كلية العلوم و التقنيات فاس
ⵜⴰⵙⵉⵡⴰⵍⵜ ⵏ ⵜⵎⴰⵙⵙⴰⵏⵉⵏ ⴷ ⵜⵉⵏⵉⴽⵉⵏ
Faculté des Sciences et Techniques de Fès

جامعة سيدي محمد بن عبد الله
Université Sidi Mohamed Ben Abdellah

Licence Sciences et Techniques (LST)

MATHEMATIQUES ET APPLICATIONS

MEMOIRE DE FIN D'ETUDES

Pour l'obtention du Diplôme de Licence Sciences et Techniques

Titre

Dualité en optimisation numérique Application au problème de portefeuille en finance

- ***Réalisé par*** : **BARBARA IDRISS**
- ***Encadré par*** : **Pr. BEN AICHA KHADIJA**

Soutenu Le 9 Juin 2018 devant le jury composé de:

- **BEN AICHA KHADIJA**
- **RAHMOUNI HASSANI AZIZA**
- **EZZAKI FATIMA**

Année Universitaire 2017 / 2018

FACULTE DES SCIENCES ET TECHNIQUES FES – SAISS
B.P. 2202 – Route d'Imouzzer – FES
212 (0)5 35 61 16 86 – Fax : 212 (0)5 35 60 82 14
Site web : http://www.fst-usmba.ac.ma

REMERCIEMENT

Au terme de ce travail, je tiens à exprimer mon remerciement les plus chaleureux à mon encadrante, Mme BEN AICHA Khadija, d'avoir accepter de m' encadrer, pour sa compétence, et ses bonnes directives, qu'elle trouve ici l'expression de ma profonde gratitude

Je tiens également à exprimer toute ma gratitude et tout mon sentiment de reconnaissance aux membres du jury composé de : Pr. **Ben Aicha Khadija**, Pr. **Fatima Ezzaki**, et Pr. **Aziza Rahmouni Hassani**, pour avoir accepter d'examiner ce travail et de l'enrichir par leurs propositions.

J'adresse mes remerciements à toute personne qui a contribué d'une façon ou d'une autre à la réalisation de ce travail.

Finalement, je ne peux oublier de remercier profondément les êtres les plus chers, mes parents, pour leur soutien, leurs encouragements et la confiance qu'ils m'accordaient depuis toujours, ainsi que mon ami El Mehdi Amhraoui pour ses conseils, et tous chers amis pour tous les bons moments passés ensembles.

MERCI.

Table des matières

Introduction **3**

1 Notions fondamentales **4**
1.1 Éléments de Topologie 4
1.1.1 la convergence des suites dans $\mathbb{R}^n$ 4
1.1.2 Ensembles ouverts,Ensembles fermés 5
1.1.3 Ensembles compacts 6
1.1.4 Notions de minimum, maximum, infimum, supremum 6
1.2 Éléments d'analyse convexe 8
1.2.1 Ensemble convexe 8
1.2.2 fonction convexe 8
1.2.3 Programme convexe 9
1.2.4 Le gradient 9

2 Analyse du problème **11**
2.1 Formulation du problème 11
2.2 Existence et unicité de solution 12
2.3 Conditions d'optimalités 15
2.4 Le Théorème de Projection sur un convexe fermé 18
2.4.1 Algorithme du gradient à pas fixe avec projection sur $\mathbb{U}$ (GPF) 19

3 Dualité Lagrangienne classique **21**
3.1 Lagrangien et Point-selle 21
3.1.1 Le Lagrangien 21
3.1.2 Points-selles du Lagrangien 22
3.1.3 Condition d'optimalité pour le Point-selle 24
3.2 Dualité 27
3.2.1 Problème Primal et Problème Dual 27
3.2.2 Relations de Kuhn-Tucker 30

4 La Méthode d'Uzawa et son application au problème du Portefeuille en finance **32**
4.1 La Méthode d'Uzawa 32
4.1.1 Algorithme d'Uzawa 33
4.1.2 L'algorithme d'Uzawa dans le cas d'une fonction quadratique 34
4.2 Problème du Portefeuille en finance 36
4.2.1 Enoncé 36
4.2.2 Modélisation 36
4.2.3 Résolution 37

Introduction

La résolution des problèmes d'optimisation est utilisée dans un grand nombre de domaines (physiques, mécaniques, chimiques, biologiques et économiques).L'homme cherche à optimiser : minimiser un coût ou bien maximiser un gain ou un profit, pour se faire il a mis en évidence certaines procédures qui permettent d'optimiser tout problème rencontré. C'était en fait un nouveau domaine de recherche en mathématiques appliquées qui a vu le jour avec la recherche opérationnelle. Le développement de l'informatique a ouvert de nouveaux horizons à la résolution de ces problèmes, et a permis un élargissement massif des champs d'application de ces techniques.

Le problème d'optimisation consacré à l'étude du minimum/maximum d'une fonction à une ou plusieurs variables sur un certain domaine de définition, de l'étude de leur existence à leur détermination , en général par la mise en œuvre d'un algorithme et par suite un programme.

Dans le présent travail, nous nous intéressons aux problèmes d'optimisation numérique avec contraintes, agissants sur des fonctionnelles non-linéaires, continues et différentiables

Le premier chapitre commence par traiter des notions mathématiques fondamentales à maîtriser avant de s'intéresser à la résolution à proprement parler du problème d'optimisation, ainsi que la notion de solution locale et globale et quelques éléments d'analyse convexe.

Dans les chapitres suivants, nous entrons dans le vif du sujet en nous intéressant à la résolution des problèmes d'optimisation non linéaire avec contraintes, tout en introduisant la méthode d'Uzawa de résolution appropriée.

Ainsi, pour terminer ce mémoire, nous allons présenter une Application au problème du Portefeuille en finance.

Chapitre 1

Notions fondamentales

1.1 Éléments de Topologie

1.1.1 la convergence des suites dans $\mathbb{R}^n$

Une suite infinie de vecteurs de $\mathbb{R}^n$: $x^1, x^2, ..., x^k, ...$ sera noter $\{x^k\}_{k\in\mathbb{N}}$ ou simplement $\{x^k\}$ lorsqu'il n'y a pas d'ambiguïté sur l'ensemble des indices .

Définition 1 *On dit que la suite $\{x^k\}$ converge vers x (ou que x la limite de la suite $\{x^k\}$) si la norme $||x^k - x||$ tend vers 0 quand $k \to +\infty$ ou, de façon équivalente,*
Si :

$$\forall \varepsilon > 0, \quad \exists N \in \mathbb{N} \quad \textit{tel que} \quad k \geq N \Rightarrow ||x^k - x|| < \varepsilon.$$

Si x est la limite de la suite $\{x^k\}$ on notera $x^k \ \to \ x$.

Théorème 1 *(Unicité de la limite)*
La limite d'une suite réelle, si elle existe, est unique.

Preuve. Supposons que $\{x^k\}$ possède deux limites $x_1, x_2 \in \mathbb{R}^n$ et montrons par l'absurde que $x_1 = x_2$.
Supposons $x_1 \neq x_2$, et posons $\varepsilon = \dfrac{||x_1 - x_2||}{2}$
Puisque on a $x^k \ \to \ x_1$ donc

$$\exists N_1 \in \mathbb{N} \quad \textit{tel que} \quad k \geq N_1 \Rightarrow ||x^k - x_1|| < \varepsilon$$

et puisque on a $x^k \ \to \ x_2$ donc

$$\exists N_2 \in \mathbb{N} \quad \textit{tel que} \quad k \geq N_2 \Rightarrow ||x^k - x_2|| < \varepsilon$$

Et on a

$$|x_1 - x_2| \leqslant |x^k - x_1| + |x^k - x_2|$$

donc

$$|x_1 - x_2| < 2\varepsilon = 2\frac{|x_1 - x_2|}{2}$$

alors

$$|x_1 - x_2| < |x_1 - x_2|$$

ce qui est absurde, donc $x_1 = x_2$. ■

1.1.2 Ensembles ouverts,Ensembles fermés

Soient $x \in \mathbb{R}^n$ *et* $r > 0$. On appelle boule ouverte de centre x et de rayon r l'ensemble :

$$B(x,r) = \{y \in \mathbb{R}^n \;\; tel\ que \;\; ||y - x|| < r\}.$$

Définition 2 ***(Ouvert, fermé)*** *Un ensemble O de $\mathbb{R}^n$ est dit ouvert si pour tout $x \in O$, il existe une boule ouverte $B(x,r)$ de centre x et de rayon r incluse dans O.*

Un ensemble F de $\mathbb{R}^n$ est dit fermé si son complémentaire est un ouvert.

Exemple.
1, Les ensembles $\emptyset$ et $\mathbb{R}^n$ sont ouverts, mais ils sont aussi fermés car leur complémentaires respectifs sont $\mathbb{R}^n$ et $\emptyset$. Ce sont d'ailleurs les seuls exemples d'ensemble à la fois ouverts et fermés.
2, Les boules ouvertes sont des ouverts.

L'interprétation géométrique de "ouvert" et "fermé" est qu'un ensemble est ouvert si et seulement si il ne contient aucun point de sa frontière. Un ensemble est fermé si et seulement si il contient tous les points de sa frontière. Les ensembles qui contiennent une partie de leur frontière mais pas l'intégralité ne sont ni ouverts ni fermés.

Proposition 1 ***(Caractérisation des fermés)*** *Soit $F \subset \mathbb{R}^n$. F est un fermé si et seulement si toute suite convergente d'éléments de F a sa limite dans F.*

Preuve. ($\Rightarrow$) Supposons F fermé et notons O son complémentaire. Soit $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de F qui converge vers $x \in \mathbb{R}^n$. Supposons par l'absurde que $x \in O$.

Comme O est un ouvert de $\mathbb{R}^n$, il existe $r > 0$ tel que : $B(x,r) \subset O$. Or $(x_n)_{n\in\mathbb{N}}$ converge vers x d'où :

$$\exists N \in \mathbb{N}, \forall n \geqslant N, \quad ||x_n - x|| < r,$$

Ceci implique $x_n \in B(x,r) \subset O$ et donc $x_n \in O$ partir d'un certain rang, ce qui est impossible car $x_n \in F$ quel que soit $n \in \mathbb{N}$ par définition.

($\Leftarrow$) Supposons que toute suite convergente de F admet une limite dans F. Montrons que O, le complémentaire de F, est ouvert. Par l'absurde, supposons que O ne soit pas ouvert, i.e. : il existe $x \in O$ tel que : $\forall r > 0$; $B(x,r) \not\subseteq O$. Autrement dit :

$$\forall r > 0 \quad B(x,r) \cap F \neq \emptyset.$$

Pour $r = \dfrac{1}{n+1}$, nous construisons une suite $(x_n)_{n\in mathbbN}$ d'éléments de F telle que

$$x_n \in B(x, \frac{1}{n+1}).$$

Ceci peut être traduit en $||x_n - x|| \leqslant \dfrac{1}{n+1}$. Il suit que x_n converge vers x. Comme F est fermé, $x \in F$, ce qui est impossible car $x \in O$. ■

1.1.3 Ensembles compacts

Définition 3 *Un sous-ensemble $K \subset \mathbb{R}^n$ est dit compact si, de toute suite infinie $\left\{x^k\right\}_{k\in\mathbb{N}}$ d'éléments de K, on peut extraire une sous-suite $\left\{x^l\right\}_{l\in L}$ $(L \subset \mathbb{N})$ converge vers un élément de K.*

La propriété suivant donne une caractérisation simple des compact dans $\mathbb{R}^n$ (et, plus généralement dans les espaces vectoriels normée de dimension finie) :

Propriété 1 *Dans $\mathbb{R}^n$ un sous-ensemble K est compact si et seulement si il est fermé et borné (contenu dans une boule de rayon $M < +\infty$).*

1.1.4 Notions de minimum, maximum, infimum, supremum

Définition 4 ***(Minorant/Majorant)*** *Soit E un sous-ensemble de $\mathbb{R}$, $m \in \mathbb{R} \cup \{-\infty, +\infty\}$ est un minorant de E ssi m est inférieur ou égal à tous les* *éléments de E tandis que $M \in \mathbb{R} \cup \{-\infty, +\infty\}$ est un majorant de E ssi il est supérieur ou égal à tous les éléments de E. Ainsi*

$$(m \text{ minorant } \Leftrightarrow \forall x \in E \quad m \leqslant x) \quad \text{et} \quad (M \text{ majorant } \Leftrightarrow \forall x \in E \quad M \geqslant x).$$

Si E admet un minorant (resp. majorant) fini alors il est dit minoré (resp. majoré).

