REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE HADJ LAKHDAR
BATNA (ALGERIE)

MEMOIRE

Présenté à la faculté des Sciences
Département de Mathématiques
Pour l'obtention du diplôme de

MAGISTER

Option : *Analyse Mathématiques Appliquée*

Par

Gasmi Bouthaina

THEME

CONTRIBUTION A L'ETUDE DES METHODES DE RESOLUTION DES PROBLEMES D'OPTIMISATION QUADRATIQUES

Soutenue le :12/02/2007

Devant le jury

Mr.B.MEZERDI	Prof.	Université de Biskra	Président
Mr.R.BENACER	Prof.	Université de Batna	Rapporteur
Mr.S.E.REBIAI	Prof.	Université de Batna	Examinateu
Mr.K.MESSAOUDI	Prof.	Université de Batna	Examinateur

Remerciement

Je remercie avant tout *ALLAH* le tout puissant de m'avoir donné la force et le courge de réaliser ce travail.

Je remercie beaucoup monsieur **Rachid Bennacer,** mon directeur de thèse, pour la qualité de son encadrement, pour sa confiance, et ses orientations judicieuses.

Je remercie le professeur **Mr. B. Mezerdi** de l'Université de Biskra, d'avoir accepté de présider le jury de soutenance.

J'aimerai exprimer ma gratitude à messieurs les professeurs **Mr. Sallah-Eddine Rebiai**, et **Khalifa Messaoudi** de l'Université Hadj Lakhdar de Batna pour avoir accepté d'être examinateurs.

Je remercie aussi beaucoup mes beaux-parents pour leur soutien tout au long de ce travail.

Table des matières

Notation :

X : l'espace euclidien $\mathbb{R}^n$, muni du produit scalaire usuel noté par $\langle .,. \rangle$ et de la norme associée $\|x\| = \sqrt{\langle x, x\rangle}$.

Y : l'espace vectoriel dual de X, et notera aussi par $\overline{\mathbb{R}} = \{+\infty, -\infty\} \cup \mathbb{R}$.

$conv\,(C)$: l'ensemble des fonctions convexes sur C a valeurs dans $\mathbb{R}$.

$DC\,(C)$: l'ensemble des fonctions DC sur C.

$C^1\,(C)$: l'ensemble des fonctions dont le gradient est localement lipchitzien sur C.

$C^{1,1}\,(C)$: le sous espace vectoriel de $C^1\,(C)$.

$LC^2\,(C)$: est le cône convexe des fonctions localement enveloppe supérieure d'une famille des fonctions de $C^2\,(C)$.

$\partial f\,(x)$: le sous-différentielle de la fonction $f\,(x)$ en le point x.

$\partial_\varepsilon f\,(x)$: le sous-différentielle approchée de la fonction $f\,(x)$ en le point x.

$ir\,(C)$: l'intérieur relatif de C.

$int\,(C)$: l'intérieur de C.

$\Gamma_\circ\,(X)$: l'ensemble des fonctions convexes, semi continue inférieurement et propres sur X.

$\langle x, y\rangle_X$: le produit scalaire dans l'espace X, qui est définie par $\langle x, y\rangle_X := y^T x$.

İntroduction générale :

Le monde est écrit en language mathématique **(Galilée)**, et nous pourrons ajouter que tous les problèmes de la vie courante, des plus simples aux plus complexes se posent comme problèmes d'optimisations. Ces problèmes peut être partagées à deux classes les quelles :

la classe des problèmes convexes, et *la classe des problèmes non convexes.*

Cette classification a poussé Rockafellar à dire :

"*The great watershed in optimization isn't between linearity or non linearity but between convexity and non convexity"[4],* où on trouve une raison simple et évidente que la plus part des problèmes d'optimisations de la vie courante sont de nature *non convexes.*

Entre ces problèmes il y a le type des problèmes de la programmation quadratique où la fonction optimiser est une forme quadratique.

Ce type de problème admet plusieurs applications dans plusieurs domaines de science et technologie, où on trouve que plusieurs problèmes non linéaires sont converges à cette forme, prend par exemple l'optimisation bilinéaire qui peut être interpritée comme un problème quadratique, et pour bien expliquer, il y a des classes spéciaux de *structured stochastic games* présentent une programmation bilinéaire interpritées à des problèmes quadratiques.

La résolution des problèmes quadratiques avec contraintes linéaires est une application très difficile, dans le cas non convexe, ça est bien dite que les problèmes quadratiques non convexes sont *NP-complete,* d'un autre côté, la recherche globale des solutions consiste une application très difficile et trés compliquée, c'est pour ça, plusieurs efforts étés proposées pour trouver des méthodes efficaces dans le but de simplifier la résolution de ce type de problème. Les méthodes de la programmation non linéaires traditionnelles connues sont toujours obtiennent des solutions locaux dans le cas des problèmes quadratiques indéfinies (non convexes). D'après plusieurs efforts, on a trouvée que dans plusieurs applications l'optimum global ou bien la bonne solution approximative du l'optimum global peut être atteint.

Dans ce travail, le problème posé est de la forme :

$$\begin{cases} \min f(x) = \frac{1}{2}x^T Q x - d^T x \\ Bx \leq c, x \geq 0 \end{cases} \tag{PQN}$$

telle que :

$$\begin{aligned} Q &= (q_{ij})\,; i = 1,2,...,n \text{ et } j = 1,2,...,n \\ d^T &= (d_1, d_2, ..., d_n) \in \mathbb{R}^n \\ B &= (b_{ij})\,; i = 1,2,...,m \text{ et } j = 1,2,...,n \\ c &= (c_1, c_1, ..., c_m) \in \mathbb{R}^m \\ x^T &= (x_1, x_2, ..., x_n) \in \mathbb{R}^n \end{aligned}$$

où la matrice Q est symétrique généralement n'est pas définie positive, si pour ça, on dit que le problème posé iç̧i est un problème de la programmation quadratique non convexe.

Ce type de problème est multitude des solutions locaux, d'un côté, en général il n'y a aucun critère qui les englobe, plusieurs efforts étaient proposées pour trouver une méthode éfficace avec une approche globale, entre ces méthodes on a choisies quatre méthodes très importantes pour résoudre globalement les problèmes de la programmation quadratique non convexes avec contraintes linéaires sous la forme de **(PQN).**

Ce mémoire compose de cinq chapitres structurés sous la forme suivante :

Le premier chapitre contient des notions d'analyse convexe et de la programmation mathématiques très utiles dans notre travail. On présente les conditions d'optimalités de type **Karuch-Kuhn-Tucker (KKT),** les conditions d'optimalités d'ordre un et d'ordre deux.

Dans le deuxième chapitre, on s'intéresse à l'étude d'une méthode de type *primal-dual,* cette méthode est appelée *la méthode* **DCA**. Ce chapitre est divisée en deux sections. La première section contient un rappelle sur la programmation **DC** où on a spécialement données le principe de la dualité avec les conditions d'optimalitées en optimisation **DC**.On présente dans la seconde section la construction de la méthode **DCA** qui est basée sur la construction de deux suites $\{x_k\}_{k\in\mathbb{N}}$ et $\{y_k\}_{k\in\mathbb{N}}$ qui sont améliorées à chaque itération dans le but de trouver leurs limites x^* et y^* respectivement comme des candidates pour être les optima locaux de problèmes primal et dual (P_{dc}) et (D_{dc}) respectivement.

Dans le troisième chapitre on présente une méthode de recherche globale appelée *la méthode de la transformation duale canonique (TDC),* qu'est efficace pour la résolution des problèmes quadratiques non convexes.Du problème primal **(PQN)** on forme un problème dual canonique sous forme d'un système algébrique plus facile pour le résoudre. On a étudié quelques exemples dans la fin du chapitre.

Le quatrième chapitre est consacrée à l'étude de *la méthode de séparation et interpolation* (Branch and Bound). Le principe de cette méthode est basé sur l'écriture de la fonction optimiser comme une forme séparable basée sur la structure propre de la matrice Q. L'approximation linéaire de la partie concave transforme notre problème **(PQN)** à un autre problème de type convexe. On présente l'erreur relatif de cette approximation entre la valeur optimale primale et la valeur optimale duale.

Dans le cinquième chapitre on présente un nouveau approche rectangulaire pour *la méthode de séparation et réduction*, dans la quelle, on donne un nouveau forme linéaire approximant inférieurement la fonction optimiser sur les n-rectangles. Cette forme détermine la borne inférieure de la valeur optimale globale associée au problème original sur chaque rectangle, puis, on explicite les techniques de partition et réduction sur un rectangle choisie. On termine ce chapitre avec l'étude de la convergence et la présentation d'un exemple d'application.

Chapitre 1

Elémments d'analyse convexe, conditions d'optimalitées.

Dans cette partie nous rappellerons brièvement quelque notions qui nous seront utiles pour la suite de notre travail. Des développements sur le sujet peuvent être trouvées dans la monographie de Rockafellar [18], et Abdelkrim Keraghel [1] avec le livre d'Ekeland [3].

1.1 Les ensembles convexes :

Définition 1.1 : On dira qu'une ensemble C est convexe si pour tout $x, y \in C$ le segment :

$$[x, y] = \{(1-t)\,x + ty \ / \ t \in [0, 1]\} \tag{1.1}$$

est contenu dans C.

Définition 1.2 : Un sous-ensemble $C \subseteq \mathbb{R}^n$ est dite strictement convexe s'il vérifier :

$$\forall x, y \in C, (x \neq y)\,,]x, y[= \{(1-\lambda)\,x + \lambda y \ / \ \lambda \in]0, 1[\} \subset int\,(C) \tag{1.2}$$

Définition 1.3 : Une partie $C \subseteq \mathbb{R}^n$ est dite fortement convexe s'il existe une constante $\gamma > 0$ telle que :

$$\frac{x_1 + x_2}{2} + y \in C; \forall x_1, x_2 \in C; \text{et } \ \|y\| \leq \gamma \|x_1 - x_2\|^2 \tag{1.3}$$

Définition 1.4 : L'enveloppe convexe d'un sous–ensemble $C \subseteq X$, notée par $CO\,(C)$, est l'ensemble des combinaisons convexes finies d'élements de C, c'est-à-dire :

$$CO\,(C) = \left\{ \sum_{i=1}^{n} \lambda_i x_i \ / \ \lambda_i \in \mathbb{R}_+, \ x_i \in C \ ; \ \forall \ i = \overline{1, n}, \text{ et } \sum_{i=1}^{n} \lambda_i = 1 \right\} \tag{1.4}$$

Définition 1.5 : L'enveloppe convexe d'un sous–ensemble $C \subseteq \mathbb{R}^n$, est l'intersection de tous les ensembles convexes contenant C, et on le note par :

$$conv\,(C)$$

d'autre terme, $conv\,(C)$ est la plus petite convexe de $\mathbb{R}^n$contenant C.

Définition 1.6 : Soit C un ensemble convexe de X, on définie la variété linéaire engendrée par C (i,e l'enveloppe affine de C) notée par $aff\,(C)$ est la plus petite partie affine de $\mathbb{R}^n$ contenant C définie par :

$$aff\,(C) = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^{n} \lambda_i x_i \ / \ \lambda_i \in \mathbb{R}, \ x_i \in C, \ \forall \ i = \overline{1, n} \text{ avec } \sum_{i=1}^{n} \lambda_i = 1 \right\} \tag{1.5}$$

Définition 1.7 : Soit $C \subseteq \mathbb{R}^n$ et $a \in \mathbb{R}^n$, on appelle translaté de C par a, l'ensemble :

$$T = C + a := \{x + a \ / \ x \in F\} \tag{1.6}$$

Remarque 1.1 : $aff\,(C)$ est le translaté d'un espace vectoriel.

Définition 1.8 : L'intérieur relatif d'un ensemble convexe C est son intérieur dans $aff\,(C)$, muni de la topologie induite de celle de X, et on le note par :

$$ir\,(C) = \{x \in C \ / \ \exists r > 0\,; B\,(x, r) \cap aff\,(C) \subset C\} \tag{1.7}$$

où :

$$B\,(x, r) = \{y : \|x - y\| \leq r\} \tag{1.8}$$

est la boule centrée en x et de rayon r.

Définition 1.9 : Soit $C \subseteq IR^n$ un ensemble convexe, et soit :

$$f : C \longrightarrow \overline{\mathbb{R}}$$

on appelle domaine de f l'ensemble définie par :

$$dom\,(f) = \{x \in C \ / \ f\,(x) < +\infty\} \tag{1.9}$$

Remarque 1.2 : Evidament on a :

$$int\,(C) \subseteq ir\,(C) \subseteq C \subseteq cl\,(C) \tag{1.10}$$

Définition 1.10 : La fonction $f : C \longrightarrow \overline{\mathbb{R}}$ est dite" *fonction propre " si elle ne prend jamais la valeur* $(-\infty)$ et si elle n'est pas identiquement égale à $(+\infty)$.

Définition 1.11 : L'épigraphe de f est l'ensemble définie par :

$$\acute{e}pi\,(f) = \{(x, \alpha) \in C \times \mathbb{R} \ / \ f\,(x) \leq \alpha\} \tag{1.11}$$

Définition 1.12 : On dit qu'une fonction $f : X \longrightarrow \overline{\mathbb{R}}$ est convexe si pour tout

$x_1, x_2 \in dom\,(f)$ et pour tout $\lambda \in [0,1]$ ona :

$$f\,(\lambda x_1 + (1-\lambda)\,x_2) \leq \lambda f\,(x_1) + (1-\lambda)\,f\,(x_2) \tag{1.12}$$

Définition 1.13 : On dit que $f : X \longrightarrow \overline{\mathbb{R}}$ est strictement convexe si pour tout $x_1, x_2 \in dom\,(f)$ avec $x_1 \neq x_2$ et pour tout $\lambda \in [0,1]$ on a :

$$f\,(\lambda x_1 + (1-\lambda)\,x_2) < \lambda f\,(x_1) + (1-\lambda)\,f\,(x_2) \tag{1.13}$$

Proposition 1.1 : *Soit $f : X \longrightarrow \overline{\mathbb{R}}$, f est convexe si et seulement si épi (f) est convexe.*

Définition 1.14 : On dira que $f : X \longrightarrow \overline{\mathbb{R}}$ est " *fortement convexe* " de module ρ sur l'ensemble convexe C s'il existe un nombre réel $\rho > 0$ telle que :

$$f\,(\lambda x_1 + (1-\lambda)\,x_2) \leq \lambda f\,(x_1) + (1-\lambda)\,f\,(x_2) - \frac{\rho}{2}\lambda\,(1-\lambda)\,\|x_1 - x_2\|^2 \tag{1.14}$$

pour tout $x_1, x_2 \in dom\,(f)$, et pour tout $\lambda \in [0,1]$.

On peut poser la conclusion suivante :

$$\begin{array}{ccc} \text{fonction } \textit{fortement convexe} & \Longrightarrow & \text{fonction } \textit{strictement convexe} \\ & \Downarrow & \\ & \text{fonction convexe} & \end{array}$$

Définition 1.15 : On a :

$$\arg\min\,(f) = \left\{ u \in X : f(u) = \inf_{x \in X} f(x) \right\} \tag{1.15}$$

et pour $\varepsilon > 0$:

$$\varepsilon - \arg\min\,(f) = \left\{ u \in X : f\,(u) \leq \inf_{w \in X} f\,(w) + \varepsilon \right\} \tag{1.16}$$

Remarque 1.3 : Si :

$$\inf_{w \in X} f(w) = -\infty$$

alors, On a :

$$\varepsilon - \arg\min(f) = \left\{u \in X : f(u) = -\varepsilon^{-1}\right\} \tag{1.17}$$

Remarque 1.4 :

$$\arg\min(f) = \bigcap_{\varepsilon \succ 0} \varepsilon - \arg\min(f) \tag{1.18}$$

Etant donné une fonction convexe et propre f sur X, donc soient les définitions suivantes :

Définition 1.16 : On dira que $y^\circ \in Y$ est un“ sous gradient ” de f sur X en $\mathring{x} \in dom(f)$ si :

$$\langle \mathring{y}, x - \mathring{x} \rangle + f(\mathring{x}) \leq f(x) \ ; \forall x \in X \tag{1.19}$$

Définition 1.17 : Le sous-différentiel de f au point $\mathring{x}$ est l'ensemble de tous les sous gradient de f au même point $\mathring{x}$, et on le note par $\partial f(\mathring{x})$.

Définition 1.18 : Le domaine du sous différentiel de la fonction f est définie par :

$$dom(\partial f) = \{x \ / \ \partial f(x) \neq \Phi\} \tag{1.20}$$

Définition 1.19 : Soit ε un réel positif, un élément $y^\circ \in Y$ est appelé “$\varepsilon-$sous différentiel” de f au point $\mathring{x}$ si :

$$\langle \mathring{y}, x - \mathring{x} \rangle + f(\mathring{x}) - \varepsilon \leq f(x) \ ;\forall x \in X \tag{1.21}$$

Remarque 1.5 : On note par $\partial_\varepsilon f(\mathring{x})$ l'ensemble de tous les $\varepsilon-$sous-différentiel de f au point $\mathring{x}$.

Définition 1.20 : Soit la fonction $f : X \longrightarrow \overline{\mathbb{R}}$, cette fonction est dite “semi continue inférieurement ” c'est-à-dire (sci) en un point $\hat{x} \in X$ si :

$$\forall (x_n) \in X : x_n \longrightarrow \hat{x}, \text{on a}, f(\hat{x}) \leq \liminf_{n \to +\infty} f(x_n) \tag{1.22}$$

Définition 1.21 : Une fonction F est coercive si :

$$\lim_{\|x\| \longrightarrow +\infty} F(x) = +\infty$$

Définition 1.22 : La conjuguée de f en $y \in Y$ est définie par :

$$f^*(y) = \sup \{\langle x, y \rangle - f(x) : x \in X\} \tag{1.23}$$

ainsi f^*est l'enveloppe supérieur des fonctions affines sur Y, telle que :

$$y \longrightarrow \langle x, y \rangle - f(x) \text{ sur } Y$$

PROPRIETES :

(1) f^*est toujours convexe.

(2) Si $f = -\infty$ alors $f^* = +\infty$.

(3) On a $(f^*)^* \leq f$ et on a l'égalité si $f \in \Gamma_\circ(X)$.

donc on a également la proposition suivante :

Proposition 1.2 :

(1) *$\partial f(x)$ est une partie convexe fermé de Y*

(2)

$$f \in \Gamma_\circ(X) \Longleftrightarrow f^* \in \Gamma_\circ(Y) \tag{1.24}$$

et dans ce cas on a $f = f^{**}$ [Fenchel.Moreau].

(3)

$$y \in \partial f(x) \Longleftrightarrow f(x) + f^*(y) = \langle x, y \rangle \tag{1.25}$$

(4) *Si $f \in \Gamma_\circ(X)$ alors on a :*

$$y \in \partial f(x) \Longleftrightarrow x \in \partial f^*(y) \tag{1.26}$$

(5) *Si $f \in \Gamma_\circ(X)$ et $\partial f(x)$ est réduit à un singleton $\{y\}$, alors f est différentiable en x et $\nabla f(x) = y$, et vice verre sa.*

(6) *Soit* $f : C \subseteq X \longrightarrow \overline{\mathbb{R}}$, *alors :*

$$[f(x^{\circ}) = \min\{f(x) \ / \ x \in X\}] \Longleftrightarrow [0 \in \partial f(x^{\circ})] \qquad (1.27)$$

1.2 Les fonctions convexes polyédrales :

Dans cette partie on va présenter quelle que ensembles spécialement polyédrales.

Définition 1.23 : Une partie convexe C est dite convexe polyédrale si :

$$C = \bigcap_{i=1}^{n} \{x \ / \ \langle a_i, x\rangle - b_i \leq 0, a_i \in Y, b_i \in \mathbb{R}\} \qquad (1.28)$$

Définition 1.24 : Une fonction $f : \mathbb{R}^n \longrightarrow \overline{\mathbb{R}}$ est dite polyédrale si son épigraphe est un ensemble polyédrale de $\mathbb{R}^{n+1}$.

Remarque 1.6 :

$$\textit{fonction polyédrale} \Longrightarrow \textit{fonction convexe}$$

Remarque 1.7 : Tout fonction plyédral apartient à $\Gamma_{\circ}(X)$.

Proposition 1.3 : *Soit* $f : X \longrightarrow \overline{\mathbb{R}}$, *une fonction convexe, alors* f *est plyédrale si et seulement si* $dom(f)$ *un ensemble convexe polyédrale, et :*

$$f(x) = \sup\{\langle a_i, x\rangle - b_i \ / \ i = \overline{1,k}\}, \forall x \in dom(f), k \leq n \qquad (1.29)$$

Proposition 1.4 : *Si* f *est une fonction polyédrale alors* f^* *l'est aussi, de plus, si* f *est partout finie alors :*

$$dom(f^*) = co\{a_i \ / \ i = \overline{1,k}\}$$

$$f^*(y) = \min\left\{\sum_{i=1}^{k} \lambda_i a_i \ / \ y = \sum_{i=1}^{k} \lambda_i a_i, \lambda_i \geq 0, \sum_{i=1}^{k} \lambda_i = 1\right\} \qquad (1.30)$$

Proposition 1.5 : *Si* f *est une" fonction polyédrale" alors* $\partial f(x)$ *est une partie convexe polyédrale non vide en tout point de* $dom(f)$.

Proposition 1.6 : *Si* $f_1, f_2, \cdots, f_s$ *sont des fonctions convexes polyédrales sur* X

telle que les ensembles convexes :

$$\left\{\{dom\,(f_i)\}\,;i=\overline{1,s}\right\}$$

ont un point commun alors :

$$\partial\,(f_1+f_2+\cdots+f_s)\,(x)=\partial f_1\,(x)+\partial f_2\,(x)+\cdots+\partial f_s\,(x) \qquad (1.31)$$

Définition 1.25 : L'ensemble de niveau $\alpha\in\mathbb{R}$:

(a) de *niveau inférieure* (large) :

$$S_\alpha\,(f)=\{x\in\mathbb{R}^n\ /\ f\,(x)\le\alpha\} \qquad (1.32)$$

(b) de *niveau supérieure :*

$$\{x\in\mathbb{R}^n\ /\ f\,(x)\ge\alpha\} \qquad (1.33)$$

Définition 1.26 : Un surface de niveau α est définie par :

$$\{x\in\mathbb{R}^n\ /\ f\,(x)\ge\alpha\}$$

Définition 1.27 : On dit qu'une fonction f est quasi-convexe si et seulement si :

$$f\,((1-\lambda)\,x+\lambda y)\le\max\{f\,(x)\,,f\,(y)\}\,,\forall\lambda\in[0,1]\,,\forall x,y\in X \qquad (1.34)$$

Géométriquement, elle est caractérisée par la convexité des ensembles de niveau inférieure (large).

Autre définition plus pratique :

Définition 1.28 : $\forall x_1,x_2\in X,\forall\lambda\in[0,1]$ on a f est quasi-convexe si et seulement si :

$$\{f\,(x_2)\ge f\,(x_1)\Longrightarrow f\,((1-\lambda)\,x_1+\lambda x_2)\le f\,(x_2)\} \qquad (1.35)$$

c'est à dire $f\,(x_2)$ reste plus grand que $f\,(x)\,;\,\forall x\in[x_1,x_2]\,.$

Proposition 1.7 :

(1) *(f est convexe) implique que (f est quasi-convexe).*

(2) *(f est strictement convexe) implique que (f est strictement quasi-convexe).*

Proposition 1.8 :(*L'INEGALITIE DE JENSON*) *Soit f une fonction définie de $\mathbb{R}^n$ vers $\overline{\mathbb{R}}$ alors :*

(1) *Cas convexe* :

$$f\left(\sum_{i=1}^{n}\lambda_i x_i\right) \leq \sum_{i=1}^{n}\lambda_i f\left(x_i\right) \tag{1.36}$$

(2) *Cas qausi-convexe :*

$$f\left(\sum_{i=1}^{n}\lambda_i x_i\right) \leq \max_{i=\overline{1,n}}\left\{f\left(x_i\right)\right\} \tag{1.37}$$

L'inconvénient de la quasi-convexité **:** Elle est insuffisante pour globaliser un minimum local, mais par contre si elle est stricte.

Définition 1.29 : La fonction indicatrice d'un ensemble $C \subseteq \mathbb{R}^n$ est définie par :

$$\Psi_c\left(x\right) = \begin{cases} 0 \quad \text{si} \quad x \in C \\ +\infty \quad \text{ailleurs} \end{cases} \tag{1.38}$$

Définition 1.30 : La régularisée de $f : \mathbb{R}^n \longrightarrow \overline{\mathbb{R}}$ est par définition la plus grande minorante (sci) de f, et on le note par $cl\left(f\right)$ ou $\overline{f}$.

Définition 1.31 : Soit $Q_{n\times n}$ une matrice réel, on dit que cette matrice est *indéfinie* si et seulement si elle admet au moins une valeur propre négative et une valeur propre positive.

Définition 1.32 : $Q_{n\times n}$ est dite *définie positive* si tous les valeurs propres de cette matrice sont positives.

Définition 1.33 : Soit $C \subseteq \mathbb{R}^n$, *le polaire de* C est l'ensemble $C^\circ \subset \mathbb{R}^n$ définie par :

$$C^\circ = \left\{y \in \mathbb{R}^n ; \langle y, x\rangle \leq 1, \text{pour tout } x \in C\right\} \tag{1.39}$$

Définition 1.34 : Un hyperplan est un ensemble de la forme :

$$H = \{x \in \mathbb{R}^n : f(x) = \alpha\} \tag{1.40}$$

où f est une fonction linéaire continue défini sur $\mathbb{R}^n$ et $\alpha \in \mathbb{R}$.

Définition 1.35 : Un polytope est l'enveloppe convexe d'un sous-ensemble finie de $\mathbb{R}^n$ noté par :

$$conv\,\{x_1, x_2, \cdots, x_{k+1}\}$$

si de plus, on a les vecteurs :

$$\left\{x_i - x_1, i = \overline{2, k+1}\right\}$$

Sont linéairement indépendantes, alors $conv\,\{x_1, x_2, \cdots, x_{k+1}\}$ est appelé un *k-simplexe* de sommets $\{x_i\}_{i=\overline{2,k+1}}$.

Remarque 1.8 : Un polytope est toujours fermé et borné (compact).

Remarque 1.9 : Le nombre des vecteurs linéairement indépendants dans $\mathbb{R}^n$, étant au plus que (n), il ne peut y avoir de simplexe dans $\mathbb{R}^n$ possède plus que $(n+1)$ sommets.

Remarque 1.10 : Dans le cas d'un "*k-simplexe* " le point $x_\circ$définie par :

$$x_\circ = \sum_{i=1}^{k+1} \lambda_i x_i \text{ avec } \lambda_i = \frac{1}{k+1}; \forall i \tag{1.41}$$

est le "*barycentre*" du simplexe.

Définition 1.36 : Soit C un convexe non vide de $\mathbb{R}^n$ et $x \in C$, le point x est dit extrémal (sommet) s'il n'est pas à l'intérieur d'un segment de droite contenu dans C, autrement dite :

$$\forall x_1, x_2 \in C, \forall \lambda \in]0,1[: x \in \lambda x_1 + (1-\lambda)\, x_2 \Longrightarrow x = x_1 \text{ ou } x = x_2 \tag{1.42}$$

Exemple 1.1 : Soit l'ensemble $C \subseteq \mathbb{R}^2$ définie par :

$$C = \left\{x \in \mathbb{R}^2 \ / \ x_1^2 + x_2^2 \leq 2, -x_1 + 2x_2 \leq 2, x_1 \geq 0, x_2 \geq 0\right\}$$

donc l'ensemble des sommets noté par :

$$E_c = \left\{ (0,0)^T, (2,0)^T, \left(\frac{2}{3}, \frac{4}{3} \right), (0,1)^T \right\}$$

Notons içi que $C = CO(E_c)$, par conséquent, C est un polytope mais ce n'est pas un simplexe.

Lemme 1.1 :

$$C \text{ simplexe} \Longrightarrow C \text{ polytope} \Longrightarrow C \text{ compacte}$$

$$C \text{ simplexe} \Longrightarrow C \text{ possède au plus } (n+1) \text{ sommets}$$

1.3 Les cônes :

Définition 1.37 : Un sous ensemble $K \subseteq \mathbb{R}^n$ est appelé "*cône*" si :

$$\mathbb{R}_+^* K \subseteq K$$

i,e :

$$\forall x \in K, \forall \lambda \in \mathbb{R}_+^* : \lambda x \in K \tag{1.43}$$

c'est la réunion des demi-droites passant par l'origine.

Définition 1.38 : Si :

$$K \cap \{-K\} = \{0\}$$

alors, K est dite "pointé"ou"saillant".

Définition 1.39 : Soit C une partie non vide de $\mathbb{R}^n$, on appelle "*cône polaire*" de C et on le note par C^* le cône convexe fermé suivant :

$$C^* = \{x \in \mathbb{R}^n \ / \ \langle x^*, x \rangle \leq 0, \forall x \in C\} \tag{1.44}$$

Remarque 1.11 : La notion de polarité est dans un sens d'une notion de *complé-*

mentarité d'ailleurs, on peut écrit pour tout cône K que :

$$K + K^* = \mathbb{R}^n$$

Définition 1.40 : Soit $C \subseteq \mathbb{R}^n$, on appelle "cône dual" ou "cône conjugué" de C l'ensemble non vide C^+ définie par :

$$C^+ = \{y \in \mathbb{R}^n \ / \ \langle y, x \rangle \geq 0, \forall x \in C\} \tag{1.45}$$

Lemme 1.2 : C^+ est un cône convexe fermé.

