République Algérienne Démocratique et Populaire

Ministère de l'enseignement supérieur et de la recherche scientifique

N° Attribué par la bibliothèque

Année univ.: 2018/2019

Optimisation sans contrainte

Mémoire présenté en vue de l'obtention du diplôme de

Master Académique

Université de Saida - Dr Moulay Tahar

Discipline : MATHEMATIQUES

Spécialité : Géométrie différentielle

par

Samiha Hachemaoui [1]

Sous la direction de

Mme Bekkouche Nouria

Soutenue le 17/07/2019 devant le jury composé de

S. Ouakkas	Université Dr Tahar Moulay - Saida	Président
N. Bekkouche	Université Dr Tahar Moulay - Saida	Encadreur
F.Z Mostefai	Université Dr Tahar Moulay - Saida	Examinateur
A. Halimi	Université Dr Tahar Moulay - Saida	Examinateur

1. e-mail : samouha1224@gmail.com

Table des matières

Introduction

L'optimisation est une branche des mathématiques cherchant à modéliser, à analyser et à résoudre analytiquement ou numériquement les problèmes qui consistent à minimiser ou maximiser une fonction sur un ensemble.

L'optimisation joue un rôle important en recherche opérationnelle (domaine à la frontière entre l'informatique, les mathématiques et l'économie), dans les mathématiques appliquées (fondamentales pour l'industrie et l'ingènierie), en analyse et en analyse numérique, en statistique pour l'estimation du maximum de vraisemblance d'une distribution, pour la recherche de stratégies dans le cadre de la théorie des jeux, ou encore en théorie du contrôle et de la commande.

Les premiers problèmes d'optimisation auraient été formulés par Euclide, au $IIIe$ siècle avant notre ère, dans son ouvrage historique éléments. Trois cent ans plus tard, Héron d'Alexandrie dans Catoptrica énonce le principe du plus court chemin dans le contexte de l'optique.

Au $XVIIe$ siècle, l'apparition du calcul déffirentiel entraîne l'invention de techniques d'optimisation, ou du moins en fait ressentir la nécessité. Newton met au point une méthode itérative permettant de trouver les extrémums locaux d'une fonction en faisant intervenir la notion de dérivée, issue de ses travaux avec Leibniz. Cette nouvelle notion permet de grandes avancées dans l'optimisation de fonctions car le problème est ramené à la recherche des racines de la dérivée.

Durant le $XVIIIe$ siècle, les travaux des mathématiciens Euler et Lagrange ménent au calcul des variations, une branche de l'analyse fonctionnelle regroupant plusieurs méthodes d'optimisation. Ce dernier invente une technique d'optimisation sous contraintes : Les multiplicateurs de Lagrange.

Le $XIXe$ siècle est marqué par l'intérêt croissant des économistes pour les mathématiques. Ceux-ci mettent en place des modèles économiques qu'il convient d'optimiser, ce qui accélère le développement des mathématiques. Depuis cette pèriode, l'optimisation est devenue un pilier des mathématiques appliquées et le foisonnement des techniques est tel qu'il ne saurait être résumé en quelques lignes.
Historiquement, le premier terme introduit fut celui de programmation linéaire, inventé par George Dantzig vers 1947.
Le terme programmation dans ce contexte ne réfère pas à la programmation informatique (bien que les ordinateurs soient largement utilisés de nos jours pour résoudre des programmes mathématiques). Il vient de l'usage du mot programme par les forces armées américaines pour établir des horaires de formation et des choix logistiques, que Dantzig étudiait à l'époque.
Dans ce mémoire on va introduire des outils mathématiques pour résoudre un problème d'optimisation dont la forme abstraite générale est la suivante :
Soit V un espace vectoriel, $\Omega \subset V$ et une application $J : V \to \mathbb{R}$, on veut résoudre le problème

$$(P) \qquad\qquad J(u^*) = \inf_{u\in\Omega} J(u)$$

On dit que (P) est un *problème d'optimisation* ou *problème de minimisation* de la fonction J sans *contrainte.*
La fonction J est appelée *fonction coût, fonction objectif ou fonction* é*conomique.*
En remplacsans J par $-J$ on transforme un problème de maximisation en problème de minimisation.

Définition 0.0.1. *Un point u^* est un minimum local de J sur Ω s'il existe un voisinage $\mathcal{V}$ de u^* tel que $J(u^*) \leq J(u)$ pour tout $u \in \Omega \cap \mathcal{V}$.*
Un point u^* est un minimum local strict de J sur Ω s'il existe un voisinage $\mathcal{V}$ de u^* tel que $J(u^*) < J(u)$ pour tout $u \in \Omega \cap \mathcal{V},\ u \neq u^*$.

Définition 0.0.2. *Une suite $(u_n)_{n\in\mathbb{N}}$ de points de Ω est une suite minimisante si $\lim\limits_{n\to+\infty} J(u_n) = \inf\limits_{u\in K} J(u)$. Une telle suite existe toujours, par définition de* inf*, par contre on ne sait rien au sujet de sa convergence éventuelle.*

Pour résoudre le problème (P), il faut se poser les questions suivantes :

• Est-ce que $\inf_{u \in K} J(u)$ existe ? c.à.d. est-ce que J est bornée inférieurement ?

• Est-ce que l'infimum est atteint dans Ω ? c.à.d. est-ce qu'il existe $u^\star \in \Omega$ vérifiant $J(u^*) = \min_{u\in K} J(u)$?

• Est-ce que $u^\star$ est unique ? Sinon, quelle est la taille de l'ensemble des solutions ?

• Est-ce que 1'on peut caractériser $u^\star$? c.à.d. peut-on trouver des conditions nécessaires pour caractériser un minimum : $\langle$*si* u *vériffie* (P), *alors* u *vériffie la propriété* $N(u)\rangle$ et / ou trouver des conditions suffisantes pour être un point optimal :
$\langle$*si* u *vériffie la propriété* $S(u)$, *alors* u *vériffie* $(P)\rangle$.

• Trouver un algorithme d'optimisation pour déterminer la, resp. les, solutions de (P).

Pour répondre à ces questions on est en particulier amené à étudier

-la structure de V : espace vectoriel, muni d'une norme, d'un produit scalaire, de dimension finie ou infinie

- les propriétés de $\Omega \subset V$: fermé, borné, convexe, . . .

- les propriétés de $J : V \to \mathbb{R}$: continuité, différentiabilité, convexité, . . .

Dans ce memoire on va donner des éléments de réponse à ces questions.

Après avoir rappelé (théorème(1.3.2)) la condition nécessaire d'extremum relatif, appelée équation d'Euler,

$$\nabla J(u) = 0$$

pour une fonction J définie sur un ensemble ouvert, on examine comment cette condition doit être modifiée lorsqu'on considère des fonctions définies sur des ensembles non ouverts particuliers. De façon plus précise, on s'intéresse au problème suivant : trouver des conditions nécessaires, et des conditions suffisantes, pour qu'un point soit un extremum relatif d'une fonction J définie sur l'espace tout entierun.

On montre ensuite comment la prise en compte des dérivées secondes permet, d'une part, de donner une deuxième condition nécessaire d'extremum et, d'autre part, de donner des conditions suffisantes, dans le cas de fonctions J définies sur un ouvert Ω d'un espace vectoriel norme V : si un point u de Ω est un minimum relatif, alors (théorème(1.3.3))

$$\nabla^2 J(u)(w,w) \geq 0 \text{ pour tout } w \in V$$

Inversement, s'il existe un α tel que

$$\alpha > 0, \quad \nabla^2 J(u)(w,w) \geq \alpha \|w\|^2 \text{ pour tout } w \in V,$$

ou bien s'il existe un voisinage B de u tel que

$$\nabla^2 J(v)(w,w) \geq 0 \text{ pour tout } v \in B, w \in V,$$

alors le point u est un minimum relatif (théorème (2.2.3)).

Il est clair que la condition nécessaire $\nabla J(u) = 0$ d'extremum relatif en un point u d'un ouvert est grossièrement fausse en général (considérer par exemple la fonction $J(v) = v$ sur un intervalle $[a,b] \subset \mathbb{R}$). Un premier exemple de situation où on peut la généraliser convenablement correspond à des ensembles U définis par des "contraintes" :$\varphi_i(v) = 0, 1 \leq i \leq m$). c'est le cas, déjà signalé, correspondant aux multiplicateurs de Lagrange (mémoire de l'éttudiant).

Un deuxième exemple est celui où les ensembles U sont convexes : si une fonction J admet un minimum relatif par rapport à un ensemble convexe U, alors (théorème (2.3.1))

$$\nabla J(u)(v-u) \geq 0 \text{ pour tout } v \in U$$

ces conditions, appelées inéquations d'Euler, étant également suffisantes pour l'existence d'un minimum si la fonction J est elle-même convexe (théorème (2.3.4)).

On notera que la prise en compte des dérivées secondes, comme la prise en compte de la convexité, permet de distinguer entre minimums et maximums, d'une part (et même parfois de préciser si ce sont, par exemple, des minimums stricts, ou encore des minimums sur l'ensemble tout entier), et d'énoncer des conditions suffisantes, d'autre part.

Afin de donner une plus grande généralité aux résultats de ce mémoire, et parce que les démonstrations sont exactement les mêmes, nous avons délibérément "abandonné" $\mathbb{R}^n$" pour nous placer dans des espaces vectoriels normes quelconques, ce qui nous conduit notamment à utiliser à plusieurs reprises la notion d'espace de Banach, c'est-à-dire d'espace vectoriel norme complet.

Chapitre 1

Notion sur le Calculs différentiels et analyse convexe

1.1 Espace vectoriels normés

Soit V un espace vectoriel sur le corps $\mathbb{R}$.

Définition 1.1.1. *Un produit scalaire sur* V, *si* $< .,. >: V \times V \to \mathbb{R}$ *est une application bilinéaire, symétrique, et définie positive, c'est-à-dire qui vérifie :*
$< u,. >: V \to \mathbb{R}$ *est linéaire pour tout* $u \in V$,
$< .,v >: V \to \mathbb{R}$ *est linéaire pour tout* $v \in V$,
$< u,v >=< v,u >$ *pour tout* $u,v \in V$,
$< v,v >= 0 \Leftrightarrow v = 0$,*et* $< v,v >\geq 0$ *pour tout* $v \in V$.

Définition 1.1.2. *L'application* $u \mapsto \|u\|$ *de* V *dans* $\mathbb{R}_+$*définie par*

$$\|v\| = \sqrt{(v,v)} \quad \text{pour tout} \quad v \in V,$$

une norme sur l'espace V telle que

Définition 1.1.3. (i) $\forall u \in V \ : \ \|u\| = 0 \Leftrightarrow u = O_E$
$(ii) \forall u \in V; \forall \lambda \in \mathbb{R} : \|\lambda u\| = |\lambda| \|u\|$
(iii) $\forall (u,\ v) \in V \times V : \|u + v\| \leq \|u\| + \|v\|$

Exemple 1.1.1. (1). *Soit* $V = \mathbb{R}^d$, *pour* $u = (u_1, \ldots, u_d) \in \mathbb{R}^d$ *on définit les normes :*

$$\|u\|_1 = \sum_{i=1}^{d} |u_i|, \ \|u\|_2 = (\sum_{i=1}^{d} |u_i|^2)^{1/2}, \ \|u\|_\infty = \max_{1\leq i\leq d} |u_i|$$

On a les inégalités classiques suivantes :

$$\forall u \in \mathbb{R}^d : \|u\|_\infty \leq \|u\|_2 \leq \|u\|_1 \ \ Vt \ \ \|u\|_1 \leq \sqrt{d}\|u\|_2 \leq d\|u\|_\infty.$$

(2). *Soit* $V = C([a, b], \mathbb{R})$, *l'espace vectoriel, de dimension infinie, des fonctions réelles continues sur l'intervalle* $[a, b]$. *Pour* $f \in V$ *on définit les normes :*

$$\|f\|_1 = \int_a^b |f(t)|dt, \ \|f\|_2 = (\int_a^b |f(t)|^2 dt)^{1/2}, \ \|f\|_\infty = \max_{[a,b]} |f(t)|$$

On a les inégalités :

$$\forall f \in C([a, b], \mathbb{R}) : \|f\|_1 \leq \sqrt{b-a}\|f\|_2 \leq (b-a)\|f\|_\infty.$$

Dans ce cas on n'a pas les inégalités inverses.

Remarque 1.1.1. *Les espaces* $(V, \|.\|_1)$ *et* $(V, \|.\|_2)$ *ne sont pas complets, tandis que* $(V, \|.\|_\infty)$ *est complet.*

Proposition 1.1.1. *Les applications*

$$\begin{array}{ccc} V\times V \to V; & \mathbb{R}\times V \to V; & V \to \mathbb{R}_+ \\ (u,v) \mapsto u+v & (\lambda,u) \mapsto \lambda u & u \mapsto \|u\| \end{array}$$

sont continues.

Proposition 1.1.2. *Soient* $(V, \|.\|_E)$ *et* $(F, \|.\|_F)$ *des e.v.n,* $\Phi : V \to F$ *une application linéaire, alors*

$$\Phi \text{ est continue} \Leftrightarrow \exists c > 0, \ \ \forall u \in V : \|\Phi(u)\|_F \leq c\|u\|_V.$$

Remarque 1.1.2. *Soient* $(V, \|.\|_V)$ *et* $(F, \|.\|_F)$ *des e.v.n. On note* $\mathcal{L}(V, F)$ *l'espace vectoriel des applications linéaires continues de* V *dans* F.

Proposition 1.1.3. *Pour $f \in \mathcal{L}(V,\ F)$ on pose $\|f\| = \sup_{\|u\|_E \leq 1} \|f(u)\|_F$. Alors :*
(i) $f \mapsto \|f\|$ est une norme sur $\mathcal{L}(V,\ F)$;
(ii) $\forall x \in E$: $\|f(u)\|_F \leq \|f\|\|u\|_E$;
(iii) $\|f\| = \sup_{\|u\|_E \leq 1} \|f(u)\|_F = \sup_{\|u\|_V = 1} \|f(u)\|_F = \sup_{u \in V, u \neq O_V} \frac{\|f(u)\|_F}{\|u\|_V}$

Propriété 1.1.1. *Application : Si $\dim V = n$ et $\dim F = m$ alors $J : V \to F$ linéaire est représenté par une matrice A de m lignes et n colonnes, on peut alors définir une norme de matrice subordonnée aux normes vectorielles par*

$$\|A\| = \sup_{\|u\|_E = 1} \|Au\|_F = \sup_{u \in E, u \neq O_V} \frac{\|Au||_F}{||u||_V}$$

Proposition 1.1.4. *Soit $(V, \|.\|_V)$ un e.v.n. On note $V' = \mathcal{L}(V, \mathbb{R})$ l'e.v.n. des formes linéaires continues sur V, alors : Une forme lineaire $J \in V'$ si et seulement si $\ker J$ est fermé dans V.*

On désigne par $\mathcal{L}(V; F)$ ou simplement $\mathcal{L}(V)$ si $V = F$, l'espace vectoriel formé par les applications linéaires continues de V dans F. C'est un espace vectoriel norme par

$$\|A\| = \sup_{\substack{u \in V \\ u \neq 0}} \frac{\|Au\|_F}{\|u\|_V}$$

et c'est un espace vectoriel complet si l'espace F est complet. Dans le cas où $V = F = \mathbb{R}^n$, la norme définie ci-dessus n'est pas autre chose qu'une norme matricielle subordonnée (à une norme vectorielle sur $\mathbb{R}^n$). Lorsque $F = \mathbb{R}$, on écrit généralement

$$\mathcal{L}(V; \mathbb{R}) = V',$$

et on appelle l'espace V' le dual de V.