Définition 5 ***(Infimum/Supremum)*** *Soit $E \subset \mathbb{R}$,*
L'infimum $inf(E) \in \mathbb{R} \cup \{-\infty, +\infty\}$ de E est le plus grand des minorants. Le supremum $sup(E) \in \mathbb{R} \cup \{-\infty, +\infty\}$ de E est le plus petit des majorants. On les note respectivement

$$inf(E) = \underset{x \in E}{inf}(x) \qquad sup(E) = \underset{x \in E}{sup}(x).$$

Définition 6 ***(Suite minimisante/Suite maximisante)*** *Soit $E \subset \mathbb{R}$, $E \neq \emptyset$*
Il existe $(x_n)_{n \in \mathbb{N}}$ et $(y_n)_{n \in \mathbb{N}}$ deux suites d'éléments de E telle que

$$\lim_{n \to +\infty} x_n = inf(E), \qquad \lim_{n \to +\infty} y_n = sup(E)$$

La suite $(x_n)_{n \in \mathbb{N}}$ est appelée "suite minimisante" et $(y_n)_{n \in \mathbb{N}}$ une "suite maximisante" de E.

Proposition 2 *Soit $F \subset \mathbb{R}$ un fermé non vide de $\mathbb{R}$.*
- *Si F est minoré alors F admet un minimum et ainsi*

$$inf(F) = min(F) \in F$$

- *Si F est majoré alors F admet un maximum et ainsi*

$$sup(F) = max(F) \in F$$

.

Preuve. Soit $(x_n)_{n \in \mathbb{N}}$ une suite minimisante

$$\lim_{n \to +\infty} x_n = inf(F),$$

Comme F est fermé, $inf(F) \in F$ et, par conséquent, $inf(F)$ est le minimum de F. La démonstration concernant le maximum est similaire. ■

Théorème 2 ***(Bolzano-Weierstrass)*** *Toute suite bornée $(x_n)_{n \in \mathbb{N}}$ de $\mathbb{R}^n$ admet une sous-suite convergente.*

1.2 Éléments d'analyse convexe

1.2.1 Ensemble convexe

Définition 7 *Un sous ensemble C est dite convexe si $\forall x, y \in C$ le segment $[x, y]$ est inclus dans C*

c'est à dire

$$\forall t \in [0,1] \quad tx + (1-t)y \in C.$$

A un ensemble convexe et B non convexe.

Exemple : $\mathbb{R}^n$ est un ensemble convexe.

Propriété 2 *1, L'intersection d'un nombre fini quelconque d'ensemble convexes est convexe.*

2, Tout sous-ensemble convexe non vide de $\mathbb{R}^n$ a un intérieur relatif non vide .

1.2.2 fonction convexe

Définition 8 *Une fonction J définie sur un sous ensemble convexe C est dite convexe si*

$$\forall x, y \in C, \quad \forall t \in [0,1] \quad J(tx + (1-t)y) \leqslant tJ(x) + (1-t)J(y),$$

La fonction est dite strictement convexe si et seulement si :

$$\forall x, y \in C, \; x \neq y, \quad \forall t \in]0,1[\quad J(tx + (1-t)y) \leqslant tJ(x) + (1-t)J(y)$$

Proposition 3 *Soit une fonction différentiable définie sur un convexe C de $\mathbb{R}^n$ alors J est convexe si et seulement si :*

$$\forall x, y \in C, \quad \nabla J(x)(y-x) \leqslant J(y) - J(x).$$

Preuve. Soit J convexe. On considère $t \in [0,1]$, et $x, y \in C$.
Puisque C est convexe : $x + t(y-x) \in C$ et en utilisant la convexité de J on a

$$J(x + t(y-x)) \leqslant (1-t)J(x) + tJ(y)$$

D'où :

$$J(y) - J(x) \geqslant \frac{J(x + t(y-x)) - J(x)}{t}$$

On conclut alors en faisant tendre t vers 0. Réciproquement, on part des deux inégalités :

$$J(y) \geqslant J(y + t(x-y) - tJ(y + t(x-y)(x-y)$$
$$J(x) \geqslant J(y + t(x - y + (1-t)J(y + t(x-y)(x-y).$$

En combinant ces deux inégalités on obtient
$J(x + t(y-x)) \leqslant (1-t)J(x) + tJ(y)$, ce qui établit la convexité. ■

Proposition 4 *Soit J une fonction convexe définie sur un ensemble convexe C Alors ;*

- *minimum local de J sur C est un minimum global.*
- *Si J est strictement convexe, il y'a au plus un minimum global (unicité du minimum global).*

1.2.3 Programme convexe

Définition 9 *On dit qu'un programme de programmation mathématique est convexe s'il consiste à minimiser une fonction convexe(resp : maximiser une fonction concave) sur un domaine convexe*

Propriété 3 *Pour un programme convexe, tout optimum local est optimum globale.*

1.2.4 Le gradient

Définition 10 *On considère $\mathbb{R}^n$ muni et de son produit scalaire canonique*

$$< (x_1, x_2, ..., x_n), (y_1, y_2, ..., y_n) > = \sum_{i=1}^{n} x_i y_i$$

On définit le gradient de J au point x^* comme étant le vecteur

$$\nabla J(x^*) = \begin{pmatrix} \dfrac{\partial J}{\partial x_1}(x^*) \\ \dfrac{\partial J}{\partial x_2}(x^*) \\ \cdot \\ \cdot \\ \\ \dfrac{\partial J}{\partial x_n}(x^*) \end{pmatrix}$$

Pour une fonction convexe J différentiable, le gradient $\nabla J(x^*)$ en tout point x^* vérifie l'inégalité fondamentale

$$J(x) \geqslant J(x^*) + \nabla J^T(x^*)(x - x^*).$$

Chapitre 2

Analyse du problème

2.1 Formulation du problème

Considérons le problème d'optimisation suivant

$$(P) \qquad \begin{cases} inf\, J(x) \\ x \in \mathbb{U} \end{cases}$$

dans lequel on minimise une fonction $J : \Omega \mapsto \mathbb{R}$ (avec Ω un ouvert de $\mathbb{R}^n$), est appelée **fonction objectif** ou **fonction coût**, sous une **ensemble des contraintes**

$$\mathbb{U} = \{x \in \mathbb{R}^n \ / \ F(x) \leq 0 \ \ et \ \ G(x) = 0\}$$

appelée aussi **ensemble des points admissibles** ou **réalisable**.

avec $F : \mathbb{R}^n \mapsto \mathbb{R}^p$ $\quad F(x) = (F_1(x), ..., F_p(x))^T$ *la fonction définissant les contraintes d'inégalités*
et $\quad G : \mathbb{R}^n \mapsto \mathbb{R}^q$, $\quad G(x) = (G_1(x), ..., G_q(x))^T$ *la fonction définissant les contraintes d'égalités*

Bien évidemment, on définit un problème de maximisation, en remplaçant $\leq$ par $\geq$ dans les contraintes d'inégalités et **inf** par **sup** dans fonction objectif.
On parlera en général de problème d'optimisation. On passe de l'un à l'autre en définissant la fonctionnelle opposée, Dans ce mémoire tous les résultats sont établis sur les problèmes de minimisation.

Résoudre le problème (P) revient à chercher des points de minimum local (ou global, c'est encore mieux !) au sens de la définition suivante :

Définition 11 *: (Minimum local/Minimum global)*

- *On appelle minimum (global) de (P) tout point $x^* \in \mathbb{U}$ vérifiant*

$$J(x^*) \leq J(x) \quad \forall x \in \mathbb{U},$$

- *On dit que $x^* \in \mathbb{U}$ est un minimum local de (P) s'il existe un voisinage V de x^* tel que*

$$J(x^*) \leq J(x) \quad \forall x \in V \cap \mathbb{U}.$$

Définition 12 *On dit qu'un problème d'optimisation est convexe si sa fonction cout et ses contraintes d'inégalités sont convexes et si ses contraintes d'égalités sont linéaires.*

2.2 Existence et unicité de solution

Dans ce paragraphe, nous nous intéressons aux propriétés de la fonction objectif J et de l'ensemble des contraintes $\mathbb{U}$ afin d'en déduire des conditions suffisantes d'existence d'un minimum de l'ensemble $J(\mathbb{U})$

Existence

Il existe principalement deux théorèmes donnant des conditions suffisantes d'existence d'un point de minimum : le premier dans le cas où l'ensemble des contraintes est fermé borné, le second pour un ensemble de contraintes fermé mais non borné.

Cas où l'ensemble $\mathbb{U}$ des contraintes est borné :

Nous nous intéressons d'abord au cas où $\mathbb{U}$ est borné. Le premier résultat est que l'image d'un fermé borné par une application continue est un fermé borné, i.e. si J est continue et $\mathbb{U}$ fermé borné, alors $J(\mathbb{U})$ est fermé borné.

Théorème 3 *(Théorème de Weierstrass)*

Soit $\mathbb{U}$ un ensemble fermé borné non vide de $\mathbb{R}^n$ et $J : \mathbb{U} \subset \mathbb{R}^n \longmapsto \mathbb{R}$ une application continue sur $\mathbb{U}$.

Alors J est bornée et atteint ses bornes. Autrement dit, il existe $x \in \mathbb{U}$ point de minimum global de J sur $\mathbb{U}$ i.e. :

$$\forall y \in \mathbb{U},\ J(x) \leq J(y)$$

Preuve :

Soit $(x_n)_{n\in\mathbb{N}}$ une suite dans $\mathbb{U}$ tel que

$$\lim_{n\to+\infty} J(x_n) = inf\, J(\mathbb{U})$$

Comme $\mathbb{U}$ est fermé borné, il existe une sous-suite extraite $(x_{\varphi(n)})_{n\in\mathbb{N}}$ qui converge vers un $x \in \mathbb{U}$, Cette suite extraite vérifie

$$x_{\varphi(n)} \longrightarrow x \quad et \quad J(x_{\varphi(n)}) \longrightarrow \inf_{y\in\mathbb{U}} J(y)$$

Or J est continue, d'où par unicité de la limite, il suit

$$J(x) = \inf_{y\in\mathbb{U}} J(y) \quad avec\ x \in \mathbb{U}$$

et J réalise son minimum sur $\mathbb{U}$. ■

Cas où l'ensemble $\mathbb{U}$ des contraintes est non borné :

Dans cette section nous nous intéressons au cas où l'ensemble des contraintes est non borné.
C'est le cas où qu'il existe un infimum mais qu'il n'est pas atteint car celui-ci se trouve à l'infini.

