République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITÉ MOHAMED KHIDER, BISKRA

FACULTÉ des SCIENCES EXACTES et des SCIENCES de la NATURE et de la VIE

DÉPARTEMENT DE MATHÉMATIQUES

Mémoire présenté en vue de l'obtention du Diplôme :

MASTER en Mathématiques

Option : **Analyse**

Par

BERIBECHE ICHRAK

Titre :

Conditions De Lagrange En Optimisation Avec Contraintes

Membres du Comité d'Examen :

Dr	HOUAS AMRANE	UMKB	Président
Dr	BELLAGOUNE ABDELGHANI	UMKB	Encadreur
Dr	ADOUANE SAIDA	UMKB	Examinatrice

juin 2022

Dédicace

Je dédie cet humble travail

A ma chère mère et à mon cher père.

A mes chères soeurs et frères

A toute ma famille.

A tous mes amis et collègues.

A tous ceux qui m'ont aidé à terminer ce mémoire.

REMERCIEMENTS

Ce mémoire est le fruit du travail et d'efforts fournis pendant une année, malgré les difficultés et les circonstances rencontrées durant le travail, les études et même dans la vie personnelle, je l'ai terminé.

En premier lieu, je voudrais remercier « Dieu » le Tout-Puissant de m'avoir aidé à réussir et de m'avoir donné le courage et la patience nécessaires pour mener à bien ce travail.

L'accomplissement de ce travail n'aurait pas été possible sans le soutien et la collaboration de nombreuses personnes que je tiens à remercier sincèrement :

Tout d'abord, je tiens à remercier .monsieur **BELLAGOUNE ABDELGHANI**, *ce lui qui a été le meilleure conseiller avec son soutien, ses conseils, ses opinions,sa disponibilité je lui adresse les plus grands remerciement et gratitude pour la qualité de son encadrement.*

Je remercie ensuite l'ensemble des membres du jury, qui m'ont fait l'honneur de bien vouloir étudier avec attention mon travail : Monsieur ***HOUAS AMRAN****; maitre de conférences classe B; pour avoir accepté d'être président de jury de ce mémoire; Madame* ***ADOUANE SAIDA****; MCB; pour avoir accepté d'examiner ce mémoire. Je leurs exprime ma profonde gratitude.*

Je tiens également à remercier tous nos enseignants du département de mathématiques de l'université de Mohamed khider, qui ont contribué à notre formation durant cette année de master et sans oublier un remerciement à tous ceux qui m'ont soutenu.

beribeche

Table des matières

Table des figures

Introduction

Euler dit :" il n'y a rien au monde qui ne se réalise sans la volonté de minimiser ou maximiser quelque chose ''.

L'optimisation est une branche des mathématiques et de l'informatique en tant que disciplines, cherchant à modéliser, à analyser et à résoudre analytiquement ou numériquement les problèmes réels : qui consiste à déterminer quelle est la ou les solution(s) (inconnues) satisfaisant un objectif quantitatif tout en respectant d'éventuelles contraintes.

Donc cela revient à Résoudre un problème d'optimisation modèle mathématique (formel) d'un problème réel.

On cherche à minimiser (ou maximiser) une fonction objectif sous des contraintes :

$x \in \mathbb{R}^n$: ensemble des variables

$f : \mathbb{R}^n \longrightarrow \mathbb{R}$: fonction cout (objectif)

$D \subset \mathbb{R}^n$:Ensemble des contraintes$g(x)$

$$PO = \left\{ \begin{array}{c} \min(\max) f(x) \\ g(x) \leq 0 \\ x \in \mathbb{R}^n \end{array} \right\}$$

Le domaine d'application d'un problème d'optimisation est l'Ingénierie, Automatique, Recherche opérationnelle, Analyse numérique, Economie/Finances, Transport, Logistique .....

L'objectif de ce mémoire est de fournir un support sur l'optimisation sous contraintes et les principales méthodes utilisées pour les résoudre.

Après une introduction générale le 1er chapitre est consacré à une introduction au calcul différentiel, le 2ème chapitre sur Conditions de Lagrange en optimisation avec contraintes et quelques algorithmes d'optimisation sous contrainteset. Dans le 3ème chapitre ; on expose un exemple économique .Et on termine par une conclusion générale.

Chapitre 1

Introduction au calcul différentiel

1.1 Généralités

L'optimisation consiste en la recherche du minimum (ou du maximum) d'une certaine quantité, appelée coût ou objectif. Dans ce chapiter, on supposera que le coût dépend de N variables réelles, rassemblées en un vecteur $x = (x_1, ..., x_N) \in \mathbb{R}^N$, et qui fournissent une valeur $J(x)$ où J est une fonction de $\mathbb{R}^N$ dans $\mathbb{R}$. En général, les variables $x_1, ..., x_N$ ne seront pas autorisées à prendre n'importe quelle valeur, mais devront satisfaire des contraintes que l'on représentera par un sous ensemble $K \subset \mathbb{R}^N$. On écrira les problèmes d'optimisation sous la forme générale suivante :

$$\inf_{x \in K} J(x) \tag{1.1}$$

On dit que problème 1.1 admet une solution s'il existe un choix de variables $x_0 \in K$ tel que $\forall x_0 \in K \qquad J(x_0) \ \leq J(x)$

On dit alors que x_0 est un minimiseur (ou point de minimum) de J sur K, et que $J(x_0)$ est un minimum de J sur K.

Exemple 1.1.1 *Un étudiant doit réviser pour ses examens. Il a* 4 *matières à passer et dispose d'une semaine de révisions,ce qui représente* 42 *heures de travail (en comptant* 6 *jours et* 7 *heures par jour). Pour* $i = 1, .., 4$*,on note* x_i *le nombre d'heures de révisions.*

Pour la matière numéro i*. L'ensemble* K *est alors décrit par :*

$$K = \left\{ \begin{array}{ccc} & & x \in \mathbb{R}^4 \\ \sum_{i=1}^{4} x_i \leq 42 & telque & \forall 1 \leq i \leq 4 \\ & & x_i \geq 0 \end{array} \right\}$$

On note $M(x)$ *la moyenne des notes* $(sur\ 20)$ *obtenues par l'étudiant après avoir révisé* x_i *heures la matière numéro* i*. L'objectif est de maximiser* $M(x)$*, ce qui revient à minimiser la différence* $20 - M(x)$*. On peut donc formuler le problème d'optimisation suivant :* $\inf_{x \in K} (20 - M(x))$

Remarque 1.1.1 *Bien sûr, dans l'exemple précédent, on ne connaît pas la formule de* $M(x)$ *de manière explicite ! De plus, il est évident que la moyenne obtenue ne dépend pas seulement du nombre d'heures de révisions, mais de beaucoup d'autres paramètres (assiduité en TD, concentration lors des révisions, qualité du sommeil la veille des épreuves...). Le choix de la fonction coût découle d'un choix de modélisation du phénomène étudié. Cependant, dans ce cours, les fonctions coût seront considérées comme des données du problème.[4]*

1.2 Rappels de calcul différentiel

On se place dans $\mathbb{R}^N$, muni de la norme euclidienne $\|-\|$ et du produit scalaire euclidien $\langle .,. \rangle$.

Pour $x \in \mathbb{R}^N$ et $R > 0$, on note $B_R(x)$ la boule ouverte de centre x et de rayon R et

$\bar{B}_R(x)$ la boule fermée correspondante :$B_R(x) = \{y \in \mathbb{R}^N, \|y - x\| < R\}$

et$\bar{B}_R(x) = \{y \in \mathbb{R}^N, \|y - x\| \leq R\}$

Définition 1.2.1 *1. Un ensemble $U \subset \mathbb{R}^N$ est* ouvert *si*

$$\forall x \in U \quad , \quad \exists R > 0 \quad , \quad B(x, R) \subset U$$

2. Un ensemble $F \subset \mathbb{R}^N$ est fermé *si son complémentaire $\mathbb{R}^N \setminus F$ est ouvert.*

Définition 1.2.2 *Soit $U \subset \mathbb{R}^N$ un ouvert et $J : U \to \mathbb{R}$ une application. Soit $u \in U$.*

1. *On dit que J est* différentiable *au point u s'il existe une application linéaire $L \in \boldsymbol{\mathcal{L}}\left(\mathbb{R}^N, \mathbb{R}\right)$ tel que pour tout $h \in \mathbb{R}^N$ et $(u + h) \in U$*

   $(J(u + h) = J(u) + L(h) + o(\|h\|)$.

   La notation $o(\|h\|)$ signifie qu'il existe une fonction $\varepsilon : \mathbb{R}^N \to \mathbb{R}$ telque $\lim\limits_{h \to 0} \varepsilon(h) = 0$, et qui permette d'écrire le reste sous la forme $o(\|h\|) = \|h\| \varepsilon(h)$ Si L existe, elle est unique, on la note : $L = DJ(u)$.

2. *Soit $d \in \mathbb{R}^N \setminus \{0\}$. On dit que J admet une* dérivèe directionnelle *dans la direction d, au point u, si l'application $t \in \mathbb{R} \to J(u + td)$ est dérivable en 0.Si c'est le cas, on note cette dérivée :*

   $$\frac{\partial J}{\partial d}(u) = \lim_{t \to 0} \frac{J(u + td) - J(u)}{t}.$$

   Si $d = e_i$ est l'un des vecteurs de base de $\mathbb{R}^N$, on appelle cette dérivée directionnelle la $i -$ème dérivée partielle *de J au point u, que l'on note :*

$$\frac{\partial J}{\partial x_i}(u) = \lim_{t \to 0} \frac{J(u + te_i) - J(u)}{t} = \lim_{t \to 0} \frac{J(u_1, ..., u_{i-1}, u_i + t, u_{i+1}, ..., u_N) - J(u_1, ..., u_N)}{t}.$$

Proposition 1.2.1 *Si $J : U \to \mathbb{R}$ est différentiable au point $u \in U$, alors elle admet des dérivées directionnelles en toute direction au point u (et en particulier,* *des dérivées partielles). De plus, sa différentielle au point u s'écrit*

$$\forall h = (h_1, ..., h_N) \in \mathbb{R}^N \quad / \quad DJ(u)(h) \quad = \quad \sum_{i=1}^{N} \frac{\partial J}{\partial x_i}(u)\, h_i.$$

En introduisant le gradient de J *au point* u*, défini par* $:\bigtriangledown J(u) = \left[\frac{\partial J}{\partial x_1}(u) \ldots . \frac{\partial J}{\partial x_N}(u)\right]^T$, *on peut écrire de manière condensée*

$$DJ(u)(h) = \langle \bigtriangledown J(u), h \rangle . \tag{1.2}$$

On montre que si J *est différentiable au point* u*, alors* $\bigtriangledown J(u)$ *est l'unique vecteur de* $\mathbb{R}^N$ *t.q. la relation 1.2 soit vérifiée.*

Remarque 1.2.1 (Calcul du gradient) *Pour calculer le gradient de* J *au point* u*, il n'est pas toujours nécessaire de calculer explicitement toutes les dérivées partielles. Une autre méthode consiste à établir un développement limité de* J *sous la forme suivante :*

$$J(u+h) = J(u) + \langle \omega, h \rangle + o(\|h\|)$$

où $\omega \in \mathbb{R}^N$ *est un certain vecteur fixé. Alors, on peut affirmer que* J *est différentiable en* u*, et que :*

$$\omega = \bigtriangledown J(u) .$$

1.2.1 La Hessienne d'une fonction

On appelle matrice hessienne $H(x_1, x_2, .., x_n)$ de f la matrice des dérivées secondes de f évaluées au point $(x_1, x_2, .., x_n)$.

$$H(x_1, x_2, .., x_n) = \begin{pmatrix} f''_{11} & f''_{12} & \cdots & f''_{1n} \\ f''_{21} & f''_{22} & \cdots & f''_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ f''_{n1} & f''_{n2} & \cdots & f''_{nn} \end{pmatrix}$$

Comme $f''_{ij} = f''_{ji}\ \forall (i,j)$, la matrice hessienne de f est une matrice symétrique d'ordre n.