De la même manière, on note $\mathcal{L}_2(V; F)$ l'espace des applications bilinéaires continues de $V \times V$ dans F, c'est-à-dire qui vérifient

$$\begin{aligned} B(\alpha_1 u_1 + \alpha_2 u_2, u) &= \alpha_1 B(u_1, u) + \alpha_2 B(u_2, u), \\ B(u, \alpha_1 u_1 + \alpha_2 u_2) &= \alpha_1 B(u, u_1) + \alpha_2 B(u, u_2), \end{aligned}$$

pour chaque $u, u_1, u_2 \in V$ et pour tout $\alpha_1, \alpha_2 \in \mathbb{R}$, et

$$\|B\|_{\mathcal{L}_2(V;F)} = \sup_{\substack{u_1,u_2\in V\\ u_1\neq 0, u_2\neq 0}} \frac{\|B(u_1,u_2)\|_F}{\|u_1\|_V \|u_2\|_V} < +\infty$$

C'est un espace vectoriel norme par l'application $\|.\|_{\mathcal{L}_2(V;F)}$ définie ci-dessus.

Remarque 1.1.3. *Soit* $(V, \|.\|)$ *un e.v. normé de dimension finie* d*, soit* $\{e_1, \ldots, e_d\}$ *une base de* V*.*
Alors application $\Phi : (\mathbb{R}^d, \|.\|_\infty) \to (V, \|.\|)$ *est une bijection continue.*

$$v = (v_1, \ldots, v_d) \mapsto \sum_{i=1}^{d} v_i e_i$$

De plus Φ^{-1} *est aussi continue,* Φ *est t donc une isomorphisme de* $\mathbb{R}^d$ *sur* V*.*

Note : Grâce à l'isomorphisme Φ on obtient de nombreux corollaires qui facilitent la manipulation des e.v. de dimension finie.

Corollaire 1.1.1. *Sur un e.v.* V *de dimension finie toutes les normes sont équivalentes :*

$$\exists c_1,\ c_2 \in \mathbb{R}_+^*,\ \forall u \in V : c_1\|u\|' \leq \|u\| \leq c_2\|u\|'$$

1.2 Espaces de Hilbert

On appelle espace préhilbertien un espace vectoriel muni d'un produit scalaire. S'il est complet pour cette norme, c'est un espace de Hilbert.
Tout espace vectoriel norme de dimension finie étant complet, l'espace $\mathbb{R}^n$ muni du produit scalaire euclidien est un exemple d'espace de Hilbert. Notons au passage l'inégalité de Schwarz :

$$|< u, v >| \leq \|u\| \, \|v\| \quad \text{pour tout} \quad u, v \in V,$$

qui sert notamment à démontrer l'inégalité triangulaire pour la norme associée au produit scalaire. L'inégalité de Cauchy-Schwarz pour le produit scalaire euclidien ou l'inégalité de Cauchy-Schwarz pour les fonctions :

$$\left|\int_I uvdx\right| \leq \left(\int_I |u|^2\, dx\right)^{\frac{1}{2}} \left(\int_I |v|^2\, dx\right)^{\frac{1}{2}}, \quad I \subset \mathbb{R},$$

en sont des cas particuliers. On remarquera que l'inégalité de Schwarz entraîne la continuité du produit scalaire, considéré comme application du produit $V \times V$ dans $\mathbb{R}^n$.

Enfin on rappelle que cette inégalité devient une égalité si, et seulement si, les deux vecteurs qui y figurent sont linéairement dépendants.

1.3 Calcul différentiel

la présente partie est consacré aux principaux définitions et résultats fondamentaux du calcul différentiel dans les espaces vectoriels normes (dérivabilité d'une fonction composée, théorème des accroissements finis, théorème des fonctions implicites, formules de Taylor). On s'est volontairement limité aux dérivées premières et secondes, puisque nous n'aurons pas l'usage de dérivées d'ordre supérieur.

1.3.1 Dérivées première et seconde d'une application

Dans tout ce paragraphe, l'écriture $f : \Omega \subset X \to Y$ signifiera systématiquement que f est une application d'un ouvert Ω d'un espace vectoriel norme X dans un espace vectoriel norme Y.

Définition 1.3.1. *On dit qu'une application $f : \Omega \subset X \to Y$ est dérivable en un point $a \in \Omega$ s'il existe un élément, noté $f'(a)$ de l'espace $\mathcal{L}(X;Y)$ tel que l'on puisse écrire*

$$f(a+h) = f(a) + f'(a)h + \|h\|\,\varepsilon(h) \qquad \lim_{h\to 0} \varepsilon(h) = 0.$$

On montre alors facilement qu'un tel élément $f'(a)$ est unique, s'il existe, et on appelle $f'(a)$ la dérivée (première) de l'application f au point a. Dans le cas où $X = Y = \mathbb{R}$, on utilise également la notation $f'(a) = \left(\dfrac{df}{dx}\right)(a)$ est également utilisée.

Remarque 1.3.1. *C'est surtout pour assurer l'unicité de la dérivée qu'il est utile de supposer que le domaine de définition de la fonction f est ouvert.*

Corollaire 1.3.1. *Une application dérivable en un point est continue en ce point. Par ailleurs, on notera que si* $f : \Omega \subset X \to Y$ *est dérivable en* $a \in \Omega$*, alors, pour tout vecteur* h *de* X*,*

$$f'(a)h = \lim_{\theta \to 0} \frac{f(a+\theta h) - f(a)}{\theta}$$

Remarque 1.3.2. *Ici, comme ailleurs, on omet de préciser qu'il faut naturellement se limiter aux valeurs de* θ *pour lesquelles les points* $a + \theta h$ *appartiennent au domaine de définition de* f *; ceci dans un souci évident d'alléger l'écriture* !

Définition 1.3.2. *On dit que l'application* $f : \Omega \subset X \to Y$ *est dérivable dans* Ω *si elle est dérivable en tout point de* Ω*. On peut alors définir une application*

$$f' : x \in \Omega \subset X \to f'(x) \in \mathcal{L}(X; Y),$$

appelée application dérivée. Si l'application dérivée $f' : \Omega \subset X \to \mathcal{L}(X; Y)$ *est continue, on dit que la fonction* $f : \Omega \subset X \to Y$ *est (une fois) continûment dérivable dans* Ω*, et on écrit, et on écrit* $f \in C^1(\Omega)$

Exemple 1.3.1. *une fonction continue affine*

$$f : x \in X \to f(x) = Cx + d \in Y$$

où $C \in \mathcal{L}(X; Y)$ *et* $d \in Y$ *est dérivable dans* X*, et*

$$f'(a) = C \text{ pour tout } a \in X,$$

puisque $f(a+h) = f(a) + Ch$*. Si* $B \in \mathcal{L}_2(X; Y)$*, l'application* f *définie par*

$$f : x \in X \to f(x) = B(x, x) \in Y,$$

est dérivable en X *et*

$$f'(a)h = B(h, a) + B(a, h),$$

puisque

$$\begin{aligned} f(a+h) &= f(a) + B(h,a) + B(a,h) + B(h,h), \\ ||B(h,a) + B(a,h)|| &\leq 2||B|| \; ||a|| \; ||h||, \\ \|B(h,h)\| &\leq \|B\| \, \|h\|^2 . \end{aligned}$$

Si l'application bilinéaire B est symétrique, c'est-à-dire si

$$B(x, y) = B(y, x) \text{ pour tout } x, y \in X,$$

la formule précédente devient

$$f'(a)h = 2B(a, h)$$

Si $Z = Z_1 \times Z_2 \times ... \times Z_p$ est un produit d'espaces vectoriels normés Z_i (l'espace Z est naturellement muni de la topologie du produit), on note

$$z = \begin{pmatrix} z_1 \\ z_2 \\ . \\ . \\ . \\ z_p \end{pmatrix}, \quad z_i \in Z_i, \quad 1 \leq i \leq p,$$

un élément quelconque de Z, cette notation étant la généralisation naturelle de la notation matricielle utilisée pour les vecteurs usuels. La donnée d'une application

$$f : \Omega \subset X \to Y = Y_1 \times Y_2 \times ... \times Y_m$$

revient à se donner m applications composantes $f_i : \Omega \subset X \to Y_i$, de telle sorte que

$$f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ . \\ . \\ . \\ f_m(x) \end{pmatrix}, \text{ pour tout } x \in \Omega$$

On établit facilement que

Proposition 1.3.1. *l'application f est dérivable en un point $a \in \Omega$ si et seulement* *si chaque application composante l'est aussi, et alors*

$$f'(a) = \begin{pmatrix} f_1'(a) \\ f_2'(a) \\ . \\ . \\ . \\ f_m'(a) \end{pmatrix}, \quad f_i'(a) \in \mathcal{L}(X;Y_i),$$

l'espace $\mathcal{L}(X;Y)$ s'identifiant de façon naturelle au produit des espaces $\mathcal{L}(X;Y_i)$.

Considérons ensuite une application

$$f : \Omega \subset X_1 \times X_2 \times ... \times X_n \to Y$$

d'un ouvert Ω d'un produit d'espaces vectoriels normés. Soit a un point de Ω, de composantes $a_1, a_2, ..., a_n$, soit $k \in \{1, 2, ..., n\}$ l'un des indices. Par définition de la topologie produit, il existe un ouvert $\Omega_k \subset X_k$ tel que tous les points de composantes $a_1, ..., a_{k-1}, x_k, a_{k+1}, ..., a_n$, appartiennent l'ouvert Ω lorsque le point x_k appartient à l'ouvert Ω_k. Par suite, on peut étudier la dérivabilité éventuelle de l'application partielle

$$f(a_1, ..., a_{k-1}, x_k, a_{k+1}, ..., a_n) : \Omega_k \subset X_k \to Y$$

Si cette application est dérivable au point $a_k \in \Omega_k$, on note

$$\partial_k f(a) \in \mathcal{L}(X_K; Y)$$

sa dérivée, appelée dérivée partielle de la fonction f au point a, par rapport à la $k^{ème}$ variable.
Si une application

$$f : \Omega \subset X_1 \times X_2 \times ... \times X_n \to Y$$

est dérivable en un point $a \in \Omega$, on établit aisément qu'elle possède des dérivées partielles par rapport à chacune des variables et que, de plus,

$$f'(a)h = \sum_{k=1}^{n} \partial_k f(a) h_k, \quad \text{pour tout } h = \begin{pmatrix} h_1 \\ h_2 \\ . \\ . \\ . \\ h_n \end{pmatrix} \in X$$

La réciproque est grossièrement inexacte : ainsi la fonction

$$f : x = (x_1, x_2) \in R^2 \to \left\{ \begin{array}{ll} 0 & \text{si } \ x_1, x_2 = 0 \\ 1 & \text{sinon} \end{array} \right\}$$

possède deux dérivées partielles au point $(0, 0)$ (car les applications partielles y sont constantes) sans être dérivable en ce point, puisqu'elle n'y est pas continue.
Supposons enfin que

$$X = X_1 \times X_2 \times ... \times X_n \quad et \quad Y = Y_1 \times Y_2 \times ... \times Y_m,$$

de sorte qu'une application $f : \Omega \subset X \to Y$ est déterminée par la donnée de m fonctions $f_i : \Omega \subset X \to Y_i$ de n variables. Alors la relation

$$k = f'(a)h, \text{ with } h = \begin{pmatrix} h_1 \\ h_2 \\ . \\ . \\ . \\ h_n \end{pmatrix} \in X \quad \text{and} \quad k = \begin{pmatrix} k_1 \\ k_2 \\ . \\ . \\ . \\ k_m \end{pmatrix} \in Y,$$

est équivaut aux relations

$$k_i = \sum_{j=1}^{n} \partial_j f_i(a) h_j, \quad 1 \leq i \leq m$$

Un cas particulier très important pour la suite est celui où
$X = \mathbb{R}^n$ et $Y = \mathbb{R}^m$. Alors, les relations précédentes s'écrivent sous la forme matricielle

$$\begin{pmatrix} k_1 \\ k_2 \\ . \\ . \\ . \\ k_m \end{pmatrix} = \begin{pmatrix} \partial_1 f_1(a) & \partial_2 f_1(a) & ... & \partial_n f_1(a) \\ \partial_1 f_2(a) & \partial_2 f_2(a) & ... & \partial_n f_2(a) \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ \partial_1 f_m(a) & \partial_2 f_m(a) & ... & \partial_n f_m(a) \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ . \\ . \\ . \\ h_n \end{pmatrix}$$

les nombres $\partial_i f_i(a)$ étant les dérivées partielles «usuelles», souvent notées $(\dfrac{\partial f_i}{\partial x_j})(a)$.
La matrice $(\partial_i f_i(a))$ représente donc l'application linéaire $f'(a) \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)$. C'est pourquoi on l'appelle, matrice dérivée de l'application f en a.
Si $m = n$, son déterminant s'appelle le jacobien de la fonction f au point a.

Remarque 1.3.3. *L'espace $\mathcal{L}(\mathbb{R})$ s'identifiant à l'espace $\mathbb{R}$, les dérivées partielles $\partial_i f(a)$ d'une application $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ peuvent être effectivement considérées comme des nombres réels.*

On notera au passage que les dérivées partielles $\partial_i f(a)$ d'une fonction $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ vérifient

$$\partial_i f(a) = f'(a) e_i,$$

e_i désignant le $i^{-ème}$ vecteur de la base canonique de $\mathbb{R}^n$.
Le résultat suivant, qui permet de calculer la dérivée d'une application composée, est constamment utilisé.

Théorème 1.3.1. *Soit $f : \Omega \subset X \to Y$ une application dérivable en un point $a \in \Omega$, et $g : \Omega' \subset Y \to Z$ une application dérivable au point $b = f(a) \in \Omega'$ Supposons que $f(\Omega) \subset \Omega'$. Alors l'application composée*

$$h = gf : \Omega \subset X \to Z$$

est est dérivable au point $a \in \Omega$ et

$$h'(a) = g'(b) f'(a)$$

Dans le cas d'applications

$$f : \Omega \subset R^n \to R, \quad g : \Omega' \subset R^m \to R^l,$$

correspond à la composition des dérivées $g'(b)$ et $f'(a)$, correspond la multiplication des matrices dérivées des applications en cause :

$$\begin{pmatrix} \partial_1 h_1(a) & \ldots & \partial_n h_1(a) \\ . & & . \\ . & & . \\ . & & . \\ \partial_1 h_l(a) & \ldots & \partial_n h_l(a) \end{pmatrix} = \begin{pmatrix} \partial_1 g_1(b) & \ldots & \partial_m g_1(b) \\ . & & . \\ . & & . \\ . & & . \\ \partial_1 g_l(b) & \ldots & \partial_m g_l(b) \end{pmatrix} \begin{pmatrix} \partial_1 f_1(a) & \ldots & \partial_n f_1(a) \\ . & & . \\ . & & . \\ . & & . \\ \partial_1 f_m(a) & \ldots & \partial_n f_m(a) \end{pmatrix}$$

ce qui peut encore s'écrire :

$$\partial_n f_m(a) = \sum_{k=1}^{m} \partial_k g_i(b) \partial_j f_k(a), \qquad 1 \leq i \leq l, \quad 1 \leq j \leq n.$$

Pour compléter ces rappels concernant les dérivées premières, nous allons énoncer deux résultats fondamentaux du calcul différentiel (théorèmes (1.3.2) et (1.3.3)). Le premier concerne l'extension de la formule des accroissements finis "usuelle" : étant donné une fonction réelle continue sur un intervalle $[a, b]$ dérivable sur l'intervalle $]a, b[$, il existe un point $c \in]a, b[$ tel que

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

Cette formule ne se généralise pas telle quelle : ainsi, l'application

$$f : t \in \mathbb{R} \to f(t) = (\cos t, \sin t) \in \mathbb{R}^2$$

est telle que $f(2\pi) - f(0) = (0, 0)$, sans que la dérivée

$$f'(t) = (-\sin t, \cos t)$$

s'annule dans l'intervalle $]0, 2\pi[$.