Définition 13 *une application $J\ :\ \mathbb{U} \subset \mathbb{R}^n \longmapsto \mathbb{R}$ est dite infinie à l'infini (ou **coercive**) sur $\mathbb{U}$ ssi*

$$\lim_{||x||\to+\infty} J(x) = +\infty \quad avec\ x \in \mathbb{U}.$$

Théorème 4 *Soient Ω un fermé non vide de $\mathbb{R}^n$ et $J : \Omega \to \mathbb{R}$ une application coercive et continue sur Ω Alors J admet un point de minimum global sur Ω, i.e. il existe $x \in \Omega$ tel que*

$$\forall y \in \Omega,\ J(x) \leq J(y)$$

Preuve : L'idée de la preuve est d'utiliser le théorème de Weierstrass (Théorème 3) en minimisant J sur un fermé borné K que nous construirons et en montrant que minimiser J sur K revient à minimiser J sur Ω

soit K un fermé borné, comme on a $\Omega \neq \emptyset$, alors $inf(\Omega) \in \mathbb{R} \cup \{-\infty\}$.
soit $A \in \mathbb{R}$, tel que $A > \underset{y\in\Omega}{inf} J(y)$,Comme J est infinie à l'infini, il existe $R_1 > 0$ tel que pour $y \in \mathbb{R}^n$

$$||y|| > R_1 \Rightarrow J(y) > A$$

De plus Ω est non vide : il existe donc $R_2 > 0$ tel que

$$B(0, R_2) \cap \Omega \neq \emptyset$$

$B(0, R_2)$ la Boule fermée de centre 0 et de rayon R_2,Choisissons $R = \max(R_1,R_2)$

$$||y|| > R \Rightarrow J(y) > A \quad et \quad B(0, R) \cap \Omega \neq \emptyset$$

On pose $K = B(0, R) \cap \Omega$, est un fermé borné non vide car $||y|| \leqslant R$ et l'intersection de deux fermés.
<u>Minimisons J sur K</u>,Comme J est continue et K est fermé borné, J atteint son minimum sur K, i.e. :

$$\exists x \in K \ / \ J(x) = \underset{y\in k}{inf} J(y) \qquad (*)$$

<u>Montrons que minimiser sur K revient à minimiser sur Ω</u>,D'une part, nous avons

$$\underset{y\in\Omega}{inf} J(y) = inf \left(\underset{y\in K}{inf} J(y), \underset{y\in\Omega\backslash K}{inf} J(y) \right)$$

D'autre part, pour $z \in \Omega$ et $z \notin K$, *on a* $||z|| > R$ soit :

$$J(z) > A > \underset{y\in\Omega}{inf} J(y)$$

Par conséquent : $\underset{y\in\Omega}{inf} J(y) < \underset{y\in\Omega\backslash K}{inf} J(y)$.Il suit

$$\underset{y\in\Omega}{inf} J(y) = \underset{y\in K}{inf} J(y)$$

et d'après (*) il existe $x \in K \subset \Omega \ / \ J(x) = \underset{y\in\Omega}{inf} J(y)$. ■

Propriété 4 *si* $F : \mathbb{R}^n \mapsto \mathbb{R}^p$ *et* $G : \mathbb{R}^n \mapsto \mathbb{R}^q$ *sont deux fonctions continues. alors* $\mathbb{U} = \{x \in \mathbb{R}^n \ / \ F(x) \leq 0 \ \ et \ \ G(x) = 0\}$ *est un ensemble fermé de* $\mathbb{R}^n$.

Ainsi , on peut conclure directement que si J, F et G sont continues et soit

- l'ensemble des contraintes $\mathbb{U} = \{x \in \mathbb{R}^n \ / \ F(x) \leq 0 \ \ et \ \ G(x) = 0\}$ est borné,
- ou J est infinie à l'infini,

  alors le problème (P) admet au moins une solution globale. Autrement dit, J admet au moins un point de minimum global sur $\mathbb{U}$.

Unicité

Remarque 1 *: si F et G sont des fonctions convexes alors l'ensemble $\mathbb{U} = \{x \in \mathbb{R}^n \ / \ F(x) \leq 0 \ \ et \ \ G(x) = 0\}$ est convexe.*

Théorème 5 *Soit $J \in C(R^n, R)$, on suppose que J est strictement convexe et $\mathbb{U}$ convexe*
Alors il existe au plus un élément x^* de $\mathbb{U}$ tel que

$$J(x^*) = \inf_{y \in \mathbb{U}} J(y)$$

Preuve : Raisonnons par absurde,

Supposons que x_1^* et x_2^* soient deux solutions du problème (P),Alors puisque J est strictement convexe

$$J\left(\frac{1}{2}x_1^* \ + \ \frac{1}{2}x_2^*\right) < \frac{1}{2}J(x_1^*) \ + \ \frac{1}{2}J(x_2^*) = \inf_{y \in \mathbb{U}} J(y)$$

D'où la contradiction.
On conclut qu'il existe au plus un élément x^* de $\mathbb{U}$ tel que $J(x^*) = \inf_{y \in \mathbb{U}} J(y)$ ■

2.3 Conditions d'optimalités

Dans cette section nous allons établir les conditions d'optimalité du premier ordre permettant de calculer les éventuels minimums de J.

Lemme 1 *soit $\mathbb{U} \subset \mathbb{R}^n$, $a \in \mathbb{R}^n$ et x^* un élément appartenant a* ***l'intérieure*** *de $\mathbb{U}$, Alors les deux assertions suivantes sont équivalentes.*

1.

$$< a, y - x^* > \geqslant 0, \qquad \forall y \in \mathbb{U} \tag{2.1}$$

2.

$$a = 0. \tag{2.2}$$

Preuve : L'implication $(2,2) \Rightarrow (2,1)$ est évidente.
Montrons l'implication inverse, soit $\omega \in \mathbb{R}^n$ avec $\omega \neq 0$, Alors il existe $\theta_0 > 0$ telle que

$$x^* + \theta\omega \in \mathbb{U} \qquad \forall\ \theta\ \in [-\theta_0, \theta_0]$$

(il suffit de prendre $\theta_0 = \dfrac{r}{2||\omega||}$ où $r > 0$ est tel que $B(x^*, r) \subset \mathbb{U}$).
Alors en prenant $x^*\ +\ \theta\omega$ à la place de y dans (2.1) on déduit

$$< a, \theta\omega >\geqslant 0, \qquad \forall\ \theta\ \in [-\theta_0, \theta_0].$$

En prenant d'abord $\theta = \theta_0$ et en suite $\theta = -\theta_0$ dans l'inégalité précédent,on déduit

$$< a, \omega > =\ 0 \quad \textit{et ceci quelque soit} \quad \omega \in \mathbb{R}^n, \omega \neq 0 \tag{2.3}$$

Il suffit de prendre $\omega = a$ pour trouver $< a, a >= 0$ donc $a = 0$. ■

Définition 14 *Soit $\mathbb{U} \subset \mathbb{R}^n$ un ensemble et $x^* \in \mathbb{U}$. On dit que $\omega \in \mathbb{R}^n$ est une **direction admissible** pour x^* en $\mathbb{U}$ s'il existe $t_0 > 0$ tel que*

$$x^* + t\omega \in \mathbb{U} \qquad \forall t \in [0, t_0].$$

Lemme 2 *soit $\Omega \subset \mathbb{R}^n$ un ouvert, $(\mathbb{U} \subset \Omega)$ $J : \Omega \longmapsto \mathbb{R}$ une fonction de classe C^1 et $x^* \in \mathbb{U}$ un point de minimum relatif de J sur $\mathbb{U}$. soit $\omega \in \mathbb{R}^n$ une direction admissible pour x^* en $\mathbb{U}$. Alors*

$$< \nabla J(x^*), \omega > \geq 0.$$

Preuve : soit V un voisinage de x^* en $\mathbb{R}^n$ tel que

$$f(y) \geqslant f(x^*), \qquad \forall y \in \mathbb{U} \cap V \tag{2.4}$$

Soit $t_1 > 0$ tel que

$$x^* + t\omega, \qquad \forall t \in [-t_1, t_1]$$

(pour $\omega \neq 0$ il suffit de prendre $t_1 = \dfrac{r}{2||\omega||}$ où $r > 0$ est tel que $B(x^*, r) \subset V$; pour $\omega = 0$ tout t_1 convient).
D'autre part on a

$$x^* + t\omega \in \mathbb{U}, \qquad \forall t \in [0, t_0]$$

où t_0 est comme dans la Définition 14.
On déduit alors de (1.4)

$$f(x^* + t\omega) - f(x^*) \geqslant 0, \qquad \forall t \in [0, min\,\{t_0, t_1\}]$$

En divisant par t et en passant à la limite pour $t \to 0$, $t > 0$, on obtient :

$$\frac{\partial f}{\partial \omega}(x^*) \geqslant 0$$

Ce qui est exactement $< \nabla f(x^*), \omega > \geqslant 0$. ■

Théorème 6 ***(condition nécessaire)***
Si J est une fonction différentiable et si $\mathbb{U}$ est convexe fermé, alors toute solution x^* de problème (P) vérifie la condition nécessaire d'optimalité
1.

$$\forall y \in \mathbb{U} \qquad < \nabla J(x^*), y - x^* > \; \geqslant \; 0 \tag{2.5}$$

(c'est ***l'inéquation d'Euler****)*
Si on plus x^* est dans l'intérieur de $\mathbb{U}$ alors la condition (2.5) est équivalent au
2.

$$\nabla J(x^*) = 0 \tag{2.6}$$

(c'est ***l'équation d'Euler****).*

Preuve : 1. C'est une conséquence immédiate du **Lemme 2** et du fait que $y - x^*$ est une direction admissible pour x^* en $\mathbb{U}$ (car $\mathbb{U}$ est convexe).

2. Cette partie est une conséquence immédiate du **Lemme 1** avec $a = \nabla J(x^*)$. ■

Théorème 7 ***(CNS dans le cas convexe)***
Supposons J convexe, différentielle et $\mathbb{U}$ convexe fermé, soit x^* un élément de $\mathbb{U}$. La condition (2.5) est nécessaire et suffisante pour que x^* soit solution de (P).

Preuve : Il reste à montrer que la condition est suffisante.
Soit x^* un élément de $\mathbb{U}$. grâce à la convexité de J,

$$J(y) \geqslant J(x^*) + < \nabla J(x^*), y - x^* > .$$

Il est alors clair que la condition (2.5) est implique que $J(y) \geqslant J(x^*) \quad \forall y \in \mathbb{U}$. ■

Remarque 2 *Soit $\mathbb{U}$ est un cône convexe fermé et si J est convexe et admet un gradient en x^* ($\nabla J(x^*) \in \mathbb{R}^n$) alors (P) admet une solution x^* caractérisée par :*

$$\begin{cases} \nabla J(x^*) y \geq 0 \\ \nabla J(x^*) x^* = 0 \end{cases} \qquad \forall y \in \mathbb{U}$$

2.4 Le Théorème de Projection sur un convexe fermé

On considère un point x dans un espace de Hilbert X et un sous-ensemble $\mathbb{U} \subset X$. Il arrive souvent dans les applications que l'on doive se contenter d'une "approximation" de x par un élément de $\mathbb{U}$, Une idée assez naturelle pour adapter les méthodes d'optimisation sans contraintes aux problème avec contrainte, est projeter à chaque étape le déplacement sur la frontière de domaine, afin de s'assurer que le nouveau point obtenu appartienne à l'ensemble des solutions réalisables.

Au sens de la distance de l'espace de Hilbert, la meilleure approximation de x est alors l'élément $y \in A$ qui minimise la distance $||y-x||$ (cette quantité mesurant l'erreur faite dans l'approximation de x par y). Lorsque $\mathbb{U}$ est convexe et fermé, on va voir que ce problème d'approximation est bien posé en ce sens qu'il y a l'existence et l'unicité de y.

Rappelons d'abord le résultat fondamental suivant :

Théorème 8 *(Théorème de projection sur un convexe fermé)*

Soit $\mathbb{U} \subset \mathbb{R}^n$ un ensemble convexe, fermé, non-vide et soit $v \in \mathbb{R}^n$ Alors il existe un unique élément on peut le noter pour plus de précision par $\mathbb{P}_{\mathbb{U}}(v)$ tel que

$$||v - \mathbb{P}_{\mathbb{U}}(v)|| = \inf_{u \in \mathbb{U}} ||v - u||$$

L'élément $\mathbb{P}_{\mathbb{U}}(v) \in \mathbb{U}$ s'appelle ***la projection de v sur $\mathbb{U}$ en*** *$\mathbb{R}^n$ En plus on a :*
a) L'élément $\mathbb{P}_{\mathbb{U}}(v) \in \mathbb{U}$ satisfait aussi

$$< v - \mathbb{P}_{\mathbb{U}}(v), u - \mathbb{P}_{\mathbb{U}}(v) > \leqslant 0 \qquad \forall u \in \mathbb{U} \tag{2.7}$$

b) Si $w \in \mathbb{U}$ tel que

$$< v - w, u - w > \leqslant 0, \qquad \forall u \in \mathbb{U} \tag{2.8}$$

Alors $w = \mathbb{P}_{\mathbb{U}}(v)$ (donc $w = \mathbb{P}_{\mathbb{U}}(v)$ est l'unique élément de $\mathbb{U}$ satisfaisant (2.7)
c) On a

$$||\mathbb{P}_{\mathbb{U}}(v_1) - \mathbb{P}_{\mathbb{U}}(v_2)|| \leqslant ||v_1 - v_2||, \qquad \forall v_1, v_2 \in \mathbb{R}^n$$

(la fonction $\mathbb{P}_{\mathbb{U}}$ n'augmente pas les distances). Ceci implique que $\mathbb{P}_{\mathbb{U}}$ est une fonction lipschitzienne, donc ***continue****.*
d) On a

$$v = \mathbb{P}_{\mathbb{U}}(v) \quad \text{si et seulement si} \quad v \in \mathbb{U}$$

e) Si $\mathbb{U}$ est le sous-espace affine fermé de $\mathbb{R}^n$ donné par $\mathbb{U} = a + \mathbb{U}_0$ avec $a \in \mathbb{R}^n$ et $\mathbb{U}_0$ un sous-espace vectoriel fermé de $\mathbb{R}^n$ alors l'inégalité (2.8) devient l'égalité

$$< v - \mathbb{P}_{\mathbb{U}}(v), u >= 0 \qquad \forall u \in \mathbb{U}_0$$

(c'est à dire $v - \mathbb{P}_{\mathbb{U}}(v) \perp \mathbb{U}_0$).