Définition 1.41 : Un cône K est dite polyédrique s'il est engendré par un nombre finie de vecteurs $a_i \in \mathbb{R}^n$ alors il va de soi que :

$$K = \left\{ x = \sum_{i=1}^{s} \lambda_i a_i \ / \ \lambda_i \geq 0; i = \overline{1, s} \right\} \tag{1.46}$$

Définition 1.42 : Soit C une partie non vide de $\mathbb{R}^n$; alors un vecteur $\overline{x} \in \mathbb{R}^n$ est dit "normal" à C au point $a \in C$ si :

$$\langle x - a, \overline{x} \rangle \leq 0; \forall x \in C \tag{1.47}$$

ou d'une manière équivalente si :

$$\langle a, \overline{x} \rangle = \sup_{x \in C} \langle x, \overline{x} \rangle$$

et on désigne par $N(C, a)$ l'ensemble des vecteurs normaux à C en a.

Définition 1.43 : Soit C une partie non vide de $\mathbb{R}^n$, un vecteur $d \in \mathbb{R}^n$ est dite "tangent à C"en a s'il existe une suite $\{x_k\}$ d'éléments de C converge vers a, et une suite $\{\lambda_k\}$ des réels strictement positifs telle que :

$$d = \lim \lambda_k (x_k - a)$$

on peut écrit :

$$a \in C \subseteq \mathbb{R}^n \Longrightarrow T(C,a) = \{d \in \mathbb{R}^n \ / \ \exists \overline{\alpha} \succ 0 : a + \alpha d \in C, \forall \alpha \in [0, \overline{\alpha}]\} \qquad (1.48)$$

où $T(C,a)$ est un cône fermé non vide $(0 \in T(C\ ,a))$ appelé "cône tangent à C en a", ou "*cône des directions admissibles*".

La définition peut être donner d'autre façon comme suit :

Définition 1.44 : Soit l'ensemble $C \subseteq \mathbb{R}^n$ et $x^\circ \in \mathbb{R}^n$, on dira que $d \in \mathbb{R}^n$est *tangent* à C en x° s'il existe deux suites $\{d_k\}_k \subset \mathbb{R}^n$ et $\{t_k\}_k \subset \mathbb{R}^n_+$ telle que :

$$d_k \longrightarrow d \text{ et } t_k \longrightarrow 0^+ \text{ avec } (x^\circ + t_k d_k) \in C \qquad (1.49)$$

L'ensemble des vecteurs tangents à C en x° noté par $T_{x^\circ}C$ est *le cône tangent* de l'ensemble C en x°.

Proposition 1.9 : *Le cône tangent est un cône fermé ; il est convexe si C est convexe.*

1.3.1 Les conditions d'optimalités :

Les conditions nécessaires d'optimalités d'ordre un :

Considérons le problème d'optimisation suivant :

$$\begin{cases} \min f(x) \\ x \in C \end{cases} \qquad (P)$$

où $C \subseteq \mathbb{R}^n$.

On a le théorème très important suivant :

Théorème 1.1 :[4] *Si x^* est un minimum local de f sur C, et si f est dérivable en x^* alors, on a :*

$$\nabla f(x^*) \in (T_{x^\circ}C)^*$$

où $(T_{x^\circ}C)^*$ désigne le dual de $T_{x^\circ}C$ ce qui se traduit par :

$$(\nabla f(x^*))^T d \geq 0; \forall d \in T_{x^\circ}C$$

Si $C \subseteq \mathbb{R}^n$ est donnée par :

$$C = \{x \in \mathbb{R}^n : g_i(x_i) = 0 \text{ si } i \in A, \text{ et } g_i(x) > 0 \text{ si } i \in B\} \tag{1.50}$$

donc, on a le problème d'optimisation suivant :

$$\begin{cases} \min f(x) \\ x \in C \end{cases}$$

où :

$$A = \{1, 2, 3, \cdots, m_1\} \text{ et } B = \{1, 2, 3, \cdots, m_2\}$$

sont deux ensembles d'indices présentent une partition de l'ensemble $\{1, 2, 3, \cdots, m\}$, et :

$$f : \mathbb{R}^n \longrightarrow \overline{\mathbb{R}} \text{ et } g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

l'ensemble C est appelé *l'ensemble admissible* de*(P)*, et on peut le donné par :

$$C = \{x \in \mathbb{R}^n : g_B(x) \geq 0, g_A(x) = 0\}$$

où g_B est un vecteur de $\mathbb{R}^{m_2}$ formé de g_i telle que $i \in B$, et g_A est le vecteur de $\mathbb{R}^{m_1}$ formé de g_i où $i \in A$, posons aussi :

$$I^\circ(x) = \{i \in B : g_i(x) = 0\}$$

l'ensemble des indices où les contraintes d'inégalités sont *actives* en x.

Définition 1.45 :[4] On appelle *cône linéairisant* de l'ensemble C en $x^* \in C$ l'ensemble :

$$(T'_{x^\circ}C) = \left\{d \in \mathbb{R}^m : \nabla g_A(x^*)^T d = 0, \nabla g_{I^\circ}(x^*)^T d \geq 0\right\} \tag{1.51}$$

Proposition 1.2 :[4]

(1) *Si $x^* \in C$ et si $g_{A \cup I^\circ(x^*)}$ est une fonction dérivable en x^*, alors, on a l'inclusion* *suivant :*

$$T_{x^\circ} \subset T^{'}_{x^\circ}$$

(2) *Si on a l'égalité :*

$$T_{x^\circ} = T'_{x^\circ}$$

alors, les contraintes de (P) sont qualifiées en x^*

le théorème suivant présente des conditions nécessaires d'optimalités pour le problème *(P).*

Théorème 1.2 :[4] *Soit* $\widetilde{x}$ *un minimum de (P), on suppose que* f *et* $g_{A\cup I^\circ(x^*)}$ *sont dérivables en* $\widetilde{x}$*, d'autre par, supposons que les contraintes* $(g_i)_i$ *soient qualifiées en* $\widetilde{x}$*, alors il existe* $\lambda^* \in \mathbb{R}^m_+$ *telle que :*

$$\begin{cases} \nabla f(\widetilde{x}) - A(\widetilde{x})^T \lambda^* = 0 \\ g_i(\widetilde{x}) = 0, \text{si } i \in A \\ g_i(\widetilde{x}) \geq 0, \text{si } i \in B \\ (\lambda_i^*)^T g_i(\widetilde{x}) = 0 \end{cases}$$

où : $A(\widetilde{x})$ *désigne la Jacobienne de* g *en* $\widetilde{x}$.

Remarque 1.12 : Si la fonction objectif de *(P)* est *convexe*, $g_A$est affine, et g_B est convexe, alors, les conditions nécessaires ci-dessus devient *des conditions suffisantes,* pour les quelles $\widetilde{x}$ être *le minimum (global) de (P).*

Application sur les problèmes quadratiques :

Soit le problème quadratique défini sous la forme suivante :

$$\begin{cases} \min f(x) = \frac{1}{2}x^T Qx + d^T x \\ Bx \leq c \text{ et } x \geq 0 \end{cases} \qquad (PQ)$$

soit :

$$\widetilde{x} \in S = \{x \in \mathbb{R}^n, Bx \leq c \text{ et } x \geq 0\}$$

posons :

$$\begin{cases} g_i(x) = B_i x - c_i \\ h_i(x) = -x_i \\ 1 \leq i \leq n \end{cases}$$

alors, le théorème de "*Karush-Kuhn-Tucker*" est donné par :

$$\widetilde{x} \text{ une solution optimale de } (P) \Longrightarrow \begin{cases} \exists \lambda \in \mathbb{R}^n_+, \exists \mu \in \mathbb{R}^n_+ : \nabla f(\widetilde{x}) + B^T \lambda - \mu = 0 \\ et \qquad B\widetilde{x} - c \leq 0 \\ et \qquad -\widetilde{x} \leq 0 \\ et \qquad \lambda^T (B\widetilde{x} - c) = 0 \\ et \qquad -\mu\widetilde{x} = 0 \end{cases}$$

$$\Longrightarrow \begin{cases} \exists \lambda \in \mathbb{R}^n_+, \exists \mu \in \mathbb{R}^n_+ : Q\widetilde{x} + d + B^T \lambda - \mu = 0 \\ et \qquad B\widetilde{x} - c \leq 0 \\ et \qquad \lambda^T (B\widetilde{x} - c) = 0 \\ et \qquad -\mu\widetilde{x} = 0 \\ et \qquad -\widetilde{x} \leq 0 \end{cases}$$

où : Q est une matrice *symétrique*, et les deux fonctions $g(.)$ et $h(.)$ sont les contraintes de *(P)*.

on peut poser des conditions nécessaires d'optimalités d'ordre deux pour le même problème (P) comme suit :

Les conditions nécessaires d'optimalités d'ordre deux : Considérant le même problème (P) avec les deux fonctions $g(.)$,$h(.)$ qui sont définées préçidament.

Définition 1.46 :[7] Le Lagrangien associée à (P) est donnée par :

$$l(x, \lambda, \mu) = f(x) + \lambda^T g(x) + \mu^T h(x)$$

d'après le théorème de *(K-K-T)*, la solution est nécessairement appartient à S, alors on peut écrit :

$$\nabla_\lambda l(\widetilde{x}, \lambda, \mu) \leq 0$$

$$\nabla_\mu l(\widetilde{x}, \lambda, \mu) \leq 0$$

où : ∇_{λ},∇_{μ}sont le gradient de la fonction *lagrangien* par rapport à μ et λ respectivement.

Lemme 1.3 :[7] Les conditions nécessaires de *(K-K-T)* prennent la forme :

$$\begin{cases} \nabla_x (\widetilde{x}, \lambda, \mu) = 0 \\ \lambda^T g(\widetilde{x}) = 0 \\ \mu^T h(\widetilde{x}) = 0 \end{cases}$$

où : ∇_xest *le gradient de la fonction lagrangien par rapport à* x.

Lemme 1.4 :[7] Si les fonctions $f, g, h \in C^2(w)$ où w est un voisinage de $\widetilde{x}$ (la solution locale), au plus de ça, si les contraintes sont *qualifiées* alors, *la hessianne* $\nabla_x^2 L(\widetilde{x}, \lambda, \mu)$ de $L(x, \lambda, \mu)$ est *semi définie positive,* et l'inégalité suivante est satisfaite :

$$y^T \nabla_x^2 L(\widetilde{x}, \lambda, \mu)\, y \geq 0, \forall y \in T(\widetilde{x}) \qquad (1.52)$$

où $T(\widetilde{x})$ est *le cône tangent en* $\widetilde{x}$.

Remarque 1.13 : La condition (1.56) est appelée *la condition nécessaire d'optimalité d'ordre deux .*

Chapitre 2

La programmation DC, la méthode DCA :

2.1 La programmation DC :

2.1.1 Les fonctions DC :

Afin d'étendre la programmation convexe à la résolution de la plus part des problèmes d'optimisations non convexes de la vie courante tout en continuant à utiliser son arsenal théorique et numérique, une nouvelle classe de fonctions fut introduite *la classe des fonctions DC*.

Définition 2.1 :[4] Soit C un ensemble convexe de X, une fonction $f : C \longrightarrow \mathbb{R} \cup \{+\infty\}$ définie sur C est dite *DC* sur$C$ si elle s'écrit sous la forme :

$$f(x) = g(x) - h(x), \forall x \in C$$

où g et h sont deux fonctions convexes sur C.

Définition 2.2 :[4] Soit C un ensemble convexe de $\mathbb{R}^n$, une fonction $f : C \longrightarrow \overline{\mathbb{R}}$ définie sur C est dite $C^2 - faible$ si f est un supremum local de la famille des fonctions de classe C^2.

Remarque 2.1 : $DC(X)$ est un sous espace contenant la classe des fonctions

C^2-*faible*, en particulier, contient l'espace $C^{1,1}(X)$.

Remarque 2.2 : $DC(X)$ est également l'espace vectoriel engendré par $conv(C)$, et sous une certaine précaution, on peut dire que $DC(X)$ est le sous espace généré par le cône convexe $\Gamma_\circ(X)$, et on écrit :

$$DC(X) = \Gamma_\circ(X) - \Gamma_\circ(X)$$

Proposition 2.1 :[11,12] *Grâce à la structure d'espace vectoriel de $DC(X)$, $DC(X)$ est également stable par rapport aux opérations usuelles utilisées en optimisation, on a donc :*

(*a*) *Une combinaison linéaire de fonctions DC sur C est une fonction DC sur C.*

(*b*) *L'enveloppe supérieure (resp :inférieure) d'un nombre finie de fonctions DC à valeurs finies sur C est une fonction DC sur C.*

(*c*) *Si f est une fonction DC à valeurs finies sur C alors les fonctions :*

$$|f|, f^+ = \max\{f, 0\}, f^- = \min\{f, 0\}$$

s*ont des fonctions DC sur C.*

Remarque 2.3 : Soit $f \in DC(X)$ et soit $f = g - h$ sa représentation DC, alors pour toute nombre finie ε on peut écrire :

$$f = (g + \varepsilon) - (h + \varepsilon)$$

qui définit une autre décomposition DC de f, ainsi, une fonction DC possède une infinité de décompositions DC.

Remarque 2.4 : Une sous classe importante de $conv(C)$ est la classe de fonctions C^2sur $\mathbb{R}^n$, cette classe particulièrement importante car les fonctions objectifs de la plupart des problèmes d'optimisations de la vie courante appartiennent à cette sous classe.

Définition 2.3 : Soit C un ouvert convexe, f est localement DC sur C si tout points

$x^{\circ} \in C$ admet un voisinage $\tilde{U}(x^{\circ})$ noté $\tilde{U}$, telle que, pour tout $x \in \tilde{U}$ on a :

$$f(x) = g_{\tilde{U}}(x) - h_{\tilde{U}}(x) \,; \text{pour tout } g_{\tilde{U}}, h_{\tilde{U}} \subset conv(C)$$

Proposition 2.2 :[11,8] *Les deux points (a) et (b) suivantes sont équivalentes :*

(a) *f est dans $LC^2(C)$.*

(b) *f est localement DC sur C avec $h_{\tilde{U}}$ une forme quadratique convexe.*

On a le théorème suivant *:*

Théorème 2.1 :[11,8] *Soit C un ouvert convexe de X, alors :*

(1) toute fonction localement DC sur C est DC sur C. Particulièrement, toute fonction de classe C^2sur C est DC sur C.

(2) on a la chaîne d'inclusion suivante :

$$conv(C) \subset LC^2(C) \subset DC(C) \tag{2.1}$$

(3) si C est en plus compact; alors $DC(C)$ est un sous espace vectoriel dense dans l'ensemble des fonctions continue sur C muni de la norme de la convergence uniforme sur C.

2.1.2 Dualité en optimisation DC :

Définition 2.4 : Un programme DC est un problème de la forme :

$$\alpha = \begin{cases} \inf f(x) = g(x) - h(x) \\ \qquad x \in X \end{cases} \tag{P_{dc}}$$

où $g(.)$,$h(.)$ sont dans $\Gamma_{\circ}(\mathbb{R}^n)$, une telle fonction DC sur X, telle que $g(.)$,$h(.)$ sont appelées ses composantes DC.

–Puisque $h \in \Gamma_{\circ}(\mathbb{R}^n)$ on a $h^{**} = h$, on peut donc écrire :

$$h(x) = \sup\{\langle x, y\rangle - h^*(y) \;\; / \; y \in Y\} \tag{2.2}$$

on utilisant la définition des fonctions conjugués, alors, on a :

$$\begin{aligned}
\alpha &= \inf \{f(x) = g(x) - h(x) : \ x \in X\} \\
&= \inf \{g(x) - \sup \{\langle x, y \rangle - h^*(y) \ / \ y \in Y\} ; \text{tq } x \in X\} \\
&= \inf \{ß(y) \ / \ y \in Y\}
\end{aligned}$$

où :

$$ß(y) = \inf \{g(x) - [\langle x, y \rangle - h^*(y)] \ / \ x \in X\} \qquad (P_{\mathrm{y}1})$$

Lemme 2.1 :[4]

Il va de soi que (P_{y}) est un problème convexe.

Preuve :

h est une fonction convexe se qui implique que h^*est une fonction convexe, d'autre part, la fonction $g(x)$ est aussi une fonction convexe donc (P_{y}) est un problème convexe. ■

D'autre part, on a :

$$\begin{aligned}
ß(y) &= \inf \{g(x) - [\langle x, y \rangle - h^*(y)] \ / \ x \in X\} \\
&= h^*(y) + \inf \{g(x) - \langle x, y \rangle \ / \ x \in X\} \\
&= h^*(y) + \inf \{-(\langle x, y \rangle - g(x)) \ / \ x \in X\} \\
&= h^*(y) - \sup \{\langle x, y \rangle - g(x) \ / \ x \in X\} \\
&= h^*(y) - g^*(y)
\end{aligned}$$

donc, il est claire que :

$$ß(y) = \begin{cases} h^*(y) - g^*(y) & \text{si } y \in dom(h^*) \\ +\infty & \text{ailleurs} \end{cases} \qquad (P_{\mathrm{y}2})$$

On obtient finalement le problème dual de (P_{dc}) :

$$\alpha = \inf \{h^*(y) - g^*(y) \ / \ y \in dom(h^*)\}$$

et grâce à la convention de *Toland* :

$$(+\infty) - (+\infty) = (+\infty)$$

on peut alors écrire :

$$\alpha = \inf \{h^* (y) - g^* (y) \ / \ y \in Y\} \tag{D_{dc}}$$

Remarque 2.5 :

$[\alpha \textit{est définit }] \Longrightarrow \begin{bmatrix} dom\,(h^*) \subset dom\,(g^*) \\ \text{et } dom\,(g) \subset dom\,(h) \end{bmatrix}.$

Remarque 2.6 :

On observe la parfaite symétrie entre le problème primal (P_{dc}) et le problème dual (D_{dc}) ; il va de soi que les résultats établis pour l'un se transposent à l'autre.

2.1.3 Les conditions d'optimalités en optimisation DC :

Définition 2.5 :[4] On note par S_p et S_D les ensembles de solutions des problèmes (P_{dc}) et (D_{dc}) respectivement, et on définie :

$$\begin{aligned} p_L &= \{x^* \in X \ / \ \partial h\,(x^*) \subset \partial g\,(x^*)\} \qquad (2.3) \\ D_L &= \{y^* \in Y \ / \ \partial g^*\,(y^*) \subset \partial h^*\,(y^*)\} \end{aligned}$$

–Le théorème suivant est celui à partir duquel *DCA est bien définie,* et avant ça, on a besoin des définitions suivantes :

Définition 2.6 : On dit qu'une point x^*est un point critique de la fonction $(g - h)$ si :

$$\partial h\,(x^*) \cap \partial g\,(x^*) \neq \phi \tag{2.4}$$

Définition 2.7 : On définie la relation $\underline{*}$ par :

$$A \underline{*} B = \{x \in X \ / \ x + B \subset A\} \tag{2.5}$$

Théorème 2.2 :[4]

(i) Transport de minima globaux *:*

$$\cup \{\partial h(x) \ / \ x \in S_p\} \subset S_D \subset dom(h^*) \tag{2.6}$$

la première inclusion se transforme en égalité si g^* est sous différentiable sur S_D, d'autre façon, si :

$$S_D \subset ir(dom(g^*))$$

ou si g^* est sous–différentiable sur $dom(h^*)$ et on écrit :

$$S_D \subset (dom(\partial g^*) \cap dom(\partial h^*)) \tag{2.7}$$

(ii) *[Si x^*est un minimum local de$(g-h)$]$\Longrightarrow$[$x^* \in p_L$].*

(iii) *[*Si *$x^* \in p_L$ et h est une fonction convexe polyédrale]$\Longrightarrow$[x^*est un minimum local de$(g-h)$]*

Soit x^* un point critique de $(g-h)$ et :

$$y^* \in \partial g(x^*) \cap \partial h(x^*) \tag{2.8}$$

et soit $\hat{U}$ un voisinage de x^* telle que :

$$\hat{U} \cap (dom(g)) \subset dom(h) \subset dom(\partial h) \tag{2.9}$$

si pour tout :

$$x \in \hat{U} \cap (dom(g))$$

il existe un $y \in \partial h(x)$ telle que :

$$h^*(y) - g^*(y) \geq h^*(y^*) - g^*(y^*) \tag{2.10}$$

alors x^* est un minimum de f, plus précisément :

$$g(x) - h(x) \geq g(x^*) - h(x^*) \text{ ; pour tout } x \in \hat{U} \cap (dom(g)) \tag{2.11}$$

(v) Transport de minima locaux *: Soit* $x^* \in dom\,(\partial h)$ *un minimum local de* $f = g - h$*, et soit* $y^* \in \partial h\,(x^*)$*donc ; si on suppose que :*

$$y^* \in int\,(dom\,(g^*)) \text{ et } \partial g^*\,(y^*) \subset \hat{U} \tag{2.12}$$

alors, si g^**est différentiable en*y^* *; alors* y^**est un minimum local de* $(h^* - g^*)$.

Remarque 2.7 :

(a) Ce théorème admet sa forme duale grâce à la symétrie observée précidament entre les deux problèmes (P_{dc}) et (D_{dc}).

(b) Notons également que tous ces résultats concernent les composantes DC g et h, et non la fonction f elle même.

Proposition 2.3 :[4]

Soit $f = g - h$ *avec* $g, h \in \Gamma_\circ\,(\mathbb{R}^n)$ *vérifiée :*

$$\begin{aligned} dom\,(g) &\subset dom\,(h) \\ ir\,(dom\,(g)) \cap ir\,(dom\,(h)) &\neq \phi \end{aligned} \tag{2.13}$$

soit $x_\circ \in dom\,(g)$*(resp :* $dom\,(h)$*) un point où la fonction* g *(resp :* h*) est continue, donc, si* f *est convexe alors :*

(1) h *est continue en* $x_\circ$.

(2)

$$\partial h\,(x_\circ) = \partial g\,(x_\circ)\,\underline{*}\,\partial h\,(x_\circ) \tag{2.14}$$

(3)

$$0 \in \partial f\,(x_\circ) \Longleftrightarrow \partial h\,(x_\circ) \subset \partial g\,(x_\circ) \tag{2.15}$$

Preuve : On a :

$$\alpha = \inf\{f\,(x) = g\,(x) - h\,(x) : x \in X\}$$

est bien définie si et seulement si :

$$dom\,(g) \subset dom\,(h)$$

alors :

$$dom\,(f) = dom\,(g-h) = dom\,(g)$$

d'autre part on a :

$$g = h + f$$

donc :

$$\partial g\,(x_\circ) = \partial\,(h+f)\,(x_\circ)$$

mais on a par hypothèse :

$$ir\,(dom\,(g)) \cap ir\,(dom\,(h)) \neq \phi$$

alors :

$$\partial f\,(x_\circ) = \partial\,(g-h)\,(x_\circ) = \partial g\,(x_\circ) - \partial h\,(x_\circ)$$

donc, on peut écrire :

$$\partial g\,(x_\circ) = \partial f\,(x_\circ) + \partial h\,(x_\circ)$$

alors, si g est continue en $x_\circ \in dom\,(g)$,on a :

$$x_\circ \in ir\,(dom\,(g)) = int\,(dom\,(g)) \subset int\,(dom\,(h))$$

ainsi l'ensemble convexe fermé $\partial h\,(x_\circ)$ est borné, on a :

$$\partial f\,(x_\circ) = \partial g\,(x_\circ) \underline{*} \partial h\,(x_\circ) = \cap \{\partial g\,(x_\circ) - \nu \;/\; \nu \in \partial h\,(x_\circ)\}$$

la troisième équivalence est immédiate. ■

Remarque 2.8 : On observe que les problèmes d'optimisations convexes peuvent s'écrire comme problème d'optimisation DC et résolus en utilisant les techniques d'optimisation DC.

Exemple 2.1 : Si on prend $f, \Psi \in \Gamma_\circ(X)$ telle que :

$$dom\,(f) \subset dom\,(\Psi)$$

et :

$$ir\,(dom\,(f)) \cap ir\,(dom\,(\Psi)) \neq \phi$$

alors, le problème d'optimisation :

$$\inf\{f(x)\,;x \in X\}$$

est équivalent au faux problème d'optimisation DC :

$$\inf\{(f+\Psi) - \Psi(x) \ \ / \ x \in X\}$$

dans le sens où ils ont la même valeur optimale et le même ensemble des solutions ; et on a aussi :

$$0 \in \partial f(\widetilde{x}) \Longleftrightarrow \partial\Psi(\widetilde{x}) \subset \partial(f+\Psi)(\widetilde{x}) = \partial f(\widetilde{x}) + \partial\Psi(\widetilde{x})$$

et l'autre implication est vraie si en plus, Ψ est continue en $\widetilde{x}$.

Proposition 2.4 :[12]

Soit :

$$f = g - h \text{ avec } g, h \in \Gamma_\circ(X)$$

vérifiant :

$$dom\,(g) \subset dom\,(h)$$

s'il existe un voisinage convexe $\hat{U}$ de x^* telle que f est finie et convexe sur$\hat{U} \subseteq X$, alors les deux relations suivantes sont équivalentes :

(i)

$$0 \in \partial(f + \chi_{\hat{U}})(x^*) \tag{2.16}$$

(ii)

$$\partial h\left(x^*\right) \subset \partial g\left(x^*\right) \qquad (2.17)$$

Théorème 2.3 :[12, 16]

Soit $f = g - h$ où $g, h \in \Gamma_\circ(X)$ alors, x^* est un minimum global de problème (P_{dc}) si et seulement si :

$$\partial_\varepsilon h\left(x^*\right) \subset \partial_\varepsilon g\left(x^*\right), \text{ pour tout } \varepsilon > 0 \qquad (2.18)$$

2.2 La méthode DCA :

La construction de *DCA* basé sur la génération de deux suites $(x_k)_k$ et $(y_k)_k$ qui sont améliorées à chaque itération de sorte que leurs limites x^* et y^* respectivement soient candidates pour être les optima locaux du (P_{dc}) et (D_{dc}) respectivement.

2.2.1 Description de la méthode :

Pour tout $x^* \in \mathbb{R}^n$ on considère le problème :

$$\inf\left\{h^*(y) - g^*(y)\,; y \in \partial h\left(x^*\right)\right\} \qquad (\mathrm{S}(x^*))$$

qu'est équivalent à la maximisation convexe de :

$$\inf\left\{\langle x^*, y\rangle - g^*(y); y \in \partial h\left(x^*\right)\right\}$$

de même pour $y^* \in \mathbb{R}^n$ fixé, on définit le problème suivant pour la dualité :

$$\inf\left\{g(x) - h(x); x \in \partial g^*(y^*)\right\} \qquad (\mathrm{T}(y^*))$$

Ce problème est équivalent à :

$$\inf\left\{\langle x, y^*\rangle - h(x); x \in \partial g^*(y^*)\right\}$$

On note par $\mathrm{S}(x^*)$, $\mathrm{T}(y^*)$ respectivement les ensembles des solutions des problèmes $(\mathrm{S}(x^*))$ et $(\mathrm{T}(y^*))$.

La forme complète de l'algorithme DC est basée sur la dualité de l'optimisation définie par (P_{dc}) et $(\mathrm{D}_{\mathrm{dc}})$, qui permet l'approximation du point :

$$(x^*, y^*) \in P_L \times D_L \tag{2.19}$$

Etant donné un point $x^\circ \in dom\,(g)$, l'algorithme construit deux suites $(x_k)_k$ et $(y_k)_k$ définie par :

$$y_k \in S\,(x_k)\,; x_{k+1} \in T\,(y_k) \tag{2.20}$$

d'autre part, l'algorithme DC complet peut être veut (considère) comme un approche de la décomposition des deux problèmes *(*P_{dc}*)* et *(*D_{dc}*)*, et en pratique, la forme suivante de l'algorithme DC simplifie est utile.

Simplification de l'algorithme :

On a déjà vue que *(DCA)* est construit deux suites $(x_k)_k$ et $(y_k)_k$, ces deux suites sont les condidats aux solution primale et duale qui sont faciles a calculer, ces deux suites sont liées par la dualité et vérifient les propriétés suivantes :

(a) les suites :

$$\{g\,(x_k) - h\,(x_k)\} \text{ et } \{g^*\,(y_k) - h^*\,(y_k)\} \tag{2.21}$$

sont décroissantes à chaque itération.

(b) si :

$$(g - h)\,(x_{k+1}) = (g - h)\,(x_k) \tag{2.22}$$

l'algorithme s'arrête à l'itération $(k+1)$ et le point x_k (resp y_k) est un point critique de $(g - h)$ (resp $(h^* - g^*)$).

(c) sinon, toute valeur d'adhérence x^*de la suite $(x_k)_k$ (resp y^* de la suite $(y_k)_k$) est un point critique de la fonction $(g - h)$ (resp $(h^* - g^*)$).

Ces suites sont générées de la manière suivante :

x_{k+1} (resp y_k) est une solution du problème convexe *(*P_{k}*)* (resp *(*D_{k}*)*) définie par :

$$\alpha_k = \inf \{g(x) - [h(x_k) + \langle x - x_k, y_k \rangle] \ / \ x \in X\} \tag{P_{k}}$$

$$\beta_k = \inf \{h^*(y) - [g^*(y_{k-1}) + \langle x_k, y - y_{k-1} \rangle] \ / \ y \in Y\} \tag{D_{k}}$$

Construction du problème (P_{k}*)* **:**

Le problème *(*P_{k}*)* est un problème convexe obtenu du problème (P_{dc}) où on remplaçant la fonction $h(x)$ par sa minorante affine définie par $y_k \in \partial h(x_k)$ comme suit :

$$\inf \{f(x) = g(x) - h(x); x \in X\} \tag{P_{dc}}$$

et le minorante affine de la fonction h en $y_k \in \partial h(x_k)$ est donné par :

$$aff_{x_k}(h) = h(x_k) + \langle y_k, x - x_k \rangle \tag{2.23}$$

grâce à la définition suivant :

$$y_k \in \partial h(x_k) \Longrightarrow h(x) - h(x_k) \geq \langle y_k, x - x_k \rangle$$

sa implique que :

$$h(x) \geq h(x_k) + \langle y_k, x - x_k \rangle \tag{2.24}$$

et à cause de ça, le problème (P_{dc}) prend la forme *(*P_{k}*)*.