Si l'on ne peut généraliser la formule, on peut par contre généraliser la majoration

$$|f(b) - f(a)| \leq \sup_{t \in]a,b[} |f'(t)| \, |b - a|$$

qui en découle (le cas où $\sup_{t\in]a,b[} |f'(t)| = +\infty$ n'est pas exclu ; il ne donne d'ailleurs aucun renseignement !). Si a et b sont deux points d'un espace vectoriel X, on utilisera les notations suivantes :

$$\begin{aligned} [a,b] &= \{x = ta + (1-t)b \in X : \quad t \in [0,1]\} \\ ]a,b[&= \{x = ta + (1-t)b \in X : \quad t \in]0,1[\} \end{aligned}$$

pour désigner le segment fermé et le segment ouvert, respectivement, d'extrémités a et b.

Théorème 1.3.2. ***(théorème des accroissements finis)****. Soit $f : \Omega \subset X \to Y$ et a et b deux points de Ω, tels que le segment $[a,b]$ soit contenu dans Ω. On suppose l'application f continue en tout point du segment fermé $[a,b]$ et dérivable en tout point du segment ouvert $]a,b[$. Alors*

$$\|f(b) - f(a)\|_Y \leq \sup_{x\in]a,b[} \|f'(x)\|_{L(X;Y)} \|b-a\|_X$$

On déduit de nombreuses conséquences de ce résultat, notamment :

Lemme 1.3.1. *Si $X = X_1 \times X_2 \times ... \times X_n$ est un produit d'espaces vectoriels normes, alors l'application f est continûment dérivable dans Ω si et seulement si*
(i) Les dérivés partiels $\partial_k f(x)$, $1 \leq k \leq n$ existent en tout point $x \in \Omega$, et
(ii) les dérivées partielles

$$\partial_k f : x \in \Omega \to \partial_k f(x) \in \mathcal{L}(X_k; Y)$$

sont continues dans Ω.

A l'aide du théorème des accroissements finis (et du théorème du point fixe), on démontre également un autre résultat fondamental du calcul différentiel. Il s'agit du théorème des fonctions implicites, qui apporte une réponse au problème suivant : Soit X_1, X_2 et Y trois espaces vectoriels normés, et $\varphi : X_1 \times X_2 \to Y$ une application donnée. Étant donné un point $b \in Y$, il se peut que, pour tout élément $x_1 \in X_1$, il existe un élément $x_2 \in X_2$ et un seul tel que $\varphi(x_1, x_2) = b$, ce qui définit une application $f : x_1 \in X_1 \to x_2 \in X_2$, appelée fonction implicite, vérifiant

$$\varphi(x_1, f(x_1)) = b \quad \text{pour tout } x_1 \in X_1.$$

Naturellement, ce genre de circonstance est exceptionnel, et il est le plus souvent illusoire d'espérer démontrer l'existence d'une fonction implicite définie sur l'ensemble X_1 tout entier. Par contre, il est plus réaliste de chercher un résultat d'existence local : supposant l'existence d'une solution particulière (a_1, a_2) de l'équation $\varphi(x_1, x_2) = b$, on cherche à quelles conditions on peut définir x_2 comme une fonction implicite de x_1 dans un voisinage convenable du point (a_1, a_2). C'est l'objet du résultat qui suit, où l'on étudie également les propriétés de continuité et de dérivabilité de la fonction implicite.
On notera aussi qu'il est essentiel d'avoir une connaissance "préalable" d'une solution particulière.

Définition 1.3.3. *Si X et Y sont deux espaces vectoriels normés, on notera*

$$I_{som}(X;Y), \text{ ou simplement } I_{som}(X) \text{ si } X = Y,$$

l'ensemble des applications linéaires continues, bijectives de X sur Y, d'applications inverses continues.

Théorème 1.3.3. ***(théorème des fonctions implicites)**. Soit $\varphi : \Omega \subset X_1 \times X_2 \to Y$ une application une fois continûment dérivable dans Ω et $(a_1, a_2) \in \Omega$ des points tels que*

$$\varphi(a_1, a_2) = b, \quad \partial_2\varphi(a_1, a_2) \in I_{som}(X_2; Y).$$

Supposons l'espace X_2 complet.
Alors, il existe un ouvert $O_1 \subset X_1$, un ouvert $O_2 \subset X_2$ et une application continue, appelée fonction implicite,

$$f : O_1 \subset X_1 \to X_2,$$

tels que $(a_1, a_2) \in O_1 \times O_2 \subset \Omega$ et $(cf.\ figure\ 1)$

$$\{(x_1, x_2) \in O_1 \times O_2 : \varphi(x_1, x_2) = b\} = \{(x_1, x_2) \in O_1 \times X_2 : x_2 = f(x_1)\}$$

De plus, l'application f est dérivable au point a_1 et

$$f'(a_1) = -\left\{\partial_2\varphi(a_1, a_2)\right\}^{-1} \partial_1\varphi(a_1, a_2).$$

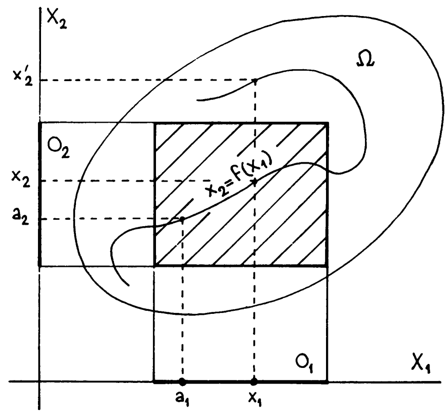


FIG. 1.1 – $figure1$:

Remarque 1.3.4. (1) *Supposant établie la dérivabilité de la fonction implicite au point a_1 la dérivée se calcule par une application du théorème* (1.3.1) *à la fonction composée*

$$h : x_1 \in O_1 \to \{\varphi(x_1, f(x_1)) - b\} = 0 \in Y.$$

Il suffit en effet d'écrire

$$0 = h'(a_1) = \partial_1\varphi(a_1, a_2) + \partial_2\varphi(a_1, a_2)f'(a_1).$$

(2) *Pour un élément donné $x_1 \in O_1$, donné, il peut naturellement exister des éléments $x_2' \in X_2$ tels que $x_2' \neq x_2, (x_1, x_2') \in \Omega$ et $\varphi(x_1, x_2') = b$, mais alors $x_2' \notin O_2$ (cf. figure 1).*

(3) *En fait, il existe un ouvert O_1' tel que $a_1 \in O_1' \subset O_1$ et tel que la fonction implicite soit dérivable en tout point de cet ouvert.*

(4) *On utilise fréquemment le corollaire suivant du théorème des fonctions implicites, appelé théorème d'inversion locale :*

Théorème 1.3.4. ***(théorème d'inversion locale)****. Soit $g : \Omega_2 \subset X_2 \to X_1$ une application une fois continûment dérivable, et $a_1 \in X_1$ des points tels que*

$$a_1 = g(a_2), \qquad g'(a_2) \in Isom(X_2; X_1)$$

On suppose l'espace X_2 complet.

Alors, il existe un ouvert $O_1 \subset X_1$, un ouvert $O_2 \subset X_2$ et une application continue $f : O_1 \subset X_1 \to X_2$, tels que $a_2 \in O_2 \subset \Omega_2$ et

$$\{(x_1, x_2) \in O_1 \times O_2 : g(x_2)\} = \{(x_1, x_2) \in O_1 \times X_2 : x_2 = f(x_1)\}$$

De plus, la fonction f est dérivable en a_1

$$f^{'}(a_1) = \{g'(a_2)\}^{-1}$$

Passons ensuite à la notion de dérivée seconde d'une application.

Définition 1.3.4. *Soit $f : \Omega \subset X \to Y$ une application dérivable dans Ω Si l'application dérivée*

$$f' : \Omega \subset X \to \mathcal{L}(X; Y)$$

est dérivable en un point $a \in \Omega$, sa dérivée, notée

$$f''(a) =^{def} (f')'(a) \in L(X; L(X; Y)),$$

est appelée la dérivée seconde de l'application f au point a, et on dit que l'application f est deux fois dérivable au point a.

Utilisant l'isomorphisme canonique entre l'espace $\mathcal{L}(X; \mathcal{L}(X; Y))$ et l'espace $\mathcal{L}_2(X; Y)$ des applications bilinéaires continues de X dans Y, on identifie la dérivée seconde à une application bilinéaire continue de X dans Y, et on écrit

$$f'(a)(h, k) = (f''(a)h)k,$$

pour tout $h, k \in X$. Grâce au théorème des accroissements finis, on démontre que la dérivée seconde est une application bilinéaire symétrique, en ce sens que

$$f''(a)(h, k) = f''(a)(k, h)$$

pour tout $h, k \in X$.
L'application f est dite deux fois dérivable dans Ω, si elle est deux fois dérivable en tout point de Ω ; il est alors possible de définir la dérivée seconde

$$f'' : \Omega \subset X \to \mathcal{L}_2(X;Y)$$

Si cette dernière fonction est continue, la fonction f est dite deux fois différentiellement en Ω, et on écrit

$$f \in C^2(\Omega).$$

Pour ce qui concerne le calcul effectif des dérivées secondes, on utilise constamment le résultat suivant, qui permet de se ramener à des calculs de dérivées premières :

Corollaire 1.3.2. *Etant donné deux vecteurs quelconques $h, k \in X$, l'élément $f''(a)(h,k) \in Y$ est égal à la dérivée au point $a \in \Omega$ de l'application $x \in \Omega \to f'(x)k \in Y$, appliquée au vecteur h.*

Exemple 1.3.2. *Considérons, par exemple, une fonction de la forme*

$$f : x \in X \to f(x) = B(x,x) + Cx + d \in Y$$

où $B \in \mathcal{L}_2(X;Y), C \in \mathcal{L}(X;Y), d \in Y$. Il a déjà été établi que cette fonction est différentiable dans X et que

$$f'(x)k = B(k,x) + B(x,k) + Ck,$$

pour chaque $x, k \in X$. Pour un vecteur fixe $k \in X$ la fonction $g : x \in X \to f'(x)k \in Y$ est ici affine ; il est donc différentiable dans X, de sorte que nous obtenons

$$f''(a)(h,k) = g'(a)h = B\,(h,k) + B(k,h),$$

appliquer le résultat donné ci-dessus.

Dans le cas particulier où $X = \mathbb{R}^n$ et $Y = \mathbb{R}$, soit $h = (h_i)$ deux vecteurs de $\mathbb{R}^n$, avec pour base (e_i). La seconde dérivée $f''(a)$ étant une fonction bilinéair

$$f''(a)(h,k) = \sum_{i,j} h_i k_j f''(a)(e_i, e_j),$$

où, par ce qui a été dit ci-dessus,

$$f''(a)(e_i, e_j) = \partial_i(\partial_j f)(a) = f''(a)(e_j, e_i).$$

D'où les chiffres

$$\partial_{ij} f(a) =^{def} \partial_i(\partial_j f)(a)$$

ne sont que les dérivées partielles «habituelles», également désignées par

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

Si $n = 1$*, la deuxième dérivée* $f''(a)$ *est également notée* $(\frac{d^2 f}{d^2 x})(a)$*, qui, dans ce cas, n'est qu'un élément de* $\mathbb{R}$*.*

Nous passons maintenant en revue diverses formules de Taylor. La première formule généralise la définition de la première dérivée d'une fonction et la seconde généralise le théorème des accroissements finis (*théorème* (1.3.2)). Quant aux troisième et quatrième formules, elles donnent une expression pour le «reste», la quatrième généralisant la formule

$$f(a+h) - f(a) = f(a) + \int_a^{a+h} f'(\theta) d\theta = \int_0^1 f'(a+th) h dt$$

qui est bien connu pour les fonctions réelles d'une variable réelle. Pour commodité, les résultats sont présentés sous la forme de deux théorèmes. On verra enfin que les formules sont données avec des hypothèses de plus en plus "fortes".

Théorème 1.3.5. ***(formules de Taylor pour les fonctions une fois dérivables)*** *Soit* $f : \Omega \subset X \to Y$ *et* $[a, a+h]$ *un segment fermé compris dans* Ω*.*
(1) *Si* f *est différentiable en* a*, alors*

$$f(a+h) = f(a) + f'(a)h + ||h||\varepsilon(h), \lim_{h \to 0} \varepsilon(h) = 0.$$

(2) ***Théorème de la valeur moyenne :*** *si* $f \in C^0(\Omega)$ *et* f *est différentiable sur* $]a, a+h[$*, alors*

$$\|f(a+h) - f(a)\| \leq \sup_{x \in]a, a+h[} \|f'(x)\| \, \|h\| \, .$$

(3) ***La formule de Taylor-Maclaurin :*** *si* $f \in C^0(\Omega)$ *,* f *est différentiable sur* $]a, a+h[$ *et* $Y = R$*, alors*

$$f(a+h) = f(a) + f'(a+\theta h)h, \quad 0 \prec \theta \prec 1$$

(4) ***Formule de Taylor avec reste intégral*** : *si $f \in C^1(\Omega)$ et Y est un espace vectoriel complet, alors*

$$f(a+h) = f(a) + \int_0^1 f'(a+th)hdt.$$

Théorème 1.3.6. ***(formules de Taylor pour les fonctions deux fois dérivables)****. Soit $f : \Omega \subset X \to Y$ et soit $[a, a+h]$ un segment fermé compris dans Ω.*

(1) ***La formule de Taylor-Young :*** *si f est dérivable dans Ω et deux fois dérivable en a, alors*

$$f(a+h) = f(a) + f'(a)h + \pm f''(a)(h,h) + \|h\|^2 \varepsilon(h) \quad \lim_{h\to 0} \varepsilon(h) = 0.$$

(2) ***Le théorème de la valeur moyenne généralisée :*** *si $f \in C^1(\Omega)$ et f est deux fois dérivable sur $]a, a+h[$, alors*

$$\|f(a+h) - f(a) - f'(a)h\| \leq \frac{1}{2} \sup_{x\in]a,a+h[} \|f''(x)\| \|h\|^2 .$$

(3) ***La formule de Taylor-Maclaurin :*** *si $f \in C^1(\Omega)$, f est deux fois différentiable sur $]a, a+h[$ et $Y = \mathbb{R}$, alors*

$$f(a+h) = f(a) + f'(a)h + \frac{1}{2} f''(a+\theta h)(h,h), \quad 0 < \theta < 1.$$

(4) ***Formule de Taylor avec reste intégral*** : *si $f \in C^2(\Omega)$ et Y est un espace complet, alors*

$$f(a+h) = f(a) + f'(a)h + \int_0^1 (1-t)\left[f''(a+th)(h,h)\right] dt.$$

Remarque 1.3.5. (1) *Alors que la formule* (1) *du* (*théorème*.1.3.5) *est exactement la définition de la dérivée première, la formule* (1) *du* (*théorème* 1.3.6) *n'est pas équivalente à la définition de la dérivabilité seconde en un point .*
(2) *Dans la formule des accroissements généralisée (2), l'expression $||f''(x)||$ désigne naturellement la norme de rélément $f''(x)(*)$ de l'espace vectoriel norme $\mathcal{L}_2(X;Y)$.*

(3) *On sait qu'il existe (au moins) un nombre* $\theta \in]0,1[$ *tel que les formules de Taylor-Maclaurin (3) soient vraies mais, en général, on n'a aucun autre renseignement sur* θ *; on rappelle au passage qu'il est indispensable de se restreindre au cas* $Y = \mathbb{R}$*.*
(4) *Pour que les formules (4) aient un sens, il faut savoir intégrer les fonctions*

$$\begin{aligned} t &\in [0,1] \to \{f'(a+th)h\} \in Y \\ t &\in [0,1] \to \{f''(a+th)(h,h)\} \in Y; \end{aligned}$$

c'est la raison pour laquelle on suppose que ces fonctions sont continues et l'espace Y *complet.*

Pour terminer, précisons quelques notations particulières aux fonctions de la forme

$$f : \Omega \subset \mathbb{R}^n \to \mathbb{R}.$$

En tout point a où cette fonction est une fois, ou deux fois, dérivable, on introduit le vecteur $\nabla f(a) \in \mathbb{R}^n$ et la matrice $\nabla^2 f(a) \in \mathcal{A}_n(\mathbb{R})$ définis respectivement par les relations

$$f'(a)h =< \bigtriangledown f(a), h > \quad \text{pour tout} \quad h \in \mathbb{R}^n$$

$$f(a)(h,k) =< \bigtriangledown^2 f(a)h, k >=>< h, \bigtriangledown^2 f(a)k > \quad \text{pour tout} \quad h,k \in \mathbb{R}^n,$$

$< .,. >$ désignant, comme d'habitude, le produit scalaire euclidien sur $\mathbb{R}^n$. Le vecteur

$$\bigtriangledown f(a) = \begin{pmatrix} \partial_1 f(a) \\ \partial_2 f(a) \\ . \\ . \\ . \\ \partial_n f(a) \end{pmatrix}$$

s'appelle le gradient de l'application f au point a, et la matrice (symétrique)

$$\nabla^2 f(a) = \begin{pmatrix} \partial_{11} f(a) & ... & \partial_{1n} f(a) \\ \partial_{21} f(a) & ... & \partial_{2n} f(a) \\ . & . & . \\ . & . & . \\ . & . & . \\ \partial_{n1} f(a) & ... & \partial_{nn} f(a) \end{pmatrix}$$

s'appelle le Hessien de l'application f au point a.