2.4.1 Algorithme du gradient à pas fixe avec projection sur $\mathbb{U}$ (GPF)

Soit $\rho > 0$ donné, on considère l'algorithme suivant :
Algorithme(GPF)
Initialisation : $x_0 \in \mathbb{U}$
Itération :

x_k connu $\qquad x_{k+1} = \mathbb{P}_{\mathbb{U}}(x_k - \rho \nabla J(x_k))$

Lemme 3 *Soit $(x_k)_k$ construite par l'algorithme (GPF). On suppose que $x_k \to x$ quand $k \to +\infty$. Alors x est solution de (P).*

Preuve : soit $\mathbb{P}_{\mathbb{U}} : \mathbb{R}^n \to \mathbb{U} \subset \mathbb{R}^n$ la projection sur $\mathbb{U}$, Alors $\mathbb{P}_{\mathbb{U}}$ est continue. Donc si $x_k \to x$ quand $k \to +\infty$ alors $x = \mathbb{P}_{\mathbb{U}}(x - \rho \nabla J(x))$ et $x \in \mathbb{U}$ (car $x_k \in \mathbb{U}$ et $\mathbb{U}$ fermé).
d'après (2.7) on a

$$< x - \rho \nabla J(x) - x, y - x > \leqslant 0$$

pour tout $y \in \mathbb{U}$,et comme $\rho > 0$,ceci entraine $< \nabla J(x), y - x > \geqslant 0$ pour tout $y \in \mathbb{U}$. or J est convexe donc $J(y) \geqslant \nabla J(x)(y - x)$ pour tout $y \in \mathbb{U}$, et donc $J(y) \geqslant J(x) \quad \forall y \in \mathbb{U}$ ce qui termine la preuve. ∎

Propriété 5 *Soient x^* une solution de (P) et $\omega = \mathbb{P}_{\mathbb{U}}(x^* - \rho \nabla J(x^*))$*

$$\begin{cases} \langle \ (x^* - \rho \nabla J(x^*) - \omega, y \ \rangle \leqslant 0 \\ \langle \ (x^* - \rho \nabla J(x^*) - \omega, \omega \ \rangle = 0 \end{cases} \qquad \forall y \in \mathbb{U}$$

Une solution de ce système (qui d'ailleurs admet une une solution unique) est $\omega = x^*$
En d'autres termes

$$x^* = \mathbb{P}_{\mathbb{U}}(x^* - \rho \nabla J(x^*)).$$

Lemme 4 *(Propriété de contraction de la projection orthogonale). Soit E un espace de Hilbert,$||.||$ la norme et $< . >$ le produit scalaire, $\mathbb{U}$ un convexe fermé non vide de E et $\mathbb{P}_{\mathbb{U}}$ la projection orthogonale sur $\mathbb{U}$, alors*

$$||\mathbb{P}_{\mathbb{U}}(x) - \mathbb{P}_{\mathbb{U}}(y)|| \leq ||x - y|| \qquad \forall (x, y) \in E^2.$$

Prouve : Comme E est un espace de Hilbert

$$||\mathbb{P}_{\mathbb{U}}(x) - \mathbb{P}_{\mathbb{U}}(y)||^2 =< \mathbb{P}_{\mathbb{U}}(x) - \mathbb{P}_{\mathbb{U}}(y), \mathbb{P}_{\mathbb{U}}(x) - \mathbb{P}_{\mathbb{U}}(y) >$$

$||\mathbb{P}_{\mathbb{U}}(x)-\mathbb{P}_{\mathbb{U}}(y)||^2 =< \mathbb{P}_{\mathbb{U}}(x)-x+x-y+y-\mathbb{P}_{\mathbb{U}}(y), \mathbb{P}_{\mathbb{U}}(x)-\mathbb{P}_{\mathbb{U}}(y)) >$
$=< \mathbb{P}_{\mathbb{U}}(x)-x, \mathbb{P}_{\mathbb{U}}(x)-\mathbb{P}_{\mathbb{U}}(y) > + < x-y, \mathbb{P}_{\mathbb{U}}(x)-\mathbb{P}_{\mathbb{U}}(y) > +$
$< y-\mathbb{P}_{\mathbb{U}}(y), \mathbb{P}_{\mathbb{U}}(x)-\mathbb{P}_{\mathbb{U}}(y) >$
Or $< \mathbb{P}_{\mathbb{U}}(x)-x, \mathbb{P}_{\mathbb{U}}(x)-\mathbb{P}_{\mathbb{U}}(y) >\geq 0$ et $< y-\mathbb{P}_{\mathbb{U}}(y), \mathbb{P}_{\mathbb{U}}(x)-\mathbb{P}_{\mathbb{U}}(y) >\geq 0$
d'où :

$$||\mathbb{P}_{\mathbb{U}}(x)-\mathbb{P}_{\mathbb{U}}(y)|| \leq < x-y, \mathbb{P}_{\mathbb{U}}(x)-\mathbb{P}_{\mathbb{U}}(y) >,$$

et donc, grâce à l'inégalité de Cauchy-Schwarz,

$||\mathbb{P}_{\mathbb{U}}(x)-\mathbb{P}_{\mathbb{U}}(y)|| \leq ||x-y|| \times ||\mathbb{P}_{\mathbb{U}}(x)-\mathbb{P}_{\mathbb{U}}(y)|| \leq ||x-y||$ ■

Théorème 9 *(Convergence de l'algorithme GPF)*
Soit $J \in C^1(\mathbb{R}^n, \mathbb{R})$, et $\mathbb{U}$ convexe fermé non vide. On suppose :
1). il existe $\alpha > 0$ tel que $< \nabla J(x) - \nabla J(y), x-y >\geq \alpha |x-y|^2$, pour tout $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n$
2). il existe $M > 0$ tel que $|\nabla J(x) - \nabla J(y)| \leq M|x-y|$ pour tout $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n$, alors
1. il existe un unique élément un unique élément $x^* \in \mathbb{U}$ solution de (P),
2.si $0 < \rho < \dfrac{2\alpha}{M^2}$, la suit (x_k) définie par l'algorithme (GPF) converge vers x^* lorsque $n \to +\infty$.

Preuve :
1. La condition 1) donne que J est strictement convexe et que $J(x) \to +\infty$ quand $|x| \to +\infty$ comme $\mathbb{U}$ est convexe fermé non vide ,il existe donc un unique x^* solution de (P).
2. On pose ,pour $x \in \mathbb{R}^n$,$h(x) = \mathbb{P}_{\mathbb{U}}(x - \rho \nabla J(x))$. On a donc $x_{k+1} = h(x_k)$. Pour montrer que la suite $(x_k)_{n\in\mathbb{N}}$ converge .il suffit donc de monter que h est strictement contractante dès que

$$0 < \rho < \frac{2\alpha}{M^2} \tag{2.9}$$

Grâce au lemme (4), on sait que $\mathbb{P}_{\mathbb{U}}$ est contractante. Or h est définie par :

$$h(x) = \mathbb{P}_{\mathbb{U}}(g(x)) \quad g(x) = x - \rho \nabla J(x).$$

On a g est strictement contractante si la condition (1.9) est vérifiée.

$$|g(x) - g(y)| \leq k|x-y|^2.$$

On en déduit que :

$$|h(x) - h(y)|^2 \leq |\mathbb{P}_{\mathbb{U}}(g(x)) - \mathbb{P}_{\mathbb{U}}(g(y))|^2 \leq k|x-y|^2.$$

L'application h est donc strictement contractante dès que $0 < \dfrac{2\alpha}{M^2}$. La suite $(x_k)_{n\in\mathbb{N}}$ converge donc bien vers $x = x^*$ ■

Chapitre 3

Dualité Lagrangienne classique

3.1 Lagrangien et Point-selle

Le chapitre précédent a permis d'étudier les conditions d'optimalité d'un problème d'optimisation avec contraintes (P) en termes de configuration d'un gradient de la fonction coût par rapport au cône orthogonal à l'ensemble admissible au point solution, Pour cette raison, nous avons qualifié ces conditions de "locales".
Cependant,les hypothèses de convexité permettent de passer facilement de considérations locales à des considérations globales, et les conditions énoncées précédemment sont alors à la fois nécessaires et suffisantes. Sans hypothèses de convexité, ces conditions locales sont en fait des conditions nécessaires seulement, et elles font appel à des hypothèses de différentiabilité.

Le présent chapitre aborde la question des conditions d'optimalités de façon globale et en termes de conditions suffisantes. On va voir en particulier que la partie primale d'un **point selle** du **Lagrangien** du problème (si un tel point selle existe) est une solution du problème d'optimisation sous contraintes indépendamment de toute considération de convexité. Évidemment, dans le cas convexe, cette théorie globale des conditions suffisantes rejoint la théorie locale des conditions nécessaires.

3.1.1 Le Lagrangien

L'intérêt du Lagrangien est de ramener le problème (P) de minimisation avec contrainte de J sur $\mathbb{U}$ à un problème sans contrainte .

Définition 15 *soit $\lambda \in \mathbb{R}^p_+$ et $\gamma \in \mathbb{R}^q$. On définit le* ***Lagrangien*** *associe au problem (P) par :*

$$\mathcal{L}(x,(\lambda,\gamma)) = J(x) + \sum_{i=1}^{p} \lambda_i F_i(x) + \sum_{j=1}^{q} \gamma_j G_j(x)$$

ou plus généralement

$$\mathcal{L}(x,(\lambda,\gamma)) = J(x) + <\lambda, F(x)>_{\mathbb{R}^p_+} + <\gamma, G(x)>_{\mathbb{R}^q}$$

Les paramètres (λ,γ) sont appelés multiplicateurs de Lagrange,le Lagrangien fait intervenir autant de multiplicateurs que de contraintes, et les multiplicateurs λ_i associés aux contraintes d'inégalité "$F_i(x) \leq 0$" sont positifs ou nuls.

Proposition 5 *Le point x^* est un minimum global de J sur $\mathbb{U}$ (i.e. solution du problème $\inf_{x\in\mathbb{U}} J(x)$) si et seulement si il est aussi une solution de*

$$\min_{x\in\mathbb{R}^n} \max_{(\lambda,\gamma)\in\mathbb{R}^p_+\times(\mathbb{R})^q} \mathcal{L}(x,(\lambda,\gamma))$$

Preuve. si $x \in \mathbb{U}$ alors $\mathcal{L}(x,(\lambda,\gamma)) = J(x) + <\lambda, F(x)>$ et le maximum de cette expression par rapport à λ est donné pour $\lambda = 0$ car $F(x) \leq 0$ et $\lambda \geq 0$.
Si $x \notin \mathbb{U}$ alors le maximum de $\mathcal{L}$ par rapport au couple (λ,γ) vaut $+\infty$.
Ainsi

$$\max_{(\lambda,\gamma)\in\mathbb{R}^p_+\times(\mathbb{R})^q} \mathcal{L}(x,(\lambda,\gamma)) = \begin{cases} J(x) & si\ x \in \mathbb{U} \\ +\infty & si\ x \notin \mathbb{U} \end{cases}$$

d'où :

$$\min_{x\in\mathbb{R}^n} \max_{(\lambda,\gamma)\in\mathbb{R}_+{}^p\times(\mathbb{R})^q} \mathcal{L}(x,(\lambda,\gamma)) = \min_{x\in\mathbb{U}} J(x)$$ ■

3.1.2 Points-selles du Lagrangien

L'idée générale de l'introduction du Lagrangien est que l'on doit maximiser $\mathcal{L}$ par rapport à sa deuxième variable et le minimiser par rapport à sa première variable, cela ressemble à la recherche d'un **point selle**.

