الجمهوريـة الجزائـريـة الديمقراطيـة الشعبيـة

République Algérienne Démocratique et Populaire

وزارة التعليـم العالـي والبحـث العلمـي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

N° Réf :...............

Centre Universitaire de Mila

Institut des Sciences et de la Technologie

Département de Mathématique et Informatique

Mémoire préparé En vue de l'obtention du diplôme de licence

en Filière: Mathématiques fondamentales

Thème
Optimisation Convexe

Préparé par :

Becheket Khadidja

Djouimaa Khawla

Houfani Yamina

Sekhri zeggar Abla

Encadré par :

Mme : BEN AOUICHA Loubna

Grade : Maitre assistant classe B

Année universitaire : 2013/2014

Remerciements

Nous remercions d'abord et avant tout le bon Dieu qui m'a donné le courage et la patience pour réaliser ce travail.

Nous remercions vivement madame L.ben aouicha d'avoir voulu proposes et assurer la direction de ce mémoire sa disponibilité son soutien ses encouragements et ses précieux conseils tout au long de ce travail.

Sans oublier tous les enseignants qui ont contribué a notre formation durant notre vie scolaire surtout les enseignant de l'institut (S-T).

Nous voudrons dire toute notre reconnaissance à nos parents pour leur dévouement sans limite et pour tout ce qu'ils nous ont donné sur tous les plants, et remercier nos familles et nos amis pour leur soutien constant.

Table des matières

Introduction Générale

L'optimisation est une branche importante des mathématiques appliquées, plusieurs domaine d'activité scientifique peut se réduire à un problème maximisation ou de minimisation.

L'importance de l'optimisation dans la résolution des problèmes pratiques en mathématiques appliquées ou dans d'autre activités de recherche est soulignées par de nombreux auteurs.

Dans ce mémoire, on commence par les définitions et les resultats préliminaires pour utilisé dans les trois chapitres suivants :

Le premier est consacré aux ensembles convexes, notion intermédiaire entre le cas

linéaire (espaces vectoriels et applications linéaires) et le cas convexe, qui permettra de traiter le sujet de maniére trés naturelle.

le second chapitre constitue une introduction aux lesquelles sont abordées progressivement en commençant par les proprietés générales (continuité, différentiabilité, et fonctions convexes d'une variable réel et de plusieurs variabeles).

le dernier chapitre est le plus importante, et il est le but de ce thème, il présente le problème d'optimisation convexe et les conditions d'optimalité dans les deux cas :

- Problème d'optimisation sans contraintes.
- Problème d'optimisation avec containtes.

Définition et resultats préliminaires :

1-Espace vectoriel

Définition 4.4.1 (Sous espace vectoriel)

Soit E un espace vectoriel.

On dit que F est un sous-espace vectoriel de E, si c'est un espace vectoriel et que $F \subset E$.

Théorème 4.4.2 (Caractérisation des sous-espaces)

Soit E un espace vectoriel.

Soit F un ensemble tel que :

i) $F \subset E$,

ii) $O_E \in F$,

iii) $\forall(\alpha, \beta) \in \mathbb{k}^2$, $\forall(x, y) \in F^2$, $\alpha x + \beta y \in F$

Alors F est un espace vectoriel, c'est un sous-espace vectoriel de E.

2-Sous-espace affine

Définition 4.5.1

Soit X un espace vectoriel réel, un sous-ensemble A de X est un sous-espace affine si :

$$\forall x \in A,\ y \in A, \forall \alpha \in \mathbb{R}, \alpha x + (1-\alpha)\, y \in A.$$

Autrement dit, un sous-espace affine contient toujours la "droite" passant par deux de ses points x et y.

3-Espace de Hilbert

Définition 4.6.1

On dit que l'espace vectoriel H muni de la forme sesquilinéaire $(\cdot|\cdot)$ est un espace de Hilbert si H est complet pour la norme $x \mapsto \sqrt{(x|x)}$.

4-Matrice positive

Définition 4.7.1

Une matrice A dans $M_{n,m}(\mathbb{R})$ est dite positive [resp. strictement positive] et on note $A \geq 0$ [resp. $A > 0$], si tous ses coefficients sont positifs ou nuls [resp. strictement positifs].

Si A, B sont deux matrices dans $M_{n,m}(\mathbb{R})$ la notation $A \geq B$ [resp. $A > B$, ou $A \leq B$, ou $A < B$] signifie que la matrice $A - B$ est positive [resp. $A - B$ est strictement positive, ou $B - A$ est positive ou $B - A$ est strictement positive].

5-Matrice semi-définie positive

Définition 4.8.1

Une matrice A, $n \times n$ est semi-définie positive si chacun des mineurs principaux de A est positive.

6-Théorème de Taylor-lagrange

Théorème 4.9.1

Soit f une fonction de classe C^p et $(p+1)$ fois dérivable sur un intervalle $I = [a, b]$ et à valeurs réelles.

Alors $\exists c \in]a, b[$ tel que :

$$f(b) = \sum_{k=0}^{p} \frac{(b-a)^k}{k!} f^{(k)}(a) + \frac{(b-a)^{p+1}}{(p+1)!} f^{(p+1)}(c).$$

7-Théorème de carathéodory

Tout point de l'enveloppe convexe d'un ensemble quelconque est représentable par une combinaison convexe d'au plus $(n+1)$ éléments de cet ensemble notamment ses point extrémaux.

Chapitre 1

Ensemble convexe

1.1 Définition et premieres propriétés

Définition 1.1.1

Soit X un espase vectoriel réel. un sous-ensemble A de X est convexe si :

$$\forall x, y \in A, \forall \alpha \in [0,1]\,; \alpha x + (1-\alpha)\,y \in A\ ;$$

Autrment dit, un sous-ensemble convexe contient toujours le segment $[x.y]$. (joignant deux de ses points x et y).

Une interprétation "optique" consiste à dire que dans une pièce convexe, deux personnes peuvent toujours s'apercevoir.

Un sous-espace affine est èvidmment convexe.

1.2 Envloppe affine et convexe

1.2.1 Enveloppe affine

Définition 1.2.1

Soit S une partie de $\mathbb{R}^n$, alors il existe une partie affine unique $F \subseteq \mathbb{R}^n$ contenant S appelée enveloppe affine de S et notée $aff\,(S)$.

C'est la plus petite partie affine de $\mathbb{R}^n$ contenant S.

$$aff\,(S) = F = \cap \{F_S / F_S \supset S\} \text{ et } F_S \text{ affine.}$$

Si : $S \neq \emptyset \Longrightarrow aff\,(S) \neq \emptyset$ (puisque $S \subseteq aff\,(S)$).

1.2.2 Enveloppe convexe

Définition 1.2.2

L'intersection de tout les ensembles convexes contenant un sous-ensemble S de $\mathbb{R}^n$ est appelèe : enveloppe convexe de S.