Remarque 1.3.6. *Alors que les première et seconde dérivées, $f'(a)$ et $f''(a)$ sont définies de manière "intrinsèque" dans les espaces $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$ et $\mathcal{L}_2(\mathbb{R}^n, \mathbb{R})$ respectivement (l'espace $\mathcal{L}_2(\mathbb{R}^n, \mathbb{R})$ étant identifié ici à l'espace $\mathcal{L}(\mathbb{R}^n)$), on ne doit pas oublier que le gradient et le Hessien en sont des représentations particulières, correspondant au produit scalaire euclidien : le choix d'un autre produit scalaire sur $\mathbb{R}^n$ correspondrait à un autre vecteur et à une autre matrice ! On notera également que le gradient est le vecteur transposé de la matrice dérivée (ici : un vecteur ligne) de l'application f en a.*

Pour illustrer ces considérations, voici trois façons équivalentes d'écrire (par exemple) la formule de Taylor-Young pour les fonctions $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ qui se différencient deux fois :

$$\begin{aligned} f(a+ft) &= f(a) + f'(a)h + \frac{1}{2} f''(a)(h,h) + \|h\|^2 \varepsilon(h), \\ f(a+h) &= f(a) + (\nabla f(a), h) + \frac{1}{2}(\nabla^2 f(a)h, h) + (h,h)\varepsilon(h), \\ f(a+ft) &= f(a) + (\nabla f(a))^T h + \frac{1}{2} h^T \nabla^2 f(a) h + h^T h \varepsilon(h). \end{aligned}$$

Chapitre 2

Problème d'optimisation sans contraintes en dimension finie

Etant donné un ensemble $U \subset \mathbb{R}^n$ et une fonction $J : U \to \mathbb{R}$; On appelle problème d'optimisation un problème noté :

$$\min_{u \in \mathbb{R}^n} J(u) \qquad (P)$$

$u \in \mathbb{R}^n$ s'appelle variable de décision, $J(u)$ s'appelle fonction économique ou fonction objectif.

Remarque 2.0.7. (1) *Le problème (P) est un problème d'optimisation sans contraintes non linéaire si J n'est pas affine.*
(2) *Le problème (P) est un problème d'optimisation sans contraintes non linéaire quadratique si*

$$J(u) = u^T A u - bu$$

où A est une matrice symetrique d'ordre $(n; n)$; u^T est le transposé du vecteur u.

Définition 2.0.5. ***Minimum local–global***
(1) *$u \in U$ est un point de minimum local de J sur U si, et seulement si*

$$\exists \eta > 0, \forall v \in U, \|v - u\| < \eta \Rightarrow J(u) \leq J(v)$$

(2) *$u \in U$ est un minimum global de J sur U si, et seulement si*

$$\forall v \in U, \quad J(u) \leq J(v)$$

Nous aurons également l'usage de la définition suivante :

Définition 2.0.6. *Soit $J : U \to \mathbb{R}$ une fonction définie sur un espace topologique U. On dit que la fonction J a un minimum relatif (ou un maximum relatif) au point $u \in U$ s'il existe un voisinage O de tel que*

$$J(u) \leq J(v) (\text{ou } J(u) \geq J(v)) \quad \text{pour chaque } v \in O.$$

S'il n'est pas nécessaire de distinguer entre maximum et minimum relatifs, on dit que la fonction J a un extremum relatif au point u.

Définition 2.0.7. *Soit $J : U \to \mathbb{R}$ une fonction définie sur un espace topologique U. On dit que la fonction J admet en un point $u \in U$ un minimum relatif strict (ou un maximum relatif strict) s'il existe un voisinage O de u tel que*

$$J(u) < J(v) \quad (\text{ou } J(u) > J(v)) \text{ pour tout } \; v \in O - \{u\}.$$

2.1 Existence et unicité de la solution du problème (P)

Examinons maintenant les questions d'existence et d'unicité de la solution du problème (P). En dimension finie, l'unicité d'une solution éventuelle est en général établie indépendamment de l'existence, le plus souvent à partir de la convexité de l'ensemble U et de la stricte convexité de la fonctionnelle J.

Pour ce qui concerne l'existence, si U est une partie fermée bornée de $V = \mathbb{R}^n$ et si la fonction $J : \mathbb{R}^n \to \mathbb{R}$ est continue, il est clair que le problème (P) a au moins une solution. mumimale.

L'unicité résulte en général de propriétés de convexité (de J et de W).

2.2 Optimisation Différentiable

2.2.1 Condition d'ordre 1

Equation d'Euler

Théorème 2.2.1. ***(Condition nécessaire pour un minimum relatif)*** *Soit Ω un ouvert d'un espace vectoriel norme V et $J : \Omega \subset V \to \mathbb{R}$ une fonction. Si la fonction J a un extremum relatif en un point $u \in \Omega$, et si elle est dérivable en ce point, alors $\nabla J(u) = 0$*

Preuve 2.2.1. *Soit v un vecteur quelconque de V. L'ensemble Ω étant ouvert, il existe un intervalle ouvert I contenant 0 tel que la fonction*

$$\varphi : t \in I \to \varphi(t) = J(u + tv)$$

soit bien définie. En appliquant le (théorème.1.3.2), nous trouvons que la fonction φ est dérivable pour $t = 0$, et

$$\varphi'(0) = \nabla J(u)v.$$

Pour fixer les idées, supposons que le point u soit un minimum relatif. Alors

$$0 \geq \lim_{t \to 0^-} \frac{\varphi(t) - \varphi(0)}{t} = \varphi'(0) = \lim_{t \to 0^+} \frac{\varphi(t) - \varphi(0)}{t} \geq 0,$$

ce qui montre que

$$\nabla J(u)v = 0.$$

Comme le vecteur v est arbitraire, il en résulte que $\nabla J(u) = 0$.

La relation $\nabla J(u) = 0$ est parfois appelée l'équation d'Euler.

Remarque 2.2.1. (1) *Le fait que Ω soit ouvert est évidemment essentiel. Considérons, par exemple, la fonction $J(v) = v$ sur l'intervalle $[0, 1]$.*
(2) *Si $V = \mathbb{R}^n$, la condition $\nabla J(u) = 0$ est équivalent au système d'équations*

$$\begin{cases} \partial_1 J(u_1, ..., u_n) = 0 \\ \qquad \vdots \\ \partial_n J(u_1, ..., u_n) = 0 \end{cases}$$

2.2.2 Conditions d'ordre 2

Condition de Legendre

Les résultats qui suivent seront énoncés pour des minimums relatifs, la prise en compte des dérivées secondes (comme celle de la convexité au paragraphe suivant) permettant en effet de préciser la nature (maximum ou minimum) des extremums considérés. Bien entendu, on pourrait énoncer des résultats analogues pour les maximums relatifs.

Théorème 2.2.2. ***(condition nécessaire de minimum relatif)***. *Soit Ω un ouvert d'un espace vectoriel norme V et $J : \Omega \subset V \to \mathbb{R}$ une fonction dérivable dans Ω, deux fois déri-vable en un point $u \in \Omega$. Si la fonction J admet un minimum relatif en u, alors*

$$\nabla^2 J(u)(w,w) \geq 0 \text{ pour tout } w \in V.$$

Preuve 2.2.2. *Soit w un vecteur non nul de V. Il existe un intervalle ouvert $I \subset \mathbb{R}$ contenant l'origine tel que*

$$t \in I \Rightarrow u + tw \Rightarrow \Omega \text{ et } J(u+tw) \geq J(u).$$

La formule de Taylor-Young et la relation $\nabla J(u) = 0$ (théorème.1.3.2) permettent d'écrire

$$0 \leq J(u+tw) - J(u) = \frac{t^2}{2}(\nabla^2 J(u)(w,w) + \varepsilon(t), \quad \lim_{t\to 0} \varepsilon(t) = 0.$$

ce qui démontre l'assertion (si $\nabla^2 J(u)(w,w)$ était < 0, il en irait de même de la différence $J(u+tw) - J(u)$ pour t suffisamment petit). |

Remarque 2.2.2. *Il n'existe pas de réciproque du résultat précédent, comme le montre l'exemple ,de la fonction $J(v) = v^3, v \in \mathbb{R}$. Pour obtenir une condition suffisante de minimum relatif, on est en effet conduit, soit à faire une hypothèse "plus forte" au point u (cas (1) du théorème qui suit), soit à supposer une propriété analogue, mais vraie dans tout un voisinage du point u (cas (2)).*

Théorème 2.2.3. ***(conditions suffisantes de minimum relatif)****. Soit Ω un ouvert d'un espace vectoriel norme V, u un point de Ω, et $J : \Omega \subset V \to \mathbb{R}$ une fonction dérivable dans Ω telle que $\nabla J(u) = 0$.*
(1) *Si la fonction J est deux fois dérivable en u et s'il existe un nombre α tel que*

$$\alpha > 0, et\ \nabla^2 J(u)(w,w) \geq \alpha ||w||^2 \ pour\ tout\ w \in V,$$

alors la fonction J admet un minimum relatif strict en u.
(2) *Si la fonction J est deux fois dérivable dans Ω, et s'il existe une boule $B \subset \Omega$ centrée en u telle que*

$$\nabla^2 J(v)(w,w) \geq 0 \ pour\ tout\ v \in B, w \in V,$$

alors la fonction J admet un minimum relatif en u.

Preuve 2.2.3. (1) *La formule de Taylor-Young permet d'écrire, pour tout vecteur w suffisamment petit,*

$$\begin{aligned} J(u+w) - J(u) &= \frac{1}{2}(\nabla^2 J(u)(w,w) + ||w||^2 \varepsilon(w)), \\ &\geq \frac{1}{2}(\alpha - \varepsilon(w)) \|w\|^2, \quad \lim_{w \to 0} \varepsilon(w) = 0, \end{aligned}$$

ce qui montre que $J(u+w) > J(u)$ pour $(u+w) \in B$, où B est une boule centrée en w, dont le rayon r est suffisamment petit pour que $\|\varepsilon(w)\| \prec \alpha$ pour $\|w\| \leq r$.
(2) *La formule de Taylor-Maclaurin montre que*

$$J(u+w) = J(u) + \frac{1}{2}\nabla^2 J(v)(w,w) \geq J(u), \quad v \in]u, u+w[, \ \ pour\ tout\ u+w \in B.$$

Il n'existe pas de réciproques aux deux assertions de l'énoncé précédent.

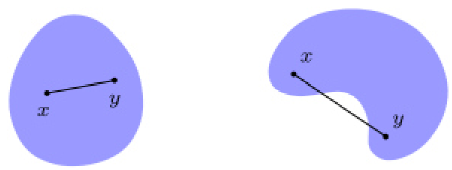
2.3 Optimisation convexe

La convexité joue un rôle très important dans la théorie classique de l'optimisation. Elle est un outil indispensable pour la recherche des conditions à la fois nécessaires et suffisantes d'optimalité.

2.3.1 Ensemble convexe

Définition 2.3.1. *Un ensemble* $U \subset \mathbb{R}^n$ *est dit convexe si pour tout couple* $(x,\ y) \in U^2$ *et* $\forall \lambda \in [0,1]$ *on a ;*

$$\lambda x + (1-\lambda)y \in U.$$



convexe non convexe

2.3.2 Propriétés des ensembles convexes

(1) . La définition d'ensemble convexe peut s'enterpréter en disant que le segment reliant x et y doit être dans U.
(2) . Soit $x_1,\ x_2,\ \ldots,\ x_k \in \mathbb{R}^n$ et t_j telle que $t_j \geq 0$ et $\sum_{j=1}^{k} t_j = 1$. Tout expréssion de la forme

$$\sum_{j=1}^{k} t_j x_j.$$

s'appelle combinaison convexe des points x_j ou barycentre.
(3) . Si U_1 et C_2 sont deux ensembles convexes de $\mathbb{R}^n$, alors $K = U_1 \cap U_2$ est convexe et

$$K = \{x \setminus x = x_1 + x_2,\quad x_1 \in U_1 \text{ et } x_2 \in U_2\}$$

est convexe.

2.3.3 Fonction convexe

Définition 2.3.2. *une application réelle* $J : U \subset V \to \mathbb{R}$ *est définie sur un ensemble convexe* U *d'un espace vectoriel* V *est dite convexe sur l'ensemble* U *si pour tous les points* $u, v \in U$ *et pour tout nombre réel positif ou nul* θ *tel que* $0 \leq \theta \leq 1$ *l'inégalité suivante est vérifiée :*

$$J(\theta u + (1-\theta)v) \leq \theta J(u) + (1-\theta)J(v)$$

et strictement convexe sur l'ensemble U si

$$u, v \in U, \ \ u \neq v \ \ et \ \ \theta \in]0, 1[$$

implique

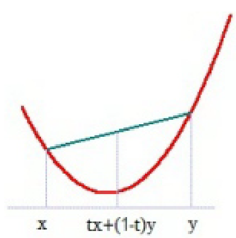
$$J(\theta u + (1-\theta)v) < \theta J(u) + (1-\theta)J(v).$$

Par exemple, une forme linéaire $f : V \rightarrow \mathbb{R}$ est convexe mais non strictement convexe ;de même une norme $||.|| : V \rightarrow \mathbb{R}$ est convexe. Naturellement, si une fonction est (strictement) convexe sur un ensemble convexe U, elle est encore (strictement) convexe sur toute partie convexe de U.

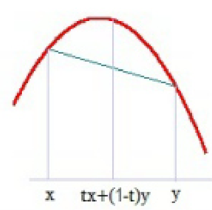
Définition 2.3.3. *Une fonction $G : U \subset V \rightarrow \mathbb{R}$ définie sur une partie convexe U d'un espace vectoriel V est dite (strictement) concave si la fonction $(-G)$ est (strictement) convexe.*

Il est claire qu'une fonction fortement convexe est strictement convexe avec une inégalité de convexité renforcée par un terme quadratique lui donnant une courbure au moins égale à θ. Et on a

$$convexit\acute{e}\ forte \Longrightarrow convexit\acute{e}\ stricte \Longrightarrow convexit\acute{e}$$



une fonction convexe



une fonction concave

Avant d'appliquer la notion de convexité aux extremums des fonctions, nous allons la caractériser à l'aide de la dérivabilité première (théorème (2.3.2)), ou seconde (théorème (2.3.3)).