Définition 16 *Soit $\mathcal{L} : U \times V \to \mathbb{R}$ une fonction à deux variables (x,p), $x \in U \subset \mathbb{R}^n$ et $p = (\lambda,\gamma) \in V \subset \mathbb{R}^p_+ \times (\mathbb{R})^q$ le point $(x^*,p^*) \in U \times V$ est dit* ***point-selle*** *ou* ***point-col*** *de $\mathcal{L}$ sur $U \times V$, si et seulement si*

$$\mathcal{L}(x^*,p) \leq \mathcal{L}(x^*,p^*) \leq \mathcal{L}(x,p^*) \quad \forall (x,p) \in U \times V$$

ou

$$\sup_{p\in V} \mathcal{L}(x^*,p) = \mathcal{L}(x^*,p^*) = \inf_{x\in U} \mathcal{L}(x,p^*)$$

Exemple : Un point-selle d'une fonction à deux variables $\mathcal{L}$ est donc un minimum en la première variable et un maximum en la deuxième variable. L'archétype de tous les points-selles est le point $(0,0)$ de la fonction de $\mathbb{R}^2$ dans $\mathbb{R}$: $\mathcal{L}(u,v) = u^2 - v^2$.

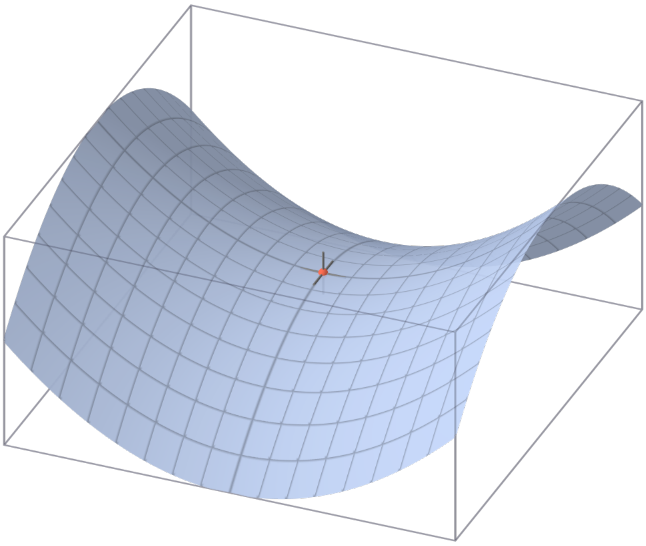

Figure 1 : point-selle $(0,0)$ de $\mathcal{L}(u,v) = u^2 - v^2$.

Proposition 6 *Si $(x^*, (\lambda^*, \gamma^*))$ est point selle se $\mathcal{L}$ alors*

$$\underset{(\lambda,\gamma)\in\mathbb{R}_+^p\times\mathbb{R}^q}{sup} \; \underset{x\in\mathbb{R}^n}{inf} \mathcal{L}(x,(\lambda,\gamma)) = \mathcal{L}(x^*,(\lambda^*,\gamma^*)) = \underset{x\in\mathbb{R}^n}{inf} \; \underset{(\lambda,\gamma)\in\mathbb{R}_+^p\times\mathbb{R}^q}{sup} \mathcal{L}(x,(\lambda,\gamma))$$

Autrement dit, on peut inverser l'ordre du sup et du inf.

Preuve. Montrons en premier que

$$\underset{(\lambda,\gamma)\in\mathbb{R}_+^p\times\mathbb{R}^q}{sup} \; \underset{x\in\mathbb{R}^n}{inf} \mathcal{L}(x,(\lambda,\gamma)) \leq \underset{x\in\mathbb{R}^n}{inf} \; \underset{(\lambda,\gamma)\in\mathbb{R}_+^p\times\mathbb{R}^q}{sup} \mathcal{L}(x,(\lambda,\gamma))$$

En effet, on a que

$$\underset{x\in\mathbb{R}^n}{inf} \mathcal{L}(x,(\lambda,\gamma)) \leq \mathcal{L}(x,(\lambda,\gamma)) \leq \underset{(\lambda,\gamma)\in\mathbb{R}_+^p\times\mathbb{R}^q}{sup} \mathcal{L}(x,(\lambda,\gamma))$$

Le membre de gauche est une fonction de (λ,γ) seulement. De même, pour le membre de droite qui est une fonction de x uniquement. Ceci implique

$$\underset{(\lambda,\gamma)\in\mathbb{R}_+^p\times\mathbb{R}^q}{sup} \; \underset{x\in\mathbb{R}^n}{inf} \mathcal{L}(x,(\lambda,\gamma)) \leq \underset{(\lambda,\gamma)\in\mathbb{R}_+^p\times\mathbb{R}^q}{sup} \mathcal{L}(x,(\lambda,\gamma)) \quad \forall x \in \mathbb{R}^n$$

et, en prenant le minimum par rapport à x,

$$\underset{(\lambda,\gamma)\in\mathbb{R}^p_+\times\mathbb{R}^q}{sup}\ \underset{x\in\mathbb{R}^n}{inf}\mathcal{L}(x,(\lambda,\gamma)) \leq \underset{x\in\mathbb{R}^n}{inf}\ \underset{(\lambda,\gamma)\in\mathbb{R}^p_+\times\mathbb{R}^q}{sup}\ \mathcal{L}(x,(\lambda,\gamma)) \qquad (1)$$

d'où le résultat.
En deuxième, montrons que

$$\underset{x\in\mathbb{R}^n}{inf}\ \underset{(\lambda,\gamma)\in\mathbb{R}^p_+\times\mathbb{R}^q}{sup}\ \mathcal{L}(x,(\lambda,\gamma)) \leq \mathcal{L}(x^*,(\lambda^*,\gamma^*))$$

En effet, selon la définition du point de selle, on a
$\underset{(\lambda,\gamma)\in\mathbb{R}^p_+\times\mathbb{R}^q}{sup}\ \mathcal{L}(x,(\lambda,\gamma)) = \mathcal{L}(x^*,(\lambda^*,\gamma^*)) \Rightarrow \mathcal{L}(x^*,(\lambda^*,\gamma^*)) \geq \underset{x\in\mathbb{R}^n}{inf}\ \underset{(\lambda,\gamma)\in\mathbb{R}^p_+\times\mathbb{R}^q}{sup}\ \mathcal{L}(x,(\lambda,\gamma))$
De même, on a

$$\underset{x\in\mathbb{R}^n}{inf}\mathcal{L}(x,(\lambda,\gamma)) = \mathcal{L}(x^*,(\lambda^*,\gamma^*)) \Rightarrow \mathcal{L}(x^*,(\lambda^*,\gamma^*)) \leq \underset{(\lambda,\gamma)\in\mathbb{R}^p_+\times\mathbb{R}^q}{sup}\ \underset{x\in\mathbb{R}^n}{inf}\mathcal{L}(x,(\lambda,\gamma))$$

Alors on obtient

$$\begin{aligned}\mathcal{L}(x^*,(\lambda^*,\gamma^*)) \leq \underset{(\lambda,\gamma)\in\mathbb{R}^p_+\times\mathbb{R}^q}{sup}\ \underset{x\in\mathbb{R}^n}{inf}\mathcal{L}(x,(\lambda,\gamma)) &\leq \underset{x\in\mathbb{R}^n}{inf}\ \underset{(\lambda,\gamma)\in\mathbb{R}^p_+\times\mathbb{R}^q}{sup}\ \mathcal{L}(x,(\lambda,\gamma)) \\ &\leq \mathcal{L}(x^*,(\lambda^*,\gamma^*))\end{aligned} \qquad (2)$$

d'où d'après (1) et (2) on a

$$\underset{(\lambda,\gamma)\in\mathbb{R}^p_+\times\mathbb{R}^q}{sup}\ \underset{x\in\mathbb{R}^n}{inf}\mathcal{L}(x,(\lambda,\gamma)) = \mathcal{L}(x^*,(\lambda^*,\gamma^*)) = \underset{x\in\mathbb{R}^n}{inf}\ \underset{(\lambda,\gamma)\in\mathbb{R}^p_+\times\mathbb{R}^q}{sup}\ \mathcal{L}(x,(\lambda,\gamma)) \qquad \blacksquare$$

Critère .(d'existence d'un point selle)
S'il existe $x \in U,\ p \in V,\ l \in \mathbb{R}$ tel que

$$\mathcal{L}(x,q) \leq l \qquad \forall q \in V$$

$$\mathcal{L}(y,p) \geq l \qquad \forall y \in U$$

Alors (x,p) est une point selle de $\mathcal{L}$.

3.1.3 Condition d'optimalité pour le Point-selle

Théorème 10 *Considérons le problème sous contraintes fonctionnelles (P) où J, F et G sont supposées différentiables. On note $\mathbb{U}$ le domaine des contraintes et $\mathcal{L}$ le Lagrangien associé à J aux contraintes $\mathbb{U}$.*

Si (x^*, p^*) est un point-selle de $\mathcal{L}$ sur $U \times V$ $(U \times V \subset \mathbb{R}^n \times (\mathbb{R}_+^p \times \mathbb{R}^q)$ alors

- *x^* est un minimum global de J sur $\mathbb{U}$.*
- *On a l'équation suivante reliant x^* et $p^* = (\lambda^*, \gamma^*)$:*

$$\nabla J(x^*) + \sum_i \lambda^* \nabla F_i(x^*) + \sum_j \gamma_j^* \nabla G_j(x^*) = 0 \quad et\ \forall j, \quad \gamma_j^* G_j(x^*) = 0 \qquad (3.1)$$

Dans ce cas p^* est appelé le " multiplicateur de Lagrange " du point-selle (x^*, p^*).

Ce théorème ne dit pas que tous les minimums globaux de J sur $\mathbb{U}$ viennent de point-selle, (ou même que les points-selles du Lagrangien existent), il dit seulement que si on trouve un point-selle (x^*, p^*), alors est un minimum global de J sur $\mathbb{U}$. En ce sens, ce théorème n'est pas un théorème de caractérisation des minimums, au sens où on donne une condition suffisante mais non nécessaire. Il va de soit que ce théorème ne dit pas que tout (x^*, p^*) qui vérifie (3.1) est forcément point-selle (ce qui entraînerait que x^* est minimum global).
Preuve. On note $p = (\lambda, \gamma)$ On rappelle que les conditions de points-selles sont

$$\mathcal{L}(x^*, p) \leq \mathcal{L}(x^*, p^*) \leq \mathcal{L}(x, p^*) \;\; \forall (x,p) \in U \times V \qquad (U \times V \subset (\mathbb{R}^n \times (\mathbb{R}_+^p \times \mathbb{R}^q))$$

En récrivant la première inégalité on obtient

$$J(x^*) + < \lambda, F(x^*) > + < \gamma, G(x^*) > \leqslant \mathcal{L}(x^*, p^*) \qquad \forall \lambda \in \mathbb{R}_+^p, \gamma \in \mathbb{R}^q$$

On a $F(x^*) \leqslant 0$ et $G(x^*) = 0$ et x^* appartient bien à $\mathbb{U}$. De plus, on obtient, en prenant $\lambda = \gamma = 0$

$$J(x^*) \leqslant J(x^*) + < \lambda^*, F(x^*) >$$

et comme λ^* est positif et $F(x^*)$ négatif, on en conclut que $< \lambda^*, F(x^* >= 0$.En récrivant la deuxième inégalité on obtient, en se limitant à des x

$$J(x^*) = \mathcal{L}(x^*, p^*) \leqslant J(x) + < \lambda^*, F(x) > \qquad \forall x \in \mathbb{U}$$

et comme λ^* est positif et $F(x^*)$ est négatif pour tout x dans $\mathbb{U}$, alors
$J(x^*) + < \lambda^*, F(x) > \leqslant J(x)$ et ainsi

$$J(x^*) \leqslant J(x) \qquad \forall x \in \mathbb{R}^n$$

On voit que x^* est un minimum sans contrainte de la fonctionnelle différentiable $g : x \mapsto \mathcal{L}(x, p^*)$ donc vérifie l'égalité d'Euler, $\nabla g(x^*) = 0$ qui est exactement la formule (3.1). On remarquera que pour l'établissement de cette formule, p^* est fixé et on ne doit dériver que par rapport à x. ■