1.2.2 Matrice Jacobienne

Soit $G = (g_1, ..., g_m)$ une fonction définie de $\mathbb{R}^N \to \mathbb{R}^N$. A tout vecteur $x = (x_1, x_2, ..., x_n)$, la fonction G associée le vecteur de fonction $(g_1(x), g_2(x), ..., g_m(x))$.
On appelle matrice *jacobienne* de G la matrice de dimension (m, n) ; $J_G(x_1, x_2, .., x_n)$ des dérivées partielles des m fonctions qui composent G :

$$J_G(x_1, x_2, .., x_n) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x) & \frac{\partial g_1}{\partial x2}(x) & \cdots & \frac{\partial g_1}{\partial x_n}(x) \\ \frac{\partial g2}{\partial x_1}(x) & \frac{\partial g2}{\partial x_2}(x) & \cdots & \frac{\partial g_2}{\partial x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1}(x) & \frac{\partial g_m}{\partial x_2}(x) & \cdots & \frac{\partial g_m}{\partial x_n}(x) \end{pmatrix}.$$

1.3 Forme quadratique

Définition 1.3.1 *Soit $A_{(n\times n)}$une matrice symétrique et $b \in \mathbb{R}^n$. On appelle forme quadratique la fonction $f : \mathbb{R}^n \to \mathbb{R}$ définie par :*

$$f(x) = \frac{1}{2}x^t Ax - b^t x.$$

Proposition 1.3.1 *Soit A une matrice carrée symétrique. Soit x un vecteur colonne quelconque. On note x^t sa transposée.*

1. *Si* $x^t Ax \geq 0 \quad , \quad \forall x \in \mathbb{R}^n \quad : \quad A \; est \; semi-définie \; positive.$
2. *Si* $x^t Ax > 0 \quad , \quad \forall x \in \mathbb{R}^n \quad : \quad A \; est \; définie \; positive.$
3. *Si* $x^t Ax \leq 0 \quad , \quad \forall x \in \mathbb{R}^n \quad : \quad A \; est \; semi-définie \; négative.$
4. *Si* $x^t Ax < 0 \quad , \quad \forall x \in \mathbb{R}^n \quad : \quad A \; est \; définie \; négative.$

Remarque 1.3.1 *Soit A une matrice carré symétrique d'ordre n.*

$*$ *A définie positive $\Leftrightarrow$ ses n mineures principaux diagonaux D_k sont > 0.*

$*$ *A semi-définie positive $\Leftrightarrow$ tous ses mineures principaux D_k sont ≥ 0.*

$*$ *A définie négative $\Leftrightarrow$ ses n mineures principaux diagonaux D_k sont alternativement $< 0(k \; impair)$ et $> 0(k \; pair)$.*

$*$ *A semi définie négative $\Leftrightarrow$ tous ses mineures principaux D_k sont alternativement $0 \leq (k \; impair)$ et $0 \geq (k \; pair)$.*

1.4 Notion de convexté

La convexité joue un role extrêmement important en optimisation.

1.4.1 Ensemble convexe

Définition 1.4.1 – *Un ensembl S de $\mathbb{R}^n$ est convexe ssi :*

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

– *Un ensemble S de $\mathbb{R}^n$ est strictement convexe ssi :*

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

Remarque 1.4.1 *La notion d'ensemble concave n'existe pas.*

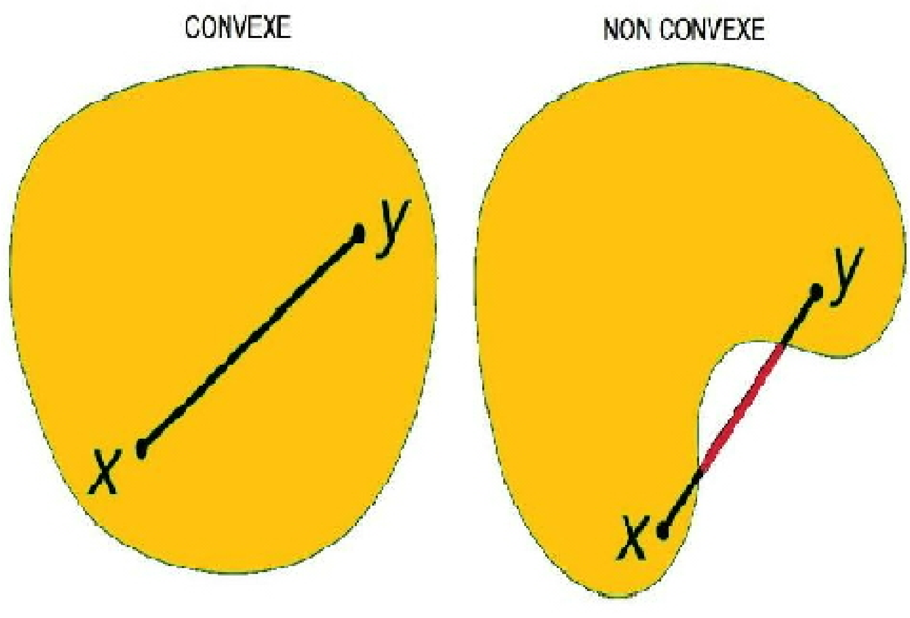

FIG. 1.1 – ensemble convexe et ensemble non convexe

Fonctions concaves, Fonction convexes

Fonction concave

– f est concave sur S ssi,$\forall (x,y) \in S^2$ et $\forall \lambda \in [0,1]$, on a :

$$f((1-\lambda)x + \lambda y) \geq (1-\lambda) f(x) + \lambda f(y).$$

– f est strictement concave sur S ssi :

$$f((1-\lambda)x + \lambda y) > (1-\lambda) f(x) + \lambda f(y).$$

Fonction convexe

– f est convexe sur S ssi,$\forall (x,y) \in S^2$ et $\forall \lambda \in [0,1]$, on a :

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

– f est strictement convexe sur S ssi :

$$f((1-\lambda)x + \lambda y) < (1-\lambda) f(x) + \lambda f(y).$$

Exemple 1.4.1 *Fonction strictement concave*

$$f(x,y) = 6 - (x-1)^2 - (y-1)^2$$

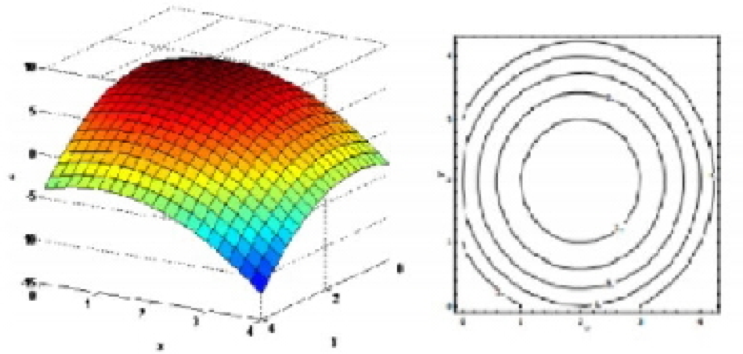

Fig. 1.2 – Fonction concave

Exemple 1.4.2 *Fonction strictement convexe*

$$f(x,y) = (x-1)^2 - (y-1)^2$$

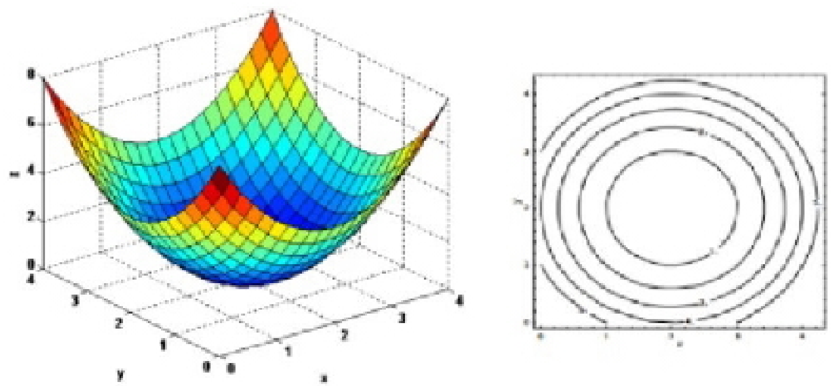

FIG. 1.3 – Fonction convexe

1.5 Extremas locaux et globaux sous contraintes

Un problème d'optimisation s'écrit en général sous la forme :

$$\begin{cases} \min\limits_{(x_1,..,x_n)} f\ (x_1,..,x_n) \\ s.c \quad contraintes \quad (g_j) \qquad \forall j \ \in \ \overrightarrow{1,...,m}. \end{cases}$$

1.5.1 Optimum local

La variable x^* est un maximum local d'une fonction f définie sur l'ensemble convexe $S \Leftrightarrow \exists \varepsilon > 0$ tel que

$$f(x) \quad \leq \quad f(x^*) \quad , \quad \forall x \in S \quad et \quad |x - x^*| \leq \varepsilon.$$

La variable x^* est un minimum local d'une fonction f définie sur l'ensemble convexe

$S \Leftrightarrow \exists\varepsilon > 0$ tel que

$$f(x) \geq f(x^*) \quad , \quad \forall x \in S \quad et \quad |x - x^*| \leq \varepsilon.$$

1.5.2 Optimum global

La variable x^* est un maximum global d'une fonction f définie sur l'ensemble convexe $S \Leftrightarrow$

$$f(x) \leq f(x^*) \quad , \quad \forall x \in S$$

La variable x^* est un minimum global d'une fonction f définie sur l'ensemble convexe $S \Leftrightarrow$

$$f(x) \geq f(x^*) \quad , \quad \forall x \in S$$

Exemple 1.5.1 *Fonction f admettant un maximum local en x^* et un maximum global en x^{**} sur l'ensemble de son domaine de définition.[8]*

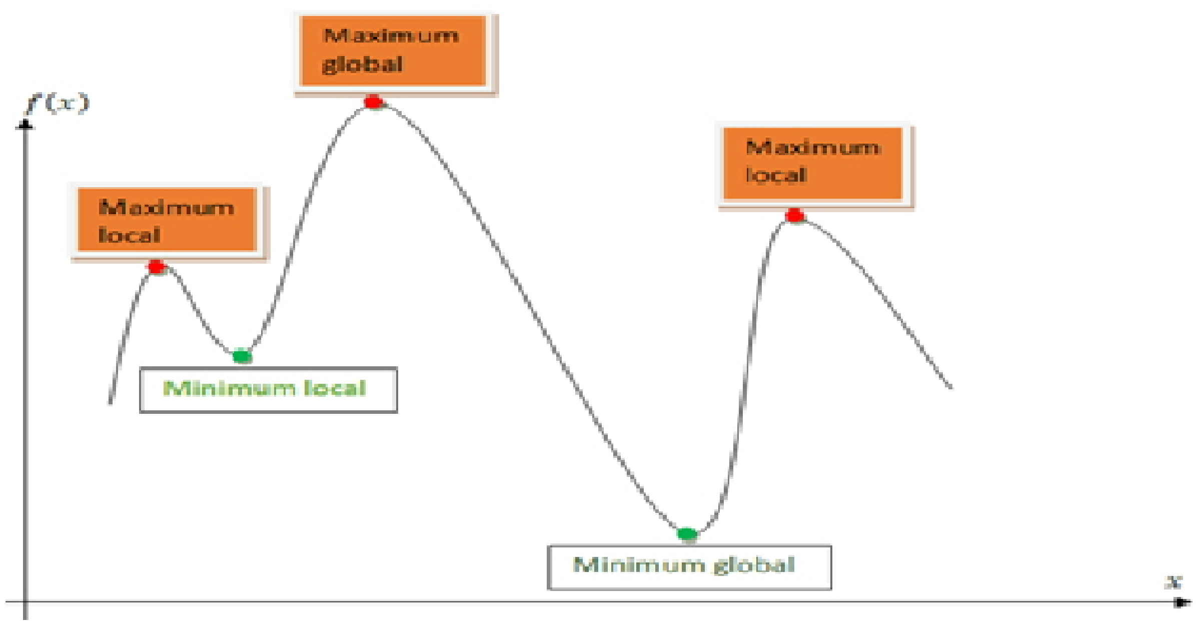

Fig. 1.4 – Extremum local et extremum global

Chapitre 2

Optimisation

Une méthode d'optimisation est une technique de résolution d'un problème équivalent (modèle), une méthode théorique dont la description permet l'élaboration d'un algorithme numériquement applicable.