Théorème 2.3.1. ***(condition nécessaire de minimum relatif sur un ensemble convexe).*** *Soit $J : \Omega \rightarrow \mathbb{R}$ une fonction définie sur un ouvert Ω d'un espace vectoriel* *norme V et U une partie convexe de Ω. Si la fonction J est dérivable en un point $u \in U$ et si elle admet en u un minimum relatif par rapport à l'ensemble U, alors*

$$\nabla J(u)(v-u) \geq 0 \quad \textit{pour tout} \quad v \in U.$$

Preuve 2.3.1. *Soit $v = u + w$ un point quelconque de l'ensemble U. Cet ensemble étant convexe, les points de la forme $u + \theta w$, $0 \leq \theta \leq 1$, sont encore dans U. La dérivabilité de la fonction J en u permet d'écrire*

$$J(u+\theta w) - J(u) = \theta(\nabla J(u)w + \varepsilon(\theta)), \qquad \lim_{\theta \to 0} \varepsilon(\theta) = 0,$$

pour tout $\theta \in [0,1]$ (le vecteur w étant fixé). Alors le nombre $\nabla J(u)w$ est nécessairement positive, sans quoi la différence $J(u+\theta w) - J(u)$ serait < 0 pour $d > 0$ suffisamment petit.

Remarque 2.3.1. (1) *Si l'ensemble U est un sous-espace vectoriel, la condition précédente devient simplement*

$$\nabla J(u)w = 0 \quad \textit{pour tout} \quad w \in U.$$

En particulier, si $U = V$, on retrouve la condition nécessaire $\nabla J(u) = 0$ du (théorème.2.2.1) .

Avant d'appliquer la notion de convexité aux extremums des fonctions, nous allons la caractériser à l'aide de la dérivabilité première (*théorème* 2.3.2), ou seconde (*théorème* 2.3.3).

Théorème 2.3.2. ***(convexité et dérivabilité première)***. *Soit $J : \Omega \subset V \to \mathbb{R}$ une fonction dérivable dans un ouvert Ω d'un espace vectoriel norme V, et U une partie convexe de Ω.*

(1) *La fonction J est convexe sur U si et seulement si*

$$J(v) \geq J(u) + \nabla J(u)(v-u) \quad \textit{pour tout} \quad u, v \in U.$$

(2) *La fonction J est strictement convexe sur U si et seulement si*

$$J(v) > J(u) + \nabla J(u)(v-u) \qquad \textit{pour tout} \quad u, v \in U, u \neq v.$$

Preuve 2.3.2. *Remarquons pour commencer que l'interprétation géométrique des conditions ci-dessus est claire* (*cf. figure* 2) *:*

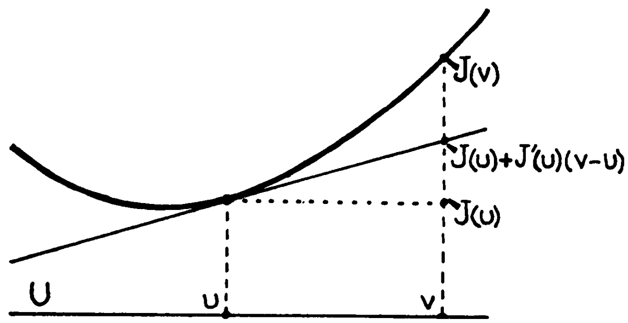


figure 2

on exprime simplement que la fonction est située "au-dessus" de son plan tangent. Soit u *et* v *deux points distincts de* U*, et* $\theta \in]0,1[$*. Si la fonction* f *est convexe,*

$$J(u+\theta(v-u)) \leq (1-\theta)J(u) + \theta J(v).$$

Par suite,

$$\frac{J(u+\theta(v-u)) - J(u)}{\theta} \leq J(v) - J(u).$$

ce qui peut encore s'écrire

$$\nabla J(u)(v-u) = \lim_{\theta \to 0} \frac{J(u+\theta(v-u)) - J(u)}{\theta} \leq J(v) - J(u).$$

Si la fonction f *est strictement convexe, le raisonnement précédent ne permet pas de conclure puisqu'on ne peut pas affirmer que l'inégalité stricte "passe à la limite". Soit alors* $\theta \in]0,1[$ *un nombre fixé. Comme*

$$u + \theta(v-u) = \frac{\omega - \theta}{\omega} u + \frac{\theta}{\omega}(u + \omega(v-u)),$$

on déduit

$$J(u+\theta(v-u)) \leq \frac{\omega - \theta}{\omega} J(u) + \frac{\theta}{\omega} J(u+\omega(v-u)), \quad 0 \leq \theta \leq \omega.$$

Par suite, si la fonction est strictement convexe,

$$\begin{aligned}\frac{J(u+\theta(v-u))-J(u)}{\theta} &\leq \frac{J(u+\omega(v-u))-J(u)}{\omega}\\ &< J(v)-J(u), \quad 0\leq\theta\leq\omega.\end{aligned}$$

puisque $\omega < 1$ *par hypothèse ; alors le passage à la limite conduit cette fois à l'inégalité stricte.*
Réciproquement, supposons que

$$J(v) \geq J(u) + \nabla J(u)(v-u) \quad \text{pour tout} \ \ u, v \in U.$$

Soit alors u *et* v *deux points distincts de* U *et* $\theta \in]0,1[$ *; on a donc, en particulier,*

$$J(v) \geq J(v+\theta(u-v)) - \theta\nabla J(v+\theta(u-v))(u-v).$$

$$J(u) \geq J(v+\theta(u-v)) + (1-\theta)\nabla J(v+\theta(u-v))(u-v),$$

et il suffit d'additionner les deux inégalités ci-dessus, multipliées respectivement par $(1-\theta)$ *et* θ*, pour obtenir*

$$J(\theta u + (1-\theta)v) \leq \theta J(u) + (1-\theta)J(v),$$

ce qui établit la convexité de la fonction /, ou la stricte convexité si les inégalités sont strictes.

Théorème 2.3.3. ***(convexité et dérivabilité seconde).*** *Soit* $J : \Omega \subset V \to \mathbb{R}$ *une fonction deux fois dérivable dans un ouvert* Ω *d'un espace vectoriel norme* V*, et* U *une partie convexe de* θ*.*
(1) *La fonction* J *est convexe sur* U *si et seulement si*

$$\nabla^2 J(u)(v-u, v-u) \geq 0 \quad \text{pour tout} \quad u, v \in U.$$

(2) *Si*

$$\nabla^2 J(u)(v-u, v-u) > 0 \qquad \text{pour tout} \quad u, v \in U, u \neq v,$$

la fonction J *est strictement convexe sur* U*.*

Preuve 2.3.3. *Soit u et v deux points distincts de U. Par application de la formule de Taylor-Maclaurin,*

$$\begin{aligned} J(v) - J(u) - \nabla J(u)(v-u) &= \frac{1}{2}\nabla^2 J(w)(v-u, v-u) \\ &= \frac{\rho^2}{2}\nabla^2 J(w)(v-w, v-w), \quad w \in (u,v), \end{aligned}$$

le nombre $\rho > 0$ étant défini par l'égalité $(v-u) = \rho(v-w)$. On déduit alors la convexité,ou la stricte convexité, par application du théorème précédent.
Pour établir la réciproque de (1)*, introduisons la fonction auxiliaire*

$$G : v \in \Omega \rightarrow G(v) = J(v) - \nabla J(u)v$$

pour un point $u \in U$ quelconque, mais considéré comme fixé dans ce qui suit. Un coup d'oeil à la ($figure$ 2) *aura d'ailleurs tôt fait de justifier cette introduction : si la fonction f est convexe, la fonction G admet en u un minimum par rapport à l'ensemble U ; en effet,*

$$G(v) - G(u) = J(v) - J(u) - \nabla J(u)(v-u) \geq 0, \quad \textit{pour tout} \;\; v \in U,$$

d'après la condition nécessaire (1) *du théorème* (2.3.2)*. La fonction G étant deux fois dérivable dans Ω et de dérivée $G'' = \nabla^2 J$, la formule de Taylor-Young donne, pour tout $v = u + w \in U$, et pour tout $t \in [0,1]$,*

$$0 \leq G(u+tw) - G(u) = \frac{t^2}{2}(\nabla^2 J(u)(w,w) + \varepsilon(t)), \lim_{t\to 0} \varepsilon(t) = 0,$$

puisque $G'(u) = 0$ par construction. Le raisonnement habituel entraîne alors $J"(u)(w,w) \geq 0$.

Remarque 2.3.2. (1) *On ne pouvait pas appliquer directement la condition nécessaire du* (*théorème*.2.2.1)*, établie pour des minimums relatifs par rapport à des ensembles ouverts.*
(2) *L'exemple de la fonction strictement convexe $J(v) = v^4, v \in \mathbb{R}$, montre qu'il ne saurait exister en général de réciproque à la condition* (2)*.*
(3) *Dans le cas particulier d'une fonctionnelle quadratique sur $\mathbb{R}^2$, la réciproque est vraie : de l'expression $J(v) = \frac{1}{2}(Av, v) - (b, v), \quad A = A^T$, on déduit en effet*

$$J(v) - J(u) - \nabla J(u)(v-u) = \frac{1}{2}(A(v-u), v-u),$$

et le (*théorème* 2.3.2) *permet de conclure. Rassemblant les divers résultats relatifs à ce type de fonction, on a donc établi qu'une fonctionnelle quadratique sur* $\mathbb{R}$ *(du type ci-dessus) est convexe si et seulement si la matrice symétrique* A *est positive, et strictement convexe si et seulement si la matrice* A *est définie positive.* |

2.3.4 Existence et unicité

La convexité est un outil précieux en optimisation. En effet, on a

Proposition 2.3.1. *Soit le problème* (P) *avec* J *convexe et* $U \subset V$ *non vide convexe. Alors*

(1). *Tout minimum local est un minimum global,*

(2). *L'ensemble des minima de* J *sur* U *est un ensemble convexe (éventuellement vide),*

(3). *Si* J *est strictement convexe, il y a au plus un minimum.*

Alors que nous n'avons envisagé jusqu'ici que des extremums relatifs, la convexité va nous permettre de nous affranchir du caractère "local" de cette propriété. C'est pourquoi nous introduisons les définitions suivantes :
Le résultat qui suit rassemble quelques propriétés constamment utilisées des minimums des fonction convexes.

Théorème 2.3.4. *Soit* U *une partie convexe d'un espace vectoriel norme* V.

(1) *Si une fonction convexe* $J : U \subset V \to \mathbb{R}$ *admet un minimum relatif en un point de* U *elle y admet en fait un minimum, c'est-à-dire par rapport à tout l'ensemble* U.

(2) *Une fonction strictement convexe* $J : U \subset V \to \mathbb{R}$ *admet au plus un minimum, et c'est un minimum strict.*

(3) *Soit* $J : \Omega \subset V \to \mathbb{R}$ *une fonction convexe définie sur un ouvert* Ω *de* V *contenant* U, *dérivable en un point* $u \in U$. *Alors la fonction* J *admet un minimum en* u *par rapport à l'ensemble* U *si et seulement si*

$$\nabla J(u)(v-u) \geq 0 \quad \textit{pour tout} \quad v \in U.$$

(4) *Si l'ensemble* U *est ouvert, la condition précédente équivaut à l'équation d'Eulerest ouvert,* $\nabla J(u) = 0$.

Preuve 2.3.4. (1) *Soit* $v = u + w$ *un point quelconque de l'ensemble* U*. D'après la convexité de la fonction* J*,*

$$J(u + \theta w) \leq (1 - \theta)J(u) + \theta J(v), \quad 0 \leq \theta \leq 1,$$

ce que l'on peut encore écrire

$$J(u + \theta w) - J(u) \leq \theta(J(v) - J(u), \quad 0 \leq \theta \leq 1.$$

Comme le point u *est un minimum relatif, il existe un nombre* θ_0 *tel que*

$$\theta_0 > 0 \quad et \quad 0 \leq J(u + \theta_0 w) - J(u),$$

ce qui entraîne $J(v) \geq J(u)$*.*
(2) *Si la fonction est strictement convexe, le même raisonnement conduit aux inégalités*

$$0 \leq J(u + \theta_0 w) - J(u) \leq \theta_0(J(v) - J(u), \quad \theta_0 > 0, \quad w \neq 0,$$

ce qui établit le caractère strict du minimum, et donc du même coup son unicité.
(3) *On a montré au (théorème.*2.3.1) *la nécessité de la condition* $\nabla J(u)(v - u) \geq 0$ *pour tout* $v \in U$ *(sans supposer* J *convexe). Pour établir qu'elle est suffisante, on remarque que*

$$J(v) - J(u) \geq \nabla J(u)(v - u) \quad pour\ tout \quad v \in U,$$

*d'après le (théorème.*2.3.2*).*
(4) *La dernière propriété est une conséquence immédiate de la propriété (3).*

Les relations "$\nabla J(u)(v - u) \geq 0$" pour tout $v \in U$ sont fréquemment appelées les inéquations d'Euler.

2.4 Optimisation quadratique

En optimisation mathématique, un problème d'optimisation quadratique est un problème d'optimisation dans lequel on minimise (ou maximise) une fonction quadratique sur un polyèdre convexe. L'optimisation quadratique est la discipline qui étudie ces

problèmes. L'optimisation linéaire peut être vue comme un cas particulier de l'optimisation quadratique.
Ce problème est NP-difficile dans le cas général. Dans le cas particulier de la minimisation d'une fonction objectif convexe, le problème est polynomial et on parle d'optimisation quadratique convexe ; une discipline déjà trés riche aux propriétés mieux connues.Définition du problème
On considère le problème d'optimisation (minimisation) quadratique dé nie positive sans contrainte définie mathématique comme ci-dessous :

Définition 2.4.1. ***(Problème d'optimisation quadratique définie positive sans contrainte)***

$$\inf_{u \in \mathbb{R}^n} J(u)$$

où la fonction objectif $J : \mathbb{R}^n \to \mathbb{R}$ *définie par est quadratique définie positive sur* $\mathbb{R}^n$ *donnée par la formulation matricielle relativement à la base canonique de* $\mathbb{R}^n$ *:*

$$\forall u := \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}^T \in \mathcal{A}_{n,1}(\mathbb{R}), \quad J(u) = \frac{1}{2} u^T A u - b^T u + c$$

- $A = (a_{ij})_{1 \le i,j \le n} \in \mathcal{A}_n(\mathbb{R})$ *symétrique définie positive ;*

- $b := \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathcal{A}_{n,1}(\mathbb{R})$

Définition 2.4.2. *Soit* A *une matrice symétrique* $n \times n$ *et* $b \in \mathbb{R}^n$*. On appelle forme quadratique la fonction* $J : \mathbb{R}^n \to \mathbb{R}$ *définie par*

$$J(u) = \frac{1}{2} u^T A u - b^T u$$

Lorsque la matrice A possède certaines propriétés, la fonction J peut prendre un nom particulier.
La propriété à laquelle nous allons nous intéresser est la positivité :

Définition 2.4.3. *Soit* A *une matrice symétrique* $n \times n$ *et* $b \in \mathbb{R}^n$*. On dit que* A *est semi-définie positive et on note* $A \geq 0$*, quand* $u^T A u \geq 0$*,* $\forall u \in \mathbb{R}^n$*.*

On dit que A est définie positive et on note $A > 0$, quand $u^T Au > 0$, $\forall u \in \mathbb{R}^n$, $u \neq 0$. Cette définition peut être reliée aux valeurs propres de la matrice A.

Propriété 2.4.1. *Soit A une matrice symétrique $n \times n$. On note $(\lambda_i)_{\,i=1...n}$ ses valeurs propres (réelles). On a les équivalences suivantes :*

$$\begin{aligned} A &\geq 0 \Longleftrightarrow \lambda_i \geq 0, \quad i = 1...n, \\ A &> 0 \Longleftrightarrow \lambda_i > 0, \quad i = 1...n. \end{aligned}$$

Lorsque la matrice A est définie positive (resp. semi-définie positive), on dira que $J(u)$ est une forme quadratique définie positive (resp. semi-définie positive). Dans le cas où A est définie positive la fonction J possède un certain nombre de propriétés. Nous nous intéressons dans un premier temps aux surfaces $J(u) = c$ où $c \in \mathbb{R}$.

