne en $(0, 0)$ est

$$D^2 J(0, 0) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$$

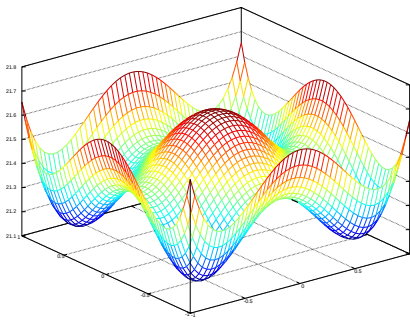
et par conséquent est définie négative. On a donc un maximum local strict en ce point. Sur le cercle de centre $(0,0)$ de rayon $\frac{1}{\sqrt{2}}$, on a

$$D^2 J(\{x^2 + y^2 = 1/2\}) = \begin{pmatrix} 12x^2 + 4y^2 - 2 & 8xy \\ 8xy & 12y^2 + 4x^2 - 2 \end{pmatrix}$$

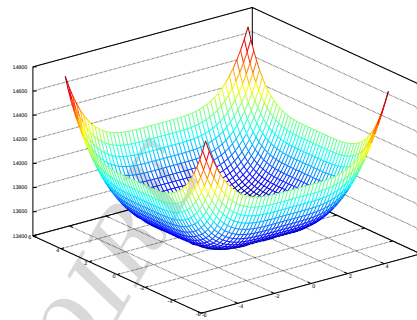
soit

$$D^2 J(\{x^2 + y^2 = 1/2\}) = \begin{pmatrix} 8x^2 & 8xy \\ 8xy & 8y^2 \end{pmatrix}$$

et a pour valeurs propres 0 et $8(x^2 + y^2) = 4$. La matrice est donc semi-définie positive...



(a) Autour des points critiques



(b) Aperçu globale

FIGURE 15 – Points critiques de la fonction $J(x, y) = x^4 + y^4 + 2x^2y^2 - x^2 - y^2$.

Exercice 2.11. Déterminer tous les points critiques de la fonction

$$J(x, y) = (x^2 + y^2 - 1)^2 + y^2$$

puis donner leurs natures.

Correction de l'exercice 2.11.

– L'équation $\nabla J(x, y) = 0$ fournit le système suivant

$$\begin{cases} 4x(x^2 + y^2 - 1) = 0, \\ 2y(2x^2 + 2y^2 - 1) = 0. \end{cases}$$

la première équation est donc nulle si et seulement si $x = 0$ ou $x^2 + y^2 = 1$. Par conséquent, on en déduit $\partial_y J(0, y) = 2y(y^2 - 1)$, soient donc les points critiques $(0, 0)$, $(0, \pm 1/\sqrt{2})$. Soit maintenant (x, y) tel que $x^2 + y^2 = 1$ alors $\partial_y J(x, y) = 2y = 0$ implique deux autres points critiques $(0, \pm 1)$. On a donc finalement que

$$\nabla J(x, y) = 0 \implies (x, y) \in \left\{ (0, 0), (0, \pm 1/\sqrt{2}), (\pm 1, 0) \right\}.$$

– La matrice hessienne est

$$D^2 J(\{x^2 + y^2 = 1/2\}) = \begin{pmatrix} 4(3x^2 + y^{-1}) & 8xy \\ 8xy & 2(2x^2 + 6y^2 - 1) \end{pmatrix}.$$

Remarquez que $J(-x, -y) = J(x, y)$, $J(-x, y) = J(x, y)$ et $J(x, -y) = J(x, y)$.
Aux points $(0, 0)$, on a

$$D^2 J = \begin{pmatrix} -4 & 0 \\ 0 & -2 \end{pmatrix}.$$

et on a donc un maximum locale.
Aux points $(0, \pm 1/\sqrt{2})$, on a

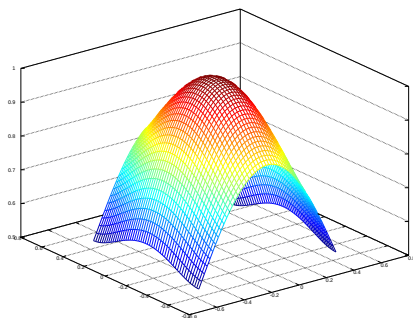
$$D^2 J = \begin{pmatrix} -2 & 0 \\ 0 & 4 \end{pmatrix}.$$