Le choix d'une technique appropriée dépend de la nature de la fonction objectif f, de sa régularité (continuité, dérivabilité), de propriétés spécifiques (parité, convexité), de la connaissance de voisinages de ses extrema, des contraintes caractérisant l'ensemble D des points admissibles (réalisables)

Processus d'optimisation

– Analyse du problème

– Modélisation et choix de méthode

– Résolution

– Interprétation des résultats ([5])

2.1 Optimisation sous contraintes

Définition 2.1.1 *Soit $E = \mathbb{R}^n$, soit $f \in C(E, \mathbb{R})$, et soit K un sous ensemble de E.On s'intéresse à la recherche de $\bar{\mu} \in K$ tel que :*

$$\begin{cases} \bar{\mu} \in K \\ f(\bar{\mu}) = \inf\limits_{K} f \end{cases} \tag{2.1}$$

ce problème est un problème de minimisation avec contrainte (ou "sous contrainte") au sens où l'on cherche μ qui minimise f en astreignant μ a être dans K. Voyons quelques exemples de ces contraintes (définies par l'ensemble K), qu'on va expliciter à l'aide des p fonctions continues, $g_i \in C(E, \mathbb{R})$, $i = \overrightarrow{1..p}$

1. **Contraintes égalités.** On pose $K = \left\{ x \in E,\ g_i(x) = 0 \ ; i = \overrightarrow{0,..,p} \right\}$.

   On verra plus loin que le problème de minimisation de f peut alors être résolu grâce au théorème des multiplicateurs de Lagrange.

2. **Contraintes inégalités.** On pose $K = \left\{ x \in E,\ g_i(x) \leq 0 \ ; i = \overrightarrow{0,..,p} \right\}$ ..Le problème de minimisation de f peut être résolu grâce au théorème de *Kuhn–Tucker.*

– Programmation linéaire. Avec un tel ensemble de contraintes K, si de plus f est linéaire, c'est-à-dire qu'il existe $b \in \mathbb{R}^n$ tel que $f(x) = b{\cdot}x$, et les fonctions g_i sont affines, c'est-à-dire qu'il existe $b_i \in \mathbb{R}^n$ et $c_i \in \mathbb{R}$ tels que $g_i(x) = b_i.x + c_i$, alors on dit qu'on a affaire ùn problème de "*programmation linéaire*". Ces problèmes sont souvent résolus numériquement à l'aide de l'algorithme de Dantzig, inventé vers 1950.[2]

– Programmation quadratique. Avec le même ensemble de contraintes K, si de plus f est quadratique, c'est-à-dire si f est de la forme $f(x) = \frac{1}{2}Ax{\cdot}x - b{\cdot}x$, et les fonctions $g_i$sont affines, alors on dit qu'on a affaire ùn problème de "*programmation* *quadratique*".

– Programmation convexe. Dans le cas où f est convexe et K est convexe, on dit qu'on a affaire ùn problème de "*programmation convexe*".

2.1.1 Existence – Unicité – Conditions d'optimalité simple

Théorème 2.1.1 *Soit* $E = \mathbb{R}^n$ *et* $f \in C(E, \mathbb{R})$

1- Si K *est un sous-ensemble fermé borné de* E*, alors il existe* $\overline{x} \in K$ *tel que :*$f(\bar{x}) = \inf\limits_K f$.

2- Si K est un sous-ensemble fermé de E, et si f est croissante à l'infini, c'est–à–dire que :

$$f(x) \to \infty \quad quand \quad |x| \to \infty \quad alors \quad \exists \bar{x} \in K \quad telque \quad f(\bar{x}) = \inf_K f.$$

Théorème 2.1.2 Démonstration

1. *Si* K *est un sous-ensemble fermé borné de* E*, comme* f *est continue, elle atteint ses bornes sur* K*, d'où l'existence de* $\bar{x}$
2. *Si* f *est croissante à l'infini, alors il existe* $R > 0$ *tel que si* $\|x\| > R$ *alors* $f(x) > f(0)$, *donc :* $\inf\limits_K f = \inf\limits_{K \cap B_R} f$
   Où B_R *désigne la boule de centre* 0 *et de rayon* R. *L'ensemble* $K \cap B_R$ *est compact, car intersection d'un fermé et d'un compact. Donc, par ce qui précède, il existe* $\bar{x} \in K$ *tel que :* $f(\bar{x}) = \inf\limits_{K \cap B_R} f = \inf\limits_{B_R} f$.

Théorème 2.1.3 (Unicité) *Soit* $E = \mathbb{R}^n$ *et* $f \in C(E, \mathbb{R})$*.On suppose que* f *est strictement convexe et que* K *est convexe. Alors il existe au plus un élément* $\bar{x}$ *de* K *tel que :*

$$f(\bar{x}) = \inf_K f.$$

Démonstration *Supposons que $\bar{x}$ et $\bar{\bar{x}}$ avec $\bar{x} \neq \bar{\bar{x}}$ Alors :*

$$f\left(\frac{1}{2}\bar{x} + \frac{1}{2}\bar{\bar{x}}\right) < \frac{1}{2}f(\bar{x}) + \frac{1}{2}f\left(\bar{\bar{x}}\right) = \inf_K f.$$

On aboutit donc à une contradiction.

Des théorèmes d'existence et d'unicité on déduit immédiatement le théorème d'existence et d'unicité suivant :

Théorème 2.1.4 (Existence et unicité) *Soit $E = \mathbb{R}^n$ et $f \in C(E, \mathbb{R})$ une fonction strictement convexe et K un sous ensemble convexe fermé de E. Si K est borné ou si f est croissante à l'infini, c'est–à–dire si :*

$$f(x) \to \infty \quad quand \quad \|x\| \to \infty,$$

alors il existe un unique élément $\bar{x}$ de K solution du problème de minimisation tel que :

$$f(\bar{x}) = \inf_K f.$$

Remarque 2.1.1 *On peut remplacer $E = \mathbb{R}^n$ par E* espace de Hilbert *de dimension infinie dans le dernier théorème, mais on a besoin dans ce cas de l'hypothèse de convexité de f pour assurer l'existence de la solution.*

Proposition 2.1.1 (Condition simple d'optimalité) *Soit $E = \mathbb{R}^n$, $f \in C(E, \mathbb{R})$ et $\bar{x} \in K$ tel que*

$$f(\bar{x}) = \inf_K f.$$

On suppose que f est différentiable en $\bar{x}$:

1. *Si $\bar{x} \in \mathring{K}$ alors $\triangledown f(\bar{x}) = 0$.*
2. *Si K est convexe, alors $\triangledown f(\bar{x}) . (x - \bar{x}) \geq 0$, $\forall x \in K$*

Démonstration

1. *Si* $\bar{x} \in \mathring{K}$, *alors il existe* $\varepsilon > 0$ *tel que :*

$$B(\bar{x}, \epsilon) \quad \subset \quad K \quad et \quad f(\bar{x}) \quad \leq \quad f(x) \quad , \quad \forall x \in B(\bar{x}, \epsilon) \quad \Rightarrow \bigtriangledown f(\bar{x}) = 0.$$

2. *Soit* $x \in K$ *Comme* $\bar{x}$ *réalise le minimum de* f *sur* K, *on a :*

$$f(\bar{x} + t(x - \bar{x})) = f(tx + (1-t)\bar{x}) \geq f(\bar{x})$$

Pour tout $t \in \left]0;1\right]$, *par convexité de* K. *On en déduit que :*

$$\frac{f(\bar{x} + t(x - \bar{x})) - f(\bar{x})}{t} \geq 0, \forall t \in \left]0;1\right]$$

En passant à la limite lorsque t *tend vers* 0 *dans cette dernière inégalité, on obtient :*

$$\bigtriangledown f(\bar{x}).(x - \bar{x}) \geq 0, \forall x \in K.$$

2.1.2 Conditions d'optimalité dans le cas de contraintes égalité

Dans tout ce paragraphe, on considèrera les hypothèses et notations suivantes :

$$\begin{aligned} & f \in C(\mathbb{R}^n, \mathbb{R}), \quad g_i \in C^1(\mathbb{R}^n, \mathbb{R}), \quad i = \overline{1...p} \\ & K = \left\{u \in \mathbb{R}^n,\ g_i(u) = 0\ , \forall i = \overline{1...p}\right\} \\ & g = (g_1, ..., g_p)^t \in C^1(\mathbb{R}^n, \mathbb{R}^p). \end{aligned} \tag{2.2}$$

Remarque 2.1.2 (Quelques rappels de calcul différentiel) *Comme* $g \in C^1(\mathbb{R}^n, \mathbb{R}^p)$, *si* $u \in \mathbb{R}^n$, *alors* $Dg(u) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$, *ce qui revient à dire, en confondant l'applica-*

tion linéaire $Dg(u)$ *avec sa matrice, que* $Dg(u) \in M_{n,p}(\mathbb{R})$. *Par définition :*

$$\mathrm{Im}\,(Dg(u)) = \{Dg(u)\,z,\ \ z \in \mathbb{R}^n\} \subset \mathbb{R}^p$$

et

$$\mathrm{rang}\,(Dg(u)) = \dim\,(\mathrm{Im}\,(Dg(u))) \leq p.$$

On rappelle de plus que :

$$Dg(u) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_p}{\partial x_1} & \cdots & \frac{\partial g_p}{\partial x_n} \end{pmatrix},$$

et que $\mathrm{rang}\,(Dg(u)) \leq \min(n,p)$. *De plus, si* $\mathrm{rang}\,(Dg(u)) = p$, *alors les vecteurs* $(Dg_i(u))_{i=1\ldots p}$ *sont linéairement indépendants dans* $\mathbb{R}^n$.

2.1.3 multiplicateurs de Lagrange

Théorème 2.1.5 *Soit* $\bar{u} \in K$ *tel que* $f(\bar{u}) = \inf\limits_K f$. *On suppose que* f *est différentiable en* $\bar{u}$ *et* $\dim\,(\mathrm{Im}\,(Dg(\bar{u})) = p, (ou\ \ \mathrm{rang}\,(Dg(\bar{u})) = p))$, *alors :*

$$\exists\,(\lambda_1,..,\lambda_n) \in \mathbb{R}^n \quad / \quad \bigtriangledown f(\bar{u}) \quad + \quad \textstyle\sum_{i=1}^p \lambda_i \bigtriangledown g_i(\bar{u}) \quad = \quad 0$$

Démonstration *Pour plus de clarté, donnons d'abord une idée* "géométrique" *de la démonstration dans le cas* $n = 2$ *et* $p = 1$. *On a dans ce cas* $f \in C^1(\mathbb{R}^2, \mathbb{R})$, *et* $K = \{(x;y) \in \mathbb{R}^2, g(x,y) = 0\}$, *et on cherche* $u \in K$ *telque* $f(\bar{x}) = \inf\limits_K f$. *Traçons dans le repère* (x,y) *la courbe* $g(x,y) = 0$, *ainsi que les courbes de niveau de* f. *Si on se* "promène" *sur la courbe* $g(x,y) = 0$, *en partant du point* P_0 *vers la droite (voir figure 2.1), on rencontre les courbes de niveau successives de* f *et on se rend compte sur le dessin que la valeur minimale que* *prend f sur la courbe $g(x,y)=0$ est atteinte lorsque cette courbe est tangente à la courbe de niveau de f : sur le dessin, ceci correspond au point P_1 où la courbe $g(x,y)=0$,est tangente à la courbe $f(x,y)=3$. Une fois qu'on a passé ce point de tangence, on peut remarquer que f augmente.*

On utilise alors le fait que si φ est une fonction continûment différentiable de $\mathbb{R}^2$dans $\mathbb{R}$. le gradient de φ est orthogonal à toute courbe de niveau de φ, c'est-à-dire toute courbe de la forme $\varphi(x,y)=c$, où $c\in\mathbb{R}$. En effet, soit $(x(t),y(t)),t\in\mathbb{R}$ un paramétrage de la courbe $g(x,y)=c$, en dérivant par rapport à t, on obtient :$\bigtriangledown g(x(t),y(t)).(x'(t),y'(t))^t=0$.