2.4.1 Propriétés des formes quadratiques définies positives

Proposition 2.4.1. *Soit A une matrice symétrique $n \times n$, définie positive et $b \in \mathbb{R}^n$. Considérons la forme quadratique*

$$J(u) = \frac{1}{2} u^T Au - b^T u.$$

On considère la famille de surfaces définie par

$$\gamma_c = \{u \in \mathbb{R}^n,\ J(c) = c\}$$

pour $c \in \mathbb{R}^n$, et on définit le vecteur $\hat{u}$ solution de $A\hat{u} = b$.
Alors γ_c est définie de la façon suivante :
– *Si $c < J(\hat{u})$ alors $\gamma_c = \varnothing$.*
– *Si $c = J(\hat{u})$ alors $\gamma_c = \{\hat{u}\}$.*
– *Si $c > J(\hat{u})$ alors γ_c est un ellipsoïde centré en $\hat{u}$.*

Preuve 2.4.1. *La matrice A étant diagonalisable, il existe une matrice P (la matrice des vecteurs propres) orthogonale telle que*

$$P^T AP = D$$

où $D = diag(\lambda_1, ..., \lambda_n)$ avec $\lambda_i > 0$. On fait le changement de variabley $= u - \hat{u}$: cela donne,

$$J(\hat{u} + y) = J(\hat{u}) + (A\hat{u} - b)^T y + \frac{1}{2} y^T A y,$$

et puisque $A\hat{u} = b$, on a

$$J(x) = J(\hat{u}) + \frac{1}{2}(u - \hat{u})^T A (u - \hat{u}).$$

On fait maintenant le changement de variable $(u - \hat{u}) = Pz$, ce qui donne

$$\begin{aligned} J(x) &= J(\hat{u}) + \frac{1}{2} z^T P^T A P z. \\ &= J(\hat{u}) + \frac{1}{2} z^T D z. \\ &= J(\hat{u}) + \frac{1}{2} \sum_{i=1}^{n} \lambda_i z_i^2 \end{aligned}$$

La surface γ_c est donc définie par

$$\gamma_c = \left\{ z \in \mathbb{R}^n, \frac{1}{2} \sum_{i=1}^{n} \lambda_i z_i^2 \ = c - J(\hat{u}) \right\}$$

Si $c - J(\hat{u}) < 0$ il est clair qu'il n'y a pas de solution à l'équation

$$\frac{1}{2} \sum_{i=1}^{n} \lambda_i z_i^2 \ = c - J(\hat{u})$$

puisque le second membre est toujours positif !

- Si $c = J(\hat{u})$ la seule solution est $z = 0$, c'est à dire $u = \hat{u}$.

- Si $c > J(\hat{u})$ l'équation définit bien un ellipsoïde, puisque les λ_i sont positifs.

Nous avons en fait démontré un résultat très intéressant qui caractérise la valeur minimale prise par $J(u)$ quand u parcourt $\mathbb{R}^n$:

Théorème 2.4.1. *Soit A une matrice symétrique $n \times n$ définie positive et $b \in \mathbb{R}^n$, et soit J la forme quadratique associée, définie par*

$$J(u) = \frac{1}{2} u^T A u - b^T u.$$

Soit $\hat{u}$ le vecteur (unique) vérifiant $A\hat{u} = b$, alors $\hat{u}$ réalise le minimum de J, c'est à dire

$$J(\hat{u}) < J(u) \quad \forall u \in \mathbb{R}^n$$

2.4.2 Gradient d'une fonction quadratique

On considère la fonction $J(u) = \frac{1}{2}u^T Au - b^T u$ où A est une matrice carrée symétrique $n \times n$. On a

$$\begin{aligned} J(u+th) &= \frac{1}{2}u^T Au + \frac{1}{2}t^2 h^T Ah + tu^T Ah + b^T(u+th) \\ &= J(u) + t\left(u^T A - b^T\right)h + \frac{1}{2}t^2 h^T Ah \end{aligned}$$

on a donc

$$\frac{J(u+th) - J(u)}{t} = \left(u^T A - b^T\right)h + \frac{1}{2}th^T Ah$$

Puisque

$$\lim_{t\to 0}\frac{1}{2}th^T Ah = 0$$

on a donc $J(u) = Ax - b$.

Matrice hessienne d'une fonction quadratique

On considère la fonction $J(u) = \frac{1}{2}u^T Au - b^T u$ où A est une matrice carrée symétrique $n \times n$. Puisque

$$\nabla f(x) = Ax - b$$

et puisque la matrice hessienne est la dérivée du gradient on a donc

$$\nabla^2 J(u) = A$$

De même, on peut montrer que l'on a.

$$\nabla^2 J(u) = \frac{1}{2}(A + A^T)$$

2.4.3 Convexité d'une fonction quadratique

Proposition 2.4.2. *soit J la forme quadratique, définie par*

$$J(u) = \frac{1}{2}u^T Au - b^T u.$$

alors J est convexe si et seulement si $A \geq 0$, et strictement convexe si et seulement si $A > 0$.

2.4.4 Solution d'un système linéaire au sens des moindres carrés

La résolution d'un système linéaire au sens des moindres carrés est un problème d'optimisation sans contraintes, correspondant aux données suivantes :

$$U = V = \mathbb{R}^n; \quad J : u \in \mathbb{R}^n \to J(u) = \frac{1}{2}\|Au - b\|_m^2 - \frac{1}{2}\|b\|_m^2$$

Comme

$$J(u) = \frac{1}{2} < A^T Au, u >_n - < A^T b, u >_m$$

il s'agit d'un problème de programmation quadratique, au sens entendu ici, seulement si la matrice symétrique $A^T A$ est définie positive. On rappelle qu'on a établi au l'existence d'une solution de ce problème dans tous les cas, compris celui où la matrice $A^T A$ est seulement positive. Lorsque la matrice $A^T A$ est définie positive, l'existence et l'unicité de la solution peuvent aussi se retrouver à partir du ($théorème$.2.3.4).

Proposition 2.4.3. *On considère la fonction quadratique définie sur* $\mathbb{R}^n$ *par* $J(u) = \frac{1}{2}u^T Au - b^T u$. *Alors si* A *est symétrique et définie positive,alors* $A^{-1}b$ *est l'unique minimum de* J.

Preuve 2.4.2. *On se donne une matrice réelle* A *de type* (m, n) *un vecteur* $b \in \mathbb{R}^m$, *et on cherche un vecteur* $u \in \mathbb{R}^n$ *tel que*

$$\|Au - b\|_m = \inf_{r \in \mathbb{R}^n} \|Au - b\|_m ,$$

où $||.||_m$ *désigne la norme euclidienne dans* $\mathbb{R}^m$. *Introduisons la fonctionnelle quadratique*

$$\begin{aligned} J(v) &= \frac{1}{2}\|Av - b\|_m^2 - \frac{1}{2}\|b\|_m^2 \\ &= \frac{1}{2}(Av, Av)_m - (c, Av)_m \\ &= \frac{1}{2}(A^T Av, v)_n - (A^T c, v)_n \quad v \in \mathbb{R}^n, \end{aligned}$$

ou $< .,. >_m$ *et* $< .,. >_n$ *désignent les produits scalaires des espaces* $\mathbb{R}^m$ *et* $\mathbb{R}^n$, *respectivement.La matrice symétrique* $A^T A$ *étant positive, la fonction* J *est convexe* (*théorème* 2.3.3).*Comme le problème considéré équivaut à la recherche d'un vecteur* $u \in \mathbb{R}^n$ *tel que*

$$J(u) = \inf_{v \in \mathbb{R}^n} J(v)$$

nous pouvons conclure du théorème (2.3.4) *que l'ensemble des solutions coïncide avec l'ensemble des solutions de l'équation*

$$\nabla J(u) = A^T A u - A^T b = 0,$$

Donc

$$\|A(u+w) - b\|_m^2 = ||Au - b||_m^2 + 2(A^T A u - A^T b, w)_n + \|Aw\|_m^2 \,,$$

qui n'est autre que la formule de Taylor

$$J(u+w) = J(u) + \nabla J(u) w + \frac{1}{2}(A^T A w, w)_n,$$

écrite pour la fonctionnelle quadratique J*, dont le Hessien est la matrice* $A^T A$.
La matrice $A^T A$ *est donc symétrique définie positive). Alors la condition nécessaire et suffisante d'optimialité est*

$$\nabla J(u) = 2(A^T A x - A^T b) = 0$$

et on retrouve bien l'équation normale $A^T A x = A^T b$.

2.4.5 Le théorème de projection ; premières conséquences

Théorème 2.4.2. ***(Théorème de projection)****. Soit* U *un sous-ensemble non vide, convexe, fermé, d'un espace de Hilbert* V*. Étant donné un élément quelconque* $w \in V$*, il existe un et un seul élément* Pw *tel que*

(1)

$$Pw \in U \text{ et } \|w - Pw\| = \inf_{v \in U} \|w - v\| \,.$$

Cet élément $Pw \in U$ *vérifie*

(2)

$$Pw \in U \quad \text{et} \quad \|w - Pw\| = \inf_{v \in U} \|w - v\| \,.$$

et, réciproquement, si un élément u vérifie

$$u \in U \quad et \quad < u - w, v - u > \geq 0 \quad pour\ tout \quad v \in U.$$

alors $u = Pw$.
L'application $P : V \to U$ *ainsi définie est telle que*
(3)

$$\|Pw_1 - Pw_2\| \leq \|w_1 - w_2\| \quad pour\ tout \quad w_1, w_2 \in V.$$

Enfin, l'application $P : V \to U \subset V$ *est linéaire si et seulement si le sous-ensemble* U *est un sous-espace vectoriel, auquel cas les inégalités* (2) *sont remplacées par les égalités*
(4)

$$< Pw - w, v >= 0 \quad pour\ tout \quad v \in U.$$

Faisons divers commentaires sur ce résultat.
(i) L'application $P : V \to U$ s'appelle l'opérateur de projection, et l'élément Pw s'appelle la projection (sur l'ensemble U) de l'élément w, l'interprétation géométrique de la relation de définition (1) étant claire à cet égard (figure 3) :

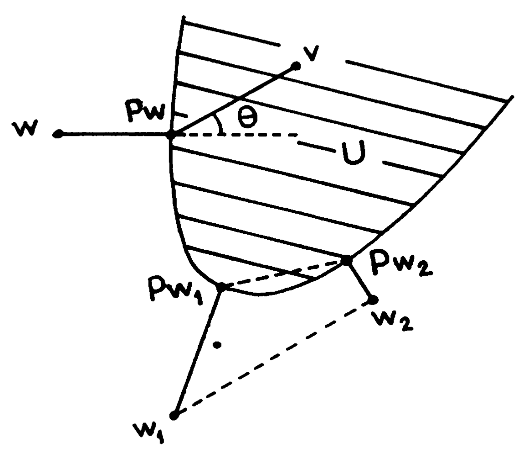


figure 3 :

l'élément "projeté" Pw est en effet l'élément de l'ensemble U "le plus voisin" du point w. De la même façon, les inégalités (2) traduisent la nécessité, intuitivement évidente, pour l'angle formé par les vecteurs $(Pw - w)$ et $(v - Pw)$ d'être $\leq \frac{\pi}{2}$ pour tous les éléments $v \in U$ $(figure\ 3)$. On notera au passage que

$$w - Pw = 0 \Leftrightarrow w \in U$$

(ii) On peut rapprocher les inégalités (2) des inéquations d'Euler $\nabla J(u)(v - u) \geq 0$ pour tout $v \in U$ du (*théorème*.2.3.4). Effectivement, leur analogie n'est pas une coïncidence : considérons en effet, pour $w \in V$ fixé, la fonction

$$J : v \in V \to J(v) = \frac{1}{2} \|w - v\|^2 = \frac{1}{2} < v, v > - < w, v > + \frac{1}{2} < w, w > .$$

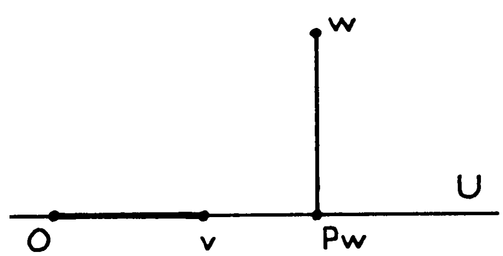
Cette fonction est dérivable, avec

$$\nabla J(v)z =< v - w, z > \quad \text{pour tout} \quad v, z \in V,$$

et (strictement) convexe (*théorème*.2.3.3). Par suite, les inégalités (2) expriment simplement la condition nécessaire et suffisante (3) du (*théorème*.2.3.4), écrite au point $u = Pw$.

(iii) L'inégalité (3) entraîne en particulier la continuité de l'opérateur de projection. On la retient parfois en disant de façon imagée que "la projection n'augmente pas les distances" (figure 3).

(iv) La condition (4) traduit l'orthogonalité (au sens défini plus loin) du vecteur $(Pw - w)$ et des vecteurs de l'ensemble V, lorsque celui-ci est un espace vectoriel. L'interprétation géométrique est encore évidente (figure 4).



$figure\ 4$:

Pour le démontrer, il suffit de vérifier la condition nécessaire et suffisante du ($théorème$.2.4.2) ; or étant donné un élément quelconque $v = (v_i)_{i=1}^n$ de l'ensemble U, la définition précédente de l'élément Pw entraîne effectivement

$$(Pw - w, v - Pw) = \sum_{i=1}^{n}((Pw)_i - w_i)(v_i - (Pw)_i) = - \sum_{i, w_i \prec 0} w_i v_i \geq 0.$$

L'extension aux ensembles de la forme

$$U = \prod_{i=1}^{n}[a_i, b_i] = \{v = (v_i)_{i=1}^n \in \mathbb{R}^n : a_i \leq v_i \leq b_i, \quad 1 \leq i \leq n\} \subset \mathbb{R}^n,$$

les cas où $a_i = -\infty$ et $a_i = +\infty$ n'étant pas exclus, n'offre aucune difficulté.
Les lecteurs vérifieront par un raisonnement analogue que l'opérateur de projection correspondant est donné par

$$(Pw)_i = \min\{\max\{w_i, a_i\} b_i\} = \left\{ \begin{array}{cl} a_i & \text{si } \ w_i \prec a_i, \\ w_i & \text{si } \ w_i \leq a_i \leq b_i, \\ b_i & \text{si } \ b_i \prec w_i. \end{array} \right\}$$

Given an element $u \in V$, the Schwarz inequality shows that the linear form

$$(u, .) : v \in V \rightarrow < u, v > \in \mathbb{R}$$

2.4.6 Application du théorème de projection

Reprenons le problème de la solution d'un système linéaire au sens des moindres carrés : trouver $u \in \mathbb{R}^n$ tel que

$$||Au - b||_m = \inf_{v \in \mathbb{R}^n} \|Av - b\|_m \, ,$$

où la matrice $A \in \mathcal{A}_{m,n}(\mathbb{R})$ et le vecteur $b \in \mathbb{R}^m$ sont donnés, et euclidienne dans $\mathbb{R}^m$.
Le sous-espace vectoriel

$$\mathrm{Im}(A) = \{Av \in \mathbb{R}^m : v \in \mathbb{R}^n\}$$

désigne la norme étant fermé (on est en dimension finie), le théorème de projection entraîne l'existence, et l'unicité, d'un élément u vérifiant

$$\tilde{u} \in \mathrm{Im}(A) \quad \text{et} \quad \|\tilde{u} - b\|_m = \inf_{\tilde{v} \in \mathrm{Im}(A)} \|\tilde{v} - b\|_m .$$

Par suite, le problème posé a toujours au moins une solution, à savoir l'un des éléments $u \in \mathbb{R}^n$ qui vérifie

$$Au = \tilde{u}.$$

Cette solution est unique si et seulement si l'application représentée par la matrice A est injective (ce qui n'est possible que si $m \geq n$), c'est-à-dire si et seulement si la matrice symétrique positive $A^T A$ est définie, ou encore si et seulement si $r(A) = n$. Dans le même esprit, la caractérisation (4) du $théorème 8.1 - 1$, à savoir

$$< \tilde{u} - b, \tilde{v} >_m = 0 \quad \text{pour tout} \quad \tilde{v} \in \mathrm{Im}(A),$$

s'écrit, en notant $<.,.>_m$ et $<.,. >_n$ les produits scalaires euclidien de $\mathbb{R}^m$ et $\mathbb{R}^n$ respectivement,

$$< Au - b, Av >_m = < A^T Au - A^T b, v >_n = 0 \quad \text{pour tout} \quad v \in \mathbb{R}^n.$$

On a ainsi établi que les équations normales

$$A^T Au = A^T b$$

ont toujours au moins une solution.