Construction du problème (D_{k}*) :*

La même chose que la construction de *(*P_{k}*)* où on remplaçant $g^*(.)$ par sa minorante affine continue en $x_k \in \partial g^*(y_{k-1})$.

Remarque 2.9 : On peut conclue le schéma de *(DCA)* suivant :

$$\boxed{\mathrm{x}_k \longrightarrow \mathrm{y}_k \in \partial \mathrm{h}(x_k)}$$

$$\boxed{\swarrow}$$

$$\boxed{x_{k+1} \in \partial g^*\left(y_k\right) \longrightarrow y_{k+1} \in \partial h x_{k+1}}$$

$$\boxed{\swarrow}$$

$$\boxed{x_{k+2} \in \partial g^*\left(y_{k+1}\right) \longrightarrow y_{k+2} \in \partial h\left(x_{k+2}\right)}$$

$$\boxed{\swarrow}$$

$$\boxed{\cdots}$$

d'après ce schéma, on peut poser l'algorithme suivant :

Algorithme :

cas :0 $x_\circ$ donné.

cas :1 Pour chaque k, trouver x_k la solution du problème convexe *($P_{\text{k-1}}$)*.

cas :2 Déterminer $y_k \in \partial h\left(x_k\right)$.

cas :3 Trouver $x_{k+1} \in \partial g^*\left(y_k\right)$.

cas :4 Si test d'arrêt vérifier *stop*, sinon : $k \longleftarrow k+1$ et *return* au *cas :1.*

Corollaire 2.1 : Dans la forme complète de l'algorithme DC on a :

$$x_{k+1} \in \arg\min\left\{g(x) - h(x); x \in \partial g^*\left(y_k\right)\right\} \tag{2.25}$$

$$y_k \in \arg\min\left\{h^*(y) - g^*(y); y \in \partial h\left(x_k\right)\right\} \tag{2.26}$$

Preuve : Puisque x_{k+1} (resp y_k) est une solution de problème convexe *(P_{k})* (resp *(D_{k})*) définie par :

$$\alpha_k = \inf\left\{g(x) - \left[h(x_k) + \langle x - x_k, y_k\rangle\right] \ / \ x \in \partial g^*\left(y_k\right)\right\} \tag{P$_\text{k}$}$$

$$\beta_k = \inf\left\{h^*(y) - \left[g^*(y_{k-1}) + \langle x_k, y - y_{k-1}\rangle\right] \ / \ y \in \partial h\left(x_k\right)\right\} \tag{D$_\text{k}$}$$

donc :

$$\inf \{g(x) - h(x); x \in \partial g^* (y_k)\} = g(x_{k+1}) - h(x_{k+1})$$

d'autre part on a :

$$\arg\min \{g(x) - h(x); x \in \partial g^* (y_k)\} = \{x^* \in \partial g^* (y_k) ; g(x^*) - h(x^*) = \inf \{g(x) - h(x); x \in \partial g^* (y_k)\}\}$$

Et puisque :

$$x_{k+1} \in T(y_k)$$

l'ensemble des solutions de problème convexe (P_{k}) alors :

$$x_{k+1} \in \arg\min \{g(x) - h(x); x \in \partial g^* (y_k)\}$$

c'est la même chose pour trouver que :

$$y_k \in \arg\min \{h^*(y) - g^*(y); y \in \partial h (x_k)\}$$

donc, la preuve est complète. ■

Lemme 2.1 : Les deux problèmes précédents sont équivalents aux problèmes suivants (respectivement) :

$$x_{k+1} \in \arg\min \{\langle x, y_k \rangle - h(x); x \in \partial g^* (y_k)\} \tag{2.27}$$

$$y_k \in \arg\min \left\{\left\langle x^k, y \right\rangle - g^*(y); y \in \partial h (x_k)\right\} \tag{2.28}$$

2.2.2 L'existence et bornitude des suites générée par (DCA) :

(DCA) sera bien définie si on peut construire les suites $(x_k)_k$ et $(y_k)_k$ à partir d'un point arbitraire $x_\circ \in X$, donc on peut poser le lemme suivant :

Lemme 2.2 :[**12, 10**] Les relations suivantes sont équivalentes :

(1) Les suites $(x_k)_k$ et $(y_k)_k$ sont bien définies.

(2)

$$dom(\partial g) \subset dom(\partial h) \text{ et } dom(\partial h^*) \subset dom(\partial g^*) \qquad (2.29)$$

Preuve : Supposons que les suites $(x_k)_k$ et $(y_k)_k$ sont bien définies, alors (DCA) est bâti (bien définie), donc, les deux problèmes (P_{k}) et (D_{k}) sont bien posées si et seulement si :

$$dom(\partial g) \subset dom(\partial h) \text{ et } dom(\partial h^*) \subset dom(\partial g^*)$$

alors, la preuve est complète. ■

Les conditions pour les quelles les suites générées par (DCA) sont bornées, sont résumées dans la proposition suivante :

Proposition 2.5 :[12, 10]

(1) *Si* $(g - h)$ *est une fonction coercive alors, la suite* $(x_k)_k$ *est bornée, et si on a :*

$$(x_k)_k \subset int(dom(h)) \qquad (2.30)$$

alors, la suite $(y_k)_k$ *est aussi bornée.*

(2) *Si* $(h^* - g^*)$ *est coercive alors, la suite* $(y_k)_k$ *est bornée, et si on a :*

$$(y_k)_k \subset int(dom(g^*)) \qquad (2.31)$$

alors, la suite $(x_k)_k$ *est aussi bornée.*

Le théorème suivant établit la convergence de (DCA) simplifier, mais on a besoins de la définition suivante :

Définition 2.6 :[15]

Soit la fonction fortement convexe $f : X \longrightarrow \overline{\mathbb{R}}$, le nombre ρ définie par :

$$\rho(f, C) = \sup\left\{\rho \geq 0 : f - \left(\frac{\rho}{2}\right)\|.\|^2 \text{ est convexe sur } C\right\} \qquad (2.32)$$

est appelée le module de la convexité forte de la fonction f sur C, et on le note par :

$$\rho(f,C) \text{ ou } \{\rho(f) \text{ si } C = X\}$$

Lemme 2.1 :[15] On dit qu'une fonction $f : X \longrightarrow \overline{\mathbb{R}}$ est fortement convexe sur $C \subseteq \mathbb{R}^n$ si $\rho(f,C) > 0$.

Théorème 2.4 :[15]

Soient :

$$f = f_1 - f_2,\ dx_k = x_{k+1} - x_k \text{ et } dy_k = y_{k+1} - y_k$$

alors, on a :

(1) *Soit* $\{\rho_i \text{ et } \rho_i^*; i = 1,2\}$ *des nombres réels non négatives telle que :*

$$0 \leq \rho_i \leq \rho(f_i)\,;\ \text{et } 0 \leq \rho_i^* \leq \rho(f_i^*)\,; i = 1,2$$

où :

$$\begin{aligned}\rho_i &= 0 \text{ si } \rho(f_i) = 0\\ \rho_i^* &= 0 \text{ si } \rho(f_i^*) = 0\end{aligned}$$

(2) *Si les deux suites* $(x_k)_k$ *et* $(y_k)_k$ *sont bien définies alors, on a :*

$$(f_1 - f_2)(x_{k+1}) \leq (f_2^* - f_1^*)(y_k) - \delta_1 \leq (f_1 - f_2)(x_k) - \delta_2 \tag{2.33}$$

avec :

$$\begin{aligned}\delta_1 &= \max\left\{\frac{\rho_2}{2}\|dx_k\|^2, \frac{\rho_2^*}{2}\|dy_k\|^2\right\}\\ \delta_2 &= \max\left\{\frac{\rho_1+\rho_2}{2}\|dx_k\|^2, \frac{\rho_1^*}{2}\|dy_{k-1}\|^2 + \frac{\rho_2}{2}\|dx_k\|^2, \frac{\rho_1^*}{2}\|dy_{k-1}\|^2 + \frac{\rho_2^*}{2}\|dy_k\|^2\right\}\end{aligned}$$

(3) *On a :*

$$(f_1 - f_2)(x_{k+1}) = (f_1 - f_2)(x_k) \tag{2.34}$$

si et seulement si :

$$x_k \in \partial f_1^*(y_k), y_k \in \partial f_2(x_{k+1}) \text{ et } (\rho_1 + \rho_2)\, dx_k = \rho_1^* dy_{k-1} = \rho_2^* dy_k = 0$$

et dans ce cas on a :

(a) $(f_1 - f_2)(x_{k+1}) = (f_2^* - f_1^*)(y_k)$.

(b) $\{x_k, x_{k+1}\}$ *sont des points critiques pour la fonction* $(f_1 - f_2)$ *avec :*

$$y_k \in (\partial f_1(x_k) \cap \partial f_2(x_k)) \cap (\partial g(x_{k+1}) \cap \partial h(x_{k+1})) \qquad (2.35)$$

(c) y_k *est un point critique pour la fonction* $(f_2^* - f_1^*)$ *et on a :*

$$[x_k, x_{k+1}] \subset (\partial f_1^*(y_k) \cap \partial f_2^*(y_k)) \qquad (2.36)$$

(d) *On a :*

$$x_{k+1} = x_k \text{ si } \rho(f_1) + \rho(f_2) > 0$$

$$y_k = y_{k-1} \text{ si } \rho(f_1^*) > 0$$

$$y_k = y_{k+1} \text{ si } \rho(f_2^*) > 0$$

(4) *Si* α *est finie alors les suites décroissantes* $\{(f_1 - f_2)(x_k)\}_k$ *et* $\{(f_2^* - f_1^*)(y_k)\}_k$ *sont convergentes et ont la même limite* $\beta \geq \alpha$*, c'est à dire que :*

$$\lim_{k \longrightarrow +\infty} (f_1 - f_2)(x_k) = \lim_{k \longrightarrow +\infty} (f_2^* - f_1^*)(y_k) = \beta \qquad (2.37)$$

d'autre part on a :

$$\begin{aligned} \text{si } \rho(f_1) + \rho(f_2) &> 0, \text{ alors } \lim_{k \longrightarrow +\infty} \{x_{k+1} - x_k\} = 0 && (2.38) \\ \text{et si } \rho(f_1^*) + \rho(f_2^*) &> 0, \text{ alors } \lim_{k \longrightarrow +\infty} \{y_{k+1} - y_k\} = 0 && (2.1) \end{aligned}$$

au plus de ça, on a :

$$\lim_{k \longrightarrow +\infty} \{f_1(x_k) + f_1^*(y_k) - \langle x_k, y_k \rangle\} = 0 \tag{2.39}$$

$$\text{et } \lim_{k \longrightarrow +\infty} \{f_2(x_{k+1}) + f_2^*(y_k) - \langle x_{k+1}, y_k \rangle\} = 0 \tag{2.2}$$

(5) *Si α est finie et si les suites $(x_k)_k$ et $(y_k)_k$ sont bornées alors pour toute valeur d'adhérence x_*de $(x_k)_k$ (resp y_* de $(y_k)_k$) il existe une valeur d'adhérence y_* de $(y_k)_k$ (resp x_*de $(x_k)_k$) telle que :*

$$y_* \in (\partial f_1(x_*) \cap \partial f_2(x_*)) \text{ et } x_* \in (\partial f_1^*(y_*) \cap \partial f_2^*(y_*)) \tag{2.40}$$

avec :

$$f_2^*(y_*) - f_1^*(y_*) = f_1(x_*) - f_2(x_*) = \beta$$

2.2.3 Comment redémarrer (DCA) pour obtenir : $x^* \in \partial f_2(x^*) \subset \partial f_1(x^*)$?

Si le *(DCA)* simplifier se termine à un point x^* pour le quelle :

$$\partial f_2(x^*) \subsetneqq \partial f_1(x^*)$$

alors, on peut réduire la valeur de la fonction objectif, en le redémarrant à partir d'un point initial $x^\circ = x^*$ avec :

$$y^\circ \in \partial f_2(x^\circ) \text{ telle que } y^\circ \notin \partial f_1(x^\circ)$$

en fait comme :

$$f_1(\widetilde{x}) + f_1^*(y^\circ) = \langle \widetilde{x}, x^\circ \rangle \leq f_2(\widetilde{x}) - f_2(x^\circ) + \langle x^\circ, y^\circ \rangle$$

et :

$$\langle x^\circ, y^\circ \rangle < f_1(x^\circ) + f_1^*(y^\circ)$$

parce que $y^\circ \notin \partial f_1(x^\circ)$, et on a :

$$f_1(\widetilde{x}) + f_1^*(y^\circ) < f_2(\widetilde{x}) - f_2(x^\circ) + f_1(x^\circ) + f_1^*(y^\circ)$$

alors :

$$f_1(\widetilde{x}) - f_2(\widetilde{x}) < f_1(x^\circ) - f_2(x^\circ)$$

si :

$$\partial f_2(\widetilde{x}) \notin \partial f_1(\widetilde{x})$$

on commençant par un nouveau point arbitraire $x^\circ = \widetilde{x}$ et on continue le travail de la même manière.

2.2.4 L'application de la méthode pour les problèmes quadratiques :

Le problème quadratique est sous la forme suivante :

$$\begin{cases} \min f(x) = \frac{1}{2}x^T Q x + b^T x \\ Bx \leq c; x \geq 0 \end{cases} \tag{PQ}$$

où :

$$\begin{cases} Q \quad \text{Une matrice symétrique} \\ B \quad \text{Une matrice symétrique} \\ c, b, x \in \mathbb{R}^n \end{cases}$$

Notons que se problème peut être s'écrire comme suit :

$$\left\{ \alpha = \inf_{x \in \mathbb{R}^n} \left\{ \frac{1}{2}x^T Q x + b^T x + \chi_E(x) \right\} \right.$$

où :

$$E = \{x \in \mathbb{R}^n;\ Bx \leq c;\ x \geq 0\}$$

avec :

$$\chi_E(x) = \begin{cases} 0 \text{ si } & x \in E \\ +\infty & \text{ailleurs} \end{cases}$$

donc ; la décomposition DC de se problème est donnée par :

$$f(x) = \frac{1}{2}x^T Q x + b^T x + \chi_E(x) = f_1(x) - f_2(x) \tag{2.41}$$

avec :

$$\begin{aligned} f_1(x) &= \frac{1}{2}x^T Q x + b^T x + \chi_E(x) + \frac{1}{2}\rho x^T x \\ f_2(x) &= \frac{1}{2}\rho x^T x \end{aligned}$$

c'est à dire que :

$$\begin{aligned} f_1(x) &= \frac{1}{2}x^T (Q + \rho I) x + b^T x + \chi_E(x) \\ f_2(x) &= \frac{1}{2}\rho x^T x \end{aligned}$$

avec, ρ est un nombre réel positif ; pour le quelle la matrice $(Q + \rho I)$ est *définie positive* ou bien *semi définie positive.*

L'algorithme DC forme simplifier :

Pour chaque itération k, x_k est une solution du problème convexe $(P_{\text{k-1}})$ donné par :

$$\alpha_k = \inf\{f(x) = f_1(x) - f_2(x); x \in \partial f_1^*(y_{k-1}); \text{ et } y_k \in \partial f_2(x_k)\} \tag{$P_{\text{k-1}}$}$$

pour notre problème on obtient :

$$y_k \in \partial f_2(x_k) \Longrightarrow (y_k = \rho x_k) \tag{2.42}$$

et x_{k+1} est la solution du problème suivant :

$$\begin{cases} \inf \{f(x) = f_1(x) - f_2(x)\} \\ x \in \partial f_1^* (y_{k-1}) \end{cases}$$

où :

$$\begin{aligned} \inf \{f_1(x) - f_2(x)\} &= \inf \left\{ \frac{1}{2} x^T (Q + \rho I) x + b^T x + \chi_E (x) - \frac{1}{2} \rho x^T x \right\} \\ &= \inf \left\{ \frac{1}{2} x^T (Q + \rho I) x + x^T \left(b - \frac{1}{2} \rho x \right) + \chi_E (x) \right\} \\ &= \inf \left\{ \frac{1}{2} x^T (Q + \rho I) x + x^T \left(b - \frac{1}{2} y_k \right) + \chi_E (x) \right\} \end{aligned}$$

donc le problème DC est définie par :

$$\inf_{x \in \partial f_1^*(y_{k-1})} \left\{ \frac{1}{2} x^T (Q + \rho I) x + x^T \left(b - \frac{1}{2} y_k \right) + \chi_E (x) \right\} \qquad (\mathrm{P_{dc}})$$

Remarque 2.10 : La décomposition DC associer au problème $(P_{\text{k-1}})$ n'est pas unique.

En effet, prend par exemple la décomposition suivante :

$$\begin{aligned} f_1(x) &= \frac{1}{2} \rho \|x\|^2 + b^T x + \chi_E (x) \\ f_2(x) &= \frac{1}{2} x^T (\rho I - Q) x \end{aligned}$$

où ρ est un nombre réel positif pour le quelle la matrice $(\rho I - Q)$ sera définie positive, et on écrit :

$$(\rho I - Q) \geq 0$$

Lemme 2.2 : Les deux fonctions $f_1(x)$ et $f_2(x)$ sont des fonctions convexes, sci et propres, alors on peut écrire :

$$f_2(x), f_1(x) \in \Gamma_\circ (\mathbb{R}^n)$$

D'après le lemme préçédent on peut interpréter l'algorithme DC simplifier au ce problème comme suit :

soit :

$$x^{\circ} \in \mathbb{R}^{n} \equiv X, k \geq 0, y_{k} = (\rho I - Q)\, x_{k} \tag{2.43}$$

et x_{k+1} est la solution du problème :

$$\min_{x\in\mathbb{R}^n} \left\{ \frac{1}{2}\rho \left\| x \right\|^{2} + b^{T}x + \chi_{E}(x) - \frac{1}{2}x^{T}(\rho I - Q)\, x \right\}$$

qu'est équivalent à :

$$\min_{x\in\mathbb{R}^n} \left\{ \frac{1}{2}\rho \left\| x \right\|^{2} + x^{T}(b - y_{k}) + \chi_{E}(x) \right\}$$

ce dernier problème est équivalent à :

$$\begin{cases} \min \left\| x - \left(\frac{y_k - b}{\rho}\right) \right\|^{2} \\ x \in E \end{cases} \tag{P_{proj}}$$

qu'est un problème du projection sur l'ensemble :

$$E = \{ x \in \mathbb{R}^{n}; Bx \leq c \}$$

d'autre façons, on dit que :

x_{k+1} est la projection de $\left(\frac{y_k - b}{\rho}\right)$ sur E, et on écrit :

$$x_{k+1} = P_{E}\left(\frac{y_{k} - b}{\rho}\right) \text{ avec } y_{k} = (\rho I - Q)\, x_{k} \tag{2.44}$$

donc :

$$x_{k+1} = P_{E}\left(\frac{((\rho I - Q)\, x_{k} - b)}{\rho}\right)$$

dans un forme plus simple :

$$x_{k+1} = P_{E}\left(x_{k} - \frac{(Qx_{k} + b)}{\rho}\right) \tag{2.45}$$

alors, x_{k+1} est la solution de le problème convexe suivant :

$$\left\| x_{k+1} - \left(x_k - \frac{(Qx_k + b)}{\rho} \right) \right\| = \min_{y \in E} \left\| y - \left(x_k - \frac{(Qx_k + b)}{\rho} \right) \right\|$$

on peut le donné sous la forme suivante :

$$\|x_{k+1} - z\| = \min_{y \in E} \|y - z\|$$

où :

$$z = \left(x_k - \frac{(Qx_k + b)}{\rho} \right)$$

ce problème est équivalent à :

$$\begin{aligned} \frac{1}{2} \|x_{k+1} - z\|^2 &= \min_{y \in E} \frac{1}{2} \|y - z\|^2 \\ &= \min_{y \in E} \left(\frac{1}{2} \langle Iy, y \rangle - \langle y, z \rangle + \|z\|^2 \right) \end{aligned}$$

d'où l'algorithme DC simplifier sera donnée par :

Algorithme :

Initialisation : $k = 0, x_0 \in \mathbb{R}^n$;

Itération k :

(1) Si : $\left(x_k - \frac{(Qx_k+b)}{\rho} \right) \in E$ alors ;

$x_{k+1} = x_k - \frac{(Qx_k+b)}{\rho}$

sinon :

$x_{k+1} = P_E \left(x_k - \frac{(Qx_k+b)}{\rho} \right)$.

(2) Si : $\|x_{k+1} - x_k\| \leq \varepsilon$; stop ;

sinon $k \longleftarrow k + 1$.

Chapitre 3

La méthode de La transformation dual canonique (TDC) :

Dans ce chapitre, on présente une méthode de recherche globale appelée *la méthode de la transformation duale canonique* (TDC).

TDC nous peut transformer le problème d'optimisation quadratique non convexe avec des contraintes linéaires à un système algébrique facile pour résoudre, au plus de ça, on peut former un problème dual canonique qui présente le problème dual parfait du problème primal dans le sens que sont admet le même ensemble des points de KKT.

On va démontrer que les solutions extrêmes locaux et globaux peut être définit par le *triality theory,* et on a donner aussi quelle que résultat expriment la relation entre le problème primal et le problème dual.

Finalement, des exemples seront étudier dans la fin de chapitre.

3.1 La transformation duale canonique dans le cas général :

Soit le problème quadratique non convexe suivant :

$$\begin{cases} \min f(x) = \frac{1}{2}x^T Qx - d^T x \\ Bx \leq c; x \geq 0 \end{cases} \tag{PQN}$$

où Q une matrice indéfinie ;

B une matrice appartienne à $\mathbb{R}^{n+m}$.

3.1.1 L'idée fondamentale de cette méthode :

Il consiste au choisir un opérateur :

$$y = \Lambda(x) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

pour le qu'elle la fonction $f(x)$ peut être écrit sous la forme canonique suivante :

$$\Phi(x, \Lambda(x)) = f(x) \text{ avec } \Phi(x, y) : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R} \tag{3.1}$$

et définie une fonction canonique en chaque variable x et y.

Définition 3.1 :[5] Soit la fonction :

$$\overline{W}(y) : \mathbb{R}^m \longrightarrow \mathbb{R}$$

on dira que $\overline{W}(y)$ est canonique sur $\mathbb{R}^m$ si la relation suivante $y^* = D\overline{W}(y)$ est inversible (*one–to–one*) pour tout $y \in \mathbb{R}^m$.

3.1.2 Description de la méthode :

On a besoin des définitions suivantes :

Définition 3.2 :[5] La conjuguée de Legendre du la fonction canonique $\overline{W}(y)$ sera définie uniquement par :

$$\begin{aligned}\overline{W}^{*}(y) &= \sup_{y\in\mathbb{R}^m}\left\{\langle y^{*},y\rangle-\overline{W}(y):y^{*}=D\overline{W}(y)\right\} \\ &= \sup_{y\in\mathbb{R}^m}\left\{y^{T}y^{*}-\overline{W}(y):y^{*}=D\overline{W}(y)\right\}\end{aligned} \qquad (3.2)$$

Sachons que la fonction $\Phi(x,\Lambda(x))$ peut être écrit sous la forme :

$$\Phi(x,y)=\overline{W}(y)-\overline{F}(y) \qquad (3.3)$$

où :

$$\overline{W}:\mathbb{R}^m\longrightarrow\mathbb{R},\overline{F}:\mathbb{R}^n\longrightarrow\mathbb{R}$$

sont des fonctions canoniques sur $\mathbb{R}^m$ et $\mathbb{R}^n$ respectivement, donc, on peut déduire la définition suivante :

Définition 3.3 :[5] On définie la fonction suivante :

$$E(x,y^{*})=\left((\Lambda(x))^{T}y^{*}-\overline{W}^{*}(y^{*})-\overline{F}(x)\right) \qquad (3.4)$$

qu'est bien définie de $\mathbb{R}^n\times\mathbb{R}^m$ vers $\mathbb{R}$.

Définition 3.4 :[5] Avec l'utilisation de la transformation duale $\Lambda-$*canonique* on définie :

$$\overline{F}^{\wedge}(y^{*})=\left\{(\Lambda(x))^{T}y^{*}-\overline{F}(x):\Lambda_t^{T}(x)y^{*}-D\overline{F}(x)=0\right\} \qquad (3.5)$$

où :

$$\Lambda_t(x)=D\Lambda(x)$$

présente *la dérivée de Gâteaux de l'opérateur* $\Lambda(x)$.

Définition 3.5 :[5] La fonction duale canonique de la fonction *non convexe* $f(x)$ peut être bien définie par :

$$f^{d}(y^{*})=\overline{F}^{\wedge}(y^{*})-\overline{W}^{*}(y^{*}) \qquad (3.6)$$

La proposition suivante nous donnera la condition pour identifier la valeur optimale duale et la valeur optimale primale ; "*there is no duality gap between the primal problem and the dual one*".

Proposition 3.1 :[5] *Si le point* (x^*, y^*) *est un point critique de la fonction* $E(x, y)$ *alors la relation suivante sera satisfaite :*

$$f^d(y^*) = f(x^*) \tag{3.7}$$

Sachons que, la transformation duale canonique a été étudier par "*Gao*" et "*Strang*" dans *les mécaniques non convexes* et *non de classe* C^∞, cette étude est pour :

$$\begin{cases} f(x) : \text{Une } \textit{fonction quadratique} \\ \overline{W}(y) : \textit{Une fonction convexe} \\ \overline{F}(x) : \textit{Une fonction linéaire} \end{cases}$$

3.1.3 Exemple :

Posons :

$$\begin{aligned} f(x) &= \frac{1}{2}\left(\frac{1}{2}|x|^2 - 1\right)^2 - d^T x \\ \overline{W}(x) &= \frac{1}{2}\left(\frac{1}{2}|x|^2 - 1\right)^2 \\ \overline{F}(x) &= d^T x \end{aligned}$$

donc :

$$x^* = D\overline{W}(x) = \left(\frac{1}{2}|x|^2 - 1\right)(|x| D|x|)$$

Cette relation n'est pas *inversible (one-to-one)*, alors, il faut étudier la conjuguée de "*Legendre*" de la fonction $\overline{W}(x)$. Alors, par définition on a :

$$\begin{aligned}\overline{W}^{*}(x^{*}) &= \sup_{x}\left\{x^{T}x^{*}-\overline{W}(x)\right\}\\ &= \sup_{x}\left\{x^{T}x^{*}-\frac{1}{2}\left(\frac{1}{2}|x|^{2}-1\right)^{2}\right\}\end{aligned}$$

Qui n'est pas unique, donc on peut pas définie la fonction duale canonique $f^{d}(x^{*})$ de $f(x)$ par la fonction $\overline{W}^{*}(x^{*})$, alors il faut trouver l'opérateur :

$$y=\Lambda(x):\mathbb{R}^{n}\longrightarrow\mathbb{R}^{m}$$

pour le quelle la fonction $f(x)$ peut s'écrire sous la forme canonique :

$$f(x)=\Phi\left(x,\Lambda\left(x\right)\right)$$

Si on prend :

$$y=\frac{1}{2}|x|^{2}$$

on trouve que :

$$\overline{\overline{W}}(y)=\frac{1}{2}(y-1)^{2}$$

Pui présente *une forme quadratique canonique* de :

$$y=\Lambda(x)=\frac{1}{2}|x|^{2}$$

D'autre part, on a :

$$\overline{\overline{W}}(y)=\overline{\overline{W}}(\Lambda(x))=\overline{W}(x)$$

et :

$$D\overline{\overline{W}}(y)=y-1$$

La dernière équation est inversible, alors on peut définir la fonction duale $f^d(y^*)$ par la conjuguée de Legendre de la fonction $\overline{\overline{W}}(y)$ qui définie par :

$$\begin{aligned}
\overline{\overline{W}}^*(y^*) &= \sup_y \left\{ y^T y^* - \overline{\overline{W}}(y) \ / \ y^* = D\overline{\overline{W}}(y) = y - 1 \right\} \\
&= \sup_y \left\{ y^T y^* - \frac{1}{2}(y-1)^2 \ / \ y^* = y - 1 \right\} \\
&= \sup_y \left\{ (y^* + 1)^T y^* - \frac{1}{2}(y^* + 1 - 1)^2 \right\} \\
&= \sup_y \left\{ (y^*)^T y^* + y^* - \frac{1}{2}(y^*)^2 \right\} \\
&= \sup_y \left\{ \frac{1}{2}(y^*)^2 + y^* \right\} : y^* = y - 1
\end{aligned}$$

On peut poser :

$$\overline{\overline{W}}^*(y^*) = \frac{1}{2}(y^*)^2 + y^*$$

Il reste de trouver la forme de la fonction linéaire $\overline{\overline{F}}^{\wedge}(y^*)$, alors on a :

$$\overline{\overline{F}}(y) = d^T y$$

Donc :

$$\begin{aligned}
\overline{\overline{F}}^{\wedge}(y^*) &= \left\{ y^T y^* - \overline{\overline{F}}(y) \ / \ y = \Lambda(x), \Lambda_t^T(x) y^* - D\overline{\overline{F}}(x) = 0 \right\} \\
&= \left\{ (\Lambda(x))^T y^* - \overline{\overline{F}}(\Lambda(x)) \ / \ y = \Lambda(x), \Lambda_t^T(x) y^* - D\overline{\overline{F}}(x) = 0 \right\}
\end{aligned}$$

mais on a :

$$\overline{\overline{F}}(\Lambda(x)) = \overline{F}(x) = d^T x$$

alors, on trouve que :

$$\begin{aligned}
\overline{\overline{F}}^{\wedge}(y^*) &= \left\{ (\Lambda(x))^T y^* - d^T x \ / \ \Lambda_t^T(x) y^* - D\overline{\overline{F}}(x) = 0 \right\} \\
&= \left\{ \frac{1}{2}|x|^2 y^* - d^T x \ / \ \Lambda_t^T(x) y^* - d = 0 \right\}
\end{aligned}$$

Finalement, on définie la fonction duale canonique de la fonction non convexe optimiser comme suit :

$$\begin{aligned} f^d(y^*) &= \overline{\overline{F}}^{\wedge}(y^*) - \overline{\overline{W}}^*(y^*) \\ &= \frac{1}{2}|x|^2 y^* - d^T x - \frac{1}{2}(y^*)^2 + y^* \;\; / \;\; y^* = y - 1 \end{aligned}$$

Le problème qui se pose içi est donné par :

$$\begin{cases} \max f^d(y^*) = \frac{1}{2}|x|^2 y^* - d^T x - \frac{1}{2}(y^*)^2 + y^* \\ \qquad y^* = y - 1 \end{cases}$$

avec $y = \frac{1}{2}|x|^2$ on peut écrire :

$$\begin{cases} \max f^d(y^*) = y^T y^* - d^T x - \frac{1}{2}(y^*)^2 + y^* \\ \qquad y^* = y - 1 \end{cases}$$

qui devient :

$$\begin{cases} \max f^d(y^*) = (y^* + 1)^T y^* - d^T x - \frac{1}{2}(y^*)^2 + y^* \\ \qquad y = y^* + 1 \end{cases} \qquad (PDC)$$

Remarque 3.1 : Remarquons que la forme duale est plus simple que la forme originale de $f(x)$.