En appliquant ceci à f et g, on en déduit qu'au point de tangence entre une courbe de niveau de f et la courbe $g(x,y)=0$, les gradients de f et g sont colinéaires. Et donc si $\nabla g(u)\neq 0$, il existe $\lambda\neq 0$tel que $\nabla f(u)=\lambda\nabla g(u)$.

Passons maintenant à la démonstration rigoureuse du théorème dans laquelle on utilise le théorème des fonctions implicites.

***Théorème 2.1.6 (des fonctions implicites)** Soient p et q des entiers naturels, soit $h\in C^1(\mathbb{R}^q\times\mathbb{R}^p,\mathbb{R}^p)$,et soient $(\bar{x},\bar{y})\in\mathbb{R}^q\times\mathbb{R}^p$ et $c\in\mathbb{R}^p$ tels que $h(\bar{x},\bar{y})=c$. On suppose que la matrice de la différentielle $D_2h(\bar{x},\bar{y})\,(\in M_p(\mathbb{R}))$ est inversible. Alors il existe $\varepsilon>0$ et $v>0$ tels que pour tout $x\in B(\bar{x},\varepsilon)$, il existe un unique $y\in B(\bar{y},v)$ te que $h(x,y)=c$. on peut ainsi définir une application ϕ de $B(\bar{x},\varepsilon)$ dans $B(\bar{y},v)$ par $\phi(x)=y$. On a $\phi(x)=\bar{y}$, $\phi\in(\mathbb{R}^p,\mathbb{R}^p)$ et*

$$D\phi(x)=-\left[D_2h(x,\phi(x))\right]^{-1}.D_1h(x,\phi(x)).$$

On revient pour compléter la preuve.

Par hypothèse, $Dg(\bar{u})\in\mathcal{L}(\mathbb{R}^n,\mathbb{R}^p)$ et $\operatorname{Im}(Dg(\bar{u}))=\mathbb{R}^p$. Donc il existe un sous espace vectoriel F de $\mathbb{R}^n$ de dimension p, tel que $Dg(\bar{u})$ soit bijective de F dans $\mathbb{R}^p$. En effet, soit $(e_1,...e_n)$ la base canonique de $\mathbb{R}^p$, alors pour tout

$i \in \{1, .., p\}$, il existe $y_i \in \mathbb{R}^n$ tel que $Dg(\bar{u})y_i = e_i$. Soit F le sous espace engendré par la famille $\{y_1, .., y_p\}$; on remarque que cette famille est libre, car si $\sum_{i=1}^{p} \lambda_i y_i = 0$, alors $\sum_{i=1}^{p} \lambda_i e_i = 0$, et donc $\lambda_i = 0, \forall i = 1, .., p$.On a ainsi montré l'existence d'un sous espace F de dimension p telle que $Dg(\bar{u})$ soit bijective (car surjective) de F dans $\mathbb{R}^p$.Il existe un sous espace vectoriel G de $\mathbb{R}^n$, tel que $\mathbb{R}^n = F \bigoplus G$. pour $v \in F$ et $w \in G$, on pose $\bar{g}(w, v) = g(v + w)$ et $\bar{f}(w, v) = f(v + w)$. On a donc $\bar{f} \in C(G \times F, \mathbb{R})$ et $\bar{g} \in C^1(G \times F, \mathbb{R})$. De plus, $D_2\bar{g}(w, v) \in \mathcal{L}(F, \mathbb{R}^p)$, et $\forall z \in F$, on a

$$D_2\bar{g}(w, v) z = Dg(v + w) z.$$

Soit $(\bar{v}, \bar{w}) \in F \times G$ tel que $\bar{u} = \bar{v} + \bar{w}$. Alors $D_2\bar{g}(\bar{w}, \bar{v}) z = Dg(\bar{u}) z$, $\forall z \in F$. L'application $D_2\bar{g}(\bar{w}, \bar{v})$ est une bijection de F sur $\mathbb{R}^p$, car, par définition de F, $Dg(\bar{u})$ est bijective de F sur $\mathbb{R}^p$.
On rappelle que $K = \{u \in \mathbb{R}^n : g(u) = 0\}$ et on définit
$\bar{K} = \{(w, v) \in G \times F, \ \bar{g}(w, v) = 0\}$.Par définition de $\bar{f}$ et de $\bar{g}$, on a :

$$\begin{cases} \forall (\bar{v}, \bar{w}) \in \bar{K} \\ \bar{f}(\bar{w}, \bar{v}) \leq f(w, v). \end{cases}$$

D'autre part, le théorème des fonctions implicites entraîne l'existence de $\varepsilon > 0$ et $v > 0$ tels que pour tout $w \in B_G(\bar{w}, \varepsilon)$,il existe un unique $v \in B_F(\bar{v}, v)$ tel que $\bar{g}(w, v) = 0$. On note $v = \phi(w)$ et on définit ainsi une application

$$\phi \in C^1(B_G(\bar{w}, \varepsilon), v \in B_F(\bar{v}, v)).$$

On déduit que :

$$\begin{cases} \forall w \in B_G(\bar{w}, \varepsilon) \\ \bar{f}(\bar{w}, \phi(\bar{w})) \leq \bar{f}(w, \phi(w)), \end{cases}$$

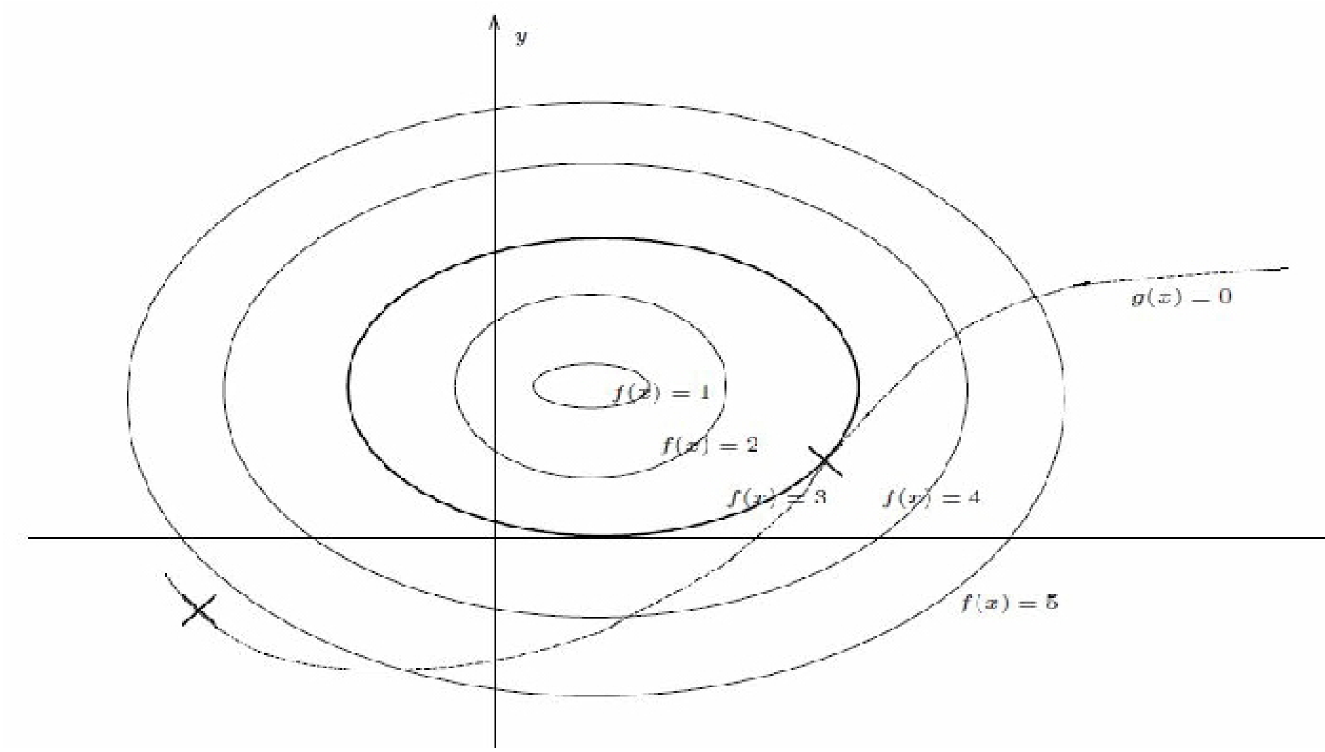

FIG. 2.1 – Interprétation géométrique des multiplicateurs de Lagrange

et donc

$$\begin{cases} \forall w \in B_G(\bar{w}, \varepsilon) \\ f(u) = f(\bar{w} + \phi(\bar{w})) \leq \bar{f}(w + \phi(w)). \end{cases}$$

En posant $\psi(w) = \bar{f}(w, \phi(w))$, on peut donc écrire :

$$\begin{cases} \forall w \in B_G(\bar{w}, \varepsilon) \\ \psi(\bar{w}) = f(\bar{w}, \phi(\bar{w})) \leq \psi(w). \end{cases}$$

On a donc, grâce à la proposition :

$$D\psi(\bar{w}) = 0. \tag{2.3}$$

Par définition de ψ, de $\bar{f}$ et de $\bar{g}$, on a :

$$D\psi(\bar{w}) = D_1\bar{f}\left(\bar{w}, \phi\left(\bar{w}\right)\right) + D_2\bar{f}\left(\bar{w}, \phi\left(\bar{w}\right)\right) D\phi\left(\bar{w}\right).$$

D'après le théorème des fonctions implicites,

$$D\phi\left(\bar{w}\right) = -\left[D_2\bar{g}\left(\bar{w}, \phi\left(\bar{w}\right)\right)\right]^{-1}.D_1\bar{g}\left(\bar{w}, \phi\left(\bar{w}\right)\right).$$

On déduit donc de (2.3) que :

$$\begin{cases} \forall w \in G \\ D_1\bar{f}\left(\bar{w}, \phi\left(\bar{w}\right)\right) w - \left[D_2\bar{g}\left(\bar{w}, \phi\left(\bar{w}\right)\right)\right]^{-1}.D_1\bar{g}\left(\bar{w}, \phi\left(\bar{w}\right)\right) w = 0. \end{cases} \tag{2.4}$$

De plus, comme

$$\left[D_2\bar{g}\left(\bar{w}, \phi\left(\bar{w}\right)\right)\right]^{-1}.D_2\bar{g}\left(\bar{w}, \phi\left(\bar{w}\right)\right) = Id,$$

On a $\forall z \in F$:

$$D_2\bar{f}\left(\bar{w}, \phi\left(\bar{w}\right)\right) z - D_2\bar{f}\left(\bar{w}, \phi\left(\bar{w}\right)\right)\left[D_2\bar{g}\left(\bar{w}, \phi\left(\bar{w}\right)\right)\right]^{-1} D_2\bar{g}\left(\bar{w}, \phi\left(\bar{w}\right)\right) = 0. \tag{2.5}$$

Soit $x \in \mathbb{R}^n$ et $(z, w) \in F \times G$ tel que $x = z + w$. En additionant 2.4 et 2.5, et en notan :

$$\Lambda = -D_2\bar{f}\left(\bar{w}, \phi\left(\bar{w}\right)\right)\left[D_2\bar{g}\left(\bar{w}, \phi\left(\bar{w}\right)\right)\right]^{-1},$$

On obtient :

$$Df\left(\bar{u}\right) x + \Lambda Dg\left(\bar{u}\right) x = 0,$$

Ce qui donne, en transposant :

$$Df\,(\bar{u}) + \sum_{i=1}^{p} \lambda_i \bigtriangledown g_i\,(\bar{u}) = 0$$

$$(\lambda_1, ..., \lambda_p) = \Lambda.$$

Remarque 2.1.3 (Utilisation pratique du théorème de Lagrange) *Soit* $f \in C^1(\mathbb{R}^n, \mathbb{R})$, $g = (g_1, ..., g_p)^t$ *avec* $g_i \in C(\mathbb{R}^n, \mathbb{R})$, *pour* $i = 1...p$, *et soit*

$K = \{u \in \mathbb{R}^n, g_i(u) = 0, i = 1..p.\}$.