et on a donc un point col.
Aux points $(\pm 1, 0)$, on a

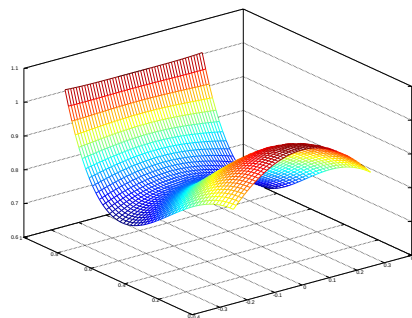
$$D^2 J = \begin{pmatrix} 8 & 0 \\ 0 & 2 \end{pmatrix}.$$

et on a donc un minimum locale.

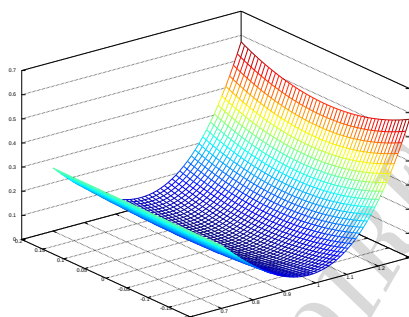
- Pour tout (x, y) , $J(x, y) \geq 0$ et on a $J(\pm 1, 0) = 0$ donc $(\pm 1, 0)$ sont des minima globaux.
Comme $\lim_{x \rightarrow +\infty} J(x, 0) = +\infty$, les maxima sont donc locaux.



(a) Autour de $(0,0)$



(b) Autour du point col $(0, 1/\sqrt{2})$



(c) Autour de $(1,0)$

FIGURE 16 – Points critiques de la fonction $J(x, y) = (x^2 + y^2 - 1)^2 + y^2$.

2.4.4 Applications à quelques exemples

Je propose ici trois applications pour clore ce chapitre du cours avant d'embrayer sur les aspects numériques. Ces exemples sont de nouveaux proposés sous formes d'exercices et peuvent être très bien traités en travaux pratiques.

Dans ce premier exemple, nous allons traiter un exemple important (déjà vu dans le cas général ?? CF LIEN) de la régression linéaire.

Exercice 2.12. On considère un nuage de points de $\mathbb{R}^2$ qu'on note $(x_i, y_i)_{1 \leq i \leq n}$. Ces données représentent des mesures physiques et il s'agit de décrire le comportement "moyen" de ce nuage. La droite obtenue s'appelle la droite de régression linéaire et constitue donc un modèle de prédiction. Il s'agit de minimiser la fonctionnelle

$$J(a, b) = \sum_{i=1}^n (y_i - ax_i - b)^2.$$

- Calculer le gradient de J et résoudre l'équation d'Euler.
- Calculer la matrice Hessienne et déterminer la nature du(des) point(s) critique(s).

Correction de l'exercice 2.12.

– on a

$$\partial_a J(a, b) = \sum_{i=1}^n 2x_i(a_x i + b - y_i) = 2a \sum_{i=1}^n x_i^2 + 2b \sum_{i=1}^n x_i - 2 \sum_{i=1}^n x_i y_i$$

et

$$\partial_b J(a, b) = \sum_{i=1}^n 2(a_x i + b - y_i) = 2a \sum_{i=1}^n x_i + 2nb - 2 \sum_{i=1}^n y_i.$$

Si on note

$$s_x = \sum_{i=1}^n x_i, \quad s_y = \sum_{i=1}^n y_i, \quad s_{xy} = \sum_{i=1}^n x_i y_i, \quad \text{et } s_{x^2} = \sum_{i=1}^n x_i^2$$

alors l'équation d'Euler s'écrit :

$$\begin{cases} s_{x^2}a + s_x b &= s_{xy} \\ s_x a + nb &= s_y \end{cases}.$$

Il existe une unique solution à ce système si et seulement si

$$d = (s_x)^2 - ns_{x^2} \neq 0$$

et la solution est alors

$$a = \frac{s_x s_y - ns_{xy}}{d} \quad \text{et} \quad b = \frac{s_{xy} s_x - s_y s_{x^2}}{d}.$$

Le couple (a, b) est un donc un point critique sous la condition $d \neq 0$.