3.2 L'application de la méthode pour les problèmes quadratiques non convexes :

Soit le problème définie plus haut par :

$$\begin{cases} \min f(x) = \frac{1}{2}x^T Q x - d^T x \\ \qquad Bx \leq c, x \geq 0 \end{cases} \qquad \text{(PQN)}$$

Où B, Q sont deux matrices symétriques;

$c, d, x \in \mathbb{R}^n$;

Alors :

comment on applique la méthode de la transformation duale canonique au se type de problèmes ?.

Pour résoudre le problème quadratique non convexe (PQN) avec l'utilisation de cette méthode *(TDC)*, il faut ajoutée la condition suivante :

$$|x|^2 \leq 2\mu \tag{3.8}$$

qu'est appelée *la condition de normalité* où $\mu > 0$ est un paramètre donné.

Remarque 3.1 : La condition préçédente est très importante, parce qu'elle nous garante l'existence de la solution globale.

On peut alors définir le problème paramétrique suivant :

$$\begin{cases} \min f(x) = \frac{1}{2}x^T Qx - d^T x \\ Bx \leq c, x \geq 0, \frac{1}{2}|x|^2 \leq \mu \end{cases}$$

qui peut transformer à :

$$\begin{cases} \min f(x) = \frac{1}{2}x^T Qx - d^T x \\ Bx \leq c, \frac{1}{2}|x|^2 \leq \mu \end{cases} \tag{P_μ}$$

où :

$$B = \left\{ B, \underleftrightarrow{\{-1, -1, \cdots, -1\}} \in \mathbb{R}^n \right\}, c = (c, 0) \tag{3.9}$$

Définition 3.5 : L'espace réalisable de (P_μ) où $\mu > 0$, et on le note par (X_μ), sera définie par :

$$(X_\mu) = \left\{ x \in \mathbb{R}^n \ / \ Bx \leq c, \frac{1}{2}|x|^2 \leq \mu \right\} \tag{3.10}$$

Corollaire 3.1 :[5] L'espace réalisable de (P_μ) est un ensemble convexe fermé de $\mathbb{R}^n$.

La proposition suivante est une conséquence de cette corollaire.

Proposition 3.2 :[5] *Le problème d'optimisation paramétrique (P_μ) admet au moins un minimum global noté par x^*_μ.*

La relation entre le problème original (PQN) et le problème paramétrique (P_μ) sera bien expliquée par la proposition suivante :

Proposition 3.3 :[5] *Soit* $\left(\mu_\circ = \frac{1}{2}\rho_\circ^2\right)$ *le rayon de l'espace réalisable* (X_μ)*, si* $\mu \geq \mu_\circ$ *alors* x_μ^* *(la solution globale de* (P_μ)*) est une solution du le problème original (PQN).*

Pour trouver la formulation duale canonique du problème paramétrique (P_μ) avec *(TDC),* on utilise la procédure standard de même méthode, et dans se cas on a :

3.2.1 La structure de l'opérateur $\Lambda(x)$ pour ce type des problèmes :

L'opérateur géométrique canonique Λ est définie par :

$$\Lambda : \mathbb{R}^n \longrightarrow \mathbb{R}^m \times \mathbb{R}$$

On peut alors définir l'opérateur $\Lambda(x)$ associée au (P_μ) par :

$$y = \Lambda(x) = \left(Bx, \frac{1}{2}|x|^2\right) = (\varepsilon, \rho) \tag{3.11}$$

Remarque 3.2 : Remarquons que l'opérateur $\Lambda(x)$ est définie comme *vecteur–valeur application* avec :

$$\begin{aligned} \varepsilon &= Bx \textit{ Est un vecteur de } \mathbb{R}^n \\ \rho &= \frac{1}{2}|x|^2 \textit{ Est un scalaire de } \mathbb{R} \end{aligned}$$

Et à cause de ça, on peut définir l'espace réalisable de (P_μ) sous une notre forme comme suit :

Définition 3.6 :[5] L'espace réalisable de (P_μ) notée par (Y_a) est un sous ensemble convexe de $Y = \mathbb{R}^m \times \mathbb{R}$ définie par :

$$(Y_a) = \{y = (\varepsilon, \rho) \in \mathbb{R}^m \times \mathbb{R} \ / \ \varepsilon \leq c, \rho \leq \mu\} \tag{3.12}$$

3.2.2 La structure de la fonction $\overline{W}(y)$:

Tous simplement, la fonction $\overline{W}(y)$ peut être définir comme étant la fonction indicatrice de l'ensemble (Y_a) définie par :

$$\overline{W}(y) : Y_a \longrightarrow \overline{\mathbb{R}}$$

$$\overline{W}(y) = \begin{cases} 0 & \text{si } y \in Y_a \\ +\infty & \text{ailleurs} \end{cases} \tag{3.13}$$

Remarque 3.3 : La convexité de la fonction $\overline{W}(y)$ sera vérifiée par la convexité de la fonction indicatrice ; puisque la fonction indicatrice est originalement convexe, et au plus de ça, elle est *semi continue inférieurement* et *propre* sur l'ensemble (Y_a).

D'après ces caractérisations, on peut définir :

Définition 3.7 : L'opérateur dual :

$$y^* \in Y^* = Y = \mathbb{R}^m \times \mathbb{R}$$

Est définie par :

$$y^* \in \partial\overline{W}(y) = \begin{cases} (\varepsilon^*, \rho^*) & \text{si } \ \varepsilon^* \geq 0, \rho^* \geq 0 \\ \phi & \text{ailleurs} \end{cases} \tag{3.14}$$

3.2.3 La structure de $\overline{W}^{\#}(y^*)$:

Par hypothèse, la matrice Q est indéfinie, alors, au lieu du définir la fonction $\overline{W}^*(y)$ *la conjuguée de Legendre* de la fonction $\overline{W}(y)$, il suffit d'utiliser la transformation de *Fenchel-Rockafllar* qu'est notée par $\overline{W}^{\#}(y^*)$, elle est appellée aussi *la conjuguée canonique.*

La définition sera donnée par :

$$\begin{aligned}
\overline{W}^{\#}(y^*) &= \sup_{y\in Y}\left\{\langle y^*, y\rangle - \overline{W}(y)\right\} \qquad (3.15)\\
&= \sup_{\varepsilon\leq c}\sup_{\rho\leq\mu}\left\{y^T y^* - \overline{W}(y)\right\};\ y\in Y_a\\
&= \sup_{\varepsilon\leq c}\sup_{\rho\leq\mu}\left\{(\varepsilon,\rho)^T(\varepsilon^*,\rho^*) - \overline{W}(y);\ y\in Y_a\right\}\\
&= \sup_{\varepsilon\leq c}\sup_{\rho\leq\mu}\left\{\varepsilon^T\varepsilon^* + \rho^T\rho^*\right\}\\
&= \begin{cases} c^T\varepsilon^* + \mu^T\rho^* & \text{si } \varepsilon^*\geq 0, \rho^*\geq 0\\ +\infty & \text{ailleurs}\end{cases}
\end{aligned}$$

Définition 3.8 : Le domaine effectif de $\overline{W}^{\#}(y^*)$ est le cône positif dans $\mathbb{R}^m\times\mathbb{R}$ définie par :

$$\begin{aligned}
Y_a^* &= dom(\overline{W}^{\#}(y^*)) \qquad (3.16)\\
&= (\varepsilon^*,\rho^*)\in\mathbb{R}^m\times\mathbb{R}; \varepsilon^*\geq 0, \rho^*\geq 0;\ \text{où } \varepsilon^*\in\mathbb{R}^m, \rho^*\in\mathbb{R}
\end{aligned}$$

Proposition 3.4 :[5] *La relation suivante* :

$$\left(y^*\in\partial\overline{W}(y)\right)\Longleftrightarrow\left(y\in\partial\overline{W}^{\#}(y^*)\right)\Longleftrightarrow\left(\overline{W}(y)+\overline{W}^{\#}(y^*)=y^Ty^*\right) \qquad (3.17)$$

qu'est appelée la relation de sup-*duality est bâti.*

Preuve : Sachons que $\overline{W}(y)$ (la fonction indicatrice de l'ensemble Y_a) est une fonction *convexe, semi continue inférieurement* et *propre* alors, on peut écrire :

$$\overline{W}(y)\in\Gamma_\circ(\mathbb{R}^n)$$

Donc ; d'après le théorème de *Fenchel-Moran* on a $\overline{W}^{\#}(y^*)$ est aussi une fonction propre,

Au plus de ça, on a :

$$y^*\in\partial\overline{W}(y)\Longrightarrow y\in\partial\overline{W}^{\#}(y^*)$$

D'autre part, on a :

$$\overline{W}(y) \in \Gamma_\circ(\mathbb{R}^m \times \mathbb{R})$$

Alors :

$$y \in \partial\overline{W}^{\#}(y^*) \Longrightarrow y^* \in \partial\overline{W}(y) \tag{3.18}$$

Donc, on peut écrire :

$$y^* \in \partial\overline{W}(y) \Longleftrightarrow y \in \partial\overline{W}^{\#}(y^*)$$

et on a aussi :

$$y \in \partial\overline{W}^{\#}(y^*) \Longleftrightarrow \overline{W}(y) + \overline{W}^{\#}(y^*) = y^T y^* \tag{3.19}$$

Finalement, d'après (*3.18*) et (*3.19*) la preuve est complète. ∎

On peut continuer notre travail sur l'espace restreint $(Y_a \times Y_a^*)$.

Lemme 3.1 :[5] La *relation de* sup-*duality* préçédente est équivalent au condition de *(KKT)* suivante :

$$Y_a \ni y \perp y^* \in Y_a^* \tag{3.20}$$

Remarque 3.3 :

(1) Le couple (y, y^*) est appelé *le couple dual canonique mesuré sur l'espace* $(Y_a \times Y_a^*)$.

(2) Les deux fonctions $\overline{W}(y)$ et $\overline{W}^{\#}(y^*)$ sont appelées *les fonctions canoniques.*

3.2.4 La structure de $\overline{F}^{\wedge}(y^*)$:

Originalement, $\overline{F}(x)$ est présentée comme une fonction linéaire donc, pour écrire la fonction optimiser $f(x)$ sous la forme canonique $\Phi(x, \Lambda(x))$ il suffit de prendre $\overline{F}(y)$ sous la forme suivante :

On a :

$$\Phi(x, \Lambda(x)) = \overline{W}(y) - \overline{F}(y) = f(x) \tag{3.21}$$

Donc, s'implique que :

$$f(x) - \overline{W}(y) = -\overline{F}(y)$$

où $y \in Y_a$, donc :

$$-\overline{F}(y) = f(x)$$

c'est à dire que :

$$\overline{F}(y) = -f(x) \tag{3.22}$$

Alors, on peut définir la conjugué Λ-*canonique* $\overline{F}^{\wedge}(y^*)$ de la fonction canonique $\overline{F}(y)$ où :

$$y^* = (\varepsilon^*, \rho^*) \in \mathbb{R}^m \times \mathbb{R}$$

un paramètre donné, telle que la matrice $(Q + \rho^* I)$ est inversible, comme suit :

$$\begin{aligned}
\overline{F}^{\wedge}(y^*) &= \sup_{y \in Y_a} \left\{ y^T y^* - \overline{F}(y) \ / \ D\overline{F}(y) = \Lambda_t^T(x)\, y^* ; x \in (X_\mu) \right\} \\
&= \sup_{y \in Y_a} \left\{ (\Lambda(x))^T y^* - \overline{F}(y) \ / \ D\overline{F}(y) = \Lambda_t^T(x)\, y^* ; x \in (X_\mu) \right\}
\end{aligned}$$

avec :

$$\begin{aligned}
\Lambda(x) &= y = \left(Bx, \frac{1}{2}|x|^2 \right) \in \mathbb{R}^m \times \mathbb{R} \\
y^* &= (\varepsilon^*, \rho^*) \in \mathbb{R}^m \times \mathbb{R}
\end{aligned}$$

Alors, on peut écrire :

$$\begin{aligned}
\overline{F}^{\wedge}(y^*) &= \sup_{y \in Y_a} \left\{ (\Lambda(x))^T y^* - \overline{F}(\Lambda(x)) \ / \ D\overline{F}(y) = \Lambda_t^T(x)\, y^* ; x \in (X_\mu) \right\} \\
&= \sup_{y \in Y_a} \left\{ \left(Bx, \frac{1}{2}|x|^2 \right)^T (\varepsilon^*, \rho^*) - d^T x + \frac{1}{2} x^T Q x \right\} \ / \ \overline{F}(\Lambda(x)) = \overline{F}(x) \\
&= \sup_{y \in Y_a} \left\{ x^T B^T \varepsilon^* + \frac{1}{2}|x|^2 \rho^* - d^T x + \frac{1}{2} x^T Q x \right\} \\
&= \sup_{y \in Y_a} \left\{ \frac{1}{2} x^T (Q + \rho^* I)\, x + x^T \left(B^T \varepsilon^* - d \right) \right\} \\
&= \sup_{y \in Y_a} \left\{ \frac{1}{2} x^T (Q + \rho^* I)\, x - \left(d - B^T \varepsilon^* \right)^T x ; \ x \in (X_\mu) \right\}
\end{aligned}$$

Et pour :

$$x = (Q + \rho^* I)^{-1} \left(d - B^T \varepsilon^* \right) \in (X_\mu) \tag{3.23}$$

on trouve que :

$$\overline{F}^{\wedge}(y^*) = \frac{1}{2}\left(d - B^T\varepsilon^*\right)^T (Q+\rho^* I)^{-1}(Q+\rho^* I)(Q+\rho^* I)^{-1}\left(d - B^T\varepsilon^*\right) \quad (3.24)$$
$$-\left(d - B^T\varepsilon^*\right)^T (Q+\rho^* I)^{-1}\left(d - B^T\varepsilon^*\right)$$
$$= \frac{-1}{2}\left(d - B^T\varepsilon^*\right)^T (Q+\rho^* I)^{-1}\left(d - B^T\varepsilon^*\right)$$

donc, on a trouver la bonne forme canonique $\overline{F}^{\wedge}(y^*)$ de $\overline{F}(y)$.

3.2.5 La structure de la fonction dual canonique $f^d(y^*)$:

On a déjà signaler que :

$$f^d(y^*) = \overline{\overline{F}}^{\wedge}(y^*) - \overline{\overline{W}}^*(y^*)$$

Donc, d'après les deux formes précédentes de $\overline{\overline{F}}^{\wedge}(y^*)et\overline{\overline{W}}^*(y^*)$ on peut trouver que :

$$f^d(y^*) = \frac{-1}{2}\left(d - B^T\varepsilon^*\right)^T (Q+\rho^* I)^{-1}\left(d - B^T\varepsilon^*\right) - c^T\varepsilon^* - \mu\rho^*;\ \text{avec } (\varepsilon^*, \rho^*) \in \mathbb{R}^m \times \mathbb{R} \quad (3.25)$$

Proposition 3.5 :[5] *Le problème dual canonique (P_μ^d) associer au problème paramétrique (P_μ) peut être définie par :*

$$\begin{cases} \max f^d(\varepsilon^*, \rho^*) \\ \varepsilon^* \geq 0, \rho^* \geq 0 \text{ et } \det(Q+\rho^* I) \neq 0 \end{cases} \quad (P_\mu^d)$$

On peut le représenter sous une notre forme comme suit :

$$\begin{cases} \operatorname{Ext} f^d(\varepsilon^*, \rho^*) \\ \varepsilon^* \geq 0, \rho^* \geq 0 \text{ et } \det(Q+\rho^* I) \neq 0 \end{cases}$$

où Ext $f(x)$ est pour trouver tout les valeurs extrêmes de la fonction $f(x)$.

Preuve : voir(Gao, 2003). ■

Le théorème suivante nous donnera l'équivalence entre le problème primal (paramétrique) (P_μ) et le problème dual canonique (P_μ^d) dans le sens que Sont admet le même ensemble des points du *(KKT).*

Théorème 3.1 :[5] *Le problème dual (P_μ^d) est un problème canonique de (P_μ) dans le sens que si :*

$$\overline{y^*} = (\overline{\varepsilon^*}, \overline{\rho^*}) \in Y_\mu^*$$

est un point de (KKT) *de (P_μ^d) alors, le vecteur qui définie par :*

$$\overline{x} = (Q + \overline{\rho^*} I)^{-1} \left(d - B^T \overline{\varepsilon^*}\right)$$

est un point de (KKT) *de (P_μ), et on a aussi :*

$$f(\overline{x}) = f^d(\overline{y^*})$$

Preuve : Supposons que :

$$\overline{y^*} = (\overline{\varepsilon^*}, \overline{\rho^*}) \in Y_\mu^*$$

est un point de *(KKT)* de (P_μ^d) alors, il satisfait les conditions suivantes :

$$0 \leq \overline{\rho^*} \perp \frac{1}{2} \left(d - B^T \overline{\varepsilon^*}\right)^T (Q + \overline{\rho^*} I)^{-2} \left(d - B^T \overline{\varepsilon^*}\right) - \mu \leq 0 \qquad (3.26)$$

qui présente la dérivé d'ordre (1) de $f^d(y^*)$ par rapport à ρ^*en $\overline{y^*} = (\overline{\varepsilon^*}, \overline{\rho^*})$.

$$0 \leq \overline{\varepsilon^*} \perp B \left(d - B^T \overline{\varepsilon^*}\right)^T (Q + \overline{\rho^*} I)^{-1} - c \leq 0 \qquad (3.27)$$

qui présente la dérivé d'ordre (1) de $f^d(y^*)$ par rapport à ε^* en $\overline{y^*} = (\overline{\varepsilon^*}, \overline{\rho^*})$, donc, si on prend :

$$\overline{x} = (Q + \overline{\rho^*} I)^{-1} \left(d - B^T \overline{\varepsilon^*}\right)$$

alors, les conditions précédentes devient :

$$\begin{aligned} 0 &\leq \overline{\rho^*} \perp \frac{1}{2} \overline{x}^T \overline{x} - \mu \leq 0 \\ 0 &\leq \overline{\varepsilon^*} \perp B\overline{x} - c \leq 0 \end{aligned}$$

qui bien dite que le point :

$$\overline{x} = (Q + \overline{\rho^*} I)^{-1} \left(d - B^T \overline{\varepsilon^*}\right)$$

est un point de *(KKT)* pour le problème paramétrique (P_μ).

D'un autre côté, les conditions précédentes peut s'écrire comme suit :

$$\begin{aligned} \overline{\rho^*}\mu &= \frac{1}{2}\overline{\rho^*}\overline{x}^T\overline{x} \\ \overline{\varepsilon^*}c &= \overline{\varepsilon^*}B\overline{x} \end{aligned} \qquad (3.28)$$

donc :

$$\begin{aligned} f^d(\overline{y^*}) &= f^d\left(\overline{\varepsilon^*}, \overline{\rho^*}\right) \\ &= \frac{-1}{2}\left(d - B^T\overline{\varepsilon^*}\right)^T (Q + \overline{\rho^*}I)^{-1}\left(d - B^T\overline{\varepsilon^*}\right) - c^T\overline{\varepsilon^*} - \mu\overline{\rho^*} \\ &= \frac{-1}{2}\left(d - B^T\overline{\varepsilon^*}\right)^T (Q + \overline{\rho^*}I)^{-1}(Q + \overline{\rho^*}I)(Q + \overline{\rho^*}I)^{-1}\left(d - B^T\overline{\varepsilon^*}\right) - c^T\overline{\varepsilon^*} - \mu\overline{\rho^*} \\ &= \frac{-1}{2}\overline{x}^T(Q + \overline{\rho^*}I)\overline{x} - \frac{1}{2}\overline{\rho^*}\overline{x}^T\overline{x} - \overline{x}^T B^T\overline{\varepsilon^*} \\ &= \frac{-1}{2}\overline{x}^T(Q + \overline{\rho^*}I)\overline{x} - \frac{1}{2}\overline{\rho^*}\overline{x}^T\overline{x} + \overline{x}^T\left(d - B^T\overline{\varepsilon^*} - d\right) \\ &= \frac{-1}{2}\overline{x}^T(Q + \overline{\rho^*}I)\overline{x} - \frac{1}{2}\overline{\rho^*}\overline{x}^T\overline{x} + \overline{x}^T(Q + \overline{\rho^*}I)(Q + \overline{\rho^*}I)^{-1}\left(d - B^T\overline{\varepsilon^*}\right) - \overline{x}^T d \\ &= \frac{-1}{2}\overline{x}^T(Q + \overline{\rho^*}I)\overline{x} - \frac{1}{2}\overline{\rho^*}\overline{x}^T\overline{x} + \overline{x}^T(Q + \overline{\rho^*}I)\overline{x} - \overline{x}^T d \\ &= \frac{1}{2}\overline{x}^T(Q + \overline{\rho^*}I)\overline{x} - \frac{1}{2}\overline{\rho^*}\overline{x}^T\overline{x} - d^T\overline{x} \\ &= \frac{1}{2}\overline{x}^T Q\overline{x} - d^T\overline{x} \\ &= f(\overline{x}) \end{aligned}$$

■

3.3 Les conditions extrêmes des points de (KKT) :

Pour bien traiter les conditions extrêmes de *(KKT)* on a besoin des définitions suivantes :

Définition 3.8 : Soit $Q \in \mathbb{R}^{n\times n}$ une matrice symétrique, on note par id; l'indice de Q, qu'est le nombre total des valeurs propres négatives distincts de la matrice Q.

Remrque 3.2 : On peut déduire que le problème quadratique est un problème non convexe si et seulement si :

$$id > 0$$

Lemme 3.2 :[**5**] Soit :

$$\left(\overline{\varepsilon_i^*}, \overline{\rho_i^*}\right) \in \mathbb{R}^m \times \mathbb{R}$$

un point de *(KKT)* pour le problème (P_μ^d) et soit Q une matrice, avec l'indice id, admet $p \leq n$ valeurs propres distincts $\{q_i\}_{i=\overline{1,p}}$ dans l'ordre suivant :

$$q_1 < q_2 < \cdots < q_{id} < 0 \leq q_{id+1} < \cdots < q_p$$

donc, si :

$$q_{id+1} = q_{id+2} = \cdots = q_p = 0$$

et le point de *(KKT)* $\overline{\rho_i^*} < -q_{id}$; alors on a :

$$\overline{\rho_i^*} < -q_{id} < \cdots < -q_1 \text{ avec } -q_1 > -q_2 > \cdots > -q_{id} > 0$$

et la matrice $\left(Q + \overline{\rho_I^*} I\right)$ sera définie négative, d'autre part ; si :

$$\overline{\rho_i^*} > -q_1 > 0$$

alors, la matrice $\left(Q + \overline{\rho_I^*} I\right)$ sera définie positive.

D'après le lemme précédent on peut poser la définition suivante :

Définition 3.9 :[5] On va définie les ensembles suivants :

$$\begin{aligned} Y^*_{\mu+} &= \left\{(\varepsilon^*, \rho^*) \in Y^*_{\mu};\ \text{telle que } (Q + \rho^* I) \text{ est définie positive}\right\} \\ Y^*_{\mu-} &= \left\{(\varepsilon^*, \rho^*) \in Y^*_{\mu};\ \text{telle que } (Q + \rho^* I) \text{ est définie négative}\right\} \\ Y^*_{\mu i} &= \left\{(\varepsilon^*, \rho^*) \in Y^*_{\mu-};\ \text{telle que } \rho^* > 0\right\} \\ X_{\mu s} &= \left\{x \in X_{\mu};\ \text{telle que } \frac{1}{2}|x|^2 = \mu\right\} \end{aligned}$$

Définition 3.10 :[18] Le lagrangien classique $L : \mathbb{R}^n \times \mathbb{R}^m_+ \longrightarrow \mathbb{R}$ associer au problème primal *(P)* est donné par :

$$L(x, v) = \frac{1}{2}x^T Q x - d^T x + v^T (Bx - b) \tag{3.29}$$

Dans le cas où la matrice Q est inversible, la fonction optimiser $f(x)$ est dite *fonction canonique*, donc, pour chaque vecteur $\widetilde{v} \in \mathbb{R}^m_+$ on peut définir la fonction duale canonique $f^d_v(\widetilde{v})$ avec l'utilisation de *la transformation duale canonique* comme suit :

$$\begin{aligned} f^d_v(\widetilde{v}) &= \max_{x \in \mathbb{R}^n} L(x, \widetilde{v}) \\ &= \max_{x \in \mathbb{R}^n} \left(\frac{1}{2}x^T Q x - d^T x + \widetilde{v}^T (Bx - b)\right) \\ &= \operatorname*{Ext}_{x \in \mathbb{R}^n} \left(\frac{-1}{2}\left(d - B^T\widetilde{v}\right)^T \left(Q^{-1}\right)\left(d - B^T\widetilde{v}\right) - b^T\widetilde{v}\right) \end{aligned} \tag{3.30}$$

où cette fonction est bien définie sur $\mathbb{R}^m_+$.

On peut alors définir le problème dual sous la forme suivante :

$$\begin{cases} \operatorname{Ext} f^d_v(\widetilde{v}) \\ \widetilde{v} \geq 0 \end{cases}$$

ou sous la forme :

$$\begin{cases} \operatorname{Ext}\left(\operatorname*{Ext}_{x \in \mathbb{R}^n}\left(\frac{-1}{2}\left(d - B^T\widetilde{v}\right)^T (Q^{-1})\left(d - B^T\widetilde{v}\right) - b^T\widetilde{v}\right)\right) \\ \widetilde{v} \geq 0 \end{cases} \tag{p^d_v}$$

Il y a une relation très forte entre les deux problèmes *(P)* et (P_v^d) sera expliquée dans le lemme suivant :

Lemme 3.3 : Le problème primal *(P)* et le problème dual (P_v^d) chaque un présente le problème dual canonique de l'autre, dans le sens que sont admet le même ensemble des points de *(KKT)* notées par $(\overline{x}, \overline{v})$, et on a aussi :

$$f_v^d(\overline{v}) = L(\overline{x}, \overline{v}) = f(\overline{x})$$

On peut résumer les conditions extrêmes dans le théorème suivant :

Théorème 3.2 :[5] *Soit Q une matrice avec l'indice $i_d > 0$ admet $\{q_i\}_{i=\overline{1,p}}$ valeurs propres distincts dans l'ordre suivant :*

$$q_1 < q_2 < \cdots < q_{id} < 0 \leq q_{id+1} < \cdots < q_p \text{ telle que } p \leq n$$

et soit $\left(\overline{\varepsilon_i^*}, \overline{\rho_i^*}\right)$ est un point de (KKT) *pour le problème dual (P_μ^d) pour un paramètre $\mu > 0$, et posons aussi :*

$$\overline{x_i} = \left(Q + \overline{\rho_i^* I}\right)^{-1}\left(d - B^T \overline{\varepsilon_i^*}\right)$$

donc, on a :

(1) *Si $\overline{\rho_i^*} > -q_1$ alors, le vecteur $\left(\overline{\varepsilon_i^*}, \overline{\rho_i^*}\right)$ est un maximum de $f^d(y^*)$ sur $Y_{\mu^+}^*$si et seulement si le vecteur $\overline{x_i}$ est un minimum de $f(x)$ sur $X_{\mu s}$, et on a aussi :*

$$f(\overline{x_i}) = \min_{x \in X_{\mu S}} f(x) = \max_{\left(\varepsilon_i^*, \rho_i^*\right) \in Y_{\mu+}^*} f^d(\varepsilon_i^*, \rho_i^*) = f^d\left(\overline{\varepsilon_i^*}, \overline{\rho_i^*}\right) \qquad (3.31)$$

(2) *Si $0 \leq \overline{\rho_i^*} < -q_{id}$ alors, le vecteur $\left(\overline{\varepsilon_i^*}, \overline{\rho_i^*}\right)$ est un maximum de $f^d(y^*)$ sur $Y_{\mu-}^*$ si et seulement si $\overline{x_i}$ est un maximum global de $f(x)$ sur $X_{\mu-}$, et on a aussi :*

$$f(\overline{x_i}) = \max_{x \in X_{\mu S}} f(x) = \max_{\left(\varepsilon_i^*, \rho_i^*\right) \in Y_{\mu-}^*} f^d(\varepsilon_i^*, \rho_i^*) = f\left(\overline{\varepsilon_i^*}, \overline{\rho_i^*}\right) \qquad (3.32)$$

(3) *Si $0 < \rho_i^* < -q_{id}$ alors, le point de*(KKT) $\left(\overline{\varepsilon_i^*}, \overline{\rho_i^*}\right)$ *est un minimum de $f^d(y^*)$ sur $Y_{\mu i}^*$ (l'ensemble ouverte) si et seulement si $\overline{x_i}$ est un minimum de $f(x)$ sur$X_{\mu s}$, et* *on a :*

$$f\left(\overline{x_i}\right)=\min_{x\in X_{\mu S}} f\left(x\right)=\min_{\left(\varepsilon_i^*,\rho_i^*\right)\in Y_{\mu i}^*} f^d\left(\varepsilon_i^*,\rho_i^*\right)=f^d\left(\overline{\varepsilon_i^*},\overline{\rho_i^*}\right) \tag{3.33}$$

Pour bien démontrer Le théorème presque précédent, il faut utiliser le lemme suivant :

Lemme 3.4 :[18] *Supposons que le vecteur* $(\overline{x},\overline{v})$ *est un point de* (KKT) *pour le problème de la programmation quadratique original, alors on a :*

(1) *Si la matrice Q est définie positive alors,* le théorème de *saddle* min-max est donné par :

$$\min_{x\in X_f}\max_{v\geq 0} L\left(x,v\right)=L\left(\overline{x},\overline{v}\right)=\max_{v\geq 0}\min_{x\in\mathbb{R}^n} L\left(x,v\right) \tag{3.34}$$

(2) *Si la matrice Q est définie négative alors,* le théorème de sup-*maximum* prend la forme :

$$\max_{x\in X_f}\max_{v\geq 0} L\left(x,v\right)=L\left(\overline{x},\overline{v}\right)=\max_{v\geq 0}\max_{x\in\mathbb{R}^n} L\left(x,v\right) \tag{3.35}$$

d'autre part, le théorème de sup-*minimum* est donné par :

$$\min_{x\in X_f}\max_{v\geq 0} L\left(x,v\right)=L\left(\overline{x},\overline{v}\right)=\min_{v\geq 0}\max_{x\in\mathbb{R}^n} L\left(x,v\right) \tag{3.36}$$

On peut alors démontrer le théorème (**3.2**).