On la note $conv\,(S)$ ou $co\,(S)$ c'est le plus petit convexe de $\mathbb{R}^n$ contenant S.

Conséquance de la définition

$\forall S \subseteq \mathbb{R}^n$, S est convexe ssi $S = conv\,(S)$.

1.3 Intérieur géométrique

Définition 1.3.1

On dit que a appartient à l'intérieur géométrique de $S \subset E$ si pour tout $d \in E$ il existe $\bar{t} > 0$ tel que $a + td \in S$ pour tout $t \in [-\bar{t}, \bar{t}]$.

L'intérieur géométrique de S est noté $intg\,(S)$.

Cette notion, indépendante de toute topologie, est liée à la structure vectorielle de E.

Il est clair que $intg\,(S) \subset S$.

Proposition 1.3.1

Si S est convexe, alors $intg\,(S)$ est convexe.

Preuve :

Soient $a, b \in intg\,(S)$, $\lambda \in [0,1]$ et $c = a + \lambda\,(b-a)$. Soit $d \in E$, alors $\bar{r}, \bar{s} > 0$ existent tels que $t \in \left]-\bar{r}, \bar{r}\right[$, alors $a + td \in S$, et $t \in \left]-\bar{s}, \bar{s}\right[$ alors $b + td \in S$.

On a :

$$a + td + \lambda\,(b + td - a - td) \in S;$$

alors

$$a + td + \lambda b + \lambda td - \lambda a - \lambda td = a + \lambda\,(b-a) + td = c + td \in S;$$

donc : $intg\,(S)$ est convexe.

1.4 Propriétés topologiques des ensembles convexes

Désignons par B, la boule unité euclidienne de $\mathbb{R}^n$

$$B = \{x \in \mathbb{R}^n / \|x\| \leq 1\} = \{x / d\,(x,0) \leq 1\}.$$

B et C sont des convexes fermés et bornés. Et nous avons :
$\bullet \forall a \in \mathbb{R}^n$, $\forall \varepsilon > 0$, la boule de centre a et rayon ε s'écrit :

$$\{x/d(x,a) \le \varepsilon\} = a + \{y/\ \|y\| \le \varepsilon\} = a + \varepsilon B.$$

$\bullet \forall C \subseteq \mathbb{R}^n$, l'ensemble des x dont la distance à C ne dépasse pas ε est :

$$\{x \in \mathbb{R}^n / \exists y \in C, \|x - y\| \le \varepsilon\} = \cup \{y + \varepsilon B / y \in C\} = C + \varepsilon B.$$

La clôture (fermeture) de C notée $\bar{C}$ ou $cl\,(C)$. Et son intérieur. Noté $\mathring{C}$ ou $int\,(C)$, peuvent donc s'écrire sous les formes suivantes :

$$cl\,(C) = \cap \{C + \varepsilon B : \varepsilon > 0\}.$$

$$int\,(C) = \{x \in C : \exists \varepsilon > 0,\ x + \varepsilon B \subseteq C\}.$$

Théorème 1.4.1
Soit C un convexe de $\mathbb{R}^n$, d'intérieur non vide,
alors :

$$\forall a \in int\,(C), \forall b \in cl\,(C) : \forall \lambda \in [0,1[, (1-\lambda)\,a + \lambda b \in int\,(C).$$

Preuve :
Prenons $\lambda \in [0,1[$ et montrons qu'il existe $\bar{\varepsilon} > 0$ tel que :
$(1-\lambda)\,a + \lambda b + \bar{\varepsilon} B \subset C$.
En effet, $b \in cl\,(C) \Longrightarrow \forall \varepsilon > 0,\ b \in C + \varepsilon B$
$\Longrightarrow b \in C + \bar{\varepsilon} B$
$\Longrightarrow (1-\lambda)\,a + \lambda b + \bar{\varepsilon} B \subseteq (1-\lambda)\,a + \lambda\,(C + \bar{\varepsilon} B) + \bar{\varepsilon} B$
$= (1-\lambda)\left[a + \frac{1+\lambda}{1-\lambda}\bar{\varepsilon} B\right] + \lambda C$.
D'autre part $a \in int\,(C) \Longrightarrow \exists \varepsilon_0 > 0/\ a + \varepsilon_0 B \subseteq C$
donc : $a + \frac{1+\lambda}{1-\lambda}\bar{\varepsilon} B \subseteq C$ si $\frac{1+\lambda}{1-\lambda}\bar{\varepsilon} \le \varepsilon_0$.
Soit $\bar{\varepsilon} = \frac{1-\lambda}{1+\lambda}\varepsilon_0 \le \varepsilon_0$ alors :
$(1-\lambda)\,a + \lambda b + \bar{\varepsilon} B \subseteq (1-\lambda)\,[a + \bar{\varepsilon} B] + \lambda C \subseteq (1-\lambda)\,C + \lambda C = C$.
Corollaire 1.4.1
Soit $C \subset E$ convexe d'intérieur non vide. Alors $int\,(C)$ et $cl\,(C)$ sont convexe,
et on a :

$$cl\,(int\,(C)) = cl\,(C)\,,\ \text{et } int\,(cl\,(C)) = int\,(C) = intg\,(C)\,.$$

Proposition 1.4.1 La fermeture d'un convexe est convexe.

Preuve :

Soient C convexe, et $a, b \in cl(C)$ on prenons $z \in int(C) \neq \emptyset$.

Alors :

$$(1-\lambda) z + \lambda b = y_\lambda \in int(C), \forall \lambda \in [0,1[.$$

Soit maintenant $\mu \in [0,1[$ (fixé), alors :

$$(1-\mu) a + \mu y_\lambda \in int(C) \subseteq C, \forall \lambda \in [0,1[;$$

Et donc : $\lim_{\lambda \longrightarrow 1} [(1-\mu) a + \mu y_\lambda] = (1-\mu) a + \mu b \in cl(C)$.

1.4.1 Intérieur relatif

Définition 1.4.1

L'intérieur relatif d'un sous ensemble non vide C de $\mathbb{R}^n$, noté $ri(C)$ est défini comme l'intérieur de C vu comme sous-ensemble de $aff(C)$.

Ce dernier étant muni de la topologie induite.

$$ri(C) = \{x \in aff(C) \,/\, \exists \varepsilon > 0, (x + \varepsilon B) \cap aff(C) \subseteq C\};$$

Evidemment : $int(C) \subseteq ri(C) \subseteq C \subseteq cl(C)$.

• L'ensemble $cl(C) \setminus ri(C) = \{x \in cl(C), x \notin ri(C)\}$ est appelé frontière relative de C.

• C est dit relativement ouvert si $ri(C) = C$.