– La matrice Hessienne de J en $(a, b) \in \mathbb{R}^2$ est donnée par

$$D^2 J(a, b) = \begin{pmatrix} s_{x^2} & s_x \\ s_x & n \end{pmatrix}.$$

Cette matrice est toujours définie positive sous la condition $d \neq 0$. Le déterminant de cette matrice correspond exactement à la quantité d . Par conséquent si $d \neq 0$ alors toutes les valeurs propres sont non nulles, i.e. la hessienne est définie positive. Donc il existe un minimum strict local de J .



FIGURE 17 – Droite de régression linéaire.

FAIRE SOUS FORMES D'UN TD/TP (LE TD ETANT ICI ET LE TP APRES LA MÉTHODE NUMERIQUE)

Le second exemple est celui de la fonctionnelle quadratique que nous avons déjà rencontré. C'est aussi un exemple classique en optimisation, notamment en optimisation numérique comme nous le verrons plus en détails dans le chapitre suivant. En générale, les propriétés d'une fonctionnelle quadratique (sous certaines conditions comme nous l'avons déjà vu ?? CF LIEN), s'approprient pour formuler un problème d'optimisation bien posé. Qui plus est, on essaye toujours en pratique de se ramener à de telles fonctionnelles. Aussi, je propose, une application numérique que nous traiterons aussi en TP.

Exercice 2.13. On veut résoudre numériquement le problème suivant :

$$\inf J(x) = \frac{1}{2}(Ax, x) - (b, x); x \in \mathbb{R}^3$$

où

$$A = \begin{pmatrix} -1 & 4 & 3 \\ -3 & 6 & 3 \\ -1 & 0 & 7 \end{pmatrix}. \quad \text{et} \quad b = \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}.$$

- Est que J est convexe, strictement convexe, coercive ?
- Existe-t-il un unique minimum ? Si oui trouvez le.

Correction de l'exercice 2.13. Les valeurs propres de la matrice A sont

$$\lambda = 6, 2, 4$$

et la matrice A est inversible. L'équation d'Euler $Ax = b$ admet alors une unique solution

$$x = -\frac{1}{3}(4, 1, 1)^t.$$

Il faut donc maintenant vérifier la nature de $D^2J(x) = A$ pour pouvoir conclure. Les valeurs propres étant toutes positives, on a un minimum local stricte.

Troisième partie

Aspects numériques

Dans cette partie, nous allons nous intéresser à la résolution numérique d'un problème d'optimisation, i.e. de manière approchée. Nous présenterons une liste d'algorithmes classiques pour lesquelles nous étudierons les critères de convergence et à quelle vitesse : c'est ce que nous appellerons ordre de convergence. Ces algorithmes sont basés sur les critères d'optimalité que nous avons vu dans partie précédente. Nous nous limiterons donc aux d'une fonctionnelle différentiable bien qu'il existe des méthodes qui traitent le cas non différentiable mais dépasse de loin les objectifs de ce cours.

1 Premières définitions

Construire un algorithme correspond à la construction d'une suite $(x_k)_{k \in \mathbb{N}}$ comme suit :

Définition 1.1. Soit $\mathcal{A}$ une application de $\mathbb{R}^n \rightarrow \mathbb{R}^n$. Un algorithme défini par cette application est une suite d'éléments de $\mathbb{R}^n$ construit par la suite :

$$\begin{cases} x_0 \in \mathbb{R}^n \text{ donné} & \text{Initialisation} \\ x_{k+1} = \mathcal{A}(x_k) \in \mathbb{R}^n \end{cases}$$

Un tel algorithme peut ne pas converger. Il est donc important de savoir sous quelles conditions un tel algorithme converge (au sens où $\lim_{k \rightarrow \infty} x_k$ existe), dans le cas contraire il ne sera d'aucune utilité pratique. On a donc pour définition

Définition 1.2. Un algorithme défini par l'application $\mathcal{A}$ converge si et seulement si la suite x_k admet une limite lorsque k tend vers l'infini.

La convergence d'un algorithme est un problème difficile et clairement important mais ce n'est pas le seul critère à garder en tête : il y a aussi le coût de calcul et la rapidité de convergence. La rapidité de convergence, qu'on appelle aussi taux de convergence, se mesure comme suit :

Définition 1.3. Soit $(x_k)_k$ une suite définie par une application $\mathcal{A}$ convergeant de limite x^* . On dit que la convergence

- est linéaire si l'erreur $e_k = \|x_k - x^*\|$ décroît linéairement :

$$\exists K \in (0, 1), \exists k_0, \forall k \geq k_0 \quad e_{k+1} \leq K e_k.$$

- est super-linéaire si l'erreur $e_k = \|x_k - x^*\|$ décroît comme :

$$e_{k+1} \leq \alpha_k e_k$$

où α_k est une suite positive convergente vers 0. Si la suite est géométrique, on dit que la convergence est géométrique.