Chapitre 3

Optimisation sans contrainte en dimension infinie

3.0.7 Existence en dimension infinie ?

Problème de compacité

Dans ce paragraphe, nous allons énoncer un résultat d'existence en dimension infinie dans le cas particulier où J satisfait une propriété de convexité forte. En général, et c'est sans grande surprise, il est bien plus difficile d'obtenir un résultat d'existence en dimension infinie. À titre d'exemple, considérons l'espace de Hilbert (de dimension infinie) des suites de carré sommable dans $\mathbb{R}$, c'est-à dire

$$\ell^2(\mathbb{R}) = \left\{ u = (u_n)_{n\in\mathbb{N}} \ / \ \sum_{n=0}^{+\infty} u_n^2 < +\infty \right\}$$

muni du produit scalaire $\langle u, v\rangle = \sum_{n=0}^{+\infty} u_n v_n$. On considère la fonctionnelle J définie par

$$\begin{array}{rcl} J & : & \ell^2(\mathbb{R}) \to \mathbb{R} \\ u & \to & \left(\|u\|^2 - 1\right)^2 + \displaystyle\sum_{n=0}^{+\infty} \frac{u_n^2}{n+1} \end{array}$$

On s'intéresse au problème d'optimisation

$$\begin{cases} \inf J(u) \\ u \in \ell^2(\mathbb{R}). \end{cases}$$

Remarquons que J est une fonctionnelle coercive (infinie à l'infini). En effet, pour tout $u \in \ell^2(\mathbb{R})$.

$$J(u) \geq \left(\|u\|^2 - 1\right)^2 \underset{\|u\| \to +\infty}{\rightarrow} +\infty$$

Cependant, le problème d'optimisation ci-dessus n'a pas de solution. Pour le vérifier, il suffit de remarquer que

$$\inf\{J(u),\ u \in \ell^2(\mathbb{R})\} = 0$$

L'existence d'un minimiseur u tel que $J(x) = 0$ étant clairement impossible, cela garantit que ce problème n'a pas de solution. Démontrons à présent que

$$\inf\{J(u),\ u \in \ell^2(\mathbb{R})\} = 0$$

On considère la suite (minimisante) $(u_n)_{n\in\mathbb{N}} \in \ell^2(\mathbb{R})^{\mathbb{N}}$ définie pour $n \in \mathbb{N}$ fixé par : $u_k^n = \delta_{k,n}$, $\forall k \in \mathbb{N}$. On vérifie alors aisément que pour tout $n \in \mathbb{N}$, $J(x^n) = \dfrac{1}{n+1}$ et la conclusion s'ensuit.

La moralité de cet exemple est que la compacité s'obtient bien plus difficilement en dimension infinie qu'en dimension finie. Bien que la suite minimisante $(u_n)_{n\in\mathbb{N}}$ soit bornée, il n'est pas possible d'en extraire une sous-suite convergente dans $\ell^2(\mathbb{R})$ (theorème de Weirstrass).

Difficulté indépendante du choix de l'espace.

Une fonction continue sur un fermé borné n'atteint pas toujours son minimum !

Exemple 3.0.1. *Exemple de non-existence : soit l'espace de Sobolev $H^1(0,1)$ muni de la norme*

$$\|v\| = \left(\int_0^1 v'(x)^2 + v(x)^2 dx\right)^{\frac{1}{2}}$$

Soit

$$J(v) = \int_0^1 \left((|v'(x)| - 1)^2 + v(x)^2 \right) dx$$

On vérifie que l'application J est continue et "infinie à l'infini", et pourtant le problème de minimisation

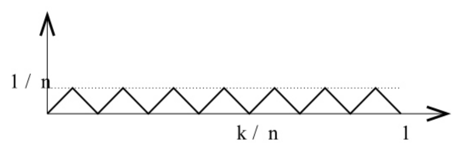
$$\inf_{v \in H^1(0,1)} J(v)$$

n'admet pas de solution. (Difficulté indépendante du choix de l'espace.)

Preuve 3.0.3. *Il n'existe aucun $v \in H^1(0,1)$ tel que $J(v) = 0$, et pourtant*

$$\inf_{v \in H^1(0,1)} J(v) = 0$$

car si on définit la suite un telle que $(u_n)' = \pm 1$



on a $J(u_n) = \int_0^1 u_n(x)^2 dx = \frac{1}{4n} \to 0$.
On voit dans cet exemple que la suite minimisante un "oscille" de plus en plus et n'est pas compacte dans $H^1(0,1)$ (bien qu'elle soit bornée).

Pour obtenir des résultats d'existence nous commençons par rappeler, avec quelques-unes de ses conséquences, le théorème de projection dans les espaces de Hilbert, dont un usage constant est fait en optimisation. Il nous permettra en particulier d'établir un certain nombre de résultats directement en dimension infinie.
Dans ce qui suit, on se place dans un espace de Hilbert V muni d'un produit scalaire $<\cdot,\cdot>$.

3.0.8 Fonction α-elliptique

Définition 3.0.4. *Soit $U \subset V$, un convexe. Une fonction $J : U \to \mathbb{R}$ est dite fortement convexe ou uniformément convexe ou α-convexe ou α-elliptique s'il existe $\alpha > 0$ tel que, pour tous $(u, v) \in U^2$, $\theta \in [0, 1]$,*

$$f(\theta u + (1-\theta)v) \leq \theta f(u) + (1-\theta)f(v) - \frac{\alpha}{2}\theta(1-\theta) \|x - y\|^2$$

Il est tout à fait clair que l'ellipticité implique la stricte convexité qui implique elle-même la convexité. On notera que la convexité correspond formellement au cas $\alpha = 0$. Bien sûr, les réciproques sont fausses.
La proposition ci-dessous examine plus précisément le lien entre "convexité et "uniforme convexi- té". Elle fournit également un critère permettant de vérifier 1'uniforme convexité d'une fonction.

Proposition 3.0.4. *Soit $J : V \to \mathbb{R}$.*
1. La fonction J est α-elliptique si et seulement si la fonction $J - \frac{\alpha}{2}\|\cdot\|^2$ est convexe.
2. On suppose que J est continue. Alors, la fonction J est α-elliptique si, et seulement si il existe $\alpha > 0$ tel que, pour tout $(u,\ v) \in V^2$,

$$J\left(\frac{u+v}{2}\right) \leq \frac{J(u)+J(v)}{2} - \frac{\alpha}{8}\|u-v\|^2$$

Preuve 3.0.4. *1. Posons $g(u) = J(u) - \frac{\alpha}{2}\|u\|^2$. En développant $\|\theta u + (1-\theta)v\|^2$ et en regroupant les termes correctement, on trouve*

$$\begin{aligned}\theta g(u) + (1-\theta)g(v) - g(\theta u + (1-\theta)v) &= \theta J(u) + (1-\theta)J(v) - J(\theta u + (1-\theta)v) \\ &\quad -\frac{\alpha}{2}\theta(1-\theta)\|u-v\|^2,\end{aligned}$$

ce qui prouve 1'équivalence annoncée.
2. Le sens direct est immédiat et s'obtient en choisissant $\theta = \frac{1}{2}$.
Le sens réciproque est un peu plus délicat. Nous allons procéder par récurrence. Pour tout $n \in \mathbb{N}$, on note $U_n = \{\xi \in [0,1],\ 2^n\xi \in \mathbb{N}\}$. Fixons u et v dans V. On appelle $\mathcal{P}_n$ la propriété : Pour tout $\theta \in U_n$, 1'inégalité

$$J(\theta u + (1-\theta)v) \leq \theta J(u) + (1-\theta)J(v) - \frac{\alpha}{2}\theta(1-\theta)\|u-v\|^2,$$

est vérifiée".
Soit $\theta \in U_{n+1}\backslash U_n$, alors $2\theta \in U_n$. Il existe $(\theta_1,\ \theta_2) \in U_n^2$ tels que $\theta_1 < \theta_2$ et $\theta = \frac{\theta_1+\theta_2}{2}$. Puisque J vérifie 1'inégalité particulière de α-convexité énoncée dans la proposition,

$$\begin{aligned}J\left(\theta u + (1-\theta)v\right) &= J\left(\frac{(\theta_1 u + (1-\theta_1)v) + (\theta_2 u + (1-\theta_2)v)}{2}\right) \\ &\leq \frac{1}{2}\left[J\left(\theta_1 u + (1-\theta_1)v\right) + J\left(\theta_2 u + (1-\theta_2)v\right)\right] - \frac{\alpha}{8}(\theta_2-\theta_1)^2\|u-v\|^2.\end{aligned}$$

Or, puisque 1'inégalité de α-ellipticité a été supposée vraie sur U_n, on en déduit

$$\begin{aligned}
J(\theta\,(u+(1-\theta)v) &\leq \frac{\theta_1 J(u)+(1-\theta_1)J(v)+\theta_2 J(u)+(1-\theta_2)J(v)}{2}\\
&\quad -\frac{\alpha}{4}(\theta_1(1-\theta_1)+\theta_2(1-\theta_2))\|u-v\|^2-\frac{\alpha}{8}(\theta_2-\theta_1)^2\|u-v\|^2\\
&= \theta J(u)+(1-\theta)J(v)-\frac{\alpha}{4}(\theta_1(1-\theta_1)+\theta_2(1-\theta_2)+\frac{1}{2}(\theta_2-\theta_1)^2)\|u-v\|^2\\
&= \theta J(u)+(1-\theta)J(v)-\frac{\alpha}{2}\theta(1-\theta)\|u-v\|^2,
\end{aligned}$$

ce qui prouve que 1'inégalité de α-ellipticité est alors valable pour tout élément de U_{n+1}. On en déduit par récurrence que 1'inégalité est valable pour $\theta \in \bigcup_{n\in\mathbb{N}} U_n$. Comme J est continue, 1'inégalité reste valable sur 1'adhérence de l'union des U_n, c'est-à-dire sur $[0,1]$.

Dans le cas où la fonction J est régulière, comme pour la convexité, il existe des caractérisations de la convexité uniforme. On peut voir ces caractérisations comme des corollaires du théorème.

Corollaire 3.0.1. ***Caractérisation des fonctions uniformément convexes dans le cas régulier***
1. Si $J : V \to \mathbb{R}$ est dérivable, on a les équivalences
(i) J est α-elliptique ;
(ii) $J(v) \geq J(u) + \langle \nabla J(u),\ v-u > +\frac{\alpha}{2}\|v-u\|^2,\ \forall (u,\ v) \in V^2$;
(iii) $\langle \nabla J(v) - \nabla J(u),\ v-u\} \geq \alpha\|v-u\|^2,\ \forall (u,\ v) \in V^2$.
2. Si $J : V \to \mathbb{R}$ est deux fois différentiable, on a les équivalences
(i) J est α-elliptique ;
(ii) $\langle HessJ(u)h,\ h >\geq \alpha\|h\|^2,\ \forall u \in V,\ \forall h \in V$.

Preuve 3.0.5. *1. Grâce à la proposition, (i) équivaut à dire que $g(u) = J(u) - \dfrac{\alpha}{2}\|u\|^2$ est convexe. or, $\nabla g(u) = \nabla J(u) - \alpha u$. En écrivant alors les conditions (i), (ii) et (iii) du théorème, on obtient exactement les conditions (ii) et (iii) du corollaire pour J.*
2. La preuve découle immédiatement du théorème, en posant comme précédemment $g(u) = J(u) - \dfrac{\alpha}{2}\|u\|^2$ et en remarquant que $Hessg(u) = HessJ(u) - \alpha I$.

Remarque 3.0.1. ***Uniformément convexe implique coercif***
Si J est α-elliptique et différentiable, en utilisant la caractérisation précédente, on obtient aisément que

$$J(u) \geq J(0) + < \nabla J(0),\ u > + \frac{\alpha}{2}\|u\|^2,$$

ce qui implique que J est coercive.

Nous sommes à présent en mesure d'établir le résultat d'existence annoncé en dimension finie.

Théorème 3.0.3. *Soit U, un convexe fermé non vide d'un espace de Hilbert V et J, une fonction α convexe continue sur U. Alors, il existe un unique minimum u^* de J sur U et on a :*

$$\|u^* - u\|^2 \leq \frac{4}{\alpha}[J(u) - J(u^*)],\ \forall v \in U.$$

En particulier, toute suite minimisante de J sur l'ensemble U converge vers u^*.

Il existe un point un peu technique dans cette démonstration qui, paradoxalement, dans beaucoup de problèmes d'optimisation, est vérifié gratuitement. En effet, il s'agit du lemme suivant :

Lemme 3.0.1. *Soit J, une fonction α-convexe sur U. Alors, il existe deux constantes $(\alpha_1,\ \alpha_2) \in \mathbb{R}_+ \times \mathbb{R}$ telles que*

$$J(u) \geq \alpha_1\|u\|^2 + \alpha_2.$$

Remarque 3.0.2. *Ce lemme est démontré dans [1]. Il utilise dans sa preuve le théorème de séparation d'un point et d'un convexe. Il assure, puisque J est "infinie à 1'infini", que J est minorée sur le convexe U, donc que $\inf\{J(u),\ u \in U\}$ est fini. Il arrive assez souvent dans la pratique que 1'on minimise des fonctionnelles naturellement positives ou minorées si bien que cette étape n'apparait plus essentielle dans ces cas.*

Preuve 3.0.6. *Démontrons à présent le théorème en admettant le lemme technique ci-dessus. On désigne par $\hat{m}$, la quantité $\inf\{J(u),\ u \in U\}$. Soit $(u_n)_{n\in\mathbb{N}}$, une suite minimisante de J sur U. Puisque J est α-elliptique, on a pour tous $(n,\ m) \in \mathbb{N}^2$,*

$$\frac{\alpha}{8}\|u_n - u_m\|^2 + J(\frac{u_n + u_m}{2}) - \hat{m} \leq \frac{1}{2}(J(u_n) - \hat{m}) + \frac{1}{2}(J(u_m) - \hat{m}).$$

Or, par définition de $\hat{m}$, $J(\frac{u_n+u_m}{2}) \geq \hat{m}$, *si bien que*

$$0 \leq \frac{\alpha}{8}\|u_n - u_m\|^2 \leq \frac{1}{2}(J(u_n) - \hat{m}) + \frac{1}{2}(J(u_m) - \hat{m}).$$

On en déduit que $(u_n)_{n\in\mathbb{N}}$ *est de Cauchy, et donc converge vers une limite* $u^* \in U$ *(*U *est fermé), qui est nécessairement le minimum de* J*, puisque* J *est continue. L'unicité découle du théorème.*
Enfin, soit $u \in U$*. Utilisons encore le caractère* α*-elliptique de* J*, on obtient :*

$$\frac{\alpha}{8}\|u^* - u\|^2 \leq \frac{J(u^*) + J(u)}{2} - J(\frac{u^* + u}{2}) \leq \frac{J(u) - J(u^*)}{2},$$

car $J(\frac{u+u^*}{2}) \geq J(u^*)$.