Théorème 1.4.2

Soit $\emptyset \neq C \subset \mathbb{R}^n$ convexe, alors $ri(C)$ est non vide. En autre on a :

$$cl(ri(C)) = cl(C), \; ri(C) = ri(cl(C));$$

et

$$a \in ri(C), \; b \in cl(C) \, et 0 < \bar{t} < 1 \Longrightarrow c = a + \bar{t}(b-a) \in ri(C).$$

Preuve :

Soit $x_0 \in C$ fixé et soit p la dimension de $aff\,(C)$. Alors il existe $x_1, x_2, ..., x_p \in C$ tels que les p vecteurs $(x_i - x_0)$, $i = 1, ..., p$ sont linéairement indépendants et

$$aff\,(C) = \left\{ x = x_0 + \sum_{i=1}^{p} t_i\,(x_i - x_0) : t_i \in \mathbb{R}, i = 1, ..., p \right\}.$$

Considérer

$$S = \left\{ x = x_0 + \sum_{i=1}^{p} t_i\,(x_i - x_0) : \sum_{i=1}^{p} t_i < 1 \text{ et } t_i > 0, i = 1, ..., p \right\}.$$

Alors S est un ouvert non vide de $aff\,(C)$ contenu dans C et donc $ri\,(C)$ est non vide.

Puisque $aff\,(C) = aff\,(cl\,(C))$ on a $cl\,(ri\,(C)) = cl\,(C)$ et $ri\,(C) = ri\,(cl\,(C))$ en concéquence du théorème.

Ce résultat ne passe pas en dimension infinie où on peut avoir en outre

$aff\,(C) \neq aff\,(cl\,(C))$.

Un contre-exemple est donné par un hyperplan non fermé C d'un espace de Hilbert H.

On obtient alors $cl\,(C) = H$, $int\,(C = \emptyset)$ et $ri\,(C) = C \neq int\,(cl\,(C)) = H$.

La notion d'intérieur relatif d'un convexe est donc une notion spécifique aux espaces vectoriels de dimension finie.

Définition 1.4.2

Etant donné $\emptyset \neq S \subset E$, on considère l'intersection de tous les convexes fermés contenant S, on obtient ainsi un convexe fermé qui est aussi le plus petit convexe fermé contenant S.

On l'appelle enveloppe convexe fermée de S et on le note $\overline{co}\,(S)$.

Proposition 1.4.2 Soit $\emptyset \neq S$. Alors, $\overline{co}\,(S) = cl\,(co\,(S))$.

Preuve :

Tout convexe fermé contenant S contient $co\,(S)$ et donc sa fermeture qui est aussi convexe.

De façon générale l'enveloppe convexe d'un fermé n'est pas fermée. Considérer l'ensemble :

$$S = \left\{ (x_1, x_2) \in \mathbb{R}^2 : x_1 > 0,\ x_1 x_2 = 1 \right\} \cup \{(0,0)\}.$$

On a alors :

$$co\,(S) = \left\{ (x_1, x_2) \in \mathbb{R}^2 : x_1 > 0,\ x_2 > 0 \right\} \cup \{(0,0)\}.$$

Cependant, en dimension finie.

Proposition 1.4.3 L'enveloppe convexe d'un fermé borné de $\mathbb{R}^n$ est fermée bornée.
Preuve :
Soit $\emptyset \neq S \subset \mathbb{R}^n$ fermé borné,

$$T = \left\{ t = (t_0, t_1, ..., t_n) \in \mathbb{R}^{n+1} : \sum_{i=0}^{n} t_i = 1,\ t_i > 0,\ i = 0, ..., n \right\}.$$

$K = S^{n+1} \times T$ et $f : K \longrightarrow \mathbb{R}^n$ définie par :
$f(x_0, x_1, ..., x_n, t_0, t_1, ..., t_n) = \sum_{i=0}^{n} t_i x_i$.

Alors K est compact, f est continue et donc $f(K)$ est compact. Le théorème de Carathéodory nous dit que $f(K) = co(K)$.

1.5 Points extrémaux

Définition 1.5.1
Etant donné $C \subset E$ convexe, $c \in C$ est dit être un point extrémal de C si :

$$c = ta + (1 - t)b \text{ avec } a, b \in C \text{ et } t \in]0, 1[\Longrightarrow a = b = c.$$

Il est clair que si C est un convexe d'intérieur non vide, tous ses points extrémaux sont contenus dans sa frontière.

En dimension finie, les points extrémaux d'un convexe sont contenus dans sa frontière relative.

Un convexe fermé peut ne pas avoir de point extrémal. C'est bien sur le cas de E.

Théorème 1.5.1 (Théorème de Krein_Milman)
Un convexe fermé borné est l'enveloppe convexe de ses points extrémaux.

1.6 Projection

Soit C un convexe fermé non vide de $\mathbb{R}^n$. Pour tout $x \in \mathbb{R}^n$ il existe un et un seul élément dans C à distance minimale de x cet élément s'appelle la projection de x sur C et est noté $p_C(x)$:

$$\begin{cases} p_C(x) \in C, \\ \|p_C(x) - x\| \leq \|c - x\| \text{ pour tout } c \in C. \end{cases}$$

Caractérisation de $p_C(x)$:
$\bar{x} \in C$ est $p_C(x)$ si et seulement si :

$$\langle x-\bar{x}, c-\bar{x}\rangle \leq 0 \text{ pour tout } c \in C.$$

Dans le cas où C est un cône convexe fermé, la caractérisation ci-dessus se simplifie quelque peu :

Soit K un cône convexe fermé de $\mathbb{R}^n$ et K° son cône polaire ; alors $\bar{x} \in K$ est $p_K(x)$ si et seulement si :

$$x-\bar{x} \in K^\circ \text{ et } \langle x-\bar{x}, \bar{x}\rangle = 0.$$

Une propriétés fondamentale de l'application $p_C : \mathbb{R}^n \longrightarrow C \subset \mathbb{R}^n$ est comme suite :

$$\|p_C(x_1) - p_C(x_2)\|^2 \leq \langle p_C(x_1) - p_C(x_2), x_1 - x_2\rangle$$

pour tout x_1 et x_2 dans $\mathbb{R}^n$.

Il en découle notamment :

$$\|p_C(x_1) - p_C(x_2)\| \leq \|x_1 - x_2\|$$

pour tout x_1 et x_2 dans $\mathbb{R}^n$.

1.7 Séparation

Tout hyperplan $H = \left\{x \in \mathbb{R}^n, a^T x = \alpha\right\}$ sépare $\mathbb{R}^n$ en deux demi espaces fermés : $H^+ = \left\{x \in \mathbb{R}^n, a^T x \geq \alpha\right\}$ et $H^- = \left\{x \in \mathbb{R}^n, a^T x \leq \alpha\right\}$.

Deux sous ensembles C_1 et C_2 non vides de $\mathbb{R}^n$ sont séparés par H si $C_1 \subseteq H^+$ et $C_2 \subseteq H^-$ ou inversement.

1.7.1 Séparation large

C_1 et C_2 sont séparés (au sens large) par H si :
$a^T x \geq \alpha \ \forall x \in C_1$ et $a^T x \leq \alpha \ \forall x \in C_2$.
L'éventualité $C_1 \cup C_2 \subset H$.

1.7.2 Séparation propre

Si de plus, $C_1 \cup C_2 \nsubseteq H$, on dit que C_1 et C_2 sont proprement séparés par H.

1.7.3 Séparation stricte

C_1 et C_2 sont strictement séparés par H si :
$a^T x > \alpha, \forall x \in C_1$ et $a^T x < \alpha, \forall x \in C_2$ ssi $C_1 \subset int\,(H^+)$ et $C_2 \subset int\,(H^-)$.

1.7.4 Séparation forte

C_1 et C_2 sont fortement séparés par H si :

$\exists \varepsilon > 0/\ a^T x \geq \alpha + \varepsilon, \forall x \in C_1$, et $a^T x \leq \alpha - \varepsilon, \forall x \in C_2$.

1.8 Polyédres convexes

Définition 1.8.1
Q est un polyédre convexe s'il est une intersection d'un nombre fini de demi espaces fermés.

Un polyédre convexe est donc de la forme :

$$Q = \{x \in \mathbb{R}^n : Ax \leq a\}.$$

Définition 1.8.2
P est un polytope convexe s'il est enveloppe convexe d'un nombre fini de points.
Un polytope convexe est donc de la forme :

$$P = \left\{ y = \sum_{i=1}^{p} \sigma_i s_i : \sum_{i=1}^{p} \sigma_i = 1, \sigma_i \geq 0, i = 1, ..., p \right\}.$$

Les s_i sont appelés sommets du polytope.

P est compact, ses points extrémaux sont des sommets mais tout sommets n'est pas extrémal.

Définition 1.8.3
D est un cône convexe polyhèdral s'il est de la forme :

$$D = \left\{ d = \sum_{j=1}^{q} \delta_j d_j : \delta_j \geq 0, j = 1, ..., q \right\}.$$

Les d_j sont appelés directions du cône convexe polyhèdral.
Proposition 1.8.1 Un cône convexe polyhèdral est fermé.

Chapitre 2

Fonction convexe

2.1 Fonctions covexes

Définition 2.1.1

On dit que $f : E \longrightarrow \bar{\mathbb{R}}$ $\left(\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty; +\infty\}\right)$ est convexe sur E si son épigraphe est convexe, f est concave si $-f$ est convexe.

Nous serons amenés à considérer la somme de deux nombres $r, s \in \bar{\mathbb{R}}$. Par convention, on prend $+\infty + s = +\infty$ pour tout $s \in \bar{\mathbb{R}}$ et $r - \infty = -\infty$ si $r < +\infty$.

La proposition suivante étant une conséquence assez directe des définitions, nous en laissons la preuve aux lecteurs.

Proposition 2.1.1

Les trois notions ci-dessous sont équivalentes :

a) $epi\,(f)$ est convexe ;

b) $\widetilde{epi}\,(f)$ est convexe ;

c) $f\,(a + t\,(b - a)) \leq f\,(a) + t\,[f\,(b) - f\,(a)]$ pour tout $a, b \in E$ et tout $t \in \left]0, 1\right[$.

Les résultats suivants sont aussi aisément obtenus, certains à partir de la définition géométrique, d'autres à partir de la caractérisation analytique.

- Si f est convexe, alors $dom\,(f)\,, S_\lambda\,(f)\,, \tilde{S}_\lambda\,(f)\,, \lambda \in \mathbb{R}$ sont convexes, tels que :

$$\begin{aligned} dom\,(f) &= \{x \in \mathbb{R}^n /\ f\,(x) < +\infty\} = \mathbb{R} \backslash\ \{x /\ f\,(x) = +\infty\} \\ &= \{x \in \mathbb{R}^n /\ \exists \lambda \in \mathbb{R},\ (x, \lambda) \in epi\,(f)\}\,. \end{aligned}$$

Et

$$S_\lambda\,(f) = epi\,(f) \cap [\mathbb{R}^n \times \{\lambda\}]\,,\ \forall \lambda \in \mathbb{R}.$$

Et

$$\tilde{S}_\lambda(f) = \{x / \ f(x) < \lambda\}.$$

- Si f, g sont convexes sur E alors $f+g$ est convexe sur E ;
- Si f est convexe sur E et si on a $\lambda \geq 0$ alors λf est convexe sur E ;
- Si les fonctions $f_i, i \in I$ sont convexes sur E, alors $\sup_{i \in I} f_i$ est convexe sur E.

Il est facile de voir que si E et F sont deux espaces vectoriel sur $\mathbb{R}$, si A est une application linéaire de E dans F, $b \in F$ et $f : F \longrightarrow \bar{R}$ est convexe sur F, alors la fonction g définie par $g(x) = f(Ax+b)$ est convexe sur E.

Soient $f : E \longrightarrow \mathbb{R}$, $k : \mathbb{R} \longrightarrow \mathbb{R}$ et $g = k \circ f$. Alors on montre facilement que :

- Si f est convexe et k convexe croissante, alors g est convexe ;
- Si f est concave et k convexe décroissante, alors g est convexe ;
- Si f est convexe et k concave décroissante, alors g est concave ;
- Si f est concave et k concave croissante, alors g est concave.

Proposition 2.1.2

Soit $\varphi : E \times F \longrightarrow \bar{\mathbb{R}}$ convexe. Soit C un convexe de $E \times F$ et soit h définie comme suit :

$$h(x) = \inf_y [\varphi(x, y) : x, y \in C].$$

Alors, h est convexe sur E.

2.2 Continuité

Il suffit de penser à la fonction indicatrice d'un sous-ensemble convexe qui passe brusquement de 0 à $+\infty$ au bord du convexe pour se convaincre qu'une fonction convexe peut ne pas être continue. Cepandent les "ennuis" se concentrent en quelque sorte au bord du domaine de la fonction, c'est-à-dire là où elle n'est pas localement bornée.

Lemme 2.2.1

Soit $f : X \longrightarrow \bar{\mathbb{R}}$ une fonction convexe sur un espace de Hilbert X et $a \in X$ un point au voisinage duquel la fonction est bornée. Alors f est continue en a.

2.3 Semi-continuité (inférieure)

Définition 2.3.1

Une fonction $f : \mathbb{R}^n \to \bar{\mathbb{R}} = [-\infty, +\infty]$ est dite semi-continue inférieurment (s.c.i) en $x^0 \in \mathbb{R}^n$ si :

$\forall \varepsilon > 0, \exists\ \delta > 0 \ / \ f(x) \geq f(x^0) - \varepsilon$, dés que $\|x - x^0\| \leq \delta$.

f est dit semi-continue superieurement**(s.c.s)** en x^0 si $-f$ est **(s.c.i)** en x^0.

Evidemment, f est continue en x^0 ssi elle est à la fois **(s.c.i)** et **(s.c.s)** en x^0.

Enfin, on dit que f est **(s.c.i)** si, elle est **(s.c.i)** en tout point $x \in \mathbb{R}^n$.

Remarque 2.3.1

(1)- On peut remplacer $\mathbb{R}^n$ dant la définition réciproque par un sous-ensemble quelconque $C \subset \mathbb{R}^n$.

(2)- La semi-continuite inférieure de f en x^0 peut être exprimée par :

$f(x^0) \leq \lim_{k \longrightarrow \infty} f\left(x^k\right)$, pour chaque suite $\left\{x^k\right\}$ convergent vers x^0 et telle que : $\lim_{k \longrightarrow \infty} f\left(x^k\right) \in \bar{\mathbb{R}}$.

Autrement dit f est **s.c.i** en x^0 ssi $f(x^0) = \lim_{x \to x^0} \inf f(x) = \lim_{\delta \to 0} \left(\inf_{\|x - x^0\| \geq \delta} f(x)\right)$.

(3)- On peut également exprimer la **s.c.i** en termes de voisinages comme c'est le cas pour la continuité. A ce propos, on dit que :

f est **s.c.i** en x^0 si $\forall \lambda \in \mathbb{R}$ vérifiant $\lambda < f(x^0)$, $\exists V$ (voisinage de x) tel que $\forall x \in V$ on ait $\lambda < f(x)$.

Exemple 2.3.1

(1)- La fonction identiquement égale à $-\infty$ est trivialement **s.c.i.**

(2)- $f(x) = \begin{cases} 1/x & , x > 0 \\ 0 & , x \leq 0 \end{cases}$ est **s.c.i** en $x^0 = 0$.

2.4 Rappel sur la différentiable

On rappelle que, par définition, une fonction $f : \mathbb{R} \longrightarrow \mathbb{R}$ est dérivable en $a \in \mathbb{R}$ s'il existe $f'(a) \in \mathbb{R}$ tel que :

$$\sigma(h) = \frac{f(a+h) - f(a) - f'(a)h}{|h|} \longrightarrow 0 \text{ lorsque } h \longrightarrow 0.$$

Cette définition s'étend aux fonctions de plusieurs variables comme suit.

Définition 2.4.1

$F : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ est dit différentiable en $a \in \mathbb{R}^n$ s'il existe une matrice $p \times n$, notée $F'(a)$, telle que :

$$\sigma(h) = \frac{1}{\|h\|}\left[F(a+h) - F(a) - F'(a)h\right] \longrightarrow 0 \text{ lorsque } h \longrightarrow 0.$$

La matrice $F'(a)$ est appelée matrice Jacobienne.

Cela revient à dire que F est différentiable en a s'il existe une matrice $F'(a)$ et une fonction $\sigma : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ telles que :

$$F(a+h) = F(a) + F'(a)h + \|h\|\sigma(h) \text{ et } \sigma(h) \longrightarrow 0 \text{ lorsque } h \longrightarrow 0.$$

Il est clair que F est continue en a lorsque F est différentiable en a. désignons par $F_1(a), ..., F_p(a)$ les p coordonnées de $F(a)$ et par $e_1, ..., e_n$ les n vecteurs de la base canonoique de $\mathbb{R}^n$. le coefficient (i, j) de $F'(a)$ est :

$$F'_{ij}(a) = \frac{\partial F_i}{\partial x_j}(a) = \lim_{x_j \longrightarrow 0} \frac{F_i(a + x_j e_j) - F_i(a)}{x_j}.$$

Ainsi la différentiabilité de F en a entraîne l'existence des dérivées partielles de F en a mais l'existence des dérivées partielles n'implique pas la différentiabilité. Cela est illustré par l'exemple suivant :

Exemple 2.4.1

Considérer $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par :

$$f(x, y) = \begin{cases} \frac{x^2 - y^2}{y - x^2} & \text{si} \quad y \neq x^2, \\ 0 & \text{si} \quad y = x^2. \end{cases}$$

Cette fonction n'est pas continue, donc non différentiable à l'origine, les deux dérivées partielles à l'origine existent et sont nulles.

2.4.1 Dérivés directionnelles

Définition 2.4.2

Soit $f : \mathbb{R}^n \longrightarrow \bar{\mathbb{R}}$ une fonction quelconque, $x^0 \in \mathbb{R}^n$ un point tel que $f(x^0)$ est finie. Alors la dérivée directionnelle de f en x^0 dans la direction $d \in \mathbb{R}^n$ est définie par :

$$f'\left(x^0, d\right) = \lim_{\lambda \longrightarrow 0^+} \frac{f(x^0 + \lambda d) - f(x^0)}{\lambda}$$

si elle existe.

($\pm\infty$ sont des valeurs admises pour la limite).

Notons que :

$$-f'\left(x^0, -d\right) = \lim_{\lambda \longrightarrow 0^-} \frac{f(x^0 + \lambda d) - f(x^0)}{\lambda}.$$

• La dérivée directionnelle est dite ≪bilatérale≫ si :

$$f'\left(x^0, d\right) = -f\left(x^0, -d\right).$$

• Si $f'(x^0, d)$ existe $\forall d \in \mathbb{R}^n$, on dira que f est directionnellement différentiable en x^0.

• S'il existe un vecteur constant $g_0 \in \mathbb{R}^n$ tel que :

$$f'\left(x^0, d\right) = \langle g_0, d\rangle, \forall d \in \mathbb{R}^n.$$

On dit que f est G-différentiable (ou différentiable au sens de gateaux) en x^0 et on écrit : $f'\left(x^0\right) = g_0$.

2.4.2 Sous différentiel

Définition 2.4.3
Si f est une fonction convexe de $\mathbb{R}^n$ et $x^0 \in domf$ et s'il existe $g \in \mathbb{R}^n$ tel que :

$$f(x) - f\left(x^0\right) \geq \langle x - x^0, g\rangle, \ \forall x \in \mathbb{R}^n.$$

Alors le vecteur g est appelé sous-gradient de f en x^0.

L'ensemble $\partial f\left(x^0\right)$ de tous les sous-gradients de f en x^0 s'appelle le sous-différentiel de f en x^0.

Il est clair que :

$$\partial f\left(x^0\right) = \left\{g \in \mathbb{R}^n / f(x) - f\left(x^0\right) \geq \langle x - x^0, g\rangle, \ \forall x \in \mathbb{R}^n\right\}$$

est un convexe fermé (eventuellement vide) de $\mathbb{R}^n$.

Si $x^0 \in \mathbb{R}^n \backslash domf$. Alors $\partial f\left(x^0\right) = \emptyset$ (par définition).

$\partial f\left(x^0\right) = \{g\}$ ssi $g = \nabla f\left(x^0\right)$ (c-à-d ssi f est différentiable en x^0).

Enfin, si $\partial f\left(x^0\right) \neq \emptyset$, on dit que : f est sous-différentiable en x^0. C'est le cas si f est contunie en x^0 et alors on montre que $\partial f\left(x^0\right)$ est compact (borné).

2.5 Fonctions covexes d'une variable réelle

Soit $\theta : \mathbb{R} \longrightarrow \bar{\mathbb{R}}$. Si θ est convexe, $dom(\theta) = \{t : \theta(t) < \infty\}$ est convexe et est donc un intervalle. L'observation suivante est fondamentale.

Proposition 2.5.1
Soit $\theta : \mathbb{R} \longrightarrow \bar{\mathbb{R}}$ une fonction convexe. S'il existe a tel que :
$\theta(a) = -\infty$ alors $\theta(x) = -\infty$ pour tout $x \in int(dom(\theta))$.

Définition 2.5.1

On dit qu'une fonction $f : E \longrightarrow \bar{\mathbb{R}}$ est propre si son domaine est non vide et si $f(x) > -\infty$ pour tout $x \in E$.

Etant donnée une fonction $f : C \longrightarrow \mathbb{R}$ avec $\emptyset \neq C \subset E$, on étend f à E tout entier en posant $f(x) = +\infty$ si $x \notin C$.

Cette fonction est propre par construction et $dom(f) = C$.

Théorème 2.5.1

Soit I un intervalle non vide de $\mathbb{R}$ et $\theta : I \longrightarrow \mathbb{R}$. Chacune des trois conditions suivantes est équivalente à la convexité de θ :

$\frac{\theta(c)-\theta(a)}{c-a} \leq \frac{\theta(c)-\theta(b)}{c-b}$ $\forall a, b, c \in I$ tels que $a < b < c$,

$\frac{\theta(b)-\theta(a)}{b-a} \leq \frac{\theta(c)-\theta(a)}{c-a}$ $\forall a, b, c \in I$ tels que $a < b < c$,

$\frac{\theta(b)-\theta(a)}{b-a} \leq \frac{\theta(c)-\theta(b)}{c-b}$ $\forall a, b, c \in I$ tels que $a < b < c$.

Preuve :

Il est clair que θ est convexe si et seulement si pour tout $a, c \in I$ et t tels que : $b = ta + (1-t)c$ et $a < b < c$ on a :

$$\theta(b) \leq t\theta(a) + (1-t)\theta(c).$$

Cette inégalité est équivalente à chacune des trois conditions suivantes :

on a :

$\theta(b) \leq t\theta(a) - t\theta(c) + \theta(c)$

$\theta(b) - \theta(c) \leq t\theta(a) - t\theta(c)$

$\theta(c) - \theta(b) \geq t[\theta(c) - \theta(a)]$

donc :

$$t[\theta(c) - \theta(a)] \leq \theta(c) - \theta(b),$$

et

$\theta(b) - \theta(a) \leq t\theta(a) - \theta(a) + (1-t)\theta(c) = (1-t)[\theta(c) - \theta(a)]$

donc :

$$\theta(b) - \theta(a) \leq (1-t)[\theta(c) - \theta(a)],$$

et

$\theta(b) + t\theta(b) \leq t\theta(a) + (1-t)\theta(c) + t\theta(b)$

$t\theta(b) - t\theta(a) \leq (1-t)\theta(c) + t\theta(b) - \theta(b)$

donc :

$$t[\theta(b) - \theta(a)] \leq (1-t)[\theta(c) - \theta(b)].$$

Théorème 2.5.2

Soit I un intervalle non vide de $\mathbb{R}$ et $\theta : I \longrightarrow \mathbb{R}$ convexe.

Soient $a, b, c \in I$ tels que $a < b < c$. Alors θ est continue en b et admet une dérivée à droite et une dérivée à gauche en ce point. En autre :

$$\frac{\theta(b) - \theta(a)}{b - a} \leq \theta'_{-}(b) \leq \theta'_{+}(b) \leq \frac{\theta(c) - \theta(b)}{c - b}.$$

Corollaire 2.5.1

Une fonction convexe d'une variable réelle est continue sur l'intérieur de son domaine.

Corollaire 2.5.2

Soit I un intervalle non vide de $\mathbb{R}$ et $\theta : I \longrightarrow \mathbb{R}$ convexe.

Alors pour tout $a, b, c \in int(I)$ tels que $a < b < c$ on a :

$$\theta'_{+}(a) \leq \frac{\theta(b) - \theta(a)}{b - a} \leq \theta'_{-}(b) \leq \theta'_{+}(b) \leq \frac{\theta(c) - \theta(b)}{c - b} \leq \theta'_{-}(c).$$

Si θ est dérivable les deux demi-dérivées coïncident. On obtient alors des caractérisations du premier ordre.

Proposition 2.5.2

Soit I un intervalle ouvert non vide de $\mathbb{R}$ et $\theta : I \longrightarrow \mathbb{R}$ dérivable sur I. Les trois conditions suivantes sont équivalentes :

(a) θ est convexe sur I ;

(b) $(t - s)\left(\theta'(t) - \theta'(s)\right) \geq 0$ pour tout $t, s \in I$;

(c) $\theta(t) \geq \theta(s) + (t - s)\theta'(s)$ pour tout $t, s \in I$.

Preuve :

Si θ est convexe, on obtient (b) et (c) à partir du corollaire précédent. Supposons (b), et soient $a, b, c \in I$ avec $a < b < c$.

Puisque θ est dérivable et θ' est croissante, r et s existe tels que :

$$a < r < b < s < c, \text{ et } \frac{\theta(b) - \theta(a)}{b - a} \leq 0'(r) \leq 0'(s) \leq \frac{\theta(c) - \theta(b)}{c - b}.$$

θ est donc convexe.

Finalement, supposons (c) vérifiée, alors pour tout $t \in I$,

$\theta(t) = \sup_{s \in I}\left[\theta(s) + (t - s)\theta'(s)\right]$ (faire $s = t$).

Ainsi θ est convexe comme étant sup de fonctions affines.

La caractérisation du second ordre de convexité est un corollaire direct de la proposition. Elle s'écrit comme suit.

Proposition 2.5.3

Soit $\theta : I \longrightarrow \mathbb{R}$ deux fois dérivable sur I intervalle ouvert non vide de $\mathbb{R}$. Alors θ est convexe sur I si et seulement si $\theta^{''}(t) \geq 0$ pour tout $t \in I$.

Nous terminons notre étude en remarquant que si $c = \bar{t}a + (1 - \bar{t})\, b$, la valeur

$\bar{t}\theta(a) + (1 - \bar{t})\,\theta(b)$ correspond à l'approximation de $\theta(c)$ donnée par l'interpolation linéaire à partir des points a et b. On a le résultat suivant :

Proposition 2.5.4

Soit $\theta : I \longrightarrow \mathbb{R}$ deux fois dérivable sur l'intervalle I de $\mathbb{R}$. Soient $a, b, c \in I$ et $\bar{t} \in \mathbb{R}$ tels que $c = \bar{t}a + (1 - \bar{t})\, b$. Désignons par J l'enveloppe convexe des trois points a, b et c. Alors il existe $d \in int\,(J)$ tel que :

$$\theta(c) = \bar{t}\theta(a) + (1 - \bar{t})\,\theta(b) + \frac{\bar{t}\,(1 - \bar{t})}{2}\,(b - a)^2\theta^{''}(d)\,.$$

2.6 Fonctions convexes de plusieurs variables

Théorème 2.6.1

Soit $C \subset \mathbb{R}^n$ un convexe d'intérieur non vide et $f : C \longrightarrow \mathbb{R}$ convexe. Alors f est continue en tout $a \in int\,(C)$.

Proposition 2.6.1

Soit C un convexe ouvert non vide de $\mathbb{R}^n$ et $f : C \longrightarrow \mathbb{R}$ différentiable sur C. Les trois conditions suivantes sont équivalentes :

(a) f est convexe sur C ;

(b) $\langle \nabla f(x) - \nabla f(y)\,, x - y \rangle \geq 0$ pour tout $x, y \in C$;

(c) $f(y) \geq f(x) + \langle \nabla f(x)\,, x - y \rangle$ pour tout $x, y \in C$.

Proposition 2.6.2

Soit C un convexe ouvert non vide de $\mathbb{R}^n$ et $f : C \longrightarrow \mathbb{R}$ deux fois différentiable sur C. Alors f est convexe sur C si et seulement si la matrice $\nabla^2 f(x)$ est semi-définie positive pour tout $x \in C$.

2.7 Stricte et forte convexité

Définition 2.7.1

Soit C un convexe de E et $f : C \longrightarrow \mathbb{R}$, on dit que f est strictement convexe sur C si :

$$f(ta + (1 - t)\, b) < tf(a) + (1 - t)\, f(b)\,,\ \forall t \in \,]0, 1[\ \text{ et } \forall a, b \in C, a \neq b.$$

- Une fonction strictement convexe est convexe.

Si f est convexe et non strictement convexe alors il existe :
$a, c \in C$, $a \neq c$, $b = \bar{t}a + (1-\bar{t})\,c$ avec $\bar{t} \in \,]0,1[$ tels que :

$$f(c) = \bar{t}f(a) + (1-\bar{t})\,f(b)\,.$$

- La fonction $\theta(t) = f(ta + (1-t)\,b)$ est convexe sur $[0,1]$ et $\theta(\bar{t}) = \bar{t}\theta(0) + (1-\bar{t})\,\theta(1)\,.$
On obtient :

$$f(ta + (1-t)\,b) = \theta(t) = t\theta(0) + (1-t)\,\theta(1) = tf(a) + (1-t)\,f(b)\ \ \forall t \in [0,1]\,,$$

d'où la caractérisation suivantes :

Proposition 2.7.1

Soit C un convexe ouvert non vide de $\mathbb{R}^n$ et $f : C \longrightarrow \mathbb{R}$ différentiable sur C.
Les trois conditions suivantes sont équivalentes :
(a) f est strictement convexe sur C ;
(b) $\langle \nabla f(x) - \nabla f(y)\,, x-y\rangle > 0$ pour tout $x, y \in C, x \neq y$;
(c) $f(y) > f(x) + \langle \nabla f(x)\,, x-y\rangle > 0$ pour tout $x, y \in C, x \neq y$.
On en déduit la condition suffisante mais non nécessaire (considérer la fonction $f(x) = x^4$) du second ordre.

Proposition 2.7.2

Soit C un convexe ouvert non vide de $\mathbb{R}^n$ et $f : C \longrightarrow \mathbb{R}$ deux fois différentiable sur C.

Si pour tout $x \in C$ la matrice $\nabla^2 f(x)$ est définie positive alors f est strictement convexe sur C.

Chapitre 3

Optimisation convexe

3.1 Rappels d'optimisation

Sous la forme générale, un problème d'optimisation s'écrit comme suit :

$$\min_x [f(x) : x \in C]. \qquad (P)$$

La fonction $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ est appelée fonction objectif, l'ensemble $\emptyset \neq C \subset \mathbb{R}^n$ est appelé l'ensemble des solutions réalisables.

On appelle solution optimale globale de (P), tout point $\bar{x} \in C$ satisfaisant

$$f(\bar{x}) \leq f(x) \;\; \forall x \in C.$$

L'ensemble des points $\bar{x}$ ainsi défini est noté $\arg\min_C f(x)$.

On dit qu'un point $\bar{x} \in C$ est solution optimale locale de (P), s'il existe un voisinage V de $\bar{x}$ tel que :

$$f(\bar{x}) \leq f(x) \;\; \forall x \in C \cap V.$$

L'ensemble des points $\bar{x}$ ainsi défini est noté $loc \min_C f(x)$.

Nous avons toujours $\arg\min_C f(x) \subseteq loc \min_C f(x)$.

Rappelons que si f est semi continue inférieurement (s.c.i) et C compact (fermé et borné) alors $\arg\min_C f(x) \neq \emptyset$.

3.2 Condition d'optimalité

3.2.1 Le cas sans contraintes

Ici $C = \mathbb{R}^n$. Le probléme est donc :

$$\min_x \left[f(x) : x \in \mathbb{R}^n \right].$$

•Conditions du premier ordre

Supposons la fonction $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ de classe C^1 sur $\mathbb{R}^n$.

Condition nécessaire (CN1)

Si f a un minimum local en $\bar{x}$ alors $\nabla f(\bar{x}) = 0$.

Condition suffisante (CS1)

Si f est convexe et $\nabla f(\bar{x}) = 0$ alors f a un minimum global en $\bar{x}$.

•Conditions du second ordre

Supposons la fonction $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ de classe C^2 sur $\mathbb{R}^n$.

Conditions nécessaires (CN2)

Si f a un minimum local en $\bar{x} \in \mathbb{R}^n$ alors $\nabla f(\bar{x}) = 0$ et la matrice des dérivées secondes $\nabla^2 f(\bar{x})$ est semi-définie positive sur $\mathbb{R}^n$.

Conditions suffisantes (CS2)

Si $\nabla f(\bar{x}) = 0$ et si la matrice $\nabla^2 f(\bar{x})$ est définie positive, alors f a un minimum local en $\bar{x}$.

•Cas des fonctions convexes

Si f est convexe (non nécessairement différentiable), alors une condition nécessaire et suffisante (CNS) pour que $\bar{x}$ soit un minimum global est $0 \in \partial f(\bar{x})$.

3.2.2 Le cas avec contraintes

On considère maintenant le problème :

$$\min_x \left[f(x) : x \in C \right].$$

Où C est un convexe de $\mathbb{R}^n$ et f est C^1 sur un ouvert contenant C. On a les conditions suivantes d'optimalité en un point $\bar{x} \in C$.

- Condition nécessaire du premier ordre :

Si f a un minimum local sur C en $\bar{x}$ alors $\langle \nabla f(x), x - \bar{x} \rangle \geq 0$ pour tout $x \in C$.

- Condition suffisante du premier ordre :

Si f est convexe et $\langle \nabla f(x), x - \bar{x} \rangle \geq 0$ pour tout $x \in C$ alors f a un minimum global sur C en $\bar{x}$.

Notons que dans les deux cas nous avons $\nabla f(\bar{x}) \in \tilde{N}(\bar{x})$ où

$$\tilde{N}(\bar{x}) = \left\{ x^* \in \mathbb{R}^n : \langle x^*, y - x \rangle \leq 0 \ \forall y \in \tilde{S}(\bar{x}) \right\}$$

est le cône normal à l'ensemble de niveau.

$$\tilde{S}(\bar{x}) = \{ y \in \mathbb{R}^n \text{ tel que } f(y) < f(\bar{x}) \}.$$

On suppose maintenant que C est défini comme suit :

$$C = \{ x \in \mathbb{R}^n : g_i(x) \leq 0, \ i = 1, ..., p; \ h_j(x) = 0, \ j = 1, ..., q \}.$$

Les contraintes $g_i(x) \leq 0$ sont appelées contraintes d'inégalités et les contraintes $h_j(x) = 0$ contraintes d'égalités.

L'obtention de conditions nécessaires d'optimalité nécessite que certaines conditions appelées conditions de qualification des contraintes soient vérifiées.

Les conditions de qualification :

•C est un polyédre convexe

C'est le cas lorsque les fonctions g_i et h_j sont affines.

•Condition de Slater

La condition de qualification des contraintes de Slater est comme suit :

- Les fonctions g_i sont convexes et les fonctions h_j sont affines.
- Il existe x^0 tel que $g_i(x^0) < 0$ et $h_j(x^0) = 0$ pour tout i, j.

•Condition de Mangasarian-Fromowitz

La condition de qualification des contraintes de Mangasarian-Fromowitz en un point $\bar{x} \in C$ est comme suit :

- Les q vecteurs $\nabla h_j(\bar{x})$ sont linéairement indépendants.
- Il existe $\bar{d}$ tel que $\langle \nabla h_j(\bar{x}), \bar{d} \rangle = 0$ pour tout j et $\langle \nabla g_i(\bar{x}), \bar{d} \rangle < 0$ pour tout $i \in I(\bar{x})$.

Le théorème suivant dit de Karush-Kuhn-Tucker-Lagrange donne une condition nécessaire d'optimalité :

Théorème 3.2.1

Supposons que les fonctions f, g_i, h_j sont C^1 dans un voisinage de $\bar{x} \in C$ et que les contraintes vérifient une des trois conditions de qualification ci-dessus.

Si f a un minimum local en $\bar{x}$ sur C alors il existe $\lambda \in \mathbb{R}^p$ et $\mu \in \mathbb{R}^q$ tels que :

$$\nabla f(\bar{x}) + \sum_{i=1}^{p} \lambda_i \nabla g_i(\bar{x}) + \sum_{j=1}^{q} \mu_j \nabla h_j(\bar{x}) = 0,$$

$$\lambda_i \geq 0, g_i(\bar{x}) \leq 0, \lambda_i g_i(\bar{x}) = 0 \ \forall i \text{ et } h_j(\bar{x}) = 0 \ \forall j.$$

Les quantités λ_i et μ_j sont appelées multiplicateurs de Karush-Kuhn-Tucker-Lagrange.

Lorsque le problème est convexe (f et g_i convexes et h_j affines) on a la condition suffisante d'optimalité :

Théorème 3.2.2

Supposons que les fonctions f, g_i sont C^1 dans un voisinage de $\bar{x} \in C$ et convexe et que les fonctions h_j sont affines. S'il existe $\lambda \in \mathbb{R}^p$ et $\mu \in \mathbb{R}^q$ tels que :

$$\nabla f(\bar{x}) + \sum_{i=1}^{p} \lambda_i \nabla g_i(\bar{x}) + \sum_{j=1}^{q} \mu_j \nabla h_j(\bar{x}) = 0,$$

$$\lambda_i \geq 0, g_i(\bar{x}) \leq 0, \lambda_i g_i(\bar{x}) = 0 \ \forall i \text{ et } h_j(\bar{x}) = 0 \ \forall j,$$

alors f a un minimum global en $\bar{x}$ sur C.

On associe au problème d'optimisation :

$$\min_x \{f(x) : g_i(x) \leq 0, \ i = 1, ..., p; \ h_j(x) = 0, \ j = 1, ..., q\},$$

la fonction $l : \mathbb{R}^n \times [0, \infty[^p \times \mathbb{R}^q \longrightarrow \mathbb{R}$ définie par :

$$l(x, \lambda, \mu) = f(x) + \sum_{i=1}^{p} \lambda_i g_i(x) + \sum_{j=1}^{q} \mu_j h_j(x).$$

appelée Lagrangien.

Le théorème suivant s'écrit en terme de Lagrangien comme suit :

Théorème 3.2.3

Supposons que les fonctions f , g_i , h_j sont C^1 dans un voisinage de $\bar{x} \in C$ et que les contraintes vérifient une condition de calification.

Si f a un minimum local en $\bar{x}$ sur C alors il existe $\lambda \in \mathbb{R}^p$ et $\mu \in \mathbb{R}^q$ tels que :

$$\nabla_x l(\bar{x}, \lambda, \mu) = 0, \nabla_\lambda l(\bar{x}, \lambda, \mu) \leq 0, \nabla_\mu l(\bar{x}, \lambda, \mu) = 0,$$

$$\lambda \geq 0, \langle \lambda, \nabla_\lambda l(\bar{x}, \lambda, \mu) \rangle = 0.$$

Conclusion Générale

La notion de convexité est une outil mathématique d'importance capital pour l'étude des problèmes d'optimisation et autres.

En effet, les ensembles et les fonctions convexes jouissant de propriétés algébriques et topologiques fréquement utilisées dans la caractérisation des solution optimales et même pour la développement des procedures permettant de calculer (numériquement) ces solutions.

Bibliographie

[1] A. Keraghel, Analyse convexe : Théorie fondamentale et exercices, Editions Dar El'houda Ain Mlila , Algérie, (2001).

[2] A. Keraghel, Programmtion mathématique différentiable, Université Ferhat Abbas, Sétif Algérie.

[3] M. Minoux, Programmation mathématique : Théorie et algorithmes, T1 et T2, Dunod, 1983.

[4] Jean-Baptiste Hiriart-Urruty, Optimisation et analyse convexe : Exercices et problèmes corrigés, avec rappels de cours, Les Ulis cedex A,France 2009.

[5] Guy Cohen, Convexité et optimisation, École Nationale des Ponts et Chaussées et INRIA, Édition 2000 (correction 2006).