Définition 17 *(qualification des contraintes).*

les contraintes $F(x) \leq 0$ *et* $G(x) = 0$ *sont dites qualifiées si il existe* $x^* \in \mathbb{U}$ *tel qu'en séparant les contraintes en deux sous-ensemble :*

$$F(x) < 0 \text{ et } G(x) = 0$$

c'est à dire que les gradients $\nabla F_i(x),\ i \in I^0(x)$ *et* $\nabla G_j(x),\ j \in \{1, 2, ..., q\}$ *sont linéairement indépendants donc il y a* ***q + card (***$I^0(x)$***)*** *vecteurs.*

où $I^0(x) = \{\in \{1, 2, ..., p\}$ *tel que* $F(x) = 0\}$*(l'ensemble des indices où les contraintes d'inégalités sont saturées en* x *).*

Proposition 7 *(Conditions nécessaires d'optimalité)*

Considérons le problème (P)*,où* $J,\ F$ *et* G *sont supposées différentiables. Si* x^* *est un minimum local de* J *sur* $\mathbb{U}$ *et si les contraintes sont qualifiées en* x^* *alors il existe des multiplicateurs* $(\lambda^*, \gamma^*) \in \mathbb{R}_+^p \times \mathbb{R}^q$ *tel que :*

1. $\nabla_x \mathcal{L}(x^*, (\lambda^*, \gamma^*)) = 0$ *soit*

$$\nabla J(x^*) + \sum_{i=1}^{p} \lambda_i^* \nabla F_i(x^*) + \sum_{j=1}^{q} \gamma_j^* \nabla G_j(x^*) = 0.$$

2. $< \lambda^* F(x^*) >= 0 \quad i.e \qquad \sum_{i=1}^{p} \lambda_i^* \nabla F_i(x^*) = 0$

   comme $F_i(x^*) \leq 0$ *et* $\lambda_i \geq 0$ *alors*

$$\lambda_i^* = 0 \quad ou \quad F_i(x^*) = 0. \qquad \forall i \in \{1, 2, ..., p\}$$

   Ces conditions sont appelées ***relations de complémentarité****.*

3. $\forall i \in \{1, 2, ..., p\} \quad F(x^*) \leq 0.$

4. $\forall j \in \{1, 2, ..., q\} \quad G(x^*) = 0.$

5. $\forall i \in \{1, 2, ..., p\} \quad \lambda_i^* \geq 0.$

Remarque 3 *On appelle point stationnaire du problème (P) tout point* x^* *vérifiant les conditions nécessaires d'optimalité du premier ordre :*

$$\begin{cases} \nabla_x \mathcal{L}(x^*, (\lambda^*, \gamma^*)) = 0 & \\ \lambda^* F_i(x^*) = 0, & i = 1, 2, ..., p \\ G_j(x^*) = 0, & j = 1, 2, ..., q \\ \lambda^* \geq 0, & i = 1, 2, ..., p \end{cases}$$

pour un certain multiplicateur $(\lambda^*, \gamma^*) \in \mathbb{R}^p_+ \times \mathbb{R}^q$.

Les relations : $\lambda_i F_i(x^*) = 0$ *sont appelées relations de complémentarité. Elles sont trivialement satisfaites par toute contrainte i active en* x^* *et indiquent que pour toute contrainte inactive, le multiplicateur* λ_i *correspondant est nul.*

On remarquera qu'il n'y a pas de condition sur les multiplicateurs γ_j^* *associés aux contraintes d'égalité.*

3.2 Dualité

Donnons un bref aperçu de la théorie de la dualité pour les problèmes d'optimisation .
Nous l'appliquerons au problème de minimisation convexe avec contrainte définit au deuxième chapitre. Nous avons associer à ce problème de minimisation un problème de recherche d'un point selle (x, λ, γ) pour le lagrangien
$\mathcal{L}(x,(\lambda,\gamma)) = J(x) + \sum_{i=1}^{p} \lambda_i F_i(x) + \sum_{j=1}^{q} \gamma_j G_j(x)$. Mais nous allons voir que, à l'existence d'un point selle (x, λ, γ) du lagrangien, on peut associer inversement deux problèmes d'optimisation (plus précisément, un problème de minimisation et un problème de maximisation), qui seront dits **duaux** l'un de l'autre. Le **problème dual** peut être utile pour la résolution du problème d'origine.

3.2.1 Problème Primal et Problème Dual

Définition 18 *Soit* $\mathcal{L}$ *un lagrangien défini sur une partie* $U \times V$ *de* $\mathbb{R}^n \times (\mathbb{R}^p_+ \times \mathbb{R}^q)$ *On suppose qu'il existe un point selle* (x, p) $(p = (\lambda, \gamma))$ *de* $\mathcal{L}$ *sur* $U \times V$

$$\forall q \in V \quad \mathcal{L}(x,q) \le \mathcal{L}(x,p) \le \mathcal{L}(y,p) \qquad \forall y \in U$$

Pour $y \in U$ *et* $q \in V$*, posons*

$$\mathcal{J}(y) = \sup_{q \in V} \mathcal{L}(y,q) \qquad\qquad \mathcal{G}(q) = \inf_{y \in U} \mathcal{L}(y,q) \tag{3.2}$$

On appelle ***problème primal*** *le probleme de minimisation*

$$\inf_{y \in U} \mathcal{J}(y), \tag{3.3}$$

Et le ***problème dual*** *le problème de maximisation*

$$\sup_{q \in V} \mathcal{G}(q). \tag{3.4}$$

Remarque 4 *Bien sûr, sans hypothèses supplémentaire , il peut arriver que $\mathcal{J}(y) = +\infty$ pour certaines valeurs de y ou que $\mathcal{G}(q) = -\infty$ pour certaines valeurs de q. Mais l'existence supposée du point selle (x, p) nous assure que les domaine de $\mathcal{J}$ et $\mathcal{G}$ ($\{y \in U,\ \mathcal{J}(y) < +\infty\}$ et $\{q \in V,\ \mathcal{G}(q) > +\infty\}$) ne sont pas vides, puisque la définition de point selle montre que $\mathcal{J}(x) = \mathcal{G}(p) = \mathcal{L}(x, p)$. Les problèmes primale et dual ont donc bien un sens*

Théorème 11 ***(de dualité)*** *le couple (x, p) est un point selle de $\mathcal{L}$ sur $U \times V$ si et seulement si*

$$\mathcal{J}(x) = \min_{y \in U} \mathcal{J}(y) = \max_{q \in V} \mathcal{G}(q) = \mathcal{G}(p). \tag{3.5}$$

Preuve. soit (x, p) un point selle de $\mathcal{L}$ sur $U \times V$.
Notons $\mathcal{L}^* = \mathcal{L}(x, p)$, Pour $y \in U$, il est clair d'après (3.2) que $\mathcal{J}(x) \geqslant \mathcal{L}(x, p)$, d'où $\mathcal{J}(x) \geqslant \mathcal{L}^*$
comme $\mathcal{J}(u) = \mathcal{L}^*$, ceci montre que

$$\mathcal{J}(u) = \inf_{y \in U} \mathcal{J}(y) = \mathcal{L}^*$$

de la même façon on trouve

$$\mathcal{G}(p) = \sup_{q \in V} \mathcal{G}(q) = \mathcal{L}^*.$$

Réciproquement, supposons que (3.5) a lieu et posons $\mathcal{L}^* = \mathcal{J}(x)$.
La définition (3.2) de $\mathcal{J}$ montre que

$$\mathcal{L}(x, q) \leqslant \mathcal{J}(x) = \mathcal{L} \qquad \forall q \in V$$

De même , on a aussi :

$$\mathcal{L}(y, p) \geqslant \mathcal{G}(p) = \mathcal{L}^* \qquad \forall y \in U,$$

et on déduit facilement que $\mathcal{L}(x, p) = \mathcal{L}$, ce qui montre que (u,p) est point selle. ■

Remarque 5 *Par la définition (3.2) de $\mathcal{J}$ et $\mathcal{G}$ On obtient*

$$\mathcal{J}(x) = \min_{y \in U} \left(\sup_{q \in V} \mathcal{L}(y, q) \right) = \max_{q \in V} \left(\inf_{y \in U} \mathcal{L}(y, q) \right) = \mathcal{G}(p). \tag{3.6}$$

Si le sup et l'inf sont atteints dans (3.6) (c'est-à-dire qu'on peut les écrire max et min, respectivement), on voit alors que (3.6) traduit la possibilité d'échanger l'ordre du min et du max appliqués au lagrangien $\mathcal{L}$. ce qui est faux si $\mathcal{L}$ n'admet pas de point selle.

Nous appliquons ce résultat de dualité au problème de minimisation avec contrainte

$$\underset{x\in\mathbb{U}}{inf}\, J(x)$$

avec J et $\mathbb{U}$ convexes,
Nous définissons le critère $J(x)$ et l'ensemble $\mathbb{U}$

$$\mathcal{J}(x) = \underset{q\in\mathbb{R}^p_+\times\mathbb{R}^q}{sup} \mathcal{L}(x,q) = \begin{cases} J(x) & si\ x \in \mathbb{U} \\ +\infty & sinon, \end{cases}$$

Ce qui montre que le problème primal est exactement le problème d'origine.

D'autre par, la fonction $\mathcal{G}(q)$ du problème dual est bien définie par (3.2),car est un problème de minimisation convexe. de plus $\mathcal{G}(q)$ est une fonction concave car elle est l'inf de fonctions affines, Par conséquent le problème dual

$$\underset{q\in\mathbb{R}^p_+\times\mathbb{R}q}{sup} \mathcal{G}(q),$$

est un problème de maximisation concave **plus simple** que le problème primal car les contraints sont linaires.

Corollaire 1 *On suppose que les fonctions J, F et G sont convexes et dérivables sur $\mathbb{R}^n$, Soit $x \in \mathbb{R}^n$ tel que $F(x) \leq 0, G(x) = 0$ et les contraintes sont qualifiées en x, Alors si x est un minimum global de $\mathcal{J}$ sur $\mathbb{R}^n$, il existe $p = (\lambda,\gamma) \in \mathbb{R}^p_+ \times \mathbb{R}^q$ tel que*

1. *p est un maximum global de $\mathcal{G}$ sur $\mathbb{R}^p_+ \times \mathbb{R}^q$,*

2. *(x,p) est un point selle du lagrangien $\mathcal{L}$ sur $\mathbb{R}^n \times (\mathbb{R}^p_+ \times \mathbb{R}^q)$*

3. *$(x,p) \in \mathbb{R}^n\times(\mathbb{R}^p_+\times\mathbb{R}^q)$ vérifie la condition d'optimalité nécessaire et suffisant.*

L'application la plus courante du Corollaire est la suivante, Supposons que le problème dual de maximisation est plus facile a résoudre que le problème primal (c'est le cas en générale car ses contraintes sont plus simples). Alors pour calculer la solution x^* du problème primal on procède en deux étapes.
Premièrement, On calcule la solution p^* du problème dual.
Deuxièmement, On dit que (x^*, p^*) est point selle du lagrangien, c'est-à-dire que l'on calcule x^*, solution de minimisation **sans contrainte**

$$\underset{x\in\mathbb{R}^n}{min} \mathcal{L}(x,p).$$

Précisons qu'avec les hypothèses faites il n'y a pas priori d'unicité des solution pour tous ces problèmes. Précisons aussi que pour obtenir l'existence du minimum x^* dans le corollaire il suffit d'ajouter une hypothèse de forte convexité ou de comportement infini à l'infini sur J.

3.2.2 Relations de Kuhn-Tucker

Les relations de Kuhn-Tucker expriment que si la ou les solutions d'un problème avec contraintes est à l'intérieur du domaine des points admissibles, le gradient de J s'annule comme c'est le cas pour les problèmes sans contrainte, et que si la ou les solutions sont sur le bord, il y a, comme pour le cas des contraintes égalités, colinéarité du gradient de J et des gradients des F_i et G_j.

Définition 19 *(contrainte active) On dit que la contrainte en inégalité F_i est active ou saturé au point x^* de $\mathbb{R}^n$ si $F(x^*) = 0$.*

On note $I^0(x^*)$ l'ensemble des contraintes actives en x^*.

Définition 20 *(point régulier)*

On dit qu'un élément x^* de $\mathbb{R}^n$ est régulier pour les contraintes F_i et $G_j$si :

- *Il est réalisable i.e $G_j(x^*) = 0$ et $F_i(x^*) \leqslant 0$*

- *Les vecteurs $\nabla F_i \quad \forall i \in I^0(x^*)$ et $\nabla G_j(x^*) \quad \forall j \in \{1, 2, ..., q\}$ sont indépendants.*

On dit aussi que x^* vérifie la contrainte de qualification.