- est d'ordre p si l'erreur $e_k = \|x_k - x^*\|$ décroît comme suit :

$$\exists K \in (0, 1), \exists k_0, \forall k \geq k_0 \quad e_{k+1} \leq K(e_k)^p.$$

Lorsque $p = 2$, on parle de convergence quadratique.

On dit que la convergence est locale si elle n'a lieu que dans un voisinage de x_0 , sinon on dit qu'elle est globale.

Il est évident que plus l'ordre est élevé et plus l'algorithme requiert moins de calculs pour aboutir au résultat.

2 Méthode du gradient

La méthode de gradient est une méthode dit de descente pour la raison suivante : on se donne une fonctionnelle J à minimiser et un point de départ x_0 . Partant de x_0 , on veut construire x_1 de manière à s'approcher du minimum de J : on veut donc que

$$J(x_1) < J(x_0).$$

On cherche alors x_1 sous la forme $x_1 = x_0 + h_1 d_1$ où d_1 est un vecteur non nul de $\mathbb{R}^n$ et $h_1 > 0$. On cherche donc (h_1, d_1) de telle manière

$$J(x_0 + h_1 d_1) < J(x_0).$$

La quantité d_1 peut ne pas exister. Lorsque ce dernier existe on dit que d_1 est une direction de descente. Quand à h_1 c'est le pas de descente. Une méthode de descente est donc construite à partir de

$$\begin{cases} x_0 \in \mathbb{R}^n \text{ donné} \\ x_{k+1} = x_k + h_k d_k, d_k \in \mathbb{R}^n \setminus \{0\}, h_k > 0 \end{cases}$$

avec

$$J(x_k + h_k d_k) < J(x_k).$$

L'idée naturelle est d'écrire un développement de Taylor aux points x_{k+1} :

$$J(x_{k+1}) = J(x_k + h_k d_k) = J(x_k) + h_k (\nabla J(x_k), d_k) + \|h_k d_k\| \varepsilon(h_k d_k)$$

et donc de prendre comme première approche $d_k = -\nabla J(x_k)$: c'est la méthode dite du gradient. Le pas h_k est choisi soit comme une constante soit comme une variable.

L'algorithme du gradient est donc :

lgo_gradient

Algorithm 1 Algorithme du gradient

Require: $\varepsilon > 0, k = 0, x_0 \in \mathbb{R}^n, h_0 > 0$

```
1: while  $\|x_{k+1} - x_k\| \geq \varepsilon$  do
2:    $x_{k+1} = x_k - h_k \nabla J(x_k)$ 
3: end while
4: return  $x_{k+1}$ 
```

Remarque 2.1. La condition $\|x_{k+1} - x_k\|$ est un critère d'arrêt. A une précision ε donné, la suite converge vers une solution approchée à ε près. On peut remplacer naturellement cette condition par $\|\nabla J(x_k)\|$ (on fait alors appelle à l'équation d'Euler vu précédemment).

L'avantage de cette algorithme est qu'il est simple à écrire mais en pratique il converge sous des hypothèses très forte et la convergence est généralement lente. En particulier, nous avons le résultat suivant

Théorème 2.1. Soit J une fonctionnelle de classe $C^1(\mathbb{R}^n, \mathbb{R})$ coercive et strictement convexe. On suppose de plus que le gradient est Lipschitz :

$$\exists K > 0; \forall (x, y) \in \mathbb{R} \times \mathbb{R}^n \quad \|\nabla J(x) - \nabla J(y)\| \leq K \|x - y\|.$$

Alors la méthode converge vers le minimum de J si $h_k \in [\alpha, \beta]$ avec $0 < \alpha < \beta < \frac{2}{K}$.