Le problème qu'on cherche à résoudre est le problème de minimisation (1.1) qu'on rappelle ici :

$$\begin{cases} \bar{\mu} \in K \\ f\,(\bar{\mu}) = \inf\limits_{K} f. \end{cases}$$

D'après le théorème des multiplicateurs de Lagrange, si $\bar{\mu}$ *est solution de (1.1) et* $\mathrm{Im}(Dg(\bar{\mu})) = \mathbb{R}^p$, *alors il existe* $(\lambda 1, ..., \lambda_p) \in \mathbb{R}^p$, *tel que* $\bar{\mu}$ *est solution du problème :*

$$\begin{cases} \frac{\partial f}{\partial x_j}(\bar{\mu}) + \sum_{i=1}^{p} \lambda_i \frac{\partial g_i}{\partial x_j}(\bar{\mu}) = 0, \; j = 1, .., n \\ g_i(\bar{\mu}) = 0, \quad i = 1, ..., p. \end{cases} \tag{2.6}$$

Le système (2.6) est un système non linéaire de de $(n+p)$ *équations et à* $(n+p)$ *inconnues* $(\bar{x}, ..., \bar{x}_n, \lambda_i ... \lambda_p)$. *Ce système sera résolu par une méthode de résolution de système non linéaire (Newton par exemple [1]).*

Remarque 2.1.4 *On vient de montrer que si* $\bar{x}$ *solution de (1.1)*
et $\mathrm{Im}(Dg(\bar{x})) = \mathbb{R}^p$, *alors* $\bar{x}$ *solution de (2.6). Par contre, si* $\bar{x}$ *est solution de (2.6), ceci n'entraîne pas que* $\bar{x}$ *est solution de (1.1).*

Exemple 2.1.1 *On considère la fonction* $f(x, y) = x^2 + y^2 - 4xy$ *soumise à la contrainte* $x^2 + y^2 = 8$. *Quels sont les extremums de cette fonctions ?*

On doit résoudre un problème d'extremum pour une fonction de deux variables soumise à une contrainte donnée sous forme d'égalité. On utilise donc la méthode du Lagrangien. Posons $\varphi(x, y) = x^2 + y^2$. Les fonctions f et φ sont des polynômes donc admettent des dérivées partielles continues de tous les ordres sur l'ouvert $U = \mathbb{R}^2$.

Régularité de la courbe de contrainte. Avant toute chose, on vérifie que la courbe de contrainte $\varphi(x; y) = 8$ est régulière. Pour cela, on montre que le système suivant n'a pas de solutions :

$$\begin{cases} \frac{\partial \varphi}{\partial x} = 0 \\ \frac{\partial \varphi}{\partial y} = 0 \\ \varphi(x; y) = 8 \end{cases} \Leftrightarrow \begin{cases} 2x = 0 \\ 2y = 0 \\ x^2 + y^2 = 8 \end{cases} \Leftrightarrow \begin{cases} x = y = 0 \\ 0 = 8 \end{cases}$$

La dernière équation est impossible, donc le système n'a pas de solution.

Points stationnaires du Lagrangien. On forme le Lagrangien

$$L(x, y, \lambda) = f(x, y) - \lambda(\varphi(x; y) - 8) = x^2 + y^2 - 4xy - \lambda\left(x^2 + y^2 - 8\right),$$

et on détermine les points stationnaires, c'est-à-dire les solutions du système suivant :

$$\begin{cases} \frac{\partial L}{\partial x} = 0 \\ \frac{\partial L}{\partial y} = 0 \\ \frac{\partial L}{\partial \lambda} = 0 \end{cases} \Leftrightarrow \begin{cases} 2x - 4y - 2\lambda x = 0 \\ 2y - 4x - 2\lambda y = 0 \\ x^2 + y^2 = 8 \end{cases}$$

Notons que, d'après les deux premières équations, si x ou y est nul, alors ils le sont tous les deux. Or, d'après la troisième équation, x et y ne peuvent pas être nuls en même temps, donc x et y sont nécessairement non nuls. On peut donc diviser par x

et y à volonté :

$$\begin{cases} \lambda = \frac{x-2y}{x} \\ \lambda = \frac{y-2x}{y} \\ x^2+y^2=8 \end{cases} \Leftrightarrow \begin{cases} \lambda = \frac{x-2y}{x} \\ \frac{x-2y}{x} = \frac{y-2x}{y} \\ x^2+y^2=8 \end{cases} \Leftrightarrow \begin{cases} \lambda = \frac{x-2y}{x} \\ xy-2y^2 = xy-2x^2 \\ x^2+y^2=8 \end{cases} \Leftrightarrow \begin{cases} \lambda = \frac{x-2y}{x} \\ x^2 = y^2 \\ x^2+y^2=8 \end{cases}$$

$$\Leftrightarrow \begin{cases} \lambda = \frac{x-2y}{x} \\ x^2 = y^2 \\ 2x^2 = 8 \end{cases} \Leftrightarrow \begin{cases} \lambda = \frac{x-2y}{x} \\ x^2 = y^2 = 4 \end{cases} \Leftrightarrow \begin{cases} x = \pm 2 \\ y = \pm 2 \\ \lambda = \frac{x-2y}{x} \end{cases}$$

On trouve donc les quatre points stationnaires suivants :

- le point $(x^*; y^*) = (2,2)$ auquel correspond $\lambda^* = -1$,

- le point $(x^*, y^*) = (-2,2)$ auquel correspond $\lambda^* = 3$,

- le point $(x^*, y^*) = (2,-2)$ auquel correspond $\lambda^* = 3$,

- le point $(x^*, y^*) = (-2,-2)$ auquel correspond $\lambda^* = -1$.

Nature des points stationnaires. Pour chacun des points stationnaires précédents, puisque les fonctions admettent des dérivées partielles d'ordre deux continues, on peut déterminer leur nature en examinant le signe de la forme quadratique hessienne sur les espaces tangents.

L'espace tangent à la courbe de contrainte $\varphi(x; y) = 8$ en $(x; y) = (x^*, y^*)$ est l'espace vectoriel T donné par

$$(u,v) \in T \Leftrightarrow \frac{\partial \varphi}{\partial x}(x^*, y^*)u + \frac{\partial \varphi}{\partial y}(x^*, y^*)v = 0 \Leftrightarrow 2x^*u + 2y^*v = 0$$

$$\Leftrightarrow x^*u + y^*v = 0.$$

La forme quadratique hessienne au point (x^*, y^*, λ^*) est donnée par :

$$\begin{aligned} Q(u;v) &= \frac{\partial^2 L}{\partial x^2}(x^*, y^*, \lambda^*)u^2 + 2\frac{\partial^2 L}{\partial x \partial y}(x^*, y^*, \lambda^*)uv + \frac{\partial^2 L}{\partial y^2}(x^*, y^*, \lambda^*)v^2 \\ &= (2 - 2\lambda^*)\, u^2 - 8uv + (2 - 2\lambda^*)\, v^2 \end{aligned}$$

Nature du point $(x^*, y^*) = (2, 2)$. Pour ce point, on a $\lambda^* = -1$,

$$(u, v) \in T \Leftrightarrow v = -u \text{ et } Q(u; v) = 4u^2 - 8uv + 4v^2.$$

Soit $(u, v) \in T$, on a $(u, v) \neq (0, 0) \Leftrightarrow u \neq 0$ et, sous cette condition,

$$Q(u, v) = Q(u, -u) = 4u^2 + 8u^2 + 4u^2 = 16u^2 > 0,$$

donc le point $(x^*, y^*) = (2, 2)$ correspond à un minimum.

Nature du point $(x^*, y^*) = (-2, -2)$. Pour ce point, on a $\lambda^* = 3$,

$$(u, v) \in T \Leftrightarrow v = u \text{ et } Q(u, v) = -4u^2 - 8uv - 4v^2.$$

Soit $(u, v) \in T$, on a $(u, v) \neq (0, 0) \Leftrightarrow u \neq 0$ et, sous cette condition,

$$Q(u, v) = Q(u, u) = -4u^2 - 8u^2 - 4u^2 = -16u^2 < 0,$$

donc le point $(x^*, y^*) = (-2, -2)$ correspond à un maximum.

Nature du point $(x^*, y^*) = (2, -2)$. Ici (ce n'est pas le cas en général), le calcul est le même que dans le cas précédent, donc le point $(x^*, y^*) = (2, -2)$ correspond à un minimum

2.1.4 Contraintes inégalités

Soit $f \in C(\mathbb{R}^n, \mathbb{R})$ et $g_i \in C^1(\mathbb{R}^n, \mathbb{R})$ $\;i = 1, ..., p$, on considère maintenant un ensemble K de la forme :

$$K = \left\{ \begin{array}{c} \forall i = 1, .., p \\ x \in \mathbb{R}^n, g_i(x) \leq 0, \end{array} \right\}$$

et on cherche à résoudre le problème de minimisation (1.1) qui sécrit :

$$\left\{ \begin{array}{c} \bar{x} \in K \\ f(\bar{x}) \leq f(x), \\ x \in K. \end{array} \right\}$$

Remarque 2.1.5 *Soit $\bar{x}$ une solution de (1.1) et supposons que $gi(\bar{x}) < 0$, pour tout $i \in \{1, ..., p\}$. Il existe alors $\varepsilon > 0$ tel que si $x \in B(\bar{x}, \varepsilon)$ alors $g_i(x) < 0$ pour tout $i = 1, ..., p$.*

On a donc $f(\bar{x}) \leq f(x) \forall x \in B(\bar{x}, \varepsilon)$. On est alors ramené à un problème de minimisation sans contrainte, et si f est différentiable en $\bar{x}$, on a donc $\nabla f(\bar{x}) = 0$.