Remarque 3.0.3. *On peut affaiblir les hypothèses du théorème précédent, en remplacant l'hypothèse de continuité de* J *par une hypothèse de semi-continuité inférieure de* J*. La démonstration reste alors inchangée, et il suffit d'écrire que*

$$J(u^*) \leq \liminf_{n\to+\infty} J(u_n).$$

Le résultat qui suit est essentiel.
En vue d'étendre dans un premier temps ce résultat au cas d'ensembles U non bornés (notamment lorsque $U = V = \mathbb{R}^n$), on introduit la notion suivante :

Proposition 3.0.5. *Soit* $J : \mathbb{R}^n \to \mathbb{R}$ *une fonction continue, coercive sur* $\mathbb{R}^n$*. i.e*

$$\lim_{\|v\|\to\infty} J(v) = +\infty$$

Alors il existe au moins un élément u *tel que*

$$(P) \qquad u \in U \qquad et \qquad J(u) = \inf_{v\in U} J(v)$$

(P) admet une solution optimale u.

Preuve 3.0.7. *Soit* $u \in \mathbb{R}^n$*. Puisque* $\lim_{\|u\|\to\infty} J(u) = +\infty$ *il existe* $N > 0$ *tel que* $\|v\| > N \implies J(v) > M$*, donc*

$$\exists N > 0, \quad J(v) \leq J(v_0) \Rightarrow \|v\| \leq N.$$

Puisque u est caractérisé par $J(u) \leq J(v)$, $\forall x \in R^n$, on a donc forcément $\|u\| \leq N$. Donc u est solution du problème

$$\min_{\|v\| \leq N} J(v)$$

et le théorème précédent s'applique, la boule $\{x \in R^n, \|v\| \leq N\}$ étant compacte.

C'est la compacité qui intervient de façon essentielle dans la démonstration du ($théorème$..8.1.1..). On peut s'en convaincre autrement par la considération la définition suivante :

Définition 3.0.5. ***(Suite minimisante)*** *On dit qu'une suite $(u_k)_{k\in\mathbb{N}}$ d'éléments de U est une suite minimisante si et seulement si*

$$u_k \in U \quad \textit{pour tout} \;\; k \geq 0, \qquad \lim_{k\to\infty} J(u_k) = \inf_{v\in U} J(v)$$

Cette suite étant nécessairement bornée, puisque la fonctionnelle J est coercive, on peut extraire une suite $(u_{k'})$ qui converge vers un élément $u \in U$ (l'ensemble U est fermé). La fonction J étant continue,

$$J(u) = \lim_{k\to\infty} J(u_k) = \inf_{v\in U} J(v),$$

ce qui fournit une nouvelle preuve de l'existence d'une solution du problème (P).
C'est d'ailleurs ce type de raisonnement qui permet d'étendre le résultat au cas de la dimension infinie, avec néanmoins des hypothèses supplémentaires, et essentielles, de convexité, aussi bien pour la fonctionnelle J que pour l'ensemble U. La démonstration reposant sur la compacité "faible" des parties convexes fermées bornées des espaces de Hilbert (parties (ii) et (iii) de la démonstration ci-dessous), nous commençons par la définition suivante :

Définition 3.0.6. *On dit qu'une suite $(u_k)_{k\geq 0}$ d'éléments d'un espace préhilbertien V converge faiblement s'il existe un élément $u \in V$ tel que*

$$\lim_{k\to\infty} (v, u_k) = (v, u) \;\; \text{pour tout} \;\; v \in V.$$

On notera que, si toute suite qui converge au sens de la norme converge faiblement, l'inverse n'est pas toujours vrai.

Théorème 3.0.4. *Soit U une partie non vide, convexe, fermée, d'un espace de Hilbert séparable V, et $J : V \to \mathbb{R}$ une fonctionnelle convexe, dérivable, coercive si l'ensemble U est non borné. Alors il existe au moins un élément u tel que*

$$u \in U \quad et \quad J(u) = \inf_{v\in U} J(v). \qquad (P)$$

Preuve 3.0.8. *(i) Comme dans le cas de la dimension finie (théorème.3.0.5), la coercivité de la fonctionnelle permet de se ramener au seul cas d'un ensemble U borné (et encore convexe puisqu'une boule est convexe).*
(ii) Considérons une suite minimisante $(u_k)_{k\geq 0}$:

$$u_k \in U \quad pour\ tout \quad k \geq 0, \quad \lim J(u_k) = \inf_{v\in U} J(v),$$

sans exclure à ce stade l'éventualité où $\inf_{v\in U} J(v) = -\infty$. La suite (u_k) étant bornée (d'après (i)), montrons qu'on peut en extraire une suite qui converge faiblement.
Soit C une constante telle que $||u_k|| \geq C$ pour tout $k \geq 0$. On note pour commencer que, si v est un élément quelconque de l'espace V, la suite de nombres réels $\{(v, u_k)\}_{k\geq 0}$ est bornée puisque $|(v, u_k)| \leq C||v||$. L'espace Fêtant supposé séparable, soit $(v_k)_{k\geq 0}$ un ensemble dénombrable dense. La suite $\{(v_1, u_k)\}_{k\geq 0}$ étant bornée, on peut en extraire une suite $\{(v_1, u_{k_1})\}_{k_1\geq 0}$ convergente ; de même, la suite $\{(v_2, u_{k1})\}_{k_2\geq 0}$ étant bornée, on peut en extraire une suite $\{(v_2, u_{k2})\}_{k_2\geq 0}$ convergente, et ainsi de suite.
Considérons la suite "diagonale" $(w_i)_{i\geq 0}$ où $W_i =^{def} u_{i_I}$ Par construction, chaque suite $\{(v_k, w_i)\}_{i\geq 0}, k \geq 0$, a une limite, qui est la limite de la suite $\{(v_k, w_{ik})\}_{ik\geq 0}$. On va montrer qu'en fait toute suite $\{(v, w_I)\}_{I\geq 0}, v \in V$ a une limite : étant donné un élément quelconque $v \in V$, soit en effet $\varepsilon \succ 0$ donné. Il existe un élément v_k tel que $||v - v_k|| \leq \varepsilon/4C$. Dans ces conditions,

$$\begin{aligned} |(v, w_I)| - |(v, w_m)| &= |(v, w_I - w_m)| \leq |(v_k, w_I - w_m)| + |(v - v_k, w_I - w_m)| \\ &\leq |(v_k, w_I) - (v_k, w_m)| + \frac{\varepsilon}{2}, \end{aligned}$$

puisque $\|w_l - wm\| \leq ||w_I||+\|w_m\| \leq 2C$. L'élément v_k étant fixé, la suite $\{(v, w_I)\}_{I\geq 0}$ converge d'après ce qui précède ; c'est donc une suite de Cauchy. Par suite, il existe un entier $I_0 = I_0(\varepsilon, v_k)$ tel que

$$I, k \geq I_0 \Rightarrow |(v_k, w_I) - (v_k, w_m)| \leq \frac{\varepsilon}{2},$$

et l'assertion est établie.
Définissons une application $f : V \to \mathbb{R}$ *par*

$$f(v) = \lim_{I\to\infty} (v, w_I) \quad \textit{pour tout} \quad v \in V.$$

C'est une application linéaire, et continue puisque

$$|(v, w_I)| \leq C \|v\| \quad \textit{pour tout} \quad I \Rightarrow |f(v)| \leq C \|v\|.$$

D'après le théorème de représentation de Riesz, il existe un élément $u \in V$ *tel que* $f(v) = (v, u)$ *pour tout* $v \in V$ *; on a donc bien établi la convergence faible de la suite extraite* $(w_j) = (u_{ij},)$ *vers l'élément* u.
(iii) Démontrons ensuite que la limite "faible" u *de la suite extraite* (w_I) *appartient à l'ensemble* U. *Notons* P *l'opérateur de projection associé à l'ensemble convexe fermé* U *; d'après le (théorème.....)(2),*

$$w_I \in U \Rightarrow (Pu - u, W_i - Pu) \geq 0 \quad \textit{pour tout integral } I.$$

La convergence faible de la suite (w_I) *vers l'élément* u *entraîne*

$$0 < \lim_{I\to\infty} (Pu - u, w_l - Pu) = (Pu - u, u - Pu) = -\|u - Pu\|^2 \leq 0,$$

et donc $u \in U$. *On a ainsi établi qu'un ensemble fermé convexe est "faiblement" fermé, c'est-à-dire que la limite "faible" d'une suite faiblement convergente de points d'un tel ensemble lui appartient,*
(iv) Montrons enfin que la fonctionnelle J *vérifie*

$$J(v) \leq \lim_{I\to\infty} \inf J(v_I),$$

pour toute suite (v_I) *convergeant faiblement vers un élément* v. *La fonction* J *étant supposée dérivable et convexe, on a en effet (théorème.2.3.2)*

$$J(v) + (\nabla J(v), v_i - v) \leq J(v_I) \quad \textit{pour tout intégrale } I.$$

et, par définition de la convergence faible,

$$\lim_{I\to\infty} (\nabla J(v), v_I) = (\nabla J(v), v),$$

ce qui établit la propriété annoncée ; on l'appelle la faible semi-continuité inférieure séquentielle de la fonctionnelle J.
(v) Il est maintenant facile de conclure : la limite faible $u \in U$ de la suite extraite (w_I) de la suite minimisante (u_k) vérifie

$$J(u) < \lim_{I\to\infty} \inf J(w_I) = \lim_{k\to\infty} J(u_k) = \inf_{v\in U} J(v).$$

Remarque 3.0.4. (1) *Le théorème reste vrai dans les espaces de Banach réflexifs, dont les espaces de Hilbert (séparables ou non) sont des cas particuliers ; de même, il reste vrai si on remplace l'hypothèse de dérivabilité de la fonction J par la seule continuité.*
(2) La réciproque de la propriété (ii) est vraie (toute suite faiblement convergente est bornée), mais elle ne peut pas s'établir de façon élémentaire.

Dans certains cas particuliers, la démonstration de l'existence d'une solution peut être notablement simplifiée, en évitant notamment tout recours à la convergence faible. Commençons par une définition :

Définition 3.0.7. *Etant donné un espace de Hilbert V, une fonction $J : V \to U$ est appelée fonctionnelle quadratique sur V si elle est de la forme*

$$J(v) = \frac{1}{2}a(v,v) - f(v),$$

où $a(.,.) : V \times V \to \mathbb{R}$ est une forme bilinéaire, continue, symétrique $(a(u,v) - a(v,u)$ pour tout $u, v \in V$ et $f : V \to \mathbb{R}$ est une forme linéaire continue. Cette définition généralise de façon naturelle celle d'une fonctionnelle quadratique sur $\mathbb{R}^n$ puisque, grâce au théorème de représentation de Riesz, il existe un opérateur $A \in \mathcal{L}(V)$ et un élément $b \in V$,tous deux définis de façon unique, tels que

$$a(u,v) - < Au, v > = < u, Av > \quad \text{pour tout } \; u, v \in V$$

$$f(v) = < b, v > \quad \text{pour tout } \; v \in V,$$

en désignant par $< .,. >$ le produit scalaire de l'espace V.

Le théorème de projection et le théorème de représentation de Riesz permettent alors d'établir simplement un résultat général d'existence pour des problèmes (P) posés avec de telles fonctionnelles. On notera que le cas $U = V$ correspond exactement à la formulation variationnelle des problèmes aux limites.

Théorème 3.0.5. *8.2.3Soit*

$$J : v \in V \to J(v) = \frac{1}{2}a(v,v) - f(v),$$

une fonctionnelle quadratique sur un espace de Hilbert V. On suppose de plus qu'il existe un nombre a tel que

$$a > 0 \quad et \quad a(v,v) \geq \alpha ||v||_V^2 \quad pour\ tout \quad v \in V.$$

Étant donné une partie non vide, convexe, fermée U de V, il existe un et un seul élément u vérifiant (P)

$$u \in U \quad et \quad J(u) = \inf_{v \in U} J(v).$$

Cet élément u vérifie également

$$a(u, v-u) \geq f(v-u) \quad pour\ tout \quad v \in U,$$

et, réciproquement, si un élément $u \in V$ vérifie les inéquations ci-dessus, c'est la solution du problème (P). Si U est un sous-espace vectoriel, les inéquations précédentes sont remplacées par les équations

$$a(u,v) - f(v) \quad pour\ tout \quad v \in U.$$

Preuve 3.0.9. *La forme bilinéaire $a(.,.)$ est également un produit scalaire sur l'espace V, la norme associée étant équivalente à la norme $||.\,||$ associée au produit scalaire $< .,. >$ de l'espace V. En effet, les hypothèses faites entraînent :*

$$\sqrt{\alpha}\, \|v\| \leq \sqrt{a(v,v)} \leq \sqrt{||a||}||v||,$$

en désignant par $||a||$ la norme (dans l'espace $\mathcal{L}_2(V,\mathbb{R})$) de l'application bilinéaire $a(v)$. La forme linéaire f étant donc encore continue pour cette nouvelle norme, le *théorème de représentation de Riesz montre qu'il existe un élément de V et un seul tel que.*

$$f(v) = a(c, v) \quad \textit{pour tout } v \in V.$$

Par suite, on peut transformer l'expression de la fonctionnelle, en l'écrivant

$$J(v) = \frac{1}{2}a(v, v) - a(c, v) = \frac{1}{2}a(v - c, v - c) - a(c, c).$$

Dans ces conditions, résoudre le problème (P) revient à chercher la projection u de l'élement c sur l'ensemble U au sens du produit scalaire $a(.,.)$. D'après le théorème de projection, il en existe une et une seule, ce qui établit l'existence et l'unicité de la solution u du problème (P). D'après le même théorème, cette solution est également caractérisée par les inéquations

$$a(u - c, v - u) \geq 0 \quad \textit{pour tout} \;\; v \in U,$$

ou par les équations

$$a(u - c, v) = 0 \quad \textit{pour tout} \;\; v \in U,$$

si U est un sous-espace vectoriel, relations qui coïncident avec celles de l'énoncé puisque

$$a(c, v) = f(v) \quad \textit{pour tout} \;\; ve \in V$$

Remarque 3.0.5. (1) *Un usage essentiel de la symétrie de la forme bilinéaire a été fait, d'une part, pour conclure que l'expression $a(.,.)$ est un produit scalaire, d'autre part, pour écrire la nouvelle expression de la fonctionnelle.*
(2) Les inéquations $a(u, v - u) \geq f(v - u)$ sont un cas particulier des inéquations d'Euler $\nabla J(u)(v-u) \geq 0$ (théorème.2.3.4) appliquées à la fonctionnelle J, de dérivée donnée par

$$\nabla J(u)v = a(u, v) - f(v) \quad \textit{pour tout} \;\; v \in V.$$

Une observation analogue (et pour cause...) a été faite à propos du théorème de projection.
(3) On a indiqué aux paragraphes 3.4 et 3.5 les raisons pour lesquelles les relations $a(u, v) = f(v)$ sont des équations "variationnelles"; c'est dans le même esprit que les relations $a(u, v - u) \geq f(v - u)$ pour tout $v \in U$ sont appelées des inéquations variationnelles.

Références

[1] P. G. Ciarlet, , Introduction to Numerical Linear Algebra and Optimization. Cambridge etc., Cambridge University Press 1989.

[2] G. Allaire , Analyse numérique et optimisation, éditions de l'école Polytechnique, 2005.

[3] C. Zuily, H Queffélec , Analyse pour l'agrégation, 3ème édition, Dunod, 2007.

[4] G.B.BANTZIG, maximisation of lineair fonction of variables subject to lineaire inequalities,dans T.J.C.Koopmans,activity analysis of production and allocation NEW.YORK,1951.

[5] GLASS O.Analyse fonctionnelle et équations aux dérivées partielles,2014-2015.

[6] CULIOLI J.C., introduction à l'optimisation, ELLIPSE 1994.

[7] HIRART-URRUTY J-B, optimisation et analyse convexe.mathématique,PUF,1998.

[8] M.BERGOUNIOUX, optimisation et controle des systèmes lineaire, DOUNOD, 2001.