Démonstration. □

En générale, on utilise le cas de l'algorithme simplement avec un pas $h_k \equiv h_0$ constant. Néanmoins, on peut quand même utilisé des pas variables. Notamment la méthode du gradient à pas optimal. L'algorithme est le suivant

pas_optimale

Algorithm 2 Algorithme du gradient à pas optimal

Require: $\varepsilon > 0, k = 0, x_0 \in \mathbb{R}^n, h_0 > 0$

```
1: while  $\|x_{k+1} - x_k\| \geq \varepsilon$  do
2:    $x_{k+1} = x_k - h_k \nabla J(x_k)$  où  $h_k$  réalise le minimum de  $q(h) = J(x_k - h \nabla J(x_k))$ .
3: end while
4: return  $x_{k+1}$ 
```

t_pas_optimale_wolfe

Algorithm 3 Recherche linéaire de Wolfe

```

1:  $q_0 = q(0)$  et  $q'_0 = q'(0)$ ,  $t = h_k$ 
2:  $h_g = 0$  et  $h_d$  grand
3:  $T = 0$ 
4: while  $q(t) \geq q_0 + m_1 t q'_0$  ou  $q'(t) \leq m_2 q'_0$  do
5:   if  $q(t) \geq q_0 + m_1 t q'_0$  then
6:      $T = 1$ ,  $h_d = t$  et  $t = \frac{h_g + h_d}{2}$ 
7:   else
8:     if  $T \equiv 0$  then
9:        $h_d = 2h_d$ ,  $h_g = t$  et  $t = \frac{h_g + h_d}{2}$ 
10:    else
11:       $h_g = t$  et  $t = \frac{h_g + h_d}{2}$ 
12:    end if
13:  end if
14: end while
15:  $h_k = t$ 
16: return  $h_k$ 

```

Exemples 2.1. Bien que les conditions du théorème soit un peu forte, comme nous l'avons vu à multiples reprises, le problème, en générale, s'écrit sous la forme

$$J(x) = \frac{1}{2}(Ax, x) - (b, x).$$

$$\|\nabla J(x) - \nabla J(y)\| = \|A(x - y)\| \leq \|B\| \|x - y\|$$

Cependant, en pratique on ne calcule pas le rayon spectrale (car cela peut être coûteux). On initialise l'algorithme en testant différentes valeurs de h_0 . Dans le cas de non convergence, la solution soit explose soit oscille.

3 Méthode de Newton

En principe, une méthode de Newton n'est pas une méthode de minimisation mais de recherche de zéro d'une fonction, i.e. trouver

$$x^*; F(x^*) = 0.$$

Rappelons le principe dans un premier temps en dimension 1. On recherche le(s) zéro(s) d'une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$.

_newton_dim1

Algorithm 4 Algorithme de Newton en dimension 1

Require: $x_0 \in \mathbb{R}$

1: **while** $\|x_{k+1} - x_k\| \geq \varepsilon$ **do**

2: $x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$

3: **end while**

4: **return** x_{k+1}

en s'assurant que pour tout k $f'(x_k) \neq 0$.

Cette méthode porte aussi le nom de la méthode de la tangente. x_{k+1} est défini à partir de x_k en traçant la courbe de f au point $(x_k, f(x_k))$ et en prenant l'intersection avec l'axe des abscisses.

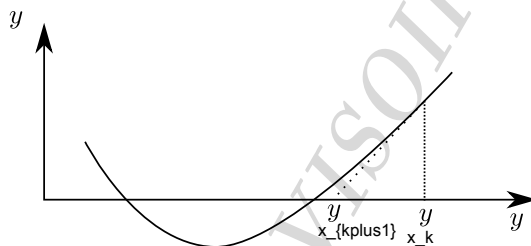


FIGURE 18 – Interprétation géométrique de la méthode de Newton

La généralisation aux cas n dimensionnel se fait sans peine. Soit $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ une fonction de classe C^1 . On suppose que l'équation

$$F(x) = 0$$

admet au moins une solution notée x^* tel que $DF(x^*)$ soit inversible.

algo_newton

Algorithm 5 Algorithme de Newton

Require: $x_0 \in \mathbb{R}$

1: **while** $\|x_{k+1} - x_k\| \geq \varepsilon$ **do**

2: $x_{k+1} = x_k - \delta_k$ avec $DF(x_k)\delta_k = F(x_k)$

3: **end while**

4: **return** x_{k+1}

Le taux de convergence et critères de convergences sont résumés dans le théorème suivant :

Théorème 3.1. Soit F une fonction de classe $C^2(\mathbb{R}^n, \mathbb{R}^n)$ admettant au moins une racine isolée x^* tel que $DF(x^*) \neq 0$. Alors

$$\exists r > 0, \forall x_0 \in B(x^*, r) \quad (x_k)_{k \in \mathbb{N}} \subset B(x^*, r) \text{ et } \lim_{k \rightarrow \infty} x_k = x^*.$$