On donne maintenant sans démonstration le théorème de Kuhn-Tücker *qui donne une caractérisation de la solution du problème (1.1).*

Théorème 2.1.7 (Kuhn–Tucker) *Soit $f \in C(\mathbb{R}^n, \mathbb{R})$, soit $g_i \in C^1(\mathbb{R}^n, \mathbb{R})$, pour $i = 1, ..., p$, et soit $K = \{x \in \mathbb{R}^n, g_i(x) \leq 0 \forall i = 1...p\}$. On suppose qu'il existe $\bar{x}$ solution de (1.1), et on pose $I(\bar{x}) = \{i \in \{1, ..., p\}; |g_i(\bar{x}) = 0\}$. On suppose que f est différentiable en $\bar{x}$ et que la famille (de $\mathbb{R}^n$) $\{\nabla g_i(\bar{x}), i \in I(\bar{x})\}$ est libre. Alors il existe une famille $(\lambda_i)_{i \in I(\bar{x})} \subset \mathbb{R}_+$ telle que :*

$$\nabla f(\bar{x}) + \sum_{i \in I(\bar{x})} \lambda_i \nabla g_i(\bar{x}) = 0.$$

Remarque 2.1.6 *1. Le théorème de Kuhn-Tucker s'applique pour des ensembles de contrainte de type inégalité. Si on a une contraite de type égalité, on peut évidemment se ramener à deux contraintes de type inégalité en remarquant que :*

$$\{h(x) = 0\} = \{h(x) \leq 0)\} \cap \{-h(x) \leq 0\}.$$

Cependant, si on pose $g_1 = h$ et $g_2 = -h$, on remarque que la famille,

$$\{\nabla g_1(\bar{x}), \nabla g_2(\bar{x})\} = \{\nabla h(\bar{x}), -\nabla h(\bar{x})\},$$

n'est pas libre. On ne peut donc pas appliquer le théorème de Kuhn-Tucker *sous la forme donnée précédemment dans ce cas (mais on peut il existe des versions du théorème de Kuhn-Tucker permettant de traiter ce cas, voir* Bonans-Saguez*).*

2. Dans la pratique, on a intérêt à écrire la conclusion du théorème de Kuhn-Tucker *(i.e. l'existence de la famille $(\lambda_i)_{i \in I(\bar{x})}$) sous la forme du système de $n + p$ équations et $2p$ inéquations à résoudre suivant :*

$$\begin{cases} \nabla f(\bar{x}) + \displaystyle\sum_{i=1}^{p} \lambda_i \nabla g_i(\bar{x}) = 0, \\ \lambda_i g_i(\bar{x}) = 0, \forall i = 1, ..., p, \\ g_i(\bar{x}) \leq 0, \forall i = 1, ..., p, \\ \lambda_i \geq 0, \forall i = 1, ..., p. \end{cases}$$

2.2 Algorithmes d'optimisation sous contraintes

2.2.1 Méthodes de gradient avec projection

On rappelle le résultat suivant de projection sur un convexe fermé :

Proposition 2.2.1 (Projection sur un convexe fermé) *Soit E un* espace de Hilbert*, muni d'une norme $\|.\|$ induite par un produit scalaire $\langle ., . \rangle$, et soit K un* convexe fermé *non vide de E. Alors, tout $x \in E$, il existe un unique $x_0 \in K$ tel que $\|x - x_0\| \leq \|x - y\|$ pour tout $y \in K$. On note $x_0 = p_k(x)$ la projection orthogonale de x sur K. On a également :*

$$x_0 = p_K(x) \quad \textit{si et seulement si} \qquad (x - x_0, x_0 - y) \geq 0 \ \forall y \in K.$$

Dans le cadre des algorithmes de minimisation avec contraintes que nous allons développer maintenant, nous considèrerons $E = \mathbb{R}^n$, $f \in C^1(\mathbb{R}^n, \mathbb{R})$ une fonction convexe, et K fermé convexe non vide. On cherche à calculer une solution approchée de $\bar{x}$, solution du problème (1.1).

Algorithme du gradient à pas fixe avec projection sur K (GPFK). Soit $\rho > 0$ donné, on considère l'algorithme suivant :

Algorithme (GPFK)

Initialisation : $x_0 \in K$

Itération :x_k *connu*

$$x_{k+1} = p_K(x_k - \rho \nabla f(x_k))$$

où p_K est la projection sur K.définie par la proposition 2.2.1 .

Lemme 2.2.1 *Soit $(x_k)_k$ construite par l'algorithme **(GPFK)**. On suppose que $x_k \to x$ quand $n \to +\infty$. Alors x est solution de 1.1.*

Démonstration Soit p_K :$\mathbb{R}_n \to K \subset \mathbb{R}_n$ la projection sur K définie par la proposition 2.2.1 . Alors p_k est continue. Donc si : $x_k \to x$ quand $n \to +\infty$ alors $x = p_K(x - \rho \nabla f(x))$ et $x \in K$ (car $x_k \in K$ et K est fermé). La caractérisation de $p_K(x - \rho\nabla f(x))$ donnée dans la proposition 2.2.1 donne alors : $(x - \rho\nabla f(x) - x/x - y) \geq 0$ pour tout $y \in K$, et comme $\rho > 0$, ceci entraîne $(\nabla f(x)/x - y) \leq 0$ pour tout $y \in K$. Or f est convexe donc $f(y) \geq f(x) + \nabla f(x)(y - x)$ pour tout $y \in K$, et donc $f(y) \geq f(x)$ pour tout $y \in K$, ce qui termine la démonstration

Remarque 2.2.1 *L'algorithme du gradient projet e est difficile a mettre en oeuvre dans le cas général, car le calcul de la projection est souvent aussi difficile que le probléme initial. Cependant, cet algorithme peut étre utilisé pour des contraintes "simples", de la forme* $x \leq 0$.([2])

Théorème 2.2.1 (Convergence de l'algorithme GPFK) *Soit* $f \in C1(\mathbb{R}_n, \mathbb{R})$, *et* K *convexe fermé non vide. On suppose que :*

1. *il existe* $\alpha > 0$ *tel que* $(\nabla f(x) - \nabla f(y)|x - y) \geq \alpha|x - y|^2$, *pour tout* $(x, y) \in \mathbb{R}_n \times \mathbb{R}_n$.
2. *il existe* $M > 0$ *tel que* $|\nabla f(x) - \nabla f(y)| \leq M|x - y|$ *pour tout* $(x, y) \in \mathbb{R}_n \times \mathbb{R}_n$

   alors :

   1. *il existe un unique élément* $\bar{x} \in K$ *solution de* 1.1 .

   2. *si* $0 < \rho < \frac{2\alpha}{M^2}$, *la suite* (x_k)*définie par l'algorithme (GP F K) converge vers* $\bar{x}$ *lorsque* $n \to +\infty$.

Démonstration

1. La condition 1. donne que f est strictement convexe et que $f(x) \to +\infty$ quand $|x| \to +\infty$. Comme K est convexefermé non vide, il existe donc un unique $\bar{x}$ solution de 1.1.
2. On pose, pour $x \in \mathbb{R}_n$, $h(x) = p_K(x - \rho\nabla f(x))$. On a donc $x_k + 1 = h(x_k)$. Pour montrer que la suite $(x_k)_{k\in IN}$ converge, il suffit donc de montrer que h est strictement contractante dès que

$$0 < \rho < \frac{2\alpha}{M^2}. \tag{2.7}$$

Grâce au lemme 2.2.2 démontré plus loin, on sait que p_K est contractante. Or h est définie par :

$$h(x) = p_K(\bar{h}(x)) o\grave{u}, \bar{h}(x) = x - \rho \nabla f(x)$$

On a déjà vu que$\bar{h}$ est strictement contractante si la condition (2.7) est vérifiée , et plus précisément :

$$|\bar{h}(x) - \bar{h}(y)| \leq (1 - 2\alpha\rho + M^2\rho^2)|x - y|^2.$$

On en déduit que :

$$|h(x)-h(y)|^2 \leq |p_K(\bar{h}(x))-p_K(\bar{h}(y))|^2 \leq |\bar{h}(x)-\bar{h}(y))|^2 \leq (1-2\alpha\rho+\rho^2 M^2)|x-y|^2.$$

L'application h est donc strictement contractante dès que $0 < \frac{2\alpha}{M^2}$. La suite $(x_k)_{k\in IN}$ converge donc bien vers $x = \bar{x}$.

Lemme 2.2.2 (Propriété de contraction de la projection orthogonale) *Soit E un espace de Hilbert,$\|.\|$ la norme et $(\cdot,\cdot)$ le produit scalaire, K un convexe fermé non vide de E et p_K la projection orthogonale sur K définie par la proposition 2.2.1, alors $\|p_K(x) - p_K(y)\| \leq \|x - y\|$ pour tout $(x, y) \in E^2$.*

Démonstration Comme E est un espace de Hilbert $\|p_K(x) - p_K(y)\|^2 = (p_K(x) - p_K(y)|p_K(x) - p_K(y))$

On a donc :

$$\|p_K(x) - p_K(y)\|^2 = (p_K(x) - x + x - y + y - p_K(y)|p_K(x) - p_K(y))$$

$$= (p_K(x)-x|p_K(x)-p_K(y))E+(x-y|p_K(x)-p_K(y))+(y-p_K(y)|p_K(x)-p_K(y))$$

Or :

$$(p_K(x) - x|p_K(x) - p_K(y)) \leq 0 \quad et \quad (y - p_K(y)|p_K(x) - p_K(y)) \leq 0$$

d'où :

$$\|p_K(x) - p_K(y)\|^2 \leq (x - y|p_K(x) - p_K(y))$$

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

$$\|p_K(x) - p_K(y)\|^2 \leq \|x - y\| \, \|p_K(x) - p_K(y)\| \, .$$

ce qui permet de conclure.

Algorithme du gradient à pas optimal avec projection sur K (GPOK)

L'algorithme du gradient à pas optimal avec projection sur K s'écrit :

Initialisation $x_0 \in K$

Itération $x_k$connu

$w_k = -\nabla f(x_k)$; calculer α_k optimal dans la direction w_k

$x_k + 1 = p_K(x_k + \alpha_k w^{(k)})$.

La démonstration de convergence de cet algorithme se déduit de celle de l'algorithme à pas fixe.

Remarque 2.2.2 *On pourrait aussi utiliser un algorithme de type Quasi–Newton [1] avec projection sur K. Les algorithmes de projection sont simples à décrire, mais ils soulèvent deux questions :*

1. *Comment calcule-t-on p_K ?*
2. *Que faire si K n'est pas convexe ?*

   On peut donner une réponse à la première question dans les cas simples :

cas 1. *On suppose ici que* $K = C^{+} = \{x \in \mathbb{R}_n, x = (x_1, ..., x_k)^t, x_i \geq 0 \forall i\}$.*Si* $y \in \mathbb{R}_n, y = (y_1...y_n)^t$ *on peut montrer que :*

$$(p_K(y))_i = y_i^{+} = max(y_i, 0), \forall i \in \{1, ..., n\}$$

cas 2.

Soit $(\alpha_i)_{i=1,...,n} \subset \mathbb{R}_n$ *et* $(\beta_i)_{i=1,...,n} \subset \mathbb{R}_n$ *tels que* $\alpha_i \leq \beta_i$ *pour tout* $i = 1, ..., n$. *Si :*

$$K = \prod_{i=1,n} [\alpha_i \beta_i],$$

alors

$$(p_K(y))_i = max(\alpha_i, min(y_i, \beta_i)), \forall i = 1, ..., n$$

Dans le cas d'un convexe K *plus "compliqué", ou dans le cas où* K *n'est pas convexe, on peut utiliser des méthodes de dualité introduites dans le paragraphe suivant.*

2.2.2 Méthodes de dualité

Supposons que les hypothèses suivantes sont vérifiées :

$$\begin{cases} f \in C^1(\mathbb{R}^n, \mathbb{R}), \\ g_i \in C^1(\mathbb{R}^n, \mathbb{R}), \\ K = \{x \in \mathbb{R}^n, g_i(x) \leq 0, i = 1, ..., p\}, \\ et\ K\ est\ nonvide. \end{cases} . \tag{2.8}$$

On définit un problème "*primal*" comme étant le problème de minimisation d'origine, c'est–à–dire :

$$\begin{cases} \bar{x} \in K \\ f(\bar{x}) \leq f(x), \forall x \in K, \end{cases} \tag{2.9}$$

On définit le "*lagrangien*" comme étant la fonction L définie de $\mathbb{R}^n \times \mathbb{R}^p$ dans $\mathbb{R}$ par :

$$L(x, \lambda) = f(x) + \lambda.g(x) = f(x) + \sum_{i=1}^{p} \lambda_i g_i(x), \tag{2.10}$$

avec

$$g(x) = (g_1(x), ..., g_p(x))^t \ et \ \lambda = (\lambda_1(x), ..., \lambda_p(x))^t.$$

On note C^+l'ensemble défini par :

$$C^+ = \{\lambda \in \mathbb{R}^p, \lambda = (\lambda_1, ..., \lambda_p)^t, \lambda_i \geq 0, \forall i = 1, ..., p\}.$$

Remarque 2.2.3 *Le théorème de* Kuhn-Tucker *entraîne que si $\bar{x}$ est solution du problème primal (2.9) alors il existe $\lambda \in C^+$tel que $D_1 L(\bar{x}, \lambda) = 0$ (c'est–à–dire $Df(\bar{x}) + \lambda \cdot Dg(\bar{x}) = 0$) et $\lambda \cdot g(\bar{x}) = 0$.*
On définit alors l'application M de $\mathbb{R}^p$ dans $\mathbb{R}$ par :

$$M(\lambda) = \inf_{x \in \mathbb{R}^n} L(x, \lambda), \forall \lambda \in \mathbb{R}^p. \tag{2.11}$$