Preuve : D'après le théorème préçédent, on a déjà vue que le vecteur :

$$\overline{y_i^*}=\left(\overline{\varepsilon_i^*},\overline{\rho_i^*}\right)\in Y_\mu^*$$

est un point de *(KKT)* pour le problème (P_μ^d) si et seulement si le vecteur :

$$\overline{x_i}=\left(Q+\overline{\rho_i^*}I\right)^{-1}\left(d-B^T\overline{\varepsilon_i^*}\right)$$

est un point de *(KKT)* pour le problème paramétrique (P_μ) et :

$$f\left(\overline{x_i}\right)=L\left(\overline{x_i},\overline{\varepsilon_i^*},\overline{\rho_i^*}\right)=f^d\left(\overline{\varepsilon_i^*},\overline{\rho_i^*}\right)=E\left(\overline{x_i},\overline{\varepsilon_i^*},\overline{\rho_i^*}\right) \tag{3.37}$$

où $E\left(\overline{x_i},\overline{\varepsilon_i^*},\overline{\rho_i^*}\right)$ est la fonction de complementarité total associée au problème (P_μ), cette fonction est définie par :

$$E\left(\overline{x_i},\overline{\varepsilon_i^*},\overline{\rho_i^*}\right)=\frac{1}{2}\overline{x_i}^T\left(Q+\overline{\rho_i^*}I\right)\overline{x_i}-\overline{W}^{\#}\left(\overline{\varepsilon_i^*},\overline{\rho_i^*}\right)+\left(B\overline{x_i}\right)^T\overline{\varepsilon_i^*}-\overline{x_i}^Td$$

En effet, on a :

$$\begin{aligned}E\left(\overline{x_i},\overline{\varepsilon_i^*},\overline{\rho_i^*}\right) &= \frac{1}{2}\overline{x_i}^TQ\overline{x_i}+\frac{1}{2}\overline{\rho_i^*}\overline{x_i}^T\overline{x_i}-c^T\overline{\varepsilon_i^*}-\mu\overline{\rho_i^*}+\left(B\overline{x_i}\right)^T\overline{\varepsilon_i^*}-\overline{x_i}^Td\\ &= \frac{1}{2}\overline{x_i}^TQ\overline{x_i}-d^T\overline{x_i}\\ &= f(\overline{x_i})=L\left(\overline{x_i},\overline{\varepsilon_i^*},\overline{\rho_i^*}\right)\end{aligned}$$

donc :

(1) Si $\overline{\rho_i^*}>-q_1$ alors la matrice $(Q+\rho_i^*I)$ sera définie positive, d'autre part, la fonction duale canonique $f^d\left(\varepsilon_i^*,\rho_i^*\right)$ est donnée par :

$$\begin{aligned}f^d\left(\varepsilon_i^*,\rho_i^*\right) &= \frac{-1}{2}\left(d-B^T\varepsilon_i^*\right)^T\left(Q+\rho_i^*I\right)^{-1}\left(d-B^T\varepsilon_i^*\right)-c^T\varepsilon_i^*-\mu\rho_i^*\\ &= E\left(\overline{x_i},\varepsilon_i^*,\rho_i^*\right)\end{aligned}$$

Cette fonction est concave en chaque composante $\left\{\varepsilon_i^*: i=\overline{1,m}\right\},\rho_i^*\in\mathbb{R}$, et à cause de ça, on peut dire que la *fonction Lagrangien* est convexe en $x\in\mathbb{R}^n$, par contre, elle est concave en chaque composante ε^*, ρ^*, donc, on a :

$$
\begin{aligned}
f^d\left(\overline{\varepsilon_i^*}, \overline{\rho_i^*}\right) &= \max_{(\varepsilon^*,\rho^*)\in Y_{\mu+}^*} f^d\left(\varepsilon^*, \rho^*\right) \\
&= \max_{(\varepsilon^*,\rho^*)\in Y_{\mu+}^*} \min_{x\in\mathbb{R}^n} E\left(x, \varepsilon^*, \rho^*\right) \\
&= \max_{\rho^*>-q_1} \max_{\varepsilon^*\geq 0} \min_{x\in\mathbb{R}^n} \left\{\frac{1}{2}x^T\left(Q+\rho^* I\right)x + \left(Bx-c\right)^T\varepsilon^* - \mu\rho^* - x^T d\right\} \\
&= \max_{\rho^*>-q_1} \min_{x\in\mathbb{R}^n} \max_{\varepsilon^*\geq 0} \left\{\frac{1}{2}x^T\left(Q+\rho^* I\right)x + \left(Bx-c\right)^T\varepsilon^* - \mu\rho^* - x^T d\right\} \\
&= \max_{\rho^*>-q_1} \min_{x\in\mathbb{R}^n} \left\{\frac{1}{2}x^T\left(Q+\rho^* I\right)x - \mu\rho^* - d^T x\right\} \text{ tq } Bx \leq c \\
&= \max_{\rho^*>-q_1} \min_{x\in\mathbb{R}^n} \left\{\frac{1}{2}x^T Q x + \frac{1}{2}\rho^* x^T x - \mu\rho^* - d^T x\right\} \\
&= \min_{x\in\mathbb{R}^n} \left\{f(x) + \max_{\rho^*>-q_1}\left\{\frac{1}{2}\rho^* x^T x - \mu\rho^*\right\}\right\}
\end{aligned}
$$

Le problème suivant :

$$\max_{\rho^*>-q_1}\left\{\frac{1}{2}\rho^* x^T x - \mu\rho^*\right\}$$

est un problème linéaire, admet une solution dans l'intervalle ouvert $]-q_1, +\infty[$ si et seulement si on a la condition de *(KKT)* suivante :

$$0 \leq \overline{\rho^*} \perp \frac{1}{2}\overline{x}^T\overline{x} - \mu \leq 0$$

qui implique que :

$$\frac{1}{2}x^T x = \mu$$

on peut alors écrire que :

$$
\begin{aligned}
f^d\left(\overline{\varepsilon_i^*}, \overline{\rho_i^*}\right) &= \min_{x\in\mathbb{R}^n}\left\{f(x) + \max_{\rho^*\succ -q_1}\left\{\frac{1}{2}\rho^* x^T x - \mu\rho^*\right\}\right\} \\
&= \min_{x\in\mathbb{R}^n} f(x)
\end{aligned}
$$

et d'après le théorème précédent, on trouve que :

$$f\left(\overline{x_i}\right) = f^d\left(\overline{\varepsilon_i^*}, \overline{\rho_i^*}\right)$$

alors, on peut écrire :

$$f\left(\overline{x_i}\right)=\min_{x\in\mathbb{R}^n} f(x)=\max_{(\varepsilon^*,\rho^*)\in Y^*_{\mu+}} f^d\left(\varepsilon^*,\rho^*\right)=f^d\left(\overline{\varepsilon_i^*},\overline{\rho_i^*}\right)$$

(2) Si :

$$0\le\overline{\rho_i^*}<-q_{id}$$

la matrice $\left(Q+\overline{\rho_i^*}I\right)$ est définie négative, et à cause de ça, la fonction Lagrangien $E\left(\overline{x_i},\varepsilon_i^*,\rho_i^*\right)$ est concave en $x\in\mathbb{R}^n$, et concave en :

$$\varepsilon_i^*\in\mathbb{R}^m_+,\ \text{ et } \rho_i^*\in[0,-q_{id})$$

donc, si $\left(\overline{\varepsilon_i^*},\overline{\rho_i^*}\right)$ est un maximum global de la fonction $f^d\left(\varepsilon_i^*,\rho_i^*\right)$ sur $Y^*_{\mu-}$ alors, d'après le théorème de sup-*duality* de Lagrange, et le lemme (?) on a :

$$\begin{aligned}
f^d\left(\overline{\varepsilon_i^*},\overline{\rho_i^*}\right) &= \max_{(\varepsilon^*,\rho^*)\in Y^*_{\mu-}} f^d\left(\varepsilon^*,\rho^*\right)\\
&= \max_{\rho^*\in[0,-q_1)}\max_{\varepsilon^*\ge 0}\max_{x\in\mathbb{R}^n} E\left(x,\varepsilon^*,\rho^*\right)\\
&= \max_{\rho^*\in[0,-q_1)}\max_{x\in\mathbb{R}^n}\max_{\varepsilon^*\ge 0}\left\{\frac{1}{2}x^T\left(Q+\rho^*I\right)x+\left(Bx-c\right)^T\varepsilon^*-\mu\rho^*-d^Tx\right\}\\
&= \max_{\rho^*\in[0,-q_1)}\max_{x\in\mathbb{R}^n}\left\{\frac{1}{2}x^TQx+\frac{1}{2}\rho^*x^Tx-\mu\rho^*-d^Tx\right\}\ \text{ tq } Bx\le c\\
&= \max_{x\in\mathbb{R}^n}\left\{f(x)+\max_{\rho^*\in[0,-q_1)}\left\{\rho^*\left(\frac{1}{2}x^Tx-\mu\right)\right\}\right\}\\
&= \max_{x\in X_\mu}\left\{f(x)\right\}\ \text{ telle que } \frac{1}{2}x^Tx\le\mu
\end{aligned}$$

Le problème suivant :

$$\max_{\rho^*\in[0,-q_1)}\left\{\rho^*\left(\frac{1}{2}x^Tx-\mu\right)\right\}$$

admet une solution si et seulement si $\left(\frac{1}{2}x^Tx\le\mu\right)$, parce que le domaine $[0,-q_1)$ est bornée inférieurement, et pour cette solution on a :

$$\max_{\rho^*\in[0,-q_1)}\left\{\rho^*\left(\frac{1}{2}x^Tx-\mu\right)\right\}=0$$

D'autre part, on a :

$$f(\overline{x_i}) = E\left(\overline{x_i}, \varepsilon_i^*, \rho_i^*\right) = f^d\left(\overline{\varepsilon_i^*}, \overline{\rho_i^*}\right)$$

Alors, on écrit :

$$f(\overline{x_i}) = \max_{x \in X_\mu} f(x) = \max_{(\varepsilon^*, \rho^*) \in Y_{\mu-}^*} f^d\left(\varepsilon^*, \rho^*\right) = f^d\left(\overline{\varepsilon_i^*}, \overline{\rho_i^*}\right)$$

(3) Si :

$$0 < \overline{\rho_i^*} < -q_{id} \text{ et } \left(\overline{\varepsilon_i^*}, \overline{\rho_i^*}\right)$$

est un minimum global de la fonction $f^d\left(y^*\right)$ sur $Y_{\mu i}^*$, alors, d'après le théorème de sup-*duality de Lagrange* on a :

$$\begin{aligned}
f^d\left(\overline{\varepsilon_i^*}, \overline{\rho_i^*}\right) &= \min_{(\varepsilon^*, \rho^*) \in Y_{\mu i}} f^d\left(\varepsilon^*, \rho^*\right) \\
&= \min_{\varepsilon^* \geq 0} \min_{\rho^* \in (0, -q_{id})} \max_{x \in \mathbb{R}^n} E\left(x, \varepsilon^*, \rho^*\right) \\
&= \min_{\rho^* \in (0, -q_{id})} \min_{\varepsilon^* \geq 0} \max_{x \in \mathbb{R}^n} E\left(x, \varepsilon^*, \rho^*\right) \\
&= \min_{\rho^* \in (0, -q_{id})} \min_{x \in \mathbb{R}^n} \max_{\varepsilon^* \geq 0} E\left(x, \varepsilon^*, \rho^*\right) \\
&= \min_{\rho^* \in (0, -q_{id})} \min_{x \in \mathbb{R}^n} \max_{\varepsilon^* \geq 0} \left\{\frac{1}{2} x^T \left(Q + \rho^* I\right) x + (Bx - c)^T \varepsilon^* - \mu\rho^* - d^T x\right\} \\
&= \min_{\rho^* \in (0, -q_{id})} \min_{x \in \mathbb{R}^n} \left\{\frac{1}{2} x^T \left(Q + \rho^* I\right) x - \mu\rho^* - d^T x\right\} \text{ tq } Bx \leq c \\
&= \min_{x \in \mathbb{R}^n} \left\{f(x) + \min_{\rho^* \in (0, -q_{id})} \left\{\rho^* \left(\frac{1}{2} x^T x - \mu\right)\right\}\right\} \text{ tq } Bx \leq c
\end{aligned}$$

Le problème de la minimisation linéaire suivant :

$$\min_{\rho^* \in (0, -q_{id})} \left\{\rho^* \left(\frac{1}{2} x^T x - \mu\right)\right\}$$

admet une solution sur le domaine ouvert $(0, -q_{id})$ si et seulement si :

$$\frac{1}{2} x^T x = \mu$$

alors :

$$\begin{aligned} f^d\left(\overline{\varepsilon_i^*}, \overline{\rho_i^*}\right) &= \min_{x\in\mathbb{R}^n} f(x) \\ &= \min_{x\in X_\mu} f(x) \end{aligned}$$

d'autre part, et pour tout points de *(KKT)* associée au problème paramétrique (P_μ),on a :

$$f^d\left(\overline{\varepsilon_i^*}, \overline{\rho_i^*}\right) = f(x_i)$$

Et on a :

$$f\left(\overline{x_i}\right) = \min_{x\in X_\mu} f(x) = \min_{(\varepsilon^*,\rho^*)\in Y_{\mu i}^*} f^d\left(\varepsilon^*, \rho^*\right) = f^d\left(\overline{\varepsilon_i^*}, \overline{\rho_i^*}\right)$$

sachons que la concavité de la fonction Lagrangien en chaque composante ε^*et ρ^* ne signifie pas la concavité en (ε^*, ρ^*) comme un vecteur, donc, la preuve est complète. ■

Remarque 3.3 : Remarquons que :

(1) Si $\overline{x_i}$ les extremums de (P_μ) dans l'ensemble non convexe $X_{\mu s}$, alors la solution duale :

$$\overline{\rho_i^*} > -q_1 > 0$$

est un point intérieur du domaine réalisable dual Y_μ^*, parce que pour :

$$\overline{\rho_i^*} > -q_1 > 0$$

La matrice $\left(Q + \overline{\rho_i^*} I\right)$ est définie positive, alors le problème dual (P_μ^d) est un problème du *maximisation d'une fonction convexe*, donc les solutions sont des points intérieurs de Y_μ^* (i,e : le premier cas dans la démonstration de le théorème préçédent).

(2) Dans le cas préçédent, les minimums primals $\overline{x_i}$ ne sont pas toujours des points critiques de la fonction objectif primale ; par contre, les solutions duaux $\overline{\rho_i^*}$ sont des points critiques pour la fonction objectif duale, cette résultat est vérifiée par l'équation :

$$\frac{1}{2}\left(d - B^T\overline{\varepsilon^*}\right)^T \left(Q + \overline{\rho^*}I\right)^{-2}\left(d - B^T\overline{\varepsilon^*}\right) = \mu$$

Se dernier sera résout complètement par l'algorithme *MATHEMATICA*, et à cause de ça, on dira que le problème primal appartient à l'ensemble des problèmes **NP-defficile**.

(3) Le problème primal paramétrique (P_μ) est équivalent au problème original *(P)*, d'autre part, si l'ensemble réalisable (X_f) est bornée alors, on peut prendre le paramètre $\mu = \mu_\circ($ le rayon de $(X_f))$ et à cause de ça, on écrit :

$$(X_f) = (X_\mu)$$

Au plus de ça, le vecteur :

$$\overline{x_i} = \left(Q + \overline{\rho_i^* I}\right)^{-1}\left(d - B^T\overline{\varepsilon_i^*}\right)$$

le minimum de (P_μ) devient une solution (minimum) de problème original (P), mais n'est pas nécessairement globale, parce que, le problème dual paramétrique (P_μ^d) peut admet plusieurs points de *(KKT)* $\left\{\overline{\rho_i^*}\right\}$.

(4) Si l'ensemble réalisable (X_f) n'est pas bornée, alors, le problème paramétrique (P_μ) sera utilisée pour trouver les minimums globaux dans (X_f) qui n'appartiennent pas à (X_μ) ; d'autre façons, si l'ensemble (X_f) n'est pas bornée alors, le problème original (P) ne peut être existe, et d'un côté physique, le problème *(P)* ne peut être bien définie (proposer).

3.4 Exemples3.2 :

3.4.1 Exemple I :

Considérons le problème de la programmation quadratique sur un ensemble convexe sous la forme suivante :

$$\begin{cases} \min f(x_1, x_2) = \frac{1}{2}(q_1 x_1^2 + q_2 x_2^2) - d_1 x_1 - d_2 x_2 \\ \frac{1}{2}x_1 + x_2 \leq 1, x_1 \geq 0, x_2 \geq 0, \frac{1}{2}(x_1^2 + x_2^2) \leq 2 \end{cases}$$

l'espace réalisable (X_f) sera donnée par :

$$(X_f) = \left\{ x = (x_1, x_2) \in \mathbb{R}^2 : Bx \leq c, \frac{1}{2}\left(x_1^2 + x_2^2\right) \leq 2 \right\}$$

où :

$$Q = \begin{pmatrix} q_1 & 0 \\ 0 & q_2 \end{pmatrix} \in \mathbb{R}^{2\times 2}, B = \begin{pmatrix} \frac{1}{2} & 1 \\ -1 & -1 \end{pmatrix} \in \mathbb{R}^{2\times 2}, c = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \in \mathbb{R}^2, x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$$

si on utilise la norme $\|.\|_2$ alors, on trouve que le rayon de (X_f) est donnée par :

$$r_\circ = \frac{1}{2}|x|^2 = \frac{1}{2}\left(x_1^2 + x_2^2\right) = 2$$

donc :

(1) Pour $q_1 \leq 0, q_2 \leq 0$ alors, la matrice $Q \in \mathbb{R}^{2\times 2}$ est définie négative où q_1, q_2 sont les valeurs propres de cette matrice, cette résultat implique que le problème préçédent est non convexe avec une fonction objectif concave, donc, notre problème presente un minimisation concave sur un ensemble convexe (bornée), et à cause de ça, le minimum global se produit nécessairement sur la frontière de (X_f) (les sommers). La résolution de ce type de problèmes est très difficile, parce que notre problème est *NP-defficile*, donc on applique *la méthode de la transformation duale canonique* comme suit :

trouver le problème dual canonique associée à se problème, le problème est sous la forme :

$$\begin{cases} \max f^d\left(\varepsilon_1^*, \varepsilon_2^*, \rho^*\right) = \frac{-1}{2}\left(d - B^T\varepsilon^*\right)^T \left(Q + \rho^* I\right)^{-1}\left(d - B^T\varepsilon^*\right) - \mu\rho^* - c^T\varepsilon^* \\ \varepsilon_1^* \geq 0; \varepsilon_2^* \geq 0; \rho^* \geq 0 \end{cases}$$

il suffit de choisir :

$$\rho^* \geq -\min\left\{q_1, q_2\right\}$$

alors avec :

$$B^T = \begin{pmatrix} \frac{1}{2} & -1 \\ 1 & -1 \end{pmatrix}, c^T = (1, 0), \varepsilon^* = \begin{pmatrix} \varepsilon_1^* \\ \varepsilon_2^* \end{pmatrix} \in \mathbb{R}^2$$

on trouve que :

$$\begin{aligned} f^d\left(\varepsilon_1^*, \varepsilon_2^*, \rho^*\right) &= \frac{-1}{2}\left(d - B^T\varepsilon^*\right)^T \left(Q + \rho^* I\right)^{-1}\left(d - B^T\varepsilon^*\right) - \mu\rho^* - c^T\varepsilon^* \\ &= \frac{-1}{2}\left[\frac{\left(d_1 - \frac{1}{2}\varepsilon_1^* + \varepsilon_2^*\right)^2}{\left(q_1 + \rho^*\right)} + \frac{\left(d_2 - \varepsilon_1^* + \varepsilon_2^*\right)^2}{\left(q_2 + \rho^*\right)}\right] - \mu\rho^* - \varepsilon_1^* \end{aligned}$$

donc, le problème dual canonique associée à se type des problèmes est donnée par :

$$\begin{cases} \max f^d\left(\varepsilon_1^*, \varepsilon_2^*, \rho^*\right) = \frac{-1}{2}\left[\frac{\left(d_1 - \frac{1}{2}\varepsilon_1^* + \varepsilon_2^*\right)^2}{\left(q_1 + \rho^*\right)} + \frac{\left(d_2 - \varepsilon_1^* + \varepsilon_2^*\right)^2}{\left(q_2 + \rho^*\right)}\right] - \mu\rho^* - \varepsilon_1^* \\ \varepsilon_1^* \geq 0; \varepsilon_2^* \geq 0; \rho^* \geq -\min\left\{q_1, q_2\right\} \end{cases}$$

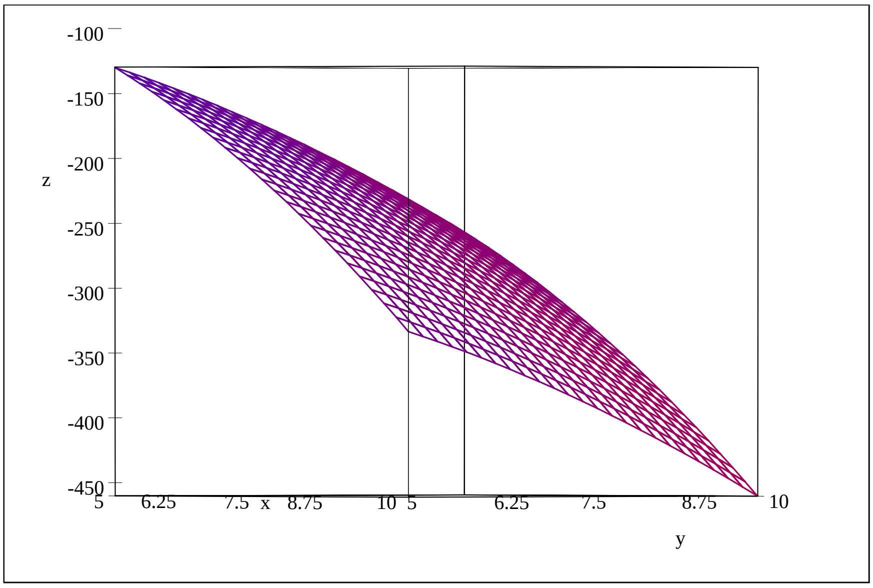
Si on prend :

$$Q = \begin{pmatrix} -5 & 0 \\ 0 & -3 \end{pmatrix} \in \mathbb{R}^{2\times 2}, d = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} \in \mathbb{R}^2$$

alors, d'après la remarque, et le fait que l'ensemble réalisable (X_μ) est convexe bornée on a :

$$\mu = \mu_\circ = \frac{1}{2}r_\circ^2 = 2 \text{ où } r_\circ = 2$$

$$\frac{1}{2}\left(-5x^2 - 3y^2\right) - 3x - 3y$$



donc, le problème dual sera donné par :

$$\begin{cases} \max f^d\left(\varepsilon_1^*, \varepsilon_2^*, \rho^*\right) = \frac{-1}{2}\left[\frac{\left(3-\frac{1}{2}\varepsilon_1^*+\varepsilon_2^*\right)^2}{(\rho^*-5)} + \frac{\left(3-\varepsilon_1^*+\varepsilon_2^*\right)^2}{(\rho^*-3)}\right] - 2\rho^* - \varepsilon_1^* \\ \varepsilon_1^* \geq 0; \varepsilon_2^* \geq 0; \rho^* \geq 5 \end{cases}$$

où :

$$\left(\rho^* \geq -\min\{-3, -5\}\right) \Longleftrightarrow \left(\rho^* \geq 5\right)$$

ce dernier problème admet une solution unique dans l'ensemble $Y_{\mu i}^* \subset Y_{\mu -}^*$, cette solution sera trouver par l'algorithme *MATHEMATICA.*

Finalement, les solutions primales seront données par :

$$\overline{x} = \left(Q + \overline{\rho^*}I\right)^{-1}\left(d - B^T\overline{\varepsilon^*}\right)$$

(2) Pour $q_1 \leq 0, q_2 \geq 0$, la matrice $Q \in \mathbb{R}^{2\times 2}$ est indéfinie où q_1, q_2 sont les valeurs propres de cette matrice, cette résultat implique que le problème préçédent est non convexe avec une fonction objectif indéfinie, donc, notre problème presente un minimisation indéfinie sur un ensemble convexe (bornée), et à cause de ça, le minimum global se produit nécessairement sur la frontière de (X_f).