De plus,

$$\exists \alpha > 0; \forall k \geq 0 \quad \|x_k - x^*\| \leq \alpha^k \|x_0 - x^*\|.$$

en outre la convergence est géométrique et quadratique

Remarque 3.1. La méthode de Newton est d'ordre 2. Cela signifie que si $e_k = \|x_{k+1} - x_k\|$ est de l'ordre de 10^{-2} à l'itération k , alors elle vaut 10^{-4} à $k+1$, puis 10^{-8} à $k+2$. Autrement quelques itérations suffisent pour que l'algorithme converge.

La convergence géométrique ici signifie que si x_0 est suffisamment proche de x^* (qu'on ne connaît pas en pratique!!!) alors la méthode converge. Ce n'est pas très grave : en pratique, nous utiliserons quelques itérations de la méthode du gradient à pas constant pour initialiser la valeur de x_0 .

Cependant l'inconvénient dans ce type d'approche c'est la résolution du système linéaire à chaque itération qui peut être coûteuse. L'idée c'est d'introduire des méthodes qui approchent l'inverse de la différentielle, ce sont les méthodes quasi-Newton que nous ne traiterons pas ici. Il existe aussi de nombreuses autres techniques.

Maintenant il s'agit de voir comment intégrer la méthode de Newton à la recherche d'extremum. En fait, la condition d'Euler, i.e. l'équation

$$\nabla J(x^*) = 0$$

est une équation de la forme

$$F(x^*) = 0.$$

Par conséquent, rechercher des extremums de J , revient à résoudre le système d'équations $\nabla J(x^*) = 0$. Il suffit donc de poser $F = \nabla J$ dans l'algorithme ^{algorithme 5.} Cette méthode permet donc de déterminer les racines qui sont potentiellement des minimums. *A posteriori*, il faudra ensuite vérifier que ce sont bien des minima.

Exercice 3.1.

Exercice 3.2 (Travaux pratiques 1). Programmer les méthodes de ce chapitre. Dans chacun des cas on tracera les fonctions et on fera une étude de la sensibilité du point de départ. On comparera les vitesses de convergence en fonction du nombre d'itérations mais aussi le temps CPU. On affichera également la trajectoire de x_k . Pour ce faire, nous les utiliserons sur les fonctions

1. $J(x, y) = x^3 - 3x^2 + 3x + y^3 + 3xy - 3y$
2. $J(x, y) = x^4 + y^4 + 2x^2y^2 - x^2 - y^2$
3. $J(x, y) = (x^2 + y^2 - 1) + y^2$
4. $J(x, y) = (1 - x)^2 + \alpha(y - x^2)^2$ pour $\alpha = 0.1, 1, 10$ et 100 .
5. $J(x, y) = \frac{1}{2}(Ax, x) - (b, x)$ avec $A = (a_{ij})$ où $a_{ij} = 2$ si $i = j$, $a_{ij} = -1$ si $i = j - 1$ et $i = j + 1$, et 0 ailleurs.

Exercice 3.3 (Travaux pratiques 2). 1. Écrire la démarche à suivre pour résoudre le problème $Ax = b$ en utilisant une méthode de descente

2. On s'intéresse à la résolution de l'équation de Laplace

$$-u''(x) = f(x) \quad x \in [0, 1]$$

avec les conditions aux bords de Dirichlet $u(0) = u(1) = 0$.

- (a) En utilisant la méthode des différences finies, montrer que ce problème s'écrit sous la forme $A_N U_n = f_N$ où $A = (a_{ij})$ est une matrice telle que $a_{ij} = 2/h^2$ si $i = j$, $a_{ij} = -1/h^2$ si $i = j-1$ et $i = j+1$, et 0 ailleurs ($h = 1/(N+1)$ le pas de discrétisation et N le nombre de points utilisés dans la méthode des différences finies).
- (b) En utilisant la question 1. appliquer la méthode du gradient à pas constant, avec recherche linéaire de Wolf et la méthode de Newton

Exercice 3.4 (TP 3 Identification de paramètres). *Rendre un rapport*

Exercice 3.5 (TP 3 Brachis). *Rendre un rapport*

VERSION PROVISOIRE