On peut donc remarquer que $M(\lambda)$réalise le minimum (en x) du problème sans contrainte, qui s'écrit, pour$\lambda \in \mathbb{R}^p$ fixé :

$$\begin{cases} x \in \mathbb{R}^n \\ L(x, \lambda) \leq L(y, \lambda), \forall x \in \mathbb{R}^n, \end{cases} \tag{2.12}$$

Lemme 2.2.3 *L'application M de $\mathbb{R}^p$ dans $\mathbb{R}$ définie par 2.11est concave (ou encore l'application -M est convexe), c'est–à–dire que pour tous $\lambda, \mu \in \mathbb{R}^p$ et pour tout $t \in]0, 1[$ on a :*

$$M(t\lambda + (1-t)\mu) \geq tM(\lambda) + (1-t)M(u).$$

Démonstration Soit $\lambda, \mu \in \mathbb{R}^p$ et $t \in]0,1[$; on veut montrer que

$$M(t\lambda + (1-t)\mu) \geq tM(\lambda) + (1-t)M(\mu).$$

Soit $x \in \mathbb{R}^n$, alors :

$$\begin{aligned} L(x, t\lambda + (1-t)\mu) &= f(x) + (t\lambda + (1-t)\mu)g(x) \\ &= tf(x) + (1-t)f(x) + (t\lambda + (1-t)\mu)g(x). \end{aligned}$$

On a donc

$$L(x, t\lambda + (1-t)\mu) = tL(x,\lambda) + (1-t)L(x,\mu).$$

Par définition de M, on en déduit que pour tout $x \in \mathbb{R}^n$:

$$L(x, t\lambda + (1-t)\mu) \geq tM(\lambda) + (1-t)M(\mu).$$

Or, toujours par définition de M :

$$M(t\lambda + (1-t)\mu) = \inf_{x \in \mathbb{R}^n} L(x, t\lambda + (1-t)\,\mu) \geq tM\,(\lambda) + (1-t)M(\mu).$$

On considère maintenant le problème d'optimisation dit "*dual*" suivant :

$$\left\{ \begin{array}{c} \mu \in C^+ \\ M\,(\mu) \geq M\,(\lambda)\,, \forall \lambda \in C^+. \end{array} \right. \tag{2.13}$$

Définition 2.2.1 *Soit* $L : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$ *et* $(x,\mu) \in \mathbb{R}^n \times C^+$. *On dit que* (x,μ) *est un* point selle *de* L *sur* $\mathbb{R}^n \times C^+$ *si*

$$L(x,\lambda) \leq L(x,\mu) \leq L(y,\mu), \forall y \in IR, \forall \lambda \in C^+.$$

Proposition 2.2.2 *Sous les hypothèses (2.8), soit L définie par $L(x, \lambda) = f(x) + \lambda g(x)$ et $(x, \mu) \in \mathbb{R}^n \times C^+$ un* point selle *de L sur $\mathbb{R}^n \times C^+$.*

Alors

1. *$\bar{x}$ est solution du problème (2.9).*
2. *μ est solution de (2.13).*
3. *$\bar{x}$ est solution du problème (2.11) avec $\lambda = \mu$.*

On admettra cette proposition.

Réciproquement, on peut montrer que (sous des hypothèses convenables sur f et g), si μ est solution de (2.13), et si $\bar{x}$ solution de (2.12) avec $\lambda = \mu$, alors $(\bar{x}, \mu)$ est un point selle de L, et donc $\bar{x}$ est solution de (2.9).

De ces résultats découle l'idée de base des méthodes de dualité : on cherche μ solution de (2.13). On obtient ensuite une solution $\bar{x}$ du problème (2.9), en cherchant $\bar{x}$ comme solution du problème (2.12) avec $\lambda = \mu$ (qui est un problème de minimisation sans contraintes). La recherche de la solution μ du problème dual (2.13) peut se faire par exemple par l'algorithme très classique d'Uzawa, que nous décrivons maintenant.

Algorithme d'Uzawa L'algorithme d'Uzawa consiste à utiliser l'algorithme du gradient à pas fixe avec projection (qu'on a appelé "**GPFK**") pour résoudre de manière itérative le problème dual (2.13). On cherche donc $\mu \in C^+$tel que $M(\mu) \geq M(\lambda)$ pour tout $\lambda \in C^+$. On se donne $\rho > 0$, et on note P_{C+} la projection sur le convexe C^+. L'algorithme (**GPFK**) pour la recherche de μ s'écrit donc :

Initialisation :

$$\mu_0 \in C^+$$

Itération :

$$\mu_{k+1} = p_{C+}(\mu_k + \rho \nabla M(\mu_k)).$$

Pour définir complètement l'algorithme d'*Uzawa*, il reste à préciser les points suivants :

1. Calcul de $\nabla M(\mu_k)$,
2. calcul de $p_{C+}(\lambda)$ pour λ dans $\mathbb{R}^n$.

On peut également s'intéresser aux propriétés de convergence de l'algorithme.

La réponse au point 2 est simple : pour $\lambda \in \mathbb{R}^n$, on calcule $p_{C+}(\lambda) = \gamma$ avec $\gamma = (\gamma_1, ..., \gamma_p)^t$ en posant $\gamma_i = max(0, \lambda_i)$ pour $i = 1, ..., p$, où $\lambda = (\lambda_1, ..., \lambda_p)^t$.

La réponse au point 1. est une conséquence de la proposition suivante (qu'on admettra ici) :

Proposition 2.2.3 *Sous les hypothèses (2.8), on suppose que pour tout $\lambda \in \mathbb{R}^n$, le problème (2.12) admet une solution unique, notée x_λ et on suppose que l'application définie de $\mathbb{R}^p$ dans $\mathbb{R}^n$ par $\lambda \to x_\lambda$ est différentiable. Alors $M(\lambda) = L(x\lambda, \lambda)$, M est différentiable en λ pour tout λ, et $\nabla M(\lambda) = g(x_\lambda)$.*
*En conséquence, pour calculer $\nabla M(\lambda)$, on est ramené à chercher x_λ solution du problème de minimisation sans contrainte (2.12). On peut dont maintenant donner le détail de l'itération générale de l'algorithme d'*Uzawa *:*
Itération de l'algorithme d'Uzawa. *Soit $\mu_k \in C^+$ connu ;*

1. *On cherche $x_k \in \mathbb{R}^n$ solution de*

$$\begin{cases} x_k \in \mathbb{R}^n \\ L(x_k, \mu_k) \leq L(x, \mu_k), \forall x \in \mathbb{R}^n \end{cases}$$

2. *On calcule $\nabla M(\mu_k) = g(x_k)$.*
3.

$$\bar{\mu}_{k+1} = \mu_k + \rho \nabla M(\mu_k) = \mu_k + \rho g(x_k) = ((\bar{\mu}_{k+1})_1, ..., (\bar{\mu}_{k+1})_p)^t.$$

4. *$\mu_{k+1} = p_{C^+}(\mu_{k+1})$, c'est–à-dire $\mu_{k+1} = ((\mu_{k+1})_1, ..., (\mu_{k+1})_p)^t$ avec $(\mu_{k+1})_i = max(0, (\mu_{k+1})_i)$ pourtout $i = 1, ..., p$.*

On a alors le résultat suivant de convergence de l'algorithme :

Proposition 2.2.4 (Convergence de l'algorithme d'Uzawa) *Sous les hypothèses (2.8), on suppose de plus que :*

1. *il existe $\alpha > 0$tel que*

$$(\nabla f(x) - \nabla f(y)) \,.\, (x - y) \geq \alpha \left|x - y\right|^2, \forall (x, y) \in (\mathbb{R}^n)^2,$$

2. *il existe $M_f > 0$,*

$$|\nabla f(x) - \nabla f(y)| \leq M_f |x - y|, \forall (x, y) \in (\mathbb{R}^n)^2,$$

3. *pour tout $\lambda \in C^+$, il existe un unique $x_\lambda \in \mathbb{R}^n$ tel que $L(x_\lambda, \lambda) \leq L(x, \lambda)$ pour tout $x \in \mathbb{R}^n$.Alors si $0 < \rho < \frac{2\alpha}{M_f^2}$, la suite $((x_k, \mu_k))_k \in \mathbb{R}^n \times C^+$donnée par l'algorithme d'*Uzawa *vérifie :*

   1. *$x_k \to \bar{x}$quand $n \to +\infty$, où $\bar{x}$ est la solution du problème (2.9),*

   2. *$(\mu_k)^n \in \mathbb{N}$ est bornée.*

Remarque 2.2.4 (Sur l'algorithme d'Uzawa) 1. *L'algorithme est très efficace si les contraintes sont affines : (i.e. si $g_i(x) = \alpha_i.x + \beta_i$ pour tout $i = 1, ..., p$,avec $\alpha_i \in \mathbb{R}^n$, $\beta_i \in \mathbb{R}$).*

2. *Pour avoir l'hypothèse 3 du théorème, il suffit que les fonctions g_i soient convexes. (On a dans ce cas existence et unicité de la solution x_λ du problème (2.12) et existence et unicité de la solution $\bar{x}$ du problème(2.10)).[3]*

Chapitre 3

Application

3.1 Le comportement du producteur

Pour caractériser le comportement du producteur, l'analyse microéconomique suppose que l'objectif principal de ce dernier consiste à rendre son profit maximal. Le profit étant définit comme la différence entre le chiffre d'affaire et les coûts, il s'écrit mathématiquement

$$\prod (K, L) = \underbrace{pQ (K, L)}_{\text{Chiffre d'affaire}} \underbrace{-wL - rK - f}_{\text{Coûts}}.$$

p le prix du bien ou service produit, $pQ(K, L)$ le chiffre d'affaire,

$-wL - rK - f$ les coûts et plus exactement : wL le coût du travail, rK le coût du capital et f la rémunération de l'ensemble des facteurs fixes de l'entreprise.

Le comportement du producteur peut alors être appréhendé de façon différentes selon qu'il rencontre ou non une contrainte sur la quantité à produire ou sur le coût qu'il peut supporter.

Le producteur contrait par son marché

Supposons que le producteur connaît le niveau maximum de production qu'il peut écouler sur le marché, il est donc contraint par les quantités et connaît à l'avance le montant de sa recette $pQ\left(K,L\right)$. Dans ce cas la maximisation du profit implique la minimisation des coûts. Donc le problème du producteur se réduit à la recherche de l'utilisation optimale des facteurs de façon à minimiser ses coûts de production. Le problème s'exprime mathématiquement sous la forme d'une minimisation de coûts sous la contrainte d'un niveau de production $\bar{Q}$.
Le programme s'écrit alors :

$$\begin{cases} \min C\left(K,L\right) = wL + rK + f \\ s.c\ Q\left(K,L\right) = \bar{Q}. \end{cases}$$

La fonction de Lagrange correspondant à ce programme est :

$$L(K,L,\lambda) = wL + rK + f - \lambda(Q(K,L) - \bar{Q})$$

Les conditions du premier ordre, qui correspondent à l'annulation des dérivées premières sont données par :

$$\begin{cases} \frac{\partial L(K,L,\lambda)}{\partial K} = 0 \Leftrightarrow r - \lambda Q_k = 0 \Rightarrow \lambda = \frac{r}{Q_k}. \\ \frac{\partial L(K,L,\lambda)}{\partial L} = 0 \Leftrightarrow w - \lambda Q_k = 0 \Rightarrow \lambda = \frac{r}{QL}. \\ \frac{\partial L(K,L,\lambda)}{\partial \lambda} = 0 \Leftrightarrow Q(K,L) - \bar{Q} = 0. \end{cases}$$

Soit

$$\frac{r}{w} = \frac{Q_K}{Q_L}$$

avec Q_K et Q_L respectivement la productivité marginale du capital et du travail. Les deux premières équations du système des **CPO** permettent de définir la condition d'optimalité du producteur