Comme on a déjà dit dans le premièr point que la résolution de ce type de problèmes est aussi très difficile, parce que notre problème appartient à l'ensemble des problèmes *NP-defficile*, donc, on applique *la méthode de la transformation duale canonique* comme suit :
trouver le problème dual canonique associée à se problème qui donné par :

$$\begin{cases} \max f^d\left(\varepsilon_1^*, \varepsilon_2^*, \rho^*\right) = \frac{-1}{2}\left[\frac{\left(d_1 - \frac{1}{2}\varepsilon_1^* + \varepsilon_2^*\right)^2}{(q_1+\rho^*)} + \frac{\left(d_2 - \varepsilon_1^* + \varepsilon_2^*\right)^2}{(q_2+\rho^*)}\right] - \mu\rho^* - \varepsilon_1^* \\ \varepsilon_1^* \geq 0; \varepsilon_2^* \geq 0; \rho^* \geq -\min\left\{q_1, q_2\right\} \end{cases}$$

si on prend :

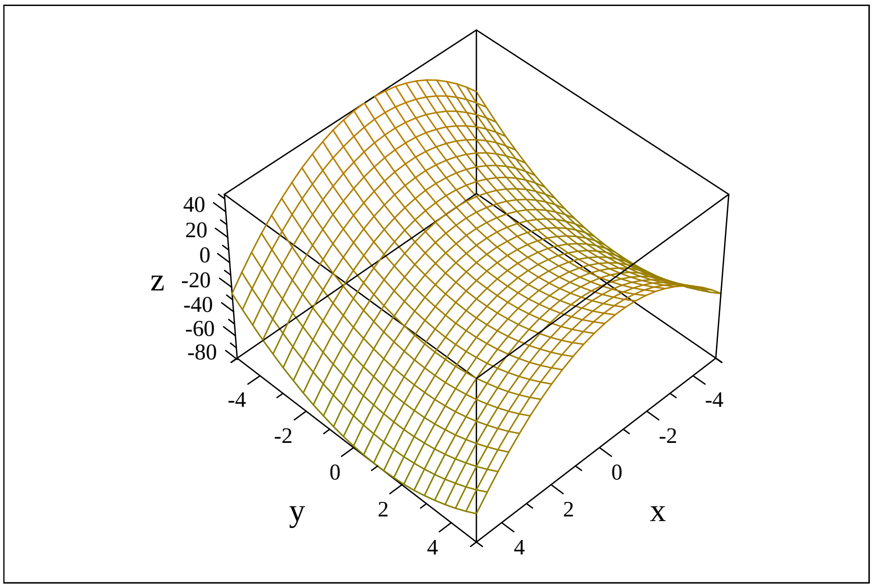
$$Q = \begin{pmatrix} -5 & 0 \\ 0 & 3 \end{pmatrix} \in \mathbb{R}^{2\times 2}, d = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} \in \mathbb{R}^2$$

alors, d'après la remarque, et le fait que l'ensemble réalisable (X_μ) est convexe bornée on a :

$$\mu = \mu_\circ = \frac{1}{2}r_\circ^2 = 2, \text{ où } r_\circ = 2$$

la fonction primale est plotée par :

$$f(x,y) = \frac{1}{2}\left(-5x^2 + 3y^2\right) - 3x - 3y$$



Graphe de $f(x_1, x_2)$ dans un plan 3 dimension rectangulaire

donc, le problème dual sera donné par :

$$\begin{cases} \max f^d(\varepsilon_1^*, \varepsilon_2^*, \rho^*) = \frac{-1}{2}\left[\frac{\left(3-\frac{1}{2}\varepsilon_1^*+\varepsilon_2^*\right)^2}{(\rho^*-5)} + \frac{\left(3-\varepsilon_1^*+\varepsilon_2^*\right)^2}{(\rho^*+3)}\right] - 2\rho^* - \varepsilon_1^* \\ \varepsilon_1^* \geq 0; \varepsilon_2^* \geq 0; \rho^* \geq 5 \end{cases}$$

où :

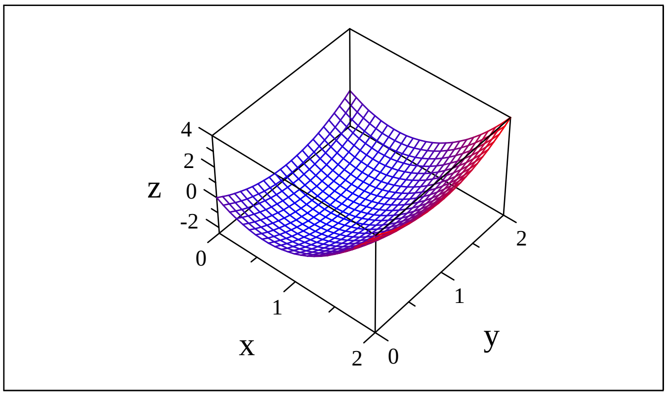
$$(\rho^* \geq -\min\{3, -5\}) \Longleftrightarrow (\rho^* \geq 5)$$

ce dernier problème admet une solution unique dans l'ensemble $Y_{\mu i}^* \subset Y_{\mu -}^*$, cette solution sera trouver par l'algorithme *MATHEMATICA*, et à cause de ça, le problème original admet une solution unique donnée par :

$$\overline{x} = \left(Q + \overline{\rho^*} I\right)^{-1}\left(d - B^T \overline{\varepsilon^*}\right)$$

(3) pour $q_1 \geq 0, q_2 \geq 0$, la matrice $Q \in \mathbb{R}^{2\times 2}$ est définie positive, cette résultat implique que le problème préçédent est un problème de la minimisation convexe sur un ensemble convexe (bornée), et à cause de ça, le calcule du minimum global est facile avec les algorithmes polynomiales ou bien par l'algorithme *MATHEMATICA*.

$$\frac{1}{2}\left(5x^2 + 3y^2\right) - 3x - 3y$$



3.4.2 Exemple II :

Soit le problème suivant :

$$\begin{cases} \min f(x) = \frac{1}{2}ax^2 - dx \\ |x| \leq r \end{cases}$$

où r est un nombre réel positif, et $a \in \mathbb{R}$; $x \in \mathbb{R}$;

donc, on a les cas suivantes :

(1) *pour* $a < 0$ alors, le problème préçédent présent une minimisation concave, et à cause de ça, la solution globale se produit sur des points extrêmes (sommet) suivantes :

$$\overline{x} = \pm r$$

d'autre part, le paramètre μ est définie dans se cas par :

$$\mu = \frac{1}{2}r^2$$

avec l'utilisation de la transformation duale canonique, la forme duale canonique est sous la forme suivante :

$$\begin{cases} \max f^d(\rho^*, \varepsilon^*) = \frac{-1}{2}\left(d - B^T\varepsilon^*\right)^T (Q + \rho^* I)^{-1}\left(d - B^T\varepsilon^*\right) - \mu\rho^* - c^T\varepsilon^* \\ \rho^* \geq 0; \varepsilon^* \geq 0; \det(Q + \rho^* I) \neq 0 \end{cases}$$

avec :

$$B^T = 0; Q = (a) \in \mathbb{R}; c^T = 0 \in \mathbb{R}$$

donc, s'implique :

$$\begin{aligned} f^d(\rho^*) &= \frac{-1}{2} d (a + \rho^*)^{-1} d - \mu\rho^* \\ &= \frac{-1}{2} \left(\frac{d^2}{(a + \rho^*)} \right) - \mu\rho^* \end{aligned}$$

où :

$$(a + \rho^*) > 0 \text{ définie positive}$$

puisque la dimension $n = 1$, la dérivée d'ordre un de la fonction $f^d(\rho^*)$ est donnée par :

$$\left(f^d\right)'(\rho^*) = \frac{1}{2} \left(\frac{d^2}{(a + \rho^*)^2} \right) - \mu$$

donc, l'équation duale algébrique admet deux racines données par :

(a) $\overline{\rho_1^*} > -a$ est un maximum unique de $f^d(\rho^*)$ sur $Y^*_{\mu+}$, qu'est équivalent de dire que le vecteur définie par :

$$\overline{x_1} = \left(\frac{d}{(a + \overline{\rho_1^*})} \right)$$

est un minimum pour la fonction $f(x)$ sur $X_{\mu s}$, et d'après *le théorème 3.2* on a :

$$f(\overline{x_1}) = \min_{x \in X_{\mu S}} f(x) = \max_{(\rho^*) \in Y^*_{\mu+}} f^d(\rho^*) = f^d\left(\overline{\rho_1^*}\right)$$

(b) $\overline{\rho_2^*} < -a$ est un minimum local pour $f^d(\rho^*)$, qui bien dite que le vecteur définie par :

$$\overline{x_2} = \left(\frac{d}{(a + \overline{\rho_2^*})} \right)$$

est un minimum pour $f(x)$, et d'après *le théorème 3.2* ona :

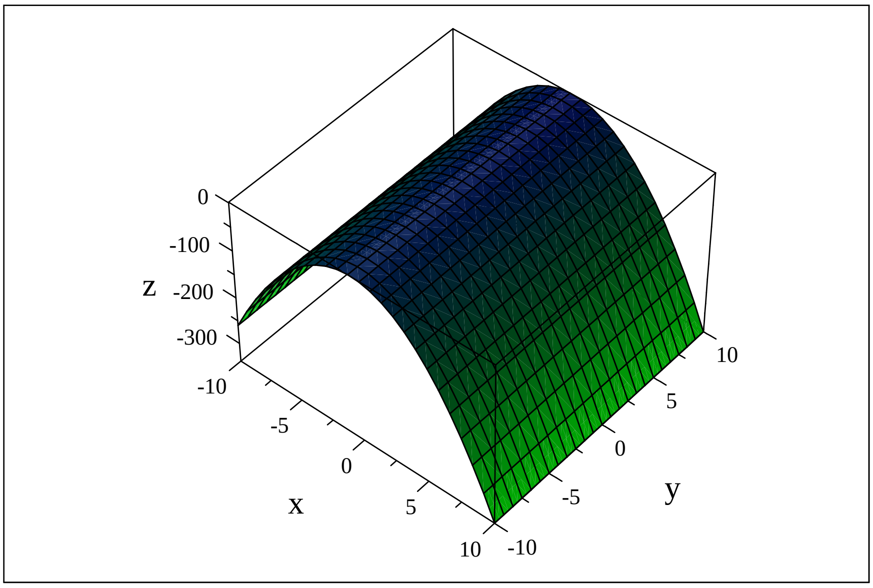
$$f(\overline{x_2}) = \min_{x \in X_{\mu S}} f(x) = \min_{\rho^* \in Y^*_{\mu i}} f^d(\rho^*) = f^d\left(\overline{\rho_2^*}\right)$$

(2) pour $a \geq 0$ alors, le problème préçédent représente une minimisation convexe, qui est plus facile pour résoudre.

Application : prend par exemple la forme suivante :

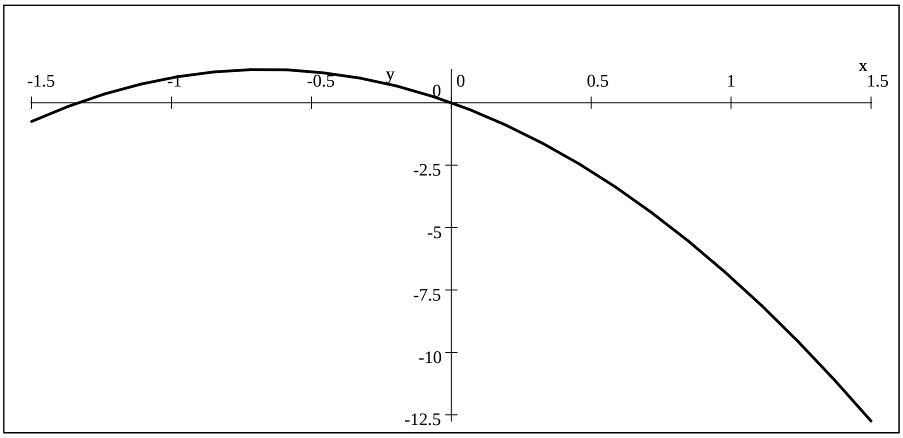
$$\begin{cases} \min f(x) = -3x^2 - 4x \\ |x| \le 1.5 \end{cases}$$

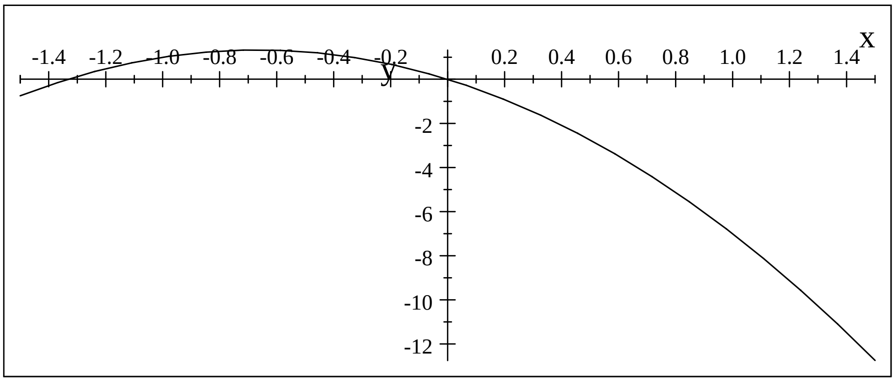
$$-3x^2 - 4x$$



où :

$$a = -6; d = 4; r = 1.5$$





$\boxed{\textit{la fonction primal}}$

Candidate(s) for extrema : $\left\{\frac{4}{3}\right\} : \frac{4}{3} = 1.3333$, at $\left\{\left[x = -\frac{2}{3}\right]\right\} = \{[x = -0.66667]\}$

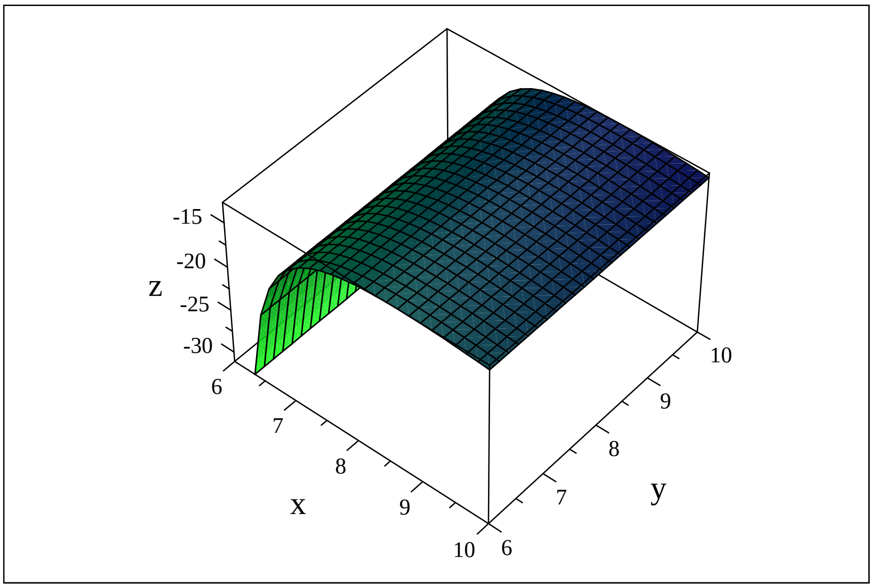
avec l'utilisation de la transformation duale canonique, la forme duale canonique est donnée par :

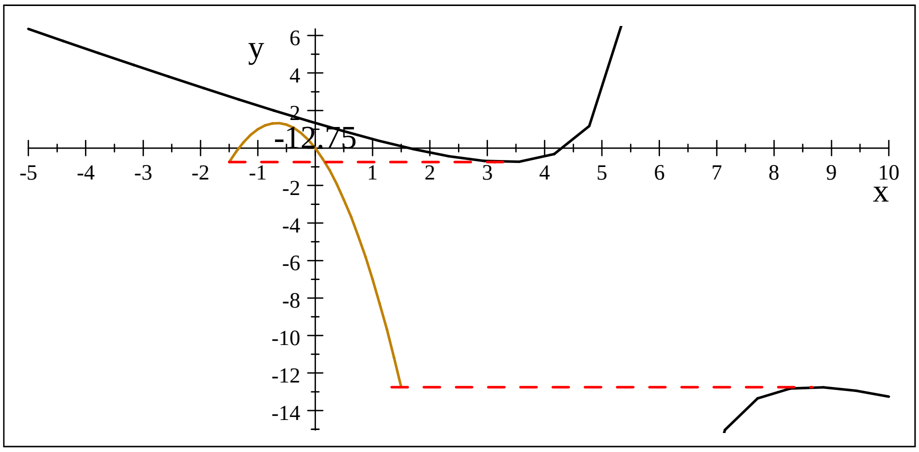
$$\begin{aligned} f^d(\rho^*) &= \frac{-1}{2} d\,(a+\rho^*)^{-1}\,d - \mu\rho^* \\ &= \frac{-1}{2}(16)(-6+\rho^*)^{-1} - \frac{1}{2}(1.5)^2\rho^* \\ &= -8(\rho^*-6)^{-1} - 1.125\rho^* \\ &= -\left(1.125\rho^* + \left(\frac{8}{\rho^*-6}\right)\right) \end{aligned}$$

et dans se cas, le problème dual canonique est définie par :

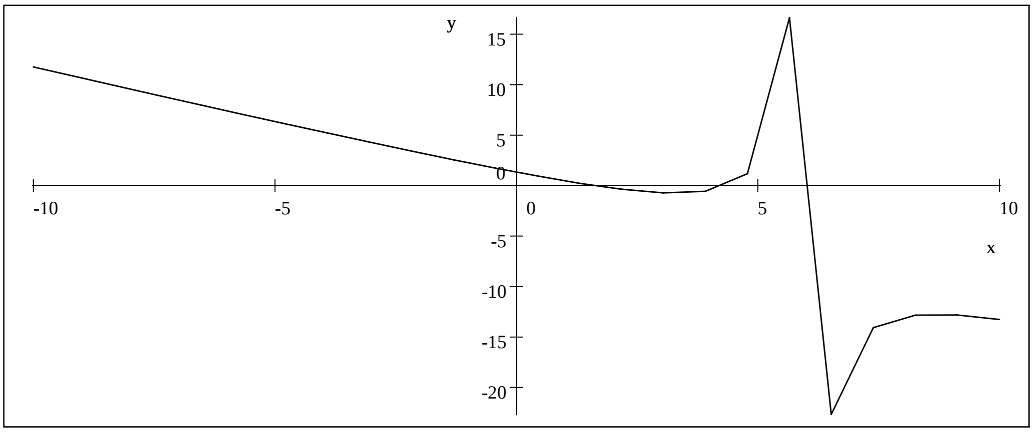
$$\left\{ \begin{array}{c} \max f^d(\rho^*) = -\left(1.125\rho^* + \left(\frac{8}{\rho^*-6}\right)\right) \\ \rho^* \geq -(a) = 6 \end{array} \right.$$

$$-\left(1.125\rho^* + \left(\frac{8}{\rho^*-6}\right)\right)$$





Candidate(s) for extrema : $\{-12.75, -0.75\}$, at $\{[\rho_1^* = 3.3333], [\rho_2^* = 8.6667]\}$



donc, on peut poser le tableau suivant :

functions	extrema	candidate(s) for extrema
primal	$-0,66667$	$1,3333$
dual	$3,3333$	$-0,75$
	$8,6667$	$-12,75$

les solutions primaux sont :

$$\begin{aligned}
\overline{x_1} &= \left(a+\overline{\rho_1^*}\right)^{-1} d \\
&= (-6+3.3333)^{-1} 4 \\
&= \frac{-4}{3.3333} \\
&= 1.4998
\end{aligned}$$

et :

$$\begin{aligned}
\overline{x_2} &= \left(a+\overline{\rho_2^*}\right)^{-1} d \\
&= \frac{4}{-6+8.6667} \\
&= \frac{4}{2.6667} \\
&= 1.5000
\end{aligned}$$

extrema duaux $\overline{\rho_i^*}$	les solutions $\overline{x_i}$	valeur primal $f\left(\overline{x_i}\right)$	valeur dual $f^d\left(\overline{\rho_i^*}\right)$
3.3333	-1.4998	-0.749	-0.75
8.6667	1.5000	-12.75	-12.75

d'après le tableau préçédent, il est bien claire que :

$$f\left(\overline{x_1}\right) = f^d\left(\overline{\rho_1^*}\right) = -12.75$$

et :

$$f\left(\overline{x_2}\right) = f^d\left(\overline{\rho_2^*}\right) = -0.75$$

et d'après les deux graphes préçédent, on a :

pour $\overline{\rho_1^*} > -a = 6$ *:*

$$f\left(\overline{x_1}\right) = \min_{x \in X_{\mu S}} f(x) = \max_{(\rho^*) \in Y^*_{\mu +}} f^d(\rho^*) = f^d\left(\overline{\rho_1^*}\right) = -12.75$$

et pour $\overline{\rho_2^*} < -a = 6$:

$$f\left(\overline{x_2}\right) = \min_{x \in X_{\mu S}} f(x) = \min_{\rho^* \in Y^*_{\mu i}} f^d(\rho^*) = f^d\left(\overline{\rho_2^*}\right) = -0.75$$

3.4.3 Exemple III :

$$f(x) = \frac{1}{2} x^T Q x - d^T x$$

où :

$$Q = \begin{pmatrix} -5 & 0 & 0 & 0 \\ 0 & -2.5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}, d^T = (3, 4, -2, 1), \mu = 2$$

La forme duale est donnée par :

$$\begin{aligned} f^d(\rho^*) &= \frac{-1}{2} d^T (Q + \rho^* I)^{-1} d - \mu \rho^* \\ &= -2.0\rho^* - \frac{8.0}{\rho^* - 2.5} - \frac{2.0}{\rho^* + 1.0} - \frac{9.0}{2.0\rho^* - 10.0} - \frac{1.0}{2.0\rho^* + 8.0} \end{aligned}$$

$$\begin{pmatrix} -5 + \rho^* & 0 & 0 & 0 \\ 0 & -2.5 + \rho^* & 0 & 0 \\ 0 & 0 & 1 + \rho^* & 0 \\ 0 & 0 & 0 & 4 + \rho^* \end{pmatrix}$$

, inverse : $\begin{pmatrix} \frac{1}{\rho-5} & 0 & 0 & 0 \\ 0 & \frac{1}{\rho-2.5} & 0 & 0 \\ 0 & 0 & \frac{1}{\rho+1} & 0 \\ 0 & 0 & 0 & \frac{1}{\rho+4} \end{pmatrix}$

$$\frac{-1}{2}(3,4,-2,1)\begin{pmatrix} \frac{1}{\rho-5} & 0 & 0 & 0 \\ 0 & \frac{1}{\rho-2.5} & 0 & 0 \\ 0 & 0 & \frac{1}{\rho+1} & 0 \\ 0 & 0 & 0 & \frac{1}{\rho+4} \end{pmatrix}\begin{pmatrix} 3 \\ 4 \\ -2 \\ 1 \end{pmatrix}$$

$$f^d(\rho^*) = \left(-\frac{2}{\rho^*+1} - \frac{1}{2(\rho^*+4)} - \frac{9}{2(\rho^*-5)} - \frac{8}{\rho^*-2.5}\right) - 2\rho^*$$

$$\frac{\partial f^d}{\partial \rho^*}(\rho^*) = -2.0 + \frac{8.0}{(\rho-2.5)^2} + \frac{2.0}{(\rho+1.0)^2} + \frac{18.0}{(2.0\rho-10.0)^2} + \frac{2.0}{(2.0\rho+8.0)^2}$$

Solution is : $\{[\rho = -2.1966]\}$;

Solution is : $\mathbb{R}\cap(\rho_1 \setminus \{-4.0,-1.0,2.5,5.0\})$ where ρ_1 is a root of $19620.0\hat{Z} - 44475.\hat{Z}^{2.0} - 24430.0\hat{Z}^{3.0} + 9286.0\hat{Z}^{4.0} + 2704.0\hat{Z}^{5.0} - 692.0\hat{Z}^{6.0} - 80.0\hat{Z}^{7.0} + 16.0\hat{Z}^{8.0} - 29825.$

, Solution is : $\{[Z = -4.5536]\}$

donc, on a trouvé que :

$$\begin{aligned} \rho_1 &= -4.5536 \\ \rho_2 &= -2.1966 \end{aligned}$$

on peut vérifier que :

$$\begin{aligned} & \left(-2.0 + \frac{8.0}{(\rho-2.5)^2} + \frac{2.0}{(\rho+1.0)^2} + \frac{18.0}{(2.0\rho-10.0)^2} + \frac{2.0}{(2.0\rho+8.0)^2}\right) \div \\ & (\rho^*+4.5536)(\rho^*+2.1966) \\ = \; & (\rho^*+3.39287)(\rho^*-6.72415) \end{aligned}$$

alors, les quatre solutions sont :

$$\begin{pmatrix} \overline{\rho_1^*} = 6.72415 \\ \overline{\rho_2^*} = -2.1966 \\ \overline{\rho_3^*} = -3.39287 \\ \overline{\rho_4^*} = -4.5536 \end{pmatrix}$$

et puisque :

$$\overline{\rho_1^*} = 6.72415 > -q_1 = 5$$

alors, d'après le théorème de base préçédent, il y a une solution primale globale définie par :

$$\begin{aligned} \overline{x} &= \left(Q + \overline{\rho_1^*} I\right)^{-1} d \\ &= \begin{pmatrix} \frac{1}{6.72415-5} & 0 & 0 & 0 \\ 0 & \frac{1}{6.72415-2.5} & 0 & 0 \\ 0 & 0 & \frac{1}{6.72415+1} & 0 \\ 0 & 0 & 0 & \frac{1}{6.72415+4} \end{pmatrix} \begin{pmatrix} 3 \\ 4 \\ -2 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 1.7400 \\ 0.94694 \\ -0.25893 \\ 9.3247 \times 10^{-2} \end{pmatrix} \end{aligned}$$

c'est à dire que, la solution global est le vecteur définie par :

$$\overline{x}^T = \left(1.7400, \quad 0.94694, \quad -0.25893, \quad 9.3247 \times 10^{-2}\right)$$

Chapitre 4

La méthode de Séparation et interpolation.

Introduction :

Soit le problème quadratique indéfinie suivant :

$$\begin{cases} \min f(x) = \frac{1}{2}x^T Q x^T - d^T x \\ \qquad Bx \leq c, x \geq 0 \end{cases} \qquad (PQI)$$

où Q est une matrice indéfinie, $B \in \mathbb{R}^{n+m}, x, d \in \mathbb{R}^n$.

On peut le inclue dans l'ensemble des problèmes *NP–difficile* ; bien que beaucoup d'algorithmes de la programmation non linéaire soient capable de trouver les optimums locaux de problème **(PQI),** mais il n'y a pas (en général) aucun critère locale qui peut vérifier l'identification entre la solution locale et la solution globale (si elle existe).

Entre ces algorithmes, il ya *l'algorithme de la programmation concave*, qui est basée sur la structure propre de la matrice indéfinie Q, il va de soi que cette algorithme est permet de transformer notre problème à un problème d'un minimisation convexe plus facile pour résoudre.

D'autre part, on a connue que les solutions de ce type des problèmes si elles existent serons produit sur des points de la frontière pour le domaine réalisable pas nécessairement des sommets, par contre les problèmes où la fonction objectif est concave ou bien quasi concave, comme elle était bien montrée par le théorème suivant :

Théorème 4.1 :[13] *La solution optimale du* ***(PQI)*** *se produit sur quelque points de frontière du domaine réalisable* Ω *telle que :*

$$\Omega = \{x \in \mathbb{R}^n : Ax \leq c; x \geq 0\} \tag{4.1}$$

Preuve : Posons $x^* \in \Omega$ être la solution globale de **(PQI)** et supposons que x^*est un point intérieur de Ω, d'autre part, et comme la matrice Q est indéfinie, il existe un valeur propre négatif λ, et $\widetilde{x}$ le vecteur propre associée à λ, donc on a :

$$\nabla f(x^*) = Qx^* - d = 0$$

alors, il existe $\varepsilon > 0$ telle que :

$$y = x^* + \varepsilon\widetilde{x} \in \Omega$$

donc, il va de soi que :

$$\begin{aligned}
f(x^* + \varepsilon\widetilde{x}) &= \frac{1}{2}(x^* + \varepsilon\widetilde{x})^T Q(x^* + \varepsilon\widetilde{x}) - d^T(x^* + \varepsilon\widetilde{x}) \\
&= \frac{1}{2}(x^*)^T Q(x^* + \varepsilon\widetilde{x}) + \frac{1}{2}\varepsilon(\widetilde{x})^T Q(x^* + \varepsilon\widetilde{x}) - d^T x^* - \varepsilon d^T\widetilde{x} \\
&= \frac{1}{2}(x^*)^T Qx^* + \frac{1}{2}\varepsilon(x^*)^T Q\widetilde{x} + \frac{1}{2}\varepsilon(\widetilde{x})^T Qx^* + \frac{1}{2}\varepsilon^2(\widetilde{x})^T Q\widetilde{x} - d^T x^* - \varepsilon d^T\widetilde{x} \\
&= f(x^*) - \varepsilon\left(d^T\widetilde{x} - \frac{1}{2}(\widetilde{x})^T Qx^* - \frac{1}{2}(x^*)^T Q\widetilde{x}\right) + \frac{1}{2}\varepsilon^2(\widetilde{x})^T Q\widetilde{x} \\
&= f(x^*) - \varepsilon\left(d^T\widetilde{x} - \frac{1}{2}(x^*)^T Q\widetilde{x} - \frac{1}{2}(x^*)^T Q\widetilde{x}\right) + \frac{1}{2}\varepsilon^2(\widetilde{x})^T Q\widetilde{x} \\
&= f(x^*) - \varepsilon\left(d^T\widetilde{x} - (x^*)^T Q\widetilde{x}\right) + \frac{1}{2}\varepsilon^2(\widetilde{x})^T Q\widetilde{x} \\
&= f(x^*) - \varepsilon\left(d^T\widetilde{x} - (Qx^*)^T\widetilde{x}\right) + \frac{1}{2}\varepsilon^2(\widetilde{x})^T Q\widetilde{x} \\
&= f(x^*) - \varepsilon\left(d^T\widetilde{x} - d^T\widetilde{x}\right) + \frac{1}{2}\varepsilon^2(\widetilde{x})^T Q\widetilde{x} \\
&= f(x^*) + \frac{1}{2}\varepsilon^2(\widetilde{x})^T Q\widetilde{x} \\
&= f(x^*) + \frac{1}{2}\lambda\varepsilon^2(\widetilde{x})^T\widetilde{x}
\end{aligned}$$

alors, avec :

$$\lambda < 0 \text{ et } \left(\widetilde{x}\right)^T \widetilde{x} = \|\widetilde{x}\|^2$$

on peut écrire :

$$f(x^* + \varepsilon\widetilde{x}) < f\left(x^*\right)$$

c'est une contradiction avec les données alors, la solution globale x^* sera produit sur quelque points de la frontière de Ω.

Exemple 4.1 :[13] Considérons le problème suivant :

$$\begin{cases} \min f(x) = x_1^2 - x_2^2 \\ \Omega = \{x \in \mathbb{R}^2 : -1 \leq x_1 \leq 1, -1 \leq x_2 \leq 1\} \end{cases}$$

qu'est la minimisation d'une fonction quadratique indéfinie sur un ensemble convexe (problème quadratique indéfinie).

On peut donner Ω sous la forme équivalente suivante :

$$\Omega = \left\{x \in \mathbb{R}^2 : Ax \leq c; x \geq 0\right\}$$

où :

$$Q = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, d = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, c = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Remarquons que :

$$f\left(-1, 1\right) = 0, f\left(-1, -1\right) = 0, f\left(1, -1\right) = 0, f\left(1, 1\right) = 0$$

et on a aussi :

$$f\left(-1, 0\right) = 1, f\left(1, 0\right) = 1, f\left(0, 1\right) = f\left(0, -1\right) = -1$$

donc, on a trouvée deux minimums globaux présentent comme deux points de la frontière de Ω (pas des sommets de Ω).

Le nombre des valeurs propres négatifs de la matrice Q peut être utilisé pour trouver le nombre des solutions locaux (globaux), comme elle était montrée par le théorème suivant :

Théorème 4.2 :[13] *Soit le problème définie par :*

$$\begin{cases} \min f(x) = \frac{1}{2}x^T Qx - d^T x \\ x \in \Omega = \{x \in \mathbb{R}^n : Ax \leq c\} \end{cases} \tag{PQN}$$

où $Q_{n\times m}$ est une matrice indéfinie, $A_{n\times m}$ est une matrice quelconque donc, si la matrice $Q_{n\times m}$ a (s) valeurs propres négatifs, alors, on a au moins (s) contraintes seront actives en chaque point solution locale (globale).