Théorème 12 *(conditions d'optimalité de Karush-Kuhn-Tucker)*

Soit J, F, G des fonctions différentiables, et x^* une solution du problème (P) et régulier pour les contraintes F et G,alors il existe $\lambda^* \in \mathbb{R}^p_+$, $\gamma^* \in \mathbb{R}^q$, tel que

$$\begin{cases} \nabla J(x^*) + \sum_{i=1}^{p} \lambda_i^* \nabla F_i(x^*) + \sum_{j=1}^{q} \gamma_j^* \nabla G_j(x^*) = 0 \\ \lambda^* F_i(x^*) = 0, \quad i = 1, 2, ..., p \\ G_j(x^*) = 0, \quad j = 1, 2, ..., q \\ \lambda^* \geq 0, \quad i = 1, 2, ..., p \end{cases} \tag{3.7}$$

Les points qui vérifient ces conditions sont appelés points KKT (Karush,Kuhn,Tucker).

Lemme 5 *Soit J, F, G des fonctions **convexes** et différentiables, et x^* est un point de $\mathbb{U}$ où les contraintes sont qualifiées, alors les propositions suivantes sont* *équivalentes :*

- x^* *est un minimum global de* J *sur* X

- *Il existe un* $p^* = (\lambda^*, \gamma^*) \in (\mathbb{R}_+^p \times \mathbb{R}^q)$ *tel que* (x^*, p^*) *est un point-selle du*
Lagrangien $\mathcal{L}$.

- *Il existe un* $p^* = (\lambda^*, \gamma^*) \in (\mathbb{R}_+^p \times \mathbb{R}^q)$ *tel que* (x^*, p^*) *vérifie les condition (3.7)*

Preuve. Comme la deuxième assertion entraîne la première (Théorème 10) et que la première entraîne la troisième (Proposition 7), il suffit de montrer que la troisième assertion entraîne la première pour conclure.

Soit un x^* tel

$$\lambda^* F_i(x^*) = 0 \text{ et } \nabla J(x^*) + \sum_i \lambda^* \nabla F_i(x^*) + \sum_j \gamma_j^* \nabla G_j(x^*) = 0 \text{ et } \forall j,\ \gamma_j^* G_j(x^*) = 0$$

Comme les F_i sont convexes et $\lambda^* \geqslant 0$ alors la fonction $L:\ x \mapsto \mathcal{L}(x, p^*)$ est convexe. Et comme

$$\nabla L(x^*) = \nabla J(x^*) + \sum_i \lambda^* \nabla F_i(x^*) + \sum_j \gamma_j^* \nabla G_j(x^*) = 0$$

alors x^* est le minimum global de L et on a bien

$$\mathcal{L}(x^*, p^*) = L(x^*) \leq L(x) = \mathcal{L}(x, p^*) \qquad \forall x \in \mathbb{R}^n$$

De plus, comme

$$< \lambda^*, F(x^*) >= 0, G(x^*) = 0 \text{ et } F(x^*) \leq 0$$

alors, pour tout $\lambda \geq 0$

$$\mathcal{L}(x^*, p) = J(x^*) + < \lambda, F(x^*) > \leq J(x^*) = \mathcal{L}(x^*, p^*)$$

Ainsi nous avons

$$\mathcal{L}(x^*, p) \leq \mathcal{L}(x^*, p^*) \leq \mathcal{L}(x, p^*) \qquad \forall x \in \mathbb{R}^n \text{ et } P = (\lambda, \gamma) \in (\mathbb{R}_+^p \times \mathbb{R}^q)$$

Ce qui prouve que (x^*, p^*) est bien un point selle. ■

Chapitre 4

La Méthode d'Uzawa et son application au problème du Portefeuille en finance

4.1 La Méthode d'Uzawa

La différentiabilité de $\mathcal{G}$ ouvre la voie aux méthodes de gradient pour le calcul de la composante (λ, γ) du point selle. Supposant que J est fortement convexe et que F et G sont lipschitziennes , L'algorithme d'Uzawa utilise la dualité introduite ci-dessus. On ramené le calcul du minimum de la fonctionnelle coût J sous l'ensemble des contraintes $\mathbb{U}$, à une double itération successive de recherche de maximums de la fonction duale $\mathcal{G}$ et de minimum du problème primal. Le maximum de $\mathcal{G}$ étant obtenu par une méthode de gradient à pas fixe.

Proposition 8 *Soit $(x^*, \lambda^*, \gamma^*) \in \mathbb{R}^n \times \mathbb{R}_+^p \times \mathbb{R}^q$ un point selle de Lagrangien $\mathcal{L}$. Pour tout $\rho > 0$, on a :*

$$(\lambda^*, \gamma^*) = \mathbb{P}_{\mathbb{R}_+^p \times \mathbb{R}^q}[(\lambda^*, \gamma^*) + \rho \nabla \mathcal{G}(\lambda^*, \gamma^*)]$$

$\mathbb{P}_{\mathbb{R}_+^p \times \mathbb{R}^q}$ étant la projection sur le cône convexe fermé $\mathbb{R}_+^p \times \mathbb{R}^q$ notez que pour

$$z = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^p \times \mathbb{R}^q \text{ tel que } x \in \mathbb{R}^p,\ y \in \mathbb{R}^q \quad \mathbb{P}_{\mathbb{R}_+^p \times \mathbb{R}^q}(z) = \begin{pmatrix} Max(0, x) \\ y \end{pmatrix}$$

Preuve. Soit (λ^*, γ^*) une solution du problème de probleme dual

$$\underset{(\lambda \times \gamma) \in \mathbb{R}_+^p \times \mathbb{R}^q}{sup} \mathcal{G}(\lambda, \gamma) = - \underset{(\lambda \times \gamma) \in \mathbb{R}_+^p \times \mathbb{R}^q}{inf} - \mathcal{G}(\lambda, \gamma)$$

D'après l'inéquation d'Euler définit au théorème 6 on a :

$$< \nabla\mathcal{G}(\lambda^*, \gamma^*), q - (\lambda^*, \gamma^*) > \geqslant 0 \qquad \forall q \in \mathbb{R}_+^p \times \mathbb{R}^q$$

Donc

$$< (\lambda^*, \gamma^*) - q, \rho\nabla\mathcal{G}(\lambda^*, \gamma^*) > \leqslant 0 \qquad \forall \rho > 0$$

Alors $\forall q \in \mathbb{R}_+^p \times \mathbb{R}^q$ *et* $\forall \rho > 0$ on a

$$< (\lambda^*, \gamma^*) - q, (\lambda^*, \gamma^*) - ((\lambda^*, \gamma^*) - \rho(-\nabla\mathcal{G}(\lambda^*, \gamma^*))) > \leqslant 0$$

Donc d'après l'unicité le la projection sur le cône convexe fermé $\mathbb{R}_+^p \times \mathbb{R}^q$, (λ^*, γ^*) est la projection sur $\mathbb{R}_+^p \times \mathbb{R}^q$ de $[(\lambda^*, \gamma^*) - \rho(-\nabla\mathcal{G}(\lambda^*, \gamma^*))]$
Donc

$$(\lambda^*, \gamma^*) = \mathbb{P}_{\mathbb{R}_+^p \times \mathbb{R}^q}[(\lambda^*, \gamma^*) + \rho\nabla\mathcal{G}(\lambda^*, \gamma^*)]. \qquad \blacksquare$$

4.1.1 Algorithme d'Uzawa

L'algorithme d'Uzawa (1958) est basée sur la formulation duale et la recherche des points selle.

L'utilisation de cette méthode suppose que la fonction duale est différentiable (au moins a l'optimum). Ce sera le cas si le minimum en x de $\mathcal{L}(x, \lambda^*, \gamma^*)$ est unique. Dans ce cas si on note $x = x(\lambda, \gamma)$ le vecteur tel que

$$\mathcal{G}(\lambda, \gamma) = \mathcal{L}(x^*, \lambda, \gamma) = J(x^*(\lambda, \gamma)) + < \lambda, F(x^*) > + < \gamma, G(x^*) >$$

$$= J(x^*(\lambda, \gamma)) + < \Lambda, \mathcal{F}(x^*) >,\ \mathcal{F}(x(\lambda, \gamma) = \begin{pmatrix} F(x(\lambda, \gamma)) \\ G(x(\lambda, \gamma)) \end{pmatrix},\ \Lambda = \begin{pmatrix} \lambda \\ \gamma \end{pmatrix}$$

Donc

$$\begin{aligned} \nabla_{(\lambda,\gamma)}\ \mathcal{G}(\lambda, \gamma) &= \nabla_{(\lambda,\gamma)}\ \mathcal{L}(x^*, \lambda, \gamma) \\ &= \mathcal{F}(x(\lambda, \gamma)) = \begin{pmatrix} F(x(\lambda, \gamma)) \\ G(x(\lambda, \gamma)) \end{pmatrix} \in \mathbb{R}^{p+q} \end{aligned}$$

On en déduit

$$\begin{cases} \lambda^* = \mathbb{P}_{\mathbb{R}_+^p}[\lambda^* + \rho F(x^*(\lambda^*, \gamma^*))] \\ \gamma^* = \mathbb{P}_{\mathbb{R}^q}[\gamma^* + \rho G(x^*(\lambda^*, \gamma^*))] \end{cases} \qquad \forall \rho > 0$$

et donc

$$\begin{cases} \lambda^* = Max(0, \lambda^* + \rho F(x^*(\lambda^*, \gamma^*))) \\ \gamma^* = \gamma^* + \rho G(x^*(\lambda^*, \gamma^*)) \end{cases} \qquad \forall \rho > 0$$

<u>D'où la méthode itérative d'uzawa</u> :

1. **Initialisation**.
   k=0 : on choisit $\lambda^0 \in \mathbb{R}^p_+$ et $\gamma^0 \in \mathbb{R}^q$.
2. **Itération k.**
   $\lambda^k = (\lambda^k_1, ..., \lambda^k_p) \in \mathbb{R}^p_+$ et $\gamma^k = (\gamma^k_1, ..., \gamma^k_q) \in \mathbb{R}^q$ sont connus.
   (a) Calcul de $x^k \in \mathbb{R}^n$ solution de :

$$(P_k) \qquad \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda^k, \gamma^k).$$

   (b) Calcul de λ^{k+1} et γ^{k+1} par les formules :

$$\begin{cases} \lambda^{k+1}_i = max(0, \lambda^k + \rho F_i(x^k)) & i \in \{1, 2, ..., p\} \\ \gamma^{k+1}_j = \gamma^k_j + \rho G(x^k) & j \in \{1, 2, ..., q\} \end{cases}$$

où $\rho > 0$, est un réel fixé (par l'utilisateur).

Théorème 13 *(la convergence) On suppose que J est C^1 et elliptique. Supposons de plus que G est affine, F est convexe de classe C^1 et lipschitziennes. On suppose de plus que le Lagrangien $\mathcal{L}$ possède un point selle $(x^*, \lambda^*, \gamma^*)$ sur $\mathbb{R}^n \times \mathbb{R}^p_+ \times \mathbb{R}^q$. Alors, il existe ρ_1 et ρ_2, avec $0 < \rho_1 < \rho_2$ tels que, pour tout $\rho \in [\rho_1, \rho_2]$, la suite $(x^k)_{k \geqslant 0}$ générée par l'algorithme d'Uzawa converge vers x^*.*

4.1.2 L'algorithme d'Uzawa dans le cas d'une fonction quadratique

Soit $J(x) = \dfrac{1}{2} < Ax, x > - < b, x >$ pour tout $x \in \mathbb{R}^n$ et A : symétrique définie positive. et l'ensemble $\mathbb{U}$ est donnée par

$$\mathbb{U} = \{x \in \mathbb{R}^n / \ C\ x \leqslant f\}$$

ou

$$\mathbb{U} = \{x \in \mathbb{R}^n / \ C\ x = f\}$$

où la matrice C est $p \times n$ et $f \in \mathbb{R}^p$.