Le producteur contraint par son budget

Le producteur peut se trouver dans une configuration alternative où il connaît son budget maximum. Les coûts ne pouvant excéder cette somme, le coût maximal de la production est connu et le recherche du profit le plus élevé possible passe par la maximisation de la recette. Le producteur va donc chercher la combinaison productive (K, L) qui maximise la volume de production tout en respectant la contrainte de coût

Le programme s'écrit alors :

$$\begin{cases} \max Q\,(K, L) \\ s.c\ \bar{C} = wL + rK + f. \end{cases}$$

Le Lagrangien est alors :

$$L(K, L, \lambda) = Q(K, L) + \lambda(\bar{C} - wL - rK - f).$$

D'où d'après les CPO :

$$\begin{cases} \frac{\partial L(K,L,\lambda)}{\partial K} = 0 \Leftrightarrow Q_k - \lambda r = 0 \Rightarrow \lambda = \frac{r}{Q_k}. \\ \frac{\partial L(K,L,\lambda)}{\partial L} = 0 \Leftrightarrow Q_L - \lambda w = 0 \Rightarrow \lambda = \frac{w}{QL}. \\ \frac{\partial L(K,L,\lambda)}{\partial \lambda} = 0 \Leftrightarrow \bar{C} - rK - wL - f = 0. \end{cases}$$

Soit

$$\frac{r}{w} = \frac{Q_K}{Q_L}.$$

3.2 Interprétation des conditions d'optimalité

La résolution des deux problèmes posés par le producteur permet d'établir qu'au point optimal, la rapport des prix des facteurs doit être égal au rapport des productivités marginales. Cette égalité s'explique simplement : la productivité marginale d'un facteur rapporté à son prix correspond à la production supplémentaire obtenue en consacrant une unité de compte à l'achat d'un facteur. Par conséquent, si cette égalité n'était pas vérifiée, reporter la dépense de cette unité de compte d'un facteur à l'autre permettrait de produire plus, ce qui signifierait que la combinaison initiale n'était pas optimale

3.2.1 Le comportement du consommateur

Le consommateur, s'il est rationnel veut maximiser son utilité sous la contrainte d'un budget limité afin de déterminer les quantités x_1 de bien 1 et x_2 de bien 2 qu'il va demander. On suppose que l'utilité de l'agent ne dépend que des quantités de bien, ici $x_1$et x_2.

La contrainte de budget est donnée par :

$$R = p_1x_1 + p_2x_2$$

avec R le revenu et $p_1x_1 + p_2x_2$ les dépenses.

Le programme de maximisation s'écrit comme suit :

$$\begin{cases} \max U\left(x_1, x_2\right) \\ s.c\ R = p_1x_1 + p_2x_2 \end{cases}$$

Nous avons précédemment résolu les programmes d'optimisation avec le Lagrangien, nous allons maintenant montrer qu'il est possible de raisonner autrement et de procéder à la **méthode dite de substitution** pour résoudre ce programme.

3.2.2 Méthode de substitution

C'est une méthode qui consiste à transformer le programme de maximisation d'une fonction à deux variables sous contrainte en un programme de maximisation d'une fonction à une variable sous contrainte. La première étape consiste à exprimer une des variables en fonction de l'autre (par exemple x_2 en fonction de x_1) à partir de la contrainte :

$$R = p_1 x_1 + p_2 x_2 \Leftrightarrow x_2 = \frac{R}{p_2} - \frac{p_1}{p_2} x_1.$$

Il faut ensuite le reporter dans la fonction d'utilité qui devient donc une fonction d'une seule variable(ici x_1).

$$U(x_1, x_2) = U(x_1, x_2(x_1)) = U\left(x_1, \frac{R}{p_2} - \frac{p_1}{p_2} x_1\right).$$

On calcul ensuite les CPO pour trouver la valeur de x_1 :

$$\frac{\partial U}{\partial x_1} = \frac{p_1}{p_2}.$$

A l'optimum du consommateur le rapport des utilités marginales doit être égales au rapport des prix. La condition d'optimalité permet alors de trouver la valeur optimale x_1^* de x_1. La valeur optimale x_2^* de x_2 s'écrit alors :

$$x_2^* = \frac{R}{p_2} - \frac{p_1}{p_2} x_1^*$$

3.3 Prenons un exemple concret dans le cas du consommateur

Soit la fonction d'utilité suivante du consommateur :

$$U = x_1^\alpha x_2^\beta$$

avec p_1 et p_2 respectivement le prix des biens 1 et 2. R le revenu.

Déterminons les fonctions de demande optimales du consommateur

On a le programme de maximisation suivant :

$$\begin{cases} \max U \\ s.c\ R = p_1 x_1 + p_2 x_2 \end{cases}$$

écriture du Lagrangien

$$L(x_1, x_2, \lambda) = 2x_1^\alpha x_2^\beta + \lambda(R - p_1 x_1 + p_2 x_2)$$

Les CPO sont données par :

$$\begin{cases} \frac{\partial L(x_1,x_2,\lambda)}{\partial x_1} = 0 \Leftrightarrow 2\alpha x_1^\alpha x_2^\beta - \lambda p_1 = 0 \Rightarrow \lambda = \frac{2\alpha x_1^\alpha x_2^\beta}{p_1}. \\ \frac{\partial L(x_1,x_2,\lambda)}{\partial x_2} = 0 \Leftrightarrow 2\beta x_1^\alpha x_2^\beta - \lambda p_2 = 0 \Rightarrow \lambda = \frac{2\beta x_1^\alpha x_2^\beta}{p_2}. \\ \frac{\partial L(x_1,x_2,\lambda)}{\partial \lambda} = 0 \Leftrightarrow R - p_1 x_1 + p_2 x_2 = 0. \end{cases}$$

On a alors :

$$\begin{cases} \lambda = \frac{2\alpha x_1^\alpha x_2^\beta}{p_1} \\ \lambda = \frac{2\beta x_1^\alpha x_2^\beta}{p_2} \end{cases}$$

Exprimons x_2 en fonction de x_1 :

$$x_2 = \frac{\beta}{\alpha} \frac{p_1 x_1}{p_2}$$

Si l'on remplace cette expression dans la contrainte nous trouverons la demande optimale de bien $1, x_1^*$:

$$x_1^* = \frac{\alpha R}{(\alpha + \beta) p_1}.$$

Cette expression dans celle de $x_2$précédente nous donne la demande optimale de bien 2 pour le consommateur, soit :

$$x_2^* = \frac{\beta R}{(\alpha + \beta) p_2}.$$

A noter que l'on arrive exactement au même résultat si l'on procède par la méthode substitution décrite précédemment.(pour plus de details voir [7])

Conclusion

Dans ce travail, nous avons étudié les problèmes d'optimisation avec contraintes. On a vu les théorèmes d'existence et d'unicité de la solution de problème. Dans le cas ou les conditions d'optimalité sont des contraintes d'égalité, le problème de minimisation sera résolu grâce au théorème des multiplicateurs (maximisation) de Lagrange. Mais dans le cas de contraintes d'inégalités, le problème de minimisation (maximisation) peut être résolu grâce au théorème de Kuhn.Tucker.

Puisque l'optimisation des systèmes permet de trouver une configuration idéale. La qualité des résultats et des prédictions dépend de l'intelligence du modèle, de l'efficacité de l'algorithme et des moyens pour le traitement numérique.

Et pour conclure, nous suggérons d'enrichir ce travail par l'optimisation sans contraint.

Bibliographie

[1] Fabin, B. (2010). *Modéles de recherche opéerationnelle.* (Page16). Universitéde Montréal.

[2] Frédéric, D. (2020) *Probléme d'Optimisation Non Linéaure avec contraintes en Tomographie de Réflexion 3D* .

[3] Herbin, R. (16 septembre 2016). *Cours analyse numérique*, Université d'Aix-Marseille.

[4] Matthier, B. (2015-2016). *cours d'optimsation.* Université Paris Diderot.

[5] Orkia, D. *Introduction aux méthodes d'optimisation.* Université Docteur Moulay Tahar-Saida.

[6] Pierre, C. (2015-2016) *cours d'Optimisation et programmation dynamique.* Université Paris Dauphine.

[7] Valérie, M. (2010-211). *cours dynamique économique : analyse des fluctuations.* Université Paris ouest Nanterre La Défense.

[mémoire]

[8] Ahlem ,B. (2019). *Algorithemd'optimisation Sous Contraintes.* Master en mathématique. Université Mohamed Khidher-Biskra.

Annexe : Abréviation et Notations

Les différentes abbréviations et notations utilisées durant cette thèse est représentées selon ce tableau

Abréviation	Notations
GPFK	gradient à pas fixe avec projection sur K
$C(E;\mathbb{R})$	espace des fonctions continues de E a $\mathbb{R}$
$\mathbb{N}$	Ensemble des Nombres Naturel
$\mathbb{R}$	Ensemble des Nombres Réelles

Résumé

L'optimisation est une branche des mathématiques et de l'informatique en tant que discipline, cherchant à modéliser, à analyser et à résoudre analytiquement ou numériquement les problèmes réels : qui consiste à déterminer quelle est la ou les solution(s) (inconnues) satisfaisant un objectif quantitatif tout en respectant d'éventuelles contraintes.

L'objectif de ce mémoire est de fournir un support sur l'optimisation sous contraintes et les principales méthodes utilisées pour les résoudre. Pour atteindre cet objectif ; on a exposé un exemple économique (le comportement du producteur et du consommateur), en utilisant la méthode des multiplicateurs de Lagrange qui permet de trouver un optimum, tout en satisfaisant les contraintes.

Pour caractériser le comportement du producteur, l'analyse microéconomique suppose que l'objectif principal de ce dernier consiste à rendre son profit maximal. Le profit étant définit comme la différence entre le chiffre d'affaire et les coûts, et pour le consommateur, s'il est rationnel veut maximiser son utilité sous la contrainte d'un budget limité.

Puisque l'optimisation des systèmes permet de trouver une configuration idéale. La qualité des résultats et des prédictions dépend de l'intelligence du modèle, de l'efficacité de l'algorithme et des moyens pour le traitement numérique.

Abstract

Optimization is a field of mathematics and computer science seeking to model, analyze and solve analytically or numerically the real problems: which consists in determining what are the solutions (unknowns) satisfying a quantitative objective while respecting any constraints.

The objective of this thesis is to provide a basis on constrained optimizations and the main methods used to solve them. To achieve this objective, we presented an example in economics (the behaviour of the producer and the consumer), while using the method of Lagrange multipliers which makes allows to find an optimum, the highest possible point, while satisfying a constraint.

To characterize the behaviour of the producer, microeconomic analysis assumes that the producer's main objective is to make maximum profit. Profit being defined as the difference between revenue and costs. And for the consumer, if he is rational he wants to maximize his benefit under the constraint of a limited budget.

Since the optimization of the systems allows to find an ideal configuration. The quality of the results and predictions depends on the intelligence of the model, the efficiency of the algorithm and the means for numerical processing.

الملخص

التحسين (الإستمثال): هو إحدى مجالات الرياضيات، الذي يسعى إلى نمذجة وتحليل وحل المشاكل الحقيقية تحليليًا أو عدديًا مع احترام القيود.

الهدف من هذه المذكرة هو تكوين أساس على مسائل التحسين والطرق الرئيسية المستخدمة لإيجاد الحلول. فلتحقيق هذا الهدف، قدمنا مثالًا في الاقتصاد (سلوك المنتج والمستهلك)، و تطبيق طريقة مضاعفات لاغرانج التي تسمح بإيجاد أقصى نقطة ممكنة، مع تلبية قيود.

ولوصف سلوك المنتج، التحليل الاقتصادي يفترض أن الهدف الرئيسي للمنتج هو تحقيق أقصى قدر من الربح مع أقل تكاليف . أما بالنسبة للمستهلك، فمنطقيا فإنه يريد تعظيم فائدته تحت قيود ميزانيته

نظرًا لأن تحسين الأنظمة يسمح بإيجاد حل مثالي. فتعتمد جودة النتائج والتنبؤات على ذكاء النموذج وكفاءة الخوارزمية ووسائل المعالجة العددية.