4.1 Description de La méthode :

On peut transformer un problème quadratique à un autre problème où la fonction objective est séparable (par l'utilisation de la transformation affine), où on obtient un problème de la programmation quadratique indéfinie avec des contraintes linéaires définie par :

$$\begin{cases} \min \Psi(x, y) = f(x) - g(y) \\ Ax + By + c \leq 0, x \geq 0, y \geq 0 \end{cases} \tag{$PQNS$}$$

l'ensemble réalisable associer à se problème est donnée par :

$$\Omega = \left\{(x, y) \in \mathbb{R}^{n+k}; Ax + By + c \leq 0, x \geq 0, y \geq 0\right\} \tag{4.2}$$

où :

$$f(x) = d^T x - \frac{1}{2}x^T Qx$$
$$g(y) = b^T y$$

avec :

Q est une matrice réel symétrique et indéfinie appartient à $\mathbb{R}^{n\times n}$ et :

$$A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{m\times k}, c \in \mathbb{R}^m, b \in \mathbb{R}^k, d \in \mathbb{R}^n$$

Généralement, on peut poser le problème **(PQI)** comme suit :

$$\begin{cases} \min \Psi(x,y) = f(x) - g(y) \\ Ax + By + c \leq 0, x \in X, y \in Y \end{cases}$$

où :

$$X, Y \text{ sont deux polyèdres dans } \mathbb{R}^n \text{ et } \mathbb{R}^k \text{ respectivement}$$

$$f(x) : \text{est une fonction quadratique indéfinie}$$

$$g(y) : \text{est une fonction linéaire, avec } k \text{ est plus large que } n$$

4.1.1 La transformation de la forme quadratique avec l'utilisation de sa structure propre :

Comme un résultat dans l'algèbre linéaire, la matrice Q admet n valeurs propres $\lambda_1, \lambda_2, ..., \lambda_n$, et $\{\zeta_i\}_{i=\overline{1,n}}$ sont les vecteurs propres associées ; d'autre part, supposons que la matrice Q admet s valeurs propres positives et $(n-s)$ valeurs propres négatives, alors on peut écrire que :

$$Q = \zeta D \zeta^T$$

où :

$$\xi = (\zeta_1, \zeta_2, ..., \zeta_n),$$

$$D = diag(\lambda_1, \lambda_2, ..., \lambda_n) \text{ avec } D_{ij} = \sigma_{ij} = \begin{cases} 0 \text{ si } i \neq j \\ 1 \text{ ailleurs} \end{cases}$$

d'autre part, et pour chaque $1 \leq i \leq n$ on va définir :

$$\begin{aligned} \underline{\xi_i} &= \min \left\{ \zeta_i^T x : (x,y) \in \Omega \right\} \\ \overline{\xi_i} &= \max \left\{ \zeta_i^T x : (x,y) \in \Omega \right\} \end{aligned} \qquad (MCR)$$

sont des n-problèmes linéaires[7].

Finalement, on définie la borne :

$$\beta_i = \overline{\xi_i} - \underline{\xi_i} : 1 \leq i \leq n \qquad (4.3)$$

si on return au problème original **(PQNS)**, il suffit de prendre :

$$B = 0, g(y) = 0$$

alors, on trouve que :

$$\Psi(x, y) = f(x)$$

donc, on peut former le problème **(PQNS)** comme suit :

La forme séparable équivalente sera donnée par :

$$\begin{cases} \min f(x) = d^T x - \frac{1}{2} x^T Q x \\ x \in \Omega \cap S_x \end{cases} \tag{SP}$$

où :

$$\begin{aligned} \Omega &= \{x \in \mathbb{R}^n : Ax \leq c\} \\ S_x &= \{x \in \mathbb{R}^n : 0 \leq x_i \leq \beta_i : 1 \leq i \leq n\} \end{aligned} \tag{4.4}$$

d'autre part, on donnera la fonction $f(x)$ sous la forme suivante :

$$f(x) = \sum_{i=1}^{n} \theta_i(x_i) \tag{4.5}$$

avec :

$$\theta_i(x_i) = d_i x_i - \frac{1}{2} \lambda_i x_i^2 : 1 \leq i \leq n \tag{4.6}$$

La structure séparable de la fonction $f(x)$:

On va utiliser les valeurs propres de la matrice Q pour trouver la forme parallèle séparable de la fonction non linéaire *(indéfinie)* $f(x)$, d'autre terme, on va essayer de partager la fonction $f(x)$ à deux parties *une partie concave* et l'autre *partie sera convexe.*

Définition de la partie concave : La partie concave est définie pour $\{\lambda_i \geq 0 : 1 \leq i \leq s\}$ comme suit :

$$f_1(x) = \sum_{i=1}^{s} \theta_i(x_i) = \sum_{i=1}^{s} \left(d_i x_i - \frac{1}{2} \lambda_i x_i^2 \right) \tag{4.7}$$

Définition de la partie convexe : La partie convexe est définie pour $\{\lambda_i \leq 0 : s+1 \leq i \leq n\}$ comme suit :

$$f_2(x) = \sum_{i=s+1}^{n} \theta_i(x_i) = \sum_{i=s+1}^{n} \left(d_i x_i - \frac{1}{2}\lambda_i x_i^2\right) \tag{4.8}$$

Remarque 4.1 : Comme on a trouvées les deux formes, on va essayer d'approximer la partie concave avec une fonction affine pour simplifier notre forme séparable, et pour trouver une forme plus simple, facile pour résoudre, de notre fonction objectif, et avec cette forme on va ramenées *la résolution d'un problème quadratique indéfinie* à *une résolution d'un problème d'optimisation linéaire (convexe).*

4.1.2 L'approximation linéaire et l'erreurs limites :

L'approximation linéaire :

Considérons le problème (SP), on a déjà vu que :

$$f(x) = \sum_{i=1}^{n} \theta_i(x_i)$$

donc, on donnera la définition suivante :

Définition 4.1 : L'enveloppe convexe de la fonction $f(x)$ sur le rectangle S_x notée par $\Gamma(x)$ est la plus grande minorante affine continue minore la fonction $f(x)$ sur S_x.

D'après la définition préçédente on peut déduire l'enveloppe convexe de la partie concave :

$$f_1(x) = \sum_{i=1}^{s} \theta_i(x_i) = \sum_{i=1}^{s} \left(d_i x_i - \frac{1}{2}\lambda_i x_i^2\right)$$

comme suit :

Définition 4.2 :[13] Les enveloppes convexes des fonctions $\{\theta_i(x_i) : 1 \leq i \leq n\}$ notées par $\{\gamma_i(x_i) : 1 \leq i \leq n\}$ sur les intervalles $\{[0, \beta_i] : 1 \leq i \leq n\}$ sont des fonctions affines coïncident avec les fonctions :

$$\{\theta_i(x_i) ; 1 \leq i \leq n\}$$

en :

$$x_i = 0 \text{ et } x_i = \beta_i$$

ses fonction sont appelées aussi *l'interpolation linéaire de* $\theta_i(x_i)$ *en* 0*et* β_i.

Définition 4.3 :[13] L'enveloppe convexe de la fonction concave $f_1(x)$ sur le rectangle S_x sera donnée par :

$$\Gamma(x) = \sum_{i=1}^{s} \gamma_i(x_i) = \sum_{i=1}^{s} \left(d_i - \frac{1}{2}\lambda_i \beta_i\right) x_i \qquad (4.9)$$

Initialement, l'approximation linéaire de la partie concave est obtenue, et à cause de ça, on peut trouver une forme convexe (linéaire) final enduite par le problème original **(PQI)** comme suit :

$$\begin{cases} \min\left\{f_\Gamma(x) = \Gamma(x) + f_2(x)\right\} \\ \qquad x \in \Omega \cap S_x \end{cases} \qquad (PQNA)$$

l'erreurs limites :

Il va de soi que la limite inférieure des valeurs optimales de la fonction optimiser $f(x)$ (la fonction objectif de **(PS)**) sera donnée par la valeur optimale du programme linéaire **(PQIA)**, alors on a le théorème suivant :

Théorème 4.3 :[13] *On a :*

$$|f(x) - f_\Gamma(x)| \leq \frac{1}{8} \sum_{i=1}^{s} \lambda_i \beta_i^2 \qquad (4.10)$$

Preuve : Par définition on a :

$$\begin{aligned}
f(x) &= f_1(x) + f_2(x) \\
f_\Gamma(x) &= \Gamma(x) + f_2(x) \\
f_1(x) &= \sum_{i=1}^{s} \theta_i(x_i) = \sum_{i=1}^{s} \left(d_i x_i - \frac{1}{2}\lambda_i x_i^2 \right) \\
f_2(x) &= \sum_{i=s+1}^{n} \theta_i(x_i) = \sum_{i=s+1}^{n} \left(d_i x_i - \frac{1}{2}\lambda_i x_i^2 \right) \\
\Gamma(x) &= \sum_{i=1}^{s} \gamma_i(x_i) = \sum_{i=1}^{s} \left(d_i - \frac{1}{2}\lambda_i \beta_i \right) x_i
\end{aligned}$$

alors :

$$\begin{aligned}
|f(x) - f_\Gamma(x)| &= |f_1(x) + f_2(x) - [\Gamma(x) + f_2(x)]| \\
&= |f_1(x) - \Gamma(x)| \\
&= \left| \sum_{i=1}^{s} \theta_i(x_i) - \sum_{i=1}^{s} \gamma_i(x_i) \right| \\
&= \left| \sum_{i=1}^{s} (\theta_i(x_i) - \gamma_i(x_i)) \right| \\
&\leq \sum_{i=1}^{s} |\theta_i(x_i) - \gamma_i(x_i)|
\end{aligned}$$

d'autre part on a :

$$\begin{aligned}
|\theta_i(x_i) - \gamma_i(x_i)| &= \left| \left(d_i x_i - \frac{1}{2}\lambda_i x_i^2 \right) - \left(d_i - \frac{1}{2}\lambda_i \beta_i \right) x_i \right| \\
&= \left| \frac{1}{2}\lambda_i \beta_i x_i - \frac{1}{2}\lambda_i x_i^2 \right| \\
&= \left| \frac{1}{2}\lambda_i \left(\beta_i x_i - x_i^2 \right) \right| \\
&\leq \frac{1}{2}\lambda_i |x_i(\beta_i - x_i)|
\end{aligned}$$

sachons que le maximum de $|x_i(\beta_i - x_i)|$ tend vers le point médial $x_i = \frac{\beta_i}{2}$, pour tout :

$$0 \leq x_i \leq \beta_i, 1 \leq i \leq s,$$

donc :

$$\begin{aligned}
|\theta_i(x_i) - \gamma_i(x_i)| &\leq \frac{1}{2}\lambda_i \left|\frac{\beta_i}{2}\left(\beta_i - \frac{\beta_i}{2}\right)\right| \\
&\leq \frac{1}{8}\lambda_i\beta_i^2
\end{aligned}$$

alors :

$$\begin{aligned}
|f(x) - f_\Gamma(x)| &\leq \sum_{i=1}^{s} |\theta_i(x_i) - \gamma_i(x_i)| \\
&\leq \sum_{i=1}^{s} \left(\frac{1}{8}\lambda_i\beta_i^2\right) \\
&\leq \frac{1}{8}\sum_{i=1}^{s} \lambda_i\beta_i^2
\end{aligned}$$

alors, la preuve est complète. ■

Supposons que $f^* = f(x^*)$ la valeur optimale du **(PS)**, et soit $\widetilde{x}$ être la solution optimale de **(PQIA)**, alors *qu'elle est la relation entre la valeur optimale de **(PQIA)** et la valeur optimale* $f^* = f(x^*)$ *et* $f(\widetilde{x})$ *?*.La réponse sera donnée par la proposition suivante :

Proposition 4.1 :[13] *Si $\widetilde{x}$ est la solution du problème **(PQIA)**, et f^*est l'optimum global de **(PS)**, alors on a les deux relations suivantes :*

$$0 \leq f(\widetilde{x}) - f^* \leq \frac{1}{8}\sum_{i=1}^{s} \lambda_i\beta_i^2 \qquad (4.11)$$

$$f_\Gamma(\widetilde{x}) \leq f^* \leq f(\widetilde{x}) \qquad (4.12)$$

Preuve : D'après *le théorème4.1* on a :

$$f(x) - f_\Gamma(x) \leq \frac{1}{8}\sum_{i=1}^{s} \lambda_i\beta_i^2, \forall x \in \Omega \cap S_x$$

alors, pour $x = \widetilde{x}$ on a :

$$f(\widetilde{x}) - f_\Gamma(\widetilde{x}) \leq \frac{1}{8}\sum_{i=1}^{s} \lambda_i\beta_i^2$$

cette relation implique que :

$$f(\widetilde{x})-f^{*}+f^{*}-f_{\Gamma}\left(\widetilde{x}\right)\leq\frac{1}{8}\sum_{i=1}^{s}\lambda_{i}\beta_{i}^{2}$$

alors :

$$f(\widetilde{x})-f^{*}\leq\frac{1}{8}\sum_{i=1}^{s}\lambda_{i}\beta_{i}^{2}+f_{\Gamma}\left(\widetilde{x}\right)-f^{*}$$

c'est à dire que :

$$f(\widetilde{x})-f^{*}\leq\frac{1}{8}\sum_{i=1}^{s}\lambda_{i}\beta_{i}^{2}$$

parce que on a :

$$f_{\Gamma}\left(\widetilde{x}\right)-f^{*}\leq 0$$

d'autre part, et puisque f^{*}est l'optimum global, alors on a la relation suivante :

$$f^{*}\leq f\left(x\right),\forall x\in\Omega\cap S_{x}$$

donc, pour $x=\widetilde{x}$ on trouve :

$$f^{*}\leq f\left(\widetilde{x}\right)$$

c'est à dire que :

$$f\left(\widetilde{x}\right)-f^{*}\geq 0$$

alors, la relation $(\mathit{4.11})$ est satisfaite.

D'autre part, la fonction $f_{\Gamma}\left(x\right)$ est une minorante affine de la fonction $f\left(x\right)$, alors on peut écrire :

$$f_{\Gamma}\left(x\right)-f(x)\leq 0,\forall x\in\Omega\cap S_{x}$$

donc, pour $x=x^{*}$ on a :

$$f_{\Gamma}\left(x\right)-f(x^{*})\leq 0$$

alors, pour $x=\widetilde{x}$ on retient :

$$f_{\Gamma}\left(\widetilde{x}\right)\leq f^{*}$$

et d'après la relation (*4.11*) on a :

$$f(\widetilde{x}) \geq f^*$$

alors, la relation (*4.12*) est satisfaite. ■

Proposition 4.2 :[7] *Soit la forme $E(x)$ définie par :*

$$E(x) = f_1(x) - \Gamma(x) \tag{4.13}$$

où :

$$\Gamma(x) = \sum_{i=1}^{s} \gamma_i(x_i) = \sum_{i=1}^{s} \left(d_i - \frac{1}{2}\lambda_i \beta_i\right) x_i$$

*alors, si f^* est l'optimum global de **(PS)** et $\widetilde{x}$ est la solution de problème d'approximation **(PQNA),** on a la relation suivante :*

$$f(\widetilde{x}) - f^* \leq E(\widetilde{x}) \tag{4.14}$$

Preuve : D'après la proposition préçédente on a :

$$f_\Gamma(\widetilde{x}) \leq f^* \leq f(\widetilde{x})$$

donc :

$$f_\Gamma(\widetilde{x}) \leq f^* \leq f_1(\widetilde{x}) + f_2(\widetilde{x})$$

alors, on peut écrire :

$$\begin{aligned} f(\widetilde{x}) - f^* &= f_1(\widetilde{x}) + f_2(\widetilde{x}) - f^* \\ &= f_1(\widetilde{x}) - \Gamma(\widetilde{x}) + \Gamma(\widetilde{x}) + f_2(\widetilde{x}) - f^* \\ &= E(\widetilde{x}) + \Gamma(\widetilde{x}) + f_2(\widetilde{x}) - f^* \end{aligned}$$

d'autre part, on a :

$$f_\Gamma(\widetilde{x}) = \Gamma(\widetilde{x}) + f_2(\widetilde{x})$$

s'implique que :

$$\Gamma(\widetilde{x}) = f_\Gamma(\widetilde{x}) - f_2(\widetilde{x})$$

donc, on a :

$$\begin{aligned} f(\widetilde{x}) - f^* &= E(\widetilde{x}) + \Gamma(\widetilde{x}) + f_2(\widetilde{x}) - f^* \\ &= E(\widetilde{x}) + f_\Gamma(\widetilde{x}) - f_2(\widetilde{x}) + f_2(\widetilde{x}) - f^* \\ &= E(\widetilde{x}) + f_\Gamma(\widetilde{x}) - f^* \end{aligned}$$

et on a aussi que :

$$f_\Gamma(\widetilde{x}) - f^* \leq 0$$

alors :

$$\begin{aligned} f(\widetilde{x}) - f^* &= E(\widetilde{x}) + f_\Gamma(\widetilde{x}) - f^* \\ &\leq E(\widetilde{x}) \end{aligned}$$

alors, la preuve est complète. ■

Lemme 4.1 :[7] Si $E(\widetilde{x})$ est plus petite, alors $f(\widetilde{x})$ est une valeur approximative acceptée pour le minimum global f^*, avec $\widetilde{x}$ est le minimum global approximée.

D'après le lemme préçédent, on peut dire que notre travail est de trouver une limite de $E(\widetilde{x})$ relative au rang de $f_1(x)$ sur le rectangle S_x, dans ce cas, supposons que :

$$\begin{aligned} f_{1\max} &= \max\{f_1(x) : x \in S_x\} \\ f_{1\min} &= \min\{f_1(x) : x \in S_x\} \end{aligned} \tag{4.15}$$

alors, on a la définition suivante :

Définition 4.4 :[7] Le range de f_1 sur le rectangle S_x est l'intervalle $[f_{1\min}, f_{1\max}]$.

On a besoin d'une limite inférieure de la valeur :

$$\triangle f_1 = (f_{1\min} - f_{1\max}) \tag{4.16}$$

alors, et son perde de généralité, supposons que :

$$\lambda_1 \beta_1^2 > \lambda_i \beta_i^2 : \forall i = 1, 2, ..., s$$

d'autre part, définissons les valeurs :

$$\rho_i = \frac{\lambda_i \beta_i^2}{\lambda_1 \beta_1^2} \leq 1, \forall i = 1, 2, ..., s \tag{4.17}$$

par définition, les fonctions concaves $\theta_i(x_i)$ serons données par :

$$\begin{aligned} \theta_i(x_i) &= \left(d_i x_i - \frac{1}{2}\lambda_i x_i^2\right), \forall i = 1, 2, ..., s \\ &= \left(d_i - \frac{1}{2}\lambda_i x_i\right) x_i, \forall i = 1, 2, ..., s \end{aligned} \tag{4.18}$$

ces fonctions atteins leurs maximums au points :

$$\overline{x_i} = \frac{d_i}{\lambda_i}, \forall i = 1, 2, ..., s \tag{4.19}$$

alors, si $\overline{x_i} \in [0, \beta_i]$ alors la limite inférieure de $\triangle f_1$ dépende avec la distance :

$$\left|\overline{x_i} - \frac{\beta_i}{2}\right|, \forall i = 1, 2, ..., s$$

d'autre part,et avec $|\beta_i| \longrightarrow 1$ on peut exprimer la dépendance entre $\triangle f_1$ et $\left|\overline{x_i} - \frac{\beta_i}{2}\right|$ sous une autre forme comme suit :

$$\eta_i = \min\left\{1, \left|\overline{x_i} - \frac{\beta_i}{2}\right| = \left|\frac{2\overline{x_i}}{\beta_i} - 1\right|\right\} \tag{4.20}$$

alors, on observe que :

$$\eta_i = \left|\frac{2\overline{x_i}}{\beta_i} - 1\right| \text{ si } \overline{x_i} \in [0, \beta_i] \tag{4.21}$$

puisque si :

$$\eta_i = \left| \frac{2\overline{x_i}}{\beta_i} - 1 \right|$$

alors :

$$\frac{2\overline{x_i}}{\beta_i} - 1 < 1$$

s'implique que :

$$\frac{2\overline{x_i}}{\beta_i} < 2$$

c'est à dire :

$$0 \leq \overline{x_i} \leq \beta_i \tag{4.22}$$

D'après les données précédentes on a le lemme suivant :

Lemme 4.2 :[7] La relation suivante est satisfaite :

$$\triangle f_1 \geq \frac{1}{8}\lambda_1\beta_1 \sum_{i=1}^{s} \rho_i \left(1 + \eta_i\right)^2 =: \overline{\triangle} f_1 \tag{4.23}$$

Preuve : Pour $1 \leq i \leq s$ posons :

$$\begin{aligned} \overline{\theta_i} &= \max\left\{\theta_i\left(x_i\right) : 0 \leq x_i \leq \beta_i\right\} \qquad (4.24) \\ \underline{\theta_i} &= \min\left\{\theta_i\left(x_i\right) : 0 \leq x_i \leq \beta_i\right\} \end{aligned}$$

et on pose aussi :

$$\triangle\theta_i = \overline{\theta_i} - \underline{\theta_i} \tag{4.25}$$

on va étudier les quatre cas suivants :

$\boxed{\textit{PREMIER CAS}}$ Pour :

$$\overline{x_i} \in \left[0, \frac{\beta_i}{2}\right]$$

on a :

$$\overline{x_i} = \left(\frac{d_i}{\lambda_i}\right)$$

s'implique que :

$$d_i = \lambda_i \overline{x_i}$$

donc :

$$\begin{aligned}
\theta_i\left(\beta_i\right) &= \left(d_i\beta_i - \frac{1}{2}\lambda_i\beta_i^2\right) \\
&= \lambda_i\overline{x_i}\beta_i - \frac{1}{2}\lambda_i\beta_i^2 \\
&= \lambda_i\beta_i\left(\overline{x_i} - \frac{1}{2}\beta_i\right) \\
&\leq 0 \\
&= \theta_i\left(0\right)
\end{aligned}$$

et dans ce cas, on peut écrire :

$$\underline{\theta_i} = \min\left\{\theta_i\left(x_i\right) : 0 \leq x_i \leq \beta_i\right\} = \theta_i\left(\beta_i\right)$$

donc :

$$\overline{\theta_i} = \max\left\{\theta_i\left(x_i\right) : 0 \leq x_i \leq \beta_i\right\} = \theta_i\left(\overline{x_i}\right)$$

où :

$$\begin{aligned}
\theta_i\left(\overline{x_i}\right) &= \lambda_i\overline{x_i}^2 - \frac{1}{2}\lambda_i\overline{x_i}^2 \\
&= \frac{1}{2}\lambda_i\overline{x_i}^2
\end{aligned}$$

d'après la définition de η_i on trouve :

$$\begin{aligned}
\eta_i &= \left|\frac{2\overline{x_i}}{\beta_i} - 1\right| = \left|\overline{x_i} - \frac{\beta_i}{2}\right| \\
&= \frac{\beta_i}{2} - \overline{x_i}
\end{aligned}$$

parce que :

$$\overline{x_i} - \frac{\beta_i}{2} \leq 0$$

on peut alors écrire :

$$\eta_i = 1 - \frac{2\overline{x_i}}{\beta_i}$$

s'implique :

$$\overline{x_i} = \frac{\beta_i}{2}\left(1-\eta_i\right)$$

par conséquent :

$$\begin{aligned}
\triangle\theta_i &= \overline{\theta_i} - \underline{\theta_i} \\
&= \frac{1}{2}\lambda_i \overline{x_i}^2 - \theta_i\left(\beta_i\right) \\
&= \frac{1}{2}\lambda_i \overline{x_i}^2 - \lambda_i\beta_i\left(\overline{x_i} - \frac{1}{2}\beta_i\right) \\
&= \frac{1}{2}\lambda_i\left(\frac{\beta_i}{2}\left(1-\eta_i\right)\right)^2 - \lambda_i\beta_i\left(\frac{\beta_i}{2}\left(1-\eta_i\right) - \frac{1}{2}\beta_i\right) \\
&= \frac{1}{2}\lambda_i\frac{\beta_i^2}{4}\left(1-\eta_i\right)^2 - \frac{\lambda_i\beta_i^2}{2}\left(1-\eta_i\right) + \frac{\lambda_i\beta_i^2}{2} \\
&= \frac{1}{8}\lambda_i\beta_i^2\left(1-\eta_i\right)^2 - \frac{\lambda_i\beta_i^2}{2}\left(1-\eta_i\right) + \frac{\lambda_i\beta_i^2}{2} \\
&= \frac{1}{8}\lambda_i\beta_i^2\left[\left(1-\eta_i\right)^2 - 4\left(1-\eta_i\right) + 4\right] \\
&= \frac{1}{8}\lambda_i\beta_i^2\left[\left(1-\eta_i\right) - 2\right]^2 \\
&= \frac{1}{8}\lambda_i\beta_i^2\left(1+\eta_i\right)^2
\end{aligned}$$

$\boxed{\textit{DEUXIEME CAS}}$ *:* Pour :

$$\overline{x_i} \in \left[\frac{\beta_i}{2}, \beta_i\right]$$

on a :

$$\theta_i\left(\beta_i\right) = \lambda_i\beta_i\left(\overline{x_i} - \frac{1}{2}\beta_i\right) \geq 0$$

alors :

$$\underline{\theta_i} = \min\left\{\theta_i\left(x_i\right) : 0 \leq x_i \leq \beta_i\right\} = \theta_i\left(0\right) = 0$$

d'autre part, et d'après la définition de η_i on trouve que :

$$\overline{x_i} = \frac{\beta_i}{2}\left(1+\eta_i\right)$$

et :

$$\begin{aligned}\triangle\theta_i &= \overline{\theta_i} - \underline{\theta_i} \\ &= \theta_i\left(\overline{x_i}\right) \\ &= \frac{1}{2}\lambda_i\left(\frac{\beta_i}{2}\left(1+\eta_i\right)\right)^2 \\ &= \frac{1}{8}\lambda_i\beta_i^2\left(1+\eta_i\right)^2\end{aligned}$$

puisque on a :

$$\eta_i = \left|\overline{x_i} - \frac{\beta_i}{2}\right| \text{ avec } \overline{x_i} - \frac{\beta_i}{2} \geq 0$$

donc :

$$\eta_i = \overline{x_i} - \frac{\beta_i}{2}$$

$\boxed{\textit{TROISIEME CAS}}$ *:* Pour $\overline{x_i} < 0$; on a :

$$\eta_i = \min\left\{1, \left|\overline{x_i} - \frac{\beta_i}{2}\right|\right\} = 1$$

donc :

$$\overline{x_i} = \frac{d_i}{\lambda_i} < 0$$

s'implique que :

$$d_i < 0$$

alors, on peut trouver :

$$\begin{aligned}\overline{\theta_i} &= 0 \text{ et } \underline{\theta_i} = \theta_i\left(\beta_i\right) = \lambda_i\beta_i\left(\overline{x_i} - \frac{1}{2}\beta_i\right) \\ &\leq \frac{-1}{2}\lambda_i\beta_i^2\end{aligned}$$

où $\overline{x_i} \leq 0$ alors :

$$\begin{aligned}
\triangle\theta_i &= \overline{\theta_i} - \underline{\theta_i} \\
&\geq \frac{1}{2}\lambda_i\beta_i^2 \\
&= \frac{1}{8}\lambda_i\beta_i^2\left(1+1\right)^2 \\
&= \frac{1}{8}\lambda_i\beta_i^2\left(1+\eta_i\right)^2
\end{aligned}$$

$\boxed{\textit{QUATRIEME CAS}}$ *:* Pour $\overline{x_i} > \beta_i$, on a :

$$\underline{\theta_i} = 0 \text{ avec } \eta_i = 1$$

donc, on a :

$$\begin{aligned}
\triangle\theta_i &= \overline{\theta_i} - \underline{\theta_i} \\
&= \theta_i\left(\beta_i\right) - \theta_i\left(0\right) \\
&= \lambda_i\beta_i\left(\overline{x_i} - \frac{1}{2}\beta_i\right) \\
&\geq \lambda_i\beta_i\left(\beta_i - \frac{\beta_i}{2}\right) \\
&\geq \frac{1}{2}\lambda_i\beta_i^2 = \frac{1}{8}\lambda_i\beta_i^2\left(1+\eta_i\right)^2
\end{aligned}$$

finalement, d'après les quatre cas préçédents, on trouve que :

$$\triangle\theta_i \geq \frac{1}{8}\lambda_i\beta_i^2\left(1+\eta_i\right)^2$$

alors :

$$\begin{aligned}
\triangle f_1 &= \sum_{i=1}^{s}\triangle\theta_i \geq \frac{1}{8}\sum_{i=1}^{s}\lambda_i\beta_i^2\left(1+\eta_i\right)^2 \\
&\geq \frac{1}{8}\lambda_1\beta_1^2\sum_{i=1}^{s}\rho_i\left(1+\eta_i\right)^2
\end{aligned}$$

où ρ_i est le mesure de f_1 sur les intervalles $\{[0,\beta_i] : 1 \leq i \leq s\}$. ■

On peut trouver la borne supérieure de la fonction d'estimation $E(x)$ comme elle est expliquée dans le lemme suivant :

Lemme 4.3 :[7] On a :

$$E(x) \leq \frac{1}{8}\lambda_1\beta_1^2\sum_{i=1}^{s}\rho_i \tag{4.26}$$

Preuve : Par définition on a :

$$E(x) = f_1(x) - \Gamma(x)$$

donc, d'après *le théorème4.3* on a :

$$|f(x) - f_\Gamma(x)| \leq \frac{1}{8}\sum_{i=1}^{s}\lambda_i\beta_i^2$$

donc :

$$\begin{aligned}
|E(x)| &= |f_1(x) - \Gamma(x)| \\
&\leq \sum_{i=1}^{s}|\theta_i(x_i) - \gamma_i(x_i)| \\
&\leq \frac{1}{2}\sum_{i=1}^{s}\lambda_i\left|\frac{\beta_i}{2}\left(\beta_i - \frac{\beta_i}{2}\right)\right| \\
&\leq \frac{1}{8}\sum_{i=1}^{s}\lambda_i\beta_i^2 \\
&\leq \frac{1}{8}\lambda_1\beta_1\sum_{i=1}^{s}\rho_i
\end{aligned}$$

alors, la preuve est complète. ■

Le théorème suivant est très important :

Théorème 4.4 :[7] *On a la relation suivante :*

$$\frac{f(\widetilde{x}) - f^*}{\triangle f_1} \leq \delta(\rho,\eta) \text{ avec } \frac{1}{4} \leq \delta \leq 1 \tag{4.27}$$

Preuve : D'après *la proposition4.2* on a l'inégalité suivante :

$$f(\widetilde{x}) - f^* \leq E(\widetilde{x})$$

alors :

$$\begin{aligned}
\frac{f(\widetilde{x}) - f^*}{\triangle f_1} &\leq \frac{E(\widetilde{x})}{\triangle f_1} \\
&\leq \frac{\frac{1}{8}\lambda_1 \beta_1^2 \sum\limits_{i=1}^{s} \rho_i}{\frac{1}{8}\lambda_1 \beta_1^2 \sum\limits_{i=1}^{s} \rho_i (1+\eta_i)^2} \\
&\leq \frac{\sum\limits_{i=1}^{s} \rho_i}{\sum\limits_{i=1}^{s} \rho_i (1+\eta_i)^2} \\
&= \delta(\rho, \eta)
\end{aligned}$$

d'autre part, pour $\widetilde{x} = \frac{\beta_i}{2}$ on a :

$$\begin{aligned}
\eta_i &= \left| \frac{2\widetilde{x}_i}{\beta_i} - 1 \right| : 1 \leq i \leq s \\
&= 0
\end{aligned}$$

alors, il va de soi que :

$$\delta(\rho, \eta) = 1$$

généralement, on a :

$$\delta(\rho, \eta) \leq 1$$

et on aussi, pour :

$$\widetilde{x} \in [0, \beta_i] : 1 \leq i \leq s$$

et d'après les deux cas (1 et 2) dans la preuve de *le lemme4.2* on a :

Si $\widetilde{x}_i \leq 0$ alors $\eta_i = 1$, donc :

$$\begin{aligned}\delta(\rho,\eta) &= \frac{\sum\limits_{i=1}^{s}\rho_i}{\sum\limits_{i=1}^{s}\rho_i\left(1+\eta_i\right)^2} \\ &= \frac{1}{4}\end{aligned}$$

et si $\widetilde{x}_i \geq \beta_i$alors $\eta_i = 1$, donc on a aussi :

$$\delta(\rho,\eta) = \frac{1}{4}$$

alors, on a trouvée que :

$$\frac{1}{4} \leq \delta(\rho,\eta) \leq 1$$

. ■

4.2 La solution garantie de $\varepsilon-$approximation :

Sachons que la borne $\delta(\rho,\eta)$ n'est pas toujours la bonne borne, c'est pour ça on devise chaque intervalle $[0,\beta_i]$ à k_i sous-intervalles avec un pas égale à :

$$h_i = \left(\frac{\beta_i}{k_i}\right) \tag{4.28}$$

il va de soi que avec cette devisions, le rectangle S_x est devisée à :

$$\prod_{i=1}^{s} S_{ix} \tag{4.29}$$

sous-rectangles.

Le travail sera appliquée sur chaque sous-rectangles de S_{ix} comme il est appliquée sur S_x, donc :

pour $1 \leq i \leq s$, l'interpolation des fonctions $\theta_i(x_i)$ devient une interpolation au points :

$$x_i = jh_i \text{ avec } 0 \leq j \leq k_i \tag{4.30}$$

pour trouver les fonctions linéaires :

$$\gamma_i\left(x_i\right) \text{ sur } \left[0, \beta_i\right]$$

telle que :

$$\Gamma\left(x\right) = \sum_{i=1}^{s} \gamma_i\left(x_i\right)$$

et à cause de ça, on trouve que la borne supérieure de la fonction d'estimation sera donnée par :

$$\begin{aligned} E\left(x\right) &\leq \frac{1}{8} \sum_{i=1}^{s} \lambda_i \left(\frac{\beta_i}{k_i}\right)^2 \qquad (4.31) \\ &\leq \frac{1}{8} \sum_{i=1}^{s} \lambda_i \left(\frac{\beta_i^2}{k_i^2}\right) \\ &\leq \frac{1}{8} \lambda_1 \beta_1^2 \sum_{i=1}^{s} \left(\frac{\rho_i}{k_i^2}\right) \end{aligned}$$

Remarque 4.2 : La limite inférieure de la valeur $\triangle f_1$ du lemme précédent reste le même.

Le *théorème4.4* sera donnée sous la forme suivante :

Corollaire 4.1 : Il va de soi que l'interpolation des fonctions concaves :

$$\theta_i\left(x_i\right), 1 \leq i \leq s$$

au points :

$$x_i = jh_i$$

satisfait l'inégalité suivante :

$$\frac{f(\widetilde{x})-f^*}{\triangle f_1} \leq \frac{\sum\limits_{i=1}^{s}\left(\frac{\rho_i}{k_i^2}\right)}{\sum\limits_{i=1}^{s}\rho_i\left(1+\eta_i\right)^2} \tag{4.32}$$

Preuve : On a :

$$\begin{aligned}
E(x) &\leq \frac{1}{8}\lambda_1\beta_1^2\sum_{i=1}^{s}\left(\frac{\rho_i}{k_i^2}\right)\\
f(\widetilde{x})-f^* &\leq E(x)\\
\triangle f_1 &\geq \frac{1}{8}\lambda_1\beta_1^2\sum_{i=1}^{s}\rho_i\left(1+\eta_i\right)^2
\end{aligned}$$

donc :

$$\begin{aligned}
\frac{f(\widetilde{x})-f^*}{\triangle f_1} &\leq \frac{E(x)}{\triangle f_1}\\
&\leq \frac{\frac{1}{8}\lambda_1\beta_1^2\sum\limits_{i=1}^{s}\left(\frac{\rho_i}{k_i^2}\right)}{\frac{1}{8}\lambda_1\beta_1^2\sum\limits_{i=1}^{s}\rho_i\left(1+\eta_i\right)^2}\\
&\leq \frac{\sum\limits_{i=1}^{s}\left(\frac{\rho_i}{k_i^2}\right)}{\sum\limits_{i=1}^{s}\rho_i\left(1+\eta_i\right)^2} \leq \varepsilon
\end{aligned}$$

alors, il est possible de choisir la devision k_i, pour le quelle et pour chaque tolérance $\varepsilon > 0$, l'erreur relative dans l'approximation linéaire par morceau est bornée par ε, comme elle est bien montrer par le corollaire suivant :

Corollaire 4.2 :[7] Si pour chaque $1 \leq i \leq s$ on a :

$$k_i \geq \left(\frac{s}{\alpha}\rho_i\right)^{\frac{1}{2}} \text{ où } \alpha = \varepsilon\sum_{i=1}^{s}\rho_i\left(1+\eta_i\right)^2 \tag{4.33}$$

alors, la solution optimale $\widetilde{x}$ de problème :

$$\begin{cases} \min f_\Gamma(x) = \sum_{i=1}^{s} \gamma_i(x) + f_2(x) \\ x \in \Omega \cap S_x \end{cases} \quad \text{(PQNA)}$$

est satisfait :

$$\frac{f(\widetilde{x}) - f^*}{\triangle f_1} \leq \frac{E(\widetilde{x})}{\triangle f_1} \leq \varepsilon \qquad (4.34)$$

Preuve : On a :

$$\begin{aligned} k_i &\geq \left(\frac{s}{\alpha}\rho_i\right)^{\frac{1}{2}} \Longrightarrow k_i^2 \geq \frac{s}{\alpha}\rho_i \\ &\Longrightarrow \frac{\alpha k_i^2}{s} \geq \rho_i \\ &\Longrightarrow \frac{\alpha}{s} \geq \frac{\rho_i}{k_i^2} \\ &\Longrightarrow \sum_{i=1}^{s}\left(\frac{\rho_i}{k_i^2}\right) \leq \sum_{i=1}^{s}\left(\frac{\alpha}{s}\right) \\ &\Longrightarrow \sum_{i=1}^{s}\left(\frac{\rho_i}{k_i^2}\right) \leq \alpha \end{aligned}$$

donc, d'après le corollaire préçédent on peut écrire :

$$\begin{aligned} \frac{f(\widetilde{x}) - f^*}{\triangle f_1} &\leq \frac{E(\widetilde{x})}{\triangle f_1} \\ &\leq \frac{\sum_{i=1}^{s}\left(\frac{\rho_i}{k_i^2}\right)}{\sum_{i=1}^{s}\rho_i(1+\eta_i)^2} \\ &\leq \frac{\alpha}{\sum_{i=1}^{s}\rho_i(1+\eta_i)^2} \\ &\leq \frac{\varepsilon\sum_{i=1}^{s}\rho_i(1+\eta_i)^2}{\sum_{i=1}^{s}\rho_i(1+\eta_i)^2} \\ &\leq \varepsilon \end{aligned}$$

4.3 L'algorithme de la résolution :

On peut poser un algorithme pour trouver la solution ε-approximative associée au problème original **(PQN)**, il consiste les étapes suivantes :

(1)- Calculer les valeurs propres (λ_i) de la matrice Q, et ses vecteurs propres associer.

(2)- Résoudre le problème de la programmation linéaire " *multiple-cost-row* "**(MCR)** avec $(2s)$ cost-row.

(3)- Trouver les points de la frontière de l'ensemble (domaine) réalisable de **(PQN)**.

(4)- Calculer les valeurs de la fonction f en chaque points de la frontière.

(5)- Choisir le point x^*qui satisfait :

$$f(x^*) = \min f(x)$$

pour tout points x de l'ensemble de la frontière, puis poser :

$$IFV = f(x^*)$$

(6)- Construire la fonction linéaire $\Gamma(x)$ et résoudre le problème linéaire approximatif pour trouver la solution approximative $\widetilde{x}$.

(7)- Calculer $\triangle f_1$et $\overline{\triangle f_1}$.

(8)- Si :

$$f(\widetilde{x}) < IFV$$

alors, poser :

$$IFV = f(\widetilde{x})$$

(9)- Si :

$$IFV - f_\Gamma(\widetilde{x}) < \varepsilon\overline{\triangle f_1}$$

alors stop. $\widetilde{x}$ est une solution $\varepsilon-$approximative.

(10)- Sinon, continue avec la procédure de "*Branch* and *Bound*".

Sachons que (IFV) *incumbent function value.*

4.3.1 la procédure de“*Branch and Bound*” :

Pour continuer dans l'étape ň=10, utilisée la procédure de *“Branch and Bound”* comme suit :

(1)- Choisir un intervalle $[0, \beta_i]$, par exemple :

$$\left[0, \beta_j\right] \text{ avec } \beta_j = \min_{1 \leq i \leq s} \{\beta_i\}$$

ou bien, l'intervalle qui corresponde à :

$$\rho_j = \max_{1 \leq i \leq s} \{\rho_i\}$$

(2)- Deviser l'intervalle $\left[0, \beta_j\right]$ à deux intervalles :

$$\left[0, \frac{\beta_j}{2}\right] \text{ et } \left[\frac{\beta_j}{2}, \beta_j\right]$$

alors, avec cette application on a devisées le rectangle S_x à deux rectangles S_{1x} et S_{2x} avec le même volume.

(3)- Remplacer l'enveloppe convexe :

$$\{\gamma_i(x_i); 1 \leq i \leq s\} \text{ de } \{\theta_i(x_i); 1 \leq i \leq s\}$$

par les deux enveloppes convexes :

$$\begin{array}{l} \Gamma_{j1} \text{ de } \theta_i(x_i) \text{ sur l'intervalle } \left[0, \dfrac{\beta_j}{2}\right] \\ \Gamma_{j2} \text{ de } \theta_i(x_i) \text{ sur l'intervalle } \left[\dfrac{\beta_j}{2}, \beta_j\right] \end{array}$$

c'est à dire que :

$$\gamma_i(x_i) = \begin{cases} \Gamma_{j1}(x) \text{ si } x \in \left[0, \frac{\beta_j}{2}\right] \\ \Gamma_{j2}(x) \text{ si } x \in \left[\frac{\beta_j}{2}, \beta_j\right] \end{cases}$$

donc, on a définie les deux problèmes linéaires suivants :

$$\begin{cases} \min \gamma_i(x) = \Gamma_{j1}(x) \\ x \in S_{1x} = \left\{ x \in \mathbb{R}^n : 0 \leq x_i \leq \frac{\beta_j}{2} \right\} \end{cases} \quad (P_1)$$

$$\begin{cases} \min \gamma_i(x) = \Gamma_{j2}(x) \\ x \in S_{2x} = \left\{ x \in \mathbb{R}^n : \frac{\beta_j}{2} \leq x_i \leq \beta_j \right\} \end{cases} \quad (P_2)$$

(4)- Résoudre les deux problèmes linéaires (P_1) et (P_2), puis calculer (IFV) pour les deux problèmes.

(5)- Poser $\widetilde{x}$ qui satisfait :

$$\gamma(\widetilde{x}) = \min \{\Gamma_{j1}(\widetilde{x}), \Gamma_{j2}(\widetilde{x})\}$$

(6)- Essayer de vérifier **l'étape ň=9**.

(7)- Si on termine pas, alors on choisir un intervalle entre les deux intervalles :

$$\left[0, \frac{\beta_j}{2}\right] \text{ et } \left[\frac{\beta_j}{2}, \beta_j\right]$$

puis on travail avec la même méthode presque précédente sur lui...etc.

4.4 EXEMPLES :

Exemple 4.1 : Considérons le problème quadratique indéfini suivant :

$$\begin{cases} \min f(x) = 42x_1 - 50x_1^2 + 44x_2 - 50x_2^2 + 45x_3 - 50x_3^2 + 47x_4 + 50x_4^2 + 47x_5 + 50x_5^2 \\ x \in \Omega = \{20x_1 + 12x_2 + 11x_3 + 7x_4 + 4x_5\}, x \in R_x = \{x \in \mathbb{R}^5 : 0 \leq x_i \leq 1\} \end{cases}$$

on résout cet problème avec l'utilisation de la méthode précédente, alors on a :

$$\begin{aligned} d^T &= (42, 44, 45, 47, 47) \in \mathbb{R}^5 \\ Q &= \begin{pmatrix} 50 & 0 & 0 & 0 & 0 \\ 0 & 50 & 0 & 0 & 0 \\ 0 & 0 & 50 & 0 & 0 \\ 0 & 0 & 0 & -50 & 0 \\ 0 & 0 & 0 & 0 & -50 \end{pmatrix} \in \mathbb{R}^{5\times 5} \\ &= diag\left(\lambda_i\right) : 1 \leq i \leq 5 \end{aligned}$$

où :

$$\begin{cases} \text{les valeurs propres positives sont : } \lambda_1 = \lambda_2 = \lambda_3 = 50 \\ \text{les valeurs propres négatives sont : } \lambda_4 = \lambda_5 = -50 \end{cases}$$

On va essayer de partager notre fonction objectif à deux parties :

la partie concave est définie pour $\{\lambda_i \geq 0; 1 \leq i \leq 3\}$ comme suit :

$$f_1\left(x\right) = \sum_{i=1}^{3} \theta_i\left(x_i\right) = \sum_{i=1}^{3} \left(d_i x_i - \lambda_i x_i^2\right)$$

alors :

$$f_1\left(x\right) = \left(42 - 50x_1\right)x_1 + \left(44 - 50x_2\right)x_2 + \left(45 - 50x_3\right)x_3$$

la partie convexe est définie pour :

$$\{\lambda_i \leq 0 : 4 \leq i \leq 5\}$$

comme suit :

$$f_2\left(x\right) = \sum_{i=4}^{5} \theta_i\left(x_i\right) = \sum_{i=4}^{5} \left(d_i x_i - \lambda_i x_i^2\right)$$

alors :

$$\begin{aligned} f_2(x) &= \left(d_4x_4 - \lambda_4 x_4^2\right) + \left(d_5x_5 - \lambda_5 x_5^2\right) \\ &= (47 - (-50)\,x_4)\,x_4 + (47 - (-50)\,x_5)\,x_5 \\ &= (47 + 50x_4)\,x_4 + (47 + 50x_5)\,x_5 \\ &= 47\,(x_4 + x_5) + 50\left(x_4^2 + x_5^2\right) \end{aligned}$$

alors, construire la fonction approximative $f_\Gamma(x)$ sous la forme suivante :

$$f_\Gamma(x) = \Gamma(x) + f_2(x)$$

avec :

$$\Gamma(x) = \sum_{i=1}^{3} \gamma_i(x_i) = \sum_{i=1}^{3} (d_i - \lambda_i \beta_i)\,x_i$$

où :

$$\{\beta_i = 1; 1 \leq i \leq 5\}$$

alors :

$$\begin{aligned} \Gamma(x) &= \sum_{i=1}^{3} (d_i - \lambda_i \beta_i)\,x_i \\ &= \sum_{i=1}^{3} (d_i - \lambda_i)\,x_i \\ &= (42 - 50)\,x_1 + (44 - 50)\,x_2 + (45 - 50)\,x_3 \\ &= (-8x_1 - 6x_2 - 5x_3) \end{aligned}$$

alors, la fonction $f_\Gamma(x)$ est donnée par :

$$\begin{aligned} f_\Gamma(x) &= \Gamma(x) + f_2(x) \\ &= (-8x_1 - 6x_2 - 5x_3) + 47\,(x_4 + x_5) + 50\left(x_4^2 + x_5^2\right) \end{aligned}$$

donc, on a trouvées le problème linéaire suivant :

$$\begin{cases} \min f_{\Gamma}(x) = \Gamma(x) + f_2(x) \\ x \in \Omega \cap S_x \end{cases}$$

alors, le problème est :

$$\begin{cases} \min f_{\Gamma}(x) = (-8x_1 - 6x_2 - 5x_3) + 47(x_4 + x_5) + 50(x_4^2 + x_5^2) \\ x \in \Omega \cap S_x \end{cases}$$

où :

$$\begin{aligned} \Omega &= \left\{x \in \mathbb{R}^5 : 20x_1 + 12x_2 + 11x_3 + 7x_4 + 4x_5 - 40 \leq 0\right\} \\ S_x &= \left\{x \in \mathbb{R}^5 : 0 \leq x_i \leq 1; 1 \leq i \leq 5\right\} \end{aligned}$$

Chapitre 5

La méthode de séparation et évaluation (Branch and Bound) :

Introduction

Dans ce chapitre on va présenter un nouveau approche rectangulaire de la méthode de *Séparation et évaluation* ou *Branch and Bound* pour résoudre les problèmes de la programmation quadratiques non convexes, dans le quelle on va présenter une nouvelle fonction linéaire approximée inférieurement la fonction quadratique optimiser (i.e : la fonction objectif de notre problème d'optimisation) sur les n-rectangles, il est donnée pour déterminer la borne inférieure de la valeur optimale globale associer au problème originale sur chaque rectangle.

Pour diviser le rectangle S° à n-rectangles, on a appliquée un subdivision normale, et à chaque itération on va éliminer un rectangle dans le but de accélérer la convergence de l'algorithme proposée.

Il va de soi que avec ses tactique l'algorithme proposée sera convergente, au plus de ça, elle est effective.

Considérons le problème de la programmation quadratique non convexe suivant :

$$\begin{cases} \min f(x) = x^T Q x + d^T x \\ \qquad Ax \leq b \end{cases} \qquad (PQN)$$

où $Q = (q_{ij})_{n \times n}$ une matrice réel symétrique non positive (n'est pas définie positive) ;

$A = (a_{ij})_{m\times n}$ une matrice réel ;

$d = (d_1, d_2, ..., d_n) \in \mathbb{R}^n$;

$b = (b_1, b_2, ..., b_m) \in \mathbb{R}^m$;

le domaine réalisable du problème *(PQN)* est un polyèdre dans $\mathbb{R}^n$ définie par :

$$\left(\chi_f\right) = \{x \in \mathbb{R}^n : Ax \leq b\} \qquad (5.1)$$

on peut supposer que $\left(\chi_f\right)$ est non nul et bornée.

Grâce à la non convexité de *(PQN)*, les solutions locaux ne sont pas nécessairement des solutions globaux, et dans se cas, l'approche classique utilisée pour résoudre globalement le problème *(PQN)* sont *les méthodes* Séparation et évaluation ou *Branch and Bound.*

La différence entre ces méthodes réside dans la stratégie d'évaluation et de séparation utilisée.

5.1 Description de la méthode :

5.1.1 Bornitude de la fonction $f(x)$ sur un rectangle :

Dans les lignes qui suit nous exposerons en détail le nouveau méthode de bornitude sur un rectangle, telle que on va donnée une nouvelle fonction linéaire approximée inférieurement la fonction optimiser $f(x)$.

Cette approximation est applique avant la subdiviser de S° le rectangle original en n-rectangles S_k, qui sont définies par :

$$S_k = \left\{x \in \mathbb{R}^n ; L^k \leq x \leq U^k\right\} \qquad (5.2)$$

Remarque 5.1 : Le but de faire l'estimation précédent est pour trouver *la borne inférieure* de la valeur optimale globale associée au problème original *(PQN).*

Construction de le rectangle initial S° :

Le rectangle initial S° qui contient le domaine réalisable de *(PQN)* sera définie par :

$$S^\circ = \{x \in \mathbb{R}^n ; L^\circ \leq x \leq U^\circ\} \qquad (5.3)$$

où les deux bornes L° et U° sont calculées sous la manière suivante :

On résout les n-programmes linéaires :

$$\max\left\{x_i : x \in \left(\chi_f\right), i = 1, 2, ..., n\right\} \tag{5.4}$$

pour obtenir les valeurs $\{L_i^\circ : i = 1, 2, ..., n\}$, et :

$$\min\left\{x_i : x \in \left(\chi_f\right), i = 1, 2, ..., n\right\} \tag{5.5}$$

pour obtenir les valeurs $\{U_i^\circ : i = 1, 2, ..., n\}$, le rectangle S° peut alors s'écrire comme :

$$S^\circ = \{x \in \mathbb{R}^n; L_i^\circ \leq x_i \leq U_i^\circ, i = 1, 2, ..., n\} \tag{5.6}$$

le rectangle S° peut être simplement subdiviser à deux subrectangles par les bissection méthodes notée par :

$$S_{+1} = \left\{x \in \mathbb{R}^n : L_s^\circ \leq x_s^\circ \leq h_s \text{ et } L_j^\circ \leq x_j^\circ \leq U_j^\circ; j = 1, 2, ..., n \text{ et } j \neq s\right\} \tag{5.7}$$

$$S_{+2} = \left\{x \in \mathbb{R}^n : h_s \leq x_s^\circ \leq U_s^\circ \text{ et } L_j^\circ \leq x_j^\circ \leq U_j^\circ; j = 1, 2, ..., n \text{ et } j \neq s\right\} \tag{5.8}$$

les bissections méthodes seront données dans **la section5.5.**

Définition 5.1 :[6] On définies le nombre positif θ qui satisfait :

$$\theta \geq |\lambda_{\min}| \tag{5.9}$$

où $\lambda_{\min}$ est le minimum de la matrice Q.

Remarque 5.2 : Le nombre positif θ est définie pour le quelle la matrice $(Q + \theta I)$ sera définie positive ou bien semi définie positive.

La proposition suivante nous donne la forme de la borne inférieure de la fonction optimiser $f(x)$.

Proposition 5.1 :[6] *La borne inférieure de la fonction $f(x)$ sur le rectangle S_k est* *estimée par la fonction linéaire définie par :*

$$f(x) = \left(x-L^k\right)^T (Q+\theta I)\left(x-L^k\right) + d^T x - \theta\sum_{i=1}^{n} x_i^2 + 2\left(L^k\right)^T (Q+\theta I)\, x \tag{5.10}$$
$$-\left(L^k\right)^T (Q+\theta I)\, L^k \tag{5.1}$$

et si on utilises la borne U^k on trouves que $f(x)$ est estimée par :

$$f(x) = \left(x-U^k\right)^T (Q+\theta I)\left(x-U^k\right) + d^T x - \theta\sum_{i=1}^{n} x_i^2 + 2\left(U^k\right)^T (Q+\theta I)\, x \tag{5.11}$$
$$-\left(U^k\right)^T (Q+\theta I)\left(U^k\right) \tag{5.2}$$

Preuve : On a :

$$\begin{aligned} f(x) &= x^T Qx + d^T x \\ &= x^T Qx + d^T x + \theta x^T x - \theta x^T x \\ &= x^T (Q+\theta I)\, x + d^T x - \theta \|x\|^2 \end{aligned}$$

si on utilise la borne inférieure L^k de le rectangle S_k, on peut alors écrire :

$$\begin{aligned} f(x) &= x^T (Q+\theta I)\, x + d^T x - \theta \|x\|^2 \\ &= \left(x-L^k+L^k\right)^T (Q+\theta I)\left(x-L^k+L^k\right) + d^T x - \theta \|x\|^2 \\ &= \left(x-L^k\right)^T (Q+\theta I)\left(x-L^k\right) + \left(x-L^k\right)^T (Q+\theta I)\left(L^k\right) \\ &\quad + \left(L^k\right)^T (Q+\theta I)\left(x-L^k\right) + \left(L^k\right)^T (Q+\theta I)\left(L^k\right) + d^T x - \theta \|x\|^2 \\ &= \left(x-L^k\right)^T (Q+\theta I)\left(x-L^k\right) + 2\left(L^k\right)^T (Q+\theta I)\left(x-L^k\right) \\ &\quad + d^T x - \theta \|x\|^2 + \left(L^k\right)^T (Q+\theta I)\left(L^k\right) \\ &= \left(x-L^k\right)^T (Q+\theta I)\left(x-L^k\right) + d^T x - \theta\sum_{i=1}^{n} x_i^2 + 2\left(L^k\right)^T (Q+\theta I)\, x \\ &\quad - \left(L^k\right)^T (Q+\theta I)\left(L^k\right) \end{aligned}$$

d'autre part, si on utilise la borne supérieure U^k de le rectangle S_k, on travaille sous la même manière suivante, et on trouve que :

$$\begin{aligned} f(x) &= x^T (Q+\theta I) x + d^T x - \theta \|x\|^2 \\ &= \left(x - U^k + U^k\right)^T (Q+\theta I) \left(x - U^k + U^k\right) + d^T x - \theta \|x\|^2 \end{aligned}$$

et dans ce cas, on a directement :

$$\begin{aligned} f(x) &= \left(x - U^k\right)^T (Q+\theta I) \left(x - U^k\right) + d^T x - \theta \sum_{i=1}^{n} x_i^2 + 2 \left(U^k\right)^T (Q+\theta I) x \\ &\quad - \left(U^k\right)^T (Q+\theta I) \left(U^k\right) \end{aligned}$$

donc la preuve est complète. ■

L'enveloppe convexe de la fonction non convexe optimiser $f(x)$ sur le rectangle S_k est donnée dans la définition suivante pour être la bonne estimateur linéaire du cette fonction.

Définition 5.2 :[21] Soient la fonction $f : C \subseteq \mathbb{R}^n \longrightarrow \overline{\mathbb{R}}$, et le rectangle $S^\circ \subseteq \mathbb{R}^n$ telle que $C \subseteq S^\circ \subseteq \mathbb{R}^n$, l'enveloppe convexe de la fonction f sur le rectangle S° est la fonction définie par :

$$f_i^\circ (x_i) = \delta_i x_i + \eta_i : i = 1, 2, ..., n \tag{5.12}$$

où :

$$\delta_i = \frac{f_i (U_i^\circ) - f_i (L_i^\circ)}{U_i^\circ - L_i^\circ} \text{ et } \eta_i = f_i (L_i^\circ) - \delta_i L_i^\circ : i = 1, 2, .., n \tag{5.13}$$

alors, d'après cette définition on a immédiatement la proposition suivante :

Proposition 5.2 :[6] *Soient L_j^k et U_j^k les $j^{\text{éme}}$ composante de L^k et U^k respectivement alors, l'enveloppe convexe de la fonction :*

$$h(x) = -\|x\|^2 = -\sum_{j=1}^{n} x_j^2 \tag{5.14}$$

sur l'intervalle $\left[L_j^k, U_j^k\right]$ *est la fonction linéaire définie par :*

$$\begin{aligned}
\varphi_{S_k}(x) &= \sum_{j=1}^{n}\left(-\left(U_j^k + L_j^k\right)x_j + U_j^k L_j^k\right) \qquad (5.15)\\
&= -\left(U^k + L^k\right)^T x + \left(L^k\right)^T U^k
\end{aligned}$$

Preuve : D'après la définition précédente l'enveloppe convexe de la fonction :

$$h(x) = -\|x\|^2 = -\sum_{j=1}^{n} x_j^2$$

sur le rectangle :

$$S_k = \left\{x \in \mathbb{R}^n : L_j^k \leq x_j \leq U_j^k : j = 1, 2, .., n\right\}$$

sera donnée par :

$$\begin{aligned}
\varphi_{S_k}(x) &= \sum_{j=1}^{n} \varphi_{S_k}^j(x)\\
&= \sum_{j=1}^{n}\left(\delta_j x_j + \eta_j\right)
\end{aligned}$$

avec :

$$\begin{aligned}
\delta_j &= \frac{h_j\left(U_j^k\right) - h_j\left