Proposition 9 *Soit* $\lambda \in \mathbb{R}^p$ *et* $x \in \mathbb{R}^n$ *tel que* (x, λ) *est un point selle de lagrangien* $\mathcal{L}(x,\lambda) = \dfrac{1}{2} < Ax, x > - < b, x > + < \lambda, Cx - f >$*. Pour tout* $\rho > 0$*, on :*

$$\lambda = \mathbb{P}_{\mathbb{F}}(\lambda + \rho(Cx - f))$$

$\mathbb{P}_{\mathbb{F}}$ *étant la projection de* $\mathbb{R}^p$ *sur le convexe fermé* $\mathbb{F}$ *(et rappelons que* $F = \mathbb{R}^p$ *dans le cas de contraintes d'égalité, et* $F = \mathbb{R}^p_+$ *dans le cas de contraintes d'inégalité).*

Preuve. On a

$$< \lambda - \mu, Cx - f > \geqslant 0 \qquad \forall \mu \in \mathbb{F}$$

alors

$$< \lambda - \mu, \lambda - (\lambda + \rho(Cx - f)) > \geqslant 0$$

Alors d'après la théorème de projection sur un convexe fermé (de chapitre 2) on

$$\lambda = \mathbb{P}_{\mathbb{F}}(\lambda + \rho(Cx - f)) \qquad \forall \rho > 0.$$

■

Algorithme

1. On choisit une condition initiale $\lambda_0 \in \mathbb{R}^p$, un pas $\rho > 0$ et une prècision $\eta > 0$
   On calcule x_0 solution de $Ax = b - C^T\lambda_0$.

2. **Tant que** $||\lambda_k - \lambda_{k-1}|| > \eta$ ou $||x_k - x_{k-1}|| > \eta$ on définit (x_{k+1}, λ_{k+1}) par

$$\begin{cases} \lambda_{k+1} = \mathbb{P}_{\mathbb{F}}(\lambda_k + \rho(Cx_k - f)) \\ x_{k+1} \ \ \textit{est solution de} \ \ Ax_k = b - C^T\lambda_{k+1}. \end{cases}$$

Théorème 14 *(La convergence) Supposons que A est symétrique définie positive. Si* $0 < \rho < \dfrac{2\lambda_{min}(A)}{||C||^2}$*, alors quel que soit l'élément initial* $\lambda_0 \in \mathbb{R}^p$*, la suite* (x_k) *définie par l'algorithme d'Uzawa converge vers le minimum* x*.*

Preuve. Grâce à des calculs directs et au théorème de projection sur un convexe fermé,on obtient :

$$||\lambda_{k+1} - \lambda||_2^2 \ = \ ||\mathbb{P}_{\mathbb{F}}(\lambda_k + \rho(Cx_k - f)) - \mathbb{P}_{\mathbb{F}}(\lambda + \rho(Cx - f)||_2^2$$

$$\leqslant ||\lambda_k - \lambda + \rho C(x_k - x)||_2^2$$
$$\leqslant ||\lambda_k - \lambda||_2^2 + \rho^2 ||C||^2 ||x_k - x||_2^2 + 2\rho < C^T(\lambda_k - \lambda), x_k - x >$$
$$\leqslant ||\lambda_k - \lambda||_2^2 + \rho^2 ||C||^2 ||x_k - x||_2^2 + 2\rho < A(x_k - x), x_k - x >$$
$$\leqslant ||\lambda_k - \lambda||_2^2 \; + \; (\rho^2 ||C||^2 - 2\rho\lambda_{min}(A))||x_k - x||_2^2.$$

Si $0 < \rho < \dfrac{2\lambda_{min}(A)}{||C||^2}$ en posant $\beta = 2\rho\lambda_{min}(A) - \rho^2||C||^2$, il vient que $\beta > 0$ et

$$||x_k - x||_2^2 \leqslant \frac{1}{\beta}\left(||\lambda_k - \lambda||_2^2 \; - \; ||\lambda_{k+1} - \lambda||_2^2\right).$$

Ce qui prouve que la suite $(||\lambda_k - \lambda||_2^2)_{k \geqslant 0}$ est décroissante et minorée (évidement par 0). Par conséquent, $||x_k - x||_2^2 \underset{k \to +\infty}{\longrightarrow} 0$. ■

4.2 Problème du Portefeuille en finance

4.2.1 Enoncé

Un portefeuille en finance désigne une collection d'actifs financiers détenus par un établissement ou un individu.

- Un investisseur, disposant d'un budget donnée, souhaite se constituer un portefeuille.
- Il peut choisir entre n titres différents.
- Comment constituer son portefeuille de façon à minimiser son risque et à se garantir un certain minimum de rendement espéré ?
- Une approche présentée en 1952 par Markowitz (lauréat d'un prix Nobel en économie) formule ce problème sous la forme d'un modéle quadratique.

4.2.2 Modélisation

On considère un portefeuille d'actions composé de $n > 2$ actions à risque $(a_1, ..., a_n)$. On note x_i, la proportion de l'action a_i dans le portefeuille. Le vecteur $x = (x_1, ..., x_n)^T$ représente donc la composition du portefeuille.Il est évident que x vérifie alors :

$$\sum_{i=1}^{n} x_i = 1, \quad et \quad x_i \geqslant 0 \qquad \forall i \in \{1, ..., n\}$$

Le rendement de l'action ai est modélisé par une variable aléatoire R_i de moyenne $r_i = \mathbb{E}(R_i)$.On introduit le vecteur de rendement moyen $r = (r_1, ..., r_n)^T$, et la matrice de covariance $A = (a_{i,j})_{1 \leqslant i,j \leqslant n}$ définie par la relation :

$$a_{i,j} = \mathbb{E}[(R_i - \mathbb{E}(R_i))(R_j - \mathbb{E}(R_j))], \quad \forall i, j \in \{1, 2, ..., n\}$$

La matrice A est symétrique définie positive.
Le risque du portefeuille est calculé par la fonctionnelle $\sigma(x) = \frac{1}{2} < Ax, x >$

On dit qu'un portefeuille x est *efficient* s'il assure à la fois un rendement maximal pour un risque donne ; et un risque minimal pour un rendement imposé.

Pour $\rho \in \mathbb{R}$ et $\sigma \in \mathbb{R}_+$ donnés, on définit les ensembles :

$$C_1(\rho) = \left\{ x \in \mathbb{R}^n / \ \sum_{i=1}^{n} x_i = 1 \quad et \quad \sum_{i=1}^{n} r_i x_i \geqslant \rho \right\}$$

$$C_2(\sigma) = \left\{ x \in \mathbb{R}^n / \ \sum_{i=1}^{n} x_i = 1 \quad et \quad \frac{1}{2} < xA, x >= \sigma \right\}$$

Donc on cherche à résoudre les problèmes

$$(P_1) \quad \begin{cases} min \ \frac{1}{2} < Ax, x > \\ x \in C(\rho) \end{cases} \qquad\qquad (P_2) \quad \begin{cases} Max \ \sum_{i=1}^{n} r_i x_i \\ x \in C(\sigma) \end{cases}$$

Ainsi on considère le problème (P) suivant qui rassemble les deux problèmes (P_1) et (P_2)

$$(P) \quad \begin{cases} Max(\sum_{i=1}^{n} r_i x_i - \beta \frac{1}{2} < Ax, x >) \\ SC \\ \quad \sum_{i=1}^{n} x_i = 1 \\ \quad x \geqslant 0 \end{cases}$$

Où β est donnée, plus il est grand plus on minimise le risque.

4.2.3 Résolution

Le problème de sélection des portefeuille est un problème d'optimisation quadratique avec contraintes, qu'on peut utiliser l'algorithme d'Uzawa pour le résoudre.

On a

$$\begin{cases} Max(\sum_{i=1}^{n} r_i x_i - \beta\frac{1}{2} < Ax, x >) \\ SC \\ \quad \sum_{i=1}^{n} x_i = 1 \\ \quad x \geqslant 0 \end{cases} \Leftrightarrow \begin{cases} -min(-\sum_{i=1}^{n} r_i x_i + \beta\frac{1}{2} < Ax, x >) \\ SC \\ \quad \sum_{i=1}^{n} x_i = 1 \\ \quad x \geqslant 0 \end{cases}$$

Algorithme d'Uzawa applique au MATLAB

```
function [x,labmda] =UZAWA(A,C,b,β,f,rho,eps)
[n,m]=size(C) ;
lambda=ones(n,1) ;
x=inv(A)*(-lambda'*C+b')' ;

    y=1000*ones(m,1) ;
while(norm(y-x)>eps)
y=x ;

lambda=max(zeros(n,1),lambda+rho*(C*x-f)) ;
x=inv(A)*(-lambda'*C+b')' ;
end
-(0.5*β*x'*A*x-b'*x)

end
```

Exemplaire de problème de portfeuille

On suppose que un investisseur dispose d'un budget de 50 millions dirhams et d'un protefeuille composé de cinq type d'actions,le but c'est de chercher le portion de budgets investi dans chaque titre a condition de dépasser un rendement moyenne $\sigma = 0.3$ et de minimiser le risque, et on suppose que l'étude de marché et certaines facteurs économiques permettent à l'investisseur d'estimer le rendement espéré,la variance de risque :

le titre	le rendement espiré	la variance de risque
1	0.4	99.22
2	0.5	103.01
3	0.3	98.2
4	0.55	105.5
5	0.6	103.1

de plus on suppose que les risques des titres sont indépendants au sens probabiliste on note par x_1, x_2, x_3, x_4, et x_5 les proportions de budget investi respectivement dans les titres 1, 2, 3, 4 et 5.

Modèle

$$\begin{cases} -min(-\sum_{i=1}^{5} r_i x_i + \beta \frac{1}{2} < Ax, x >) \\ SC \\ \quad \sum_{i=1}^{5} x_i = 1 \\ \quad x \geqslant 0 \end{cases}$$

$$\text{avec } A = \begin{pmatrix} 99.22 & 0 & 0 & 0 & 0 \\ 0 & 103.01 & 0 & 0 & 0 \\ 0 & 0 & 98.2 & 0 & 0 \\ 0 & 0 & 0 & 105.5 & 0 \\ 0 & 0 & 0 & 0 & 103.1 \end{pmatrix} \text{ et } \mathbb{U} = \{x \in \mathbb{R}^n / \ C\ x \leqslant f\}$$

$$\text{tel que } \sum_{i=1}^{5} x_i = 1 \Leftrightarrow \sum_{i=1}^{5} x_i \leqslant 1 \ \ et \ \ -\sum_{i=1}^{5} x_i \leqslant -1 \ \ et \ \ -x \leqslant 0$$

$$C = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 & -1 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} \text{ et } f = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{et} \quad b = r = \begin{pmatrix} 0.4 \\ 0.5 \\ 0.3 \\ 0.55 \\ 0.6 \end{pmatrix} \quad \beta = 10$$

On appliquant c'est donnés dans l'algorithme sur MATLAB on trouve :

```
>> A=[99.22 0 0 0 0;0 103.11 0 0 0;0 0 98.2 0 0 ;0 0 0 105.5 0; 0 0 0 0 103.1];
>> b=[0.4 0.5 0.3 0.55 0.6]';
>> C=[1 1 1 1 1 ;-1 -1 -1 -1 -1 ;-eye(5)];
>> f=[1 -1 0 0 0 0 0]';
>> UZAWA(A,C,b,f,1,10^-6)

ans =

 -101.2800


ans =

    0.2044
    0.1977
    0.2055
    0.1937
    0.1987

>>
```

Voila ! les portions de budget qu'on peut investir à chaque titre

le titre	le portion de budget investi
1	$50 \times 0,2044 = 10,22$
2	$50\ \times 0.1977 = 9.8850$
3	$50 \times 0.2055 = 10.2750$
4	$50 \times 0.1937 = 9.6850$
5	$50 \times 0.1987 = 9.9350$

Bibliographie

[1] M.Minoux, *Programmation mathématique*, Tome 1, Dunod (1987).

[2] Introduction à l'Optimisation Numérique, Frédéric de Gournay & Aude Rondepierre , www.math.univ-toulouse.fr .

[3] Gregoire Allaire ,Analyse numérique et optimisation, pages 324-327.

[4] Optimisation sous contraintes, Analyse numérique I, Université d'Aix-Marseille, R. Herbin, 10 novembre 2015

[5] Introduction à l'optimisation : aspects théoriques, numériques et algorithmes, Xavier ANTOINE , Pierre DREYFUSS et Yannick PRIVAT (2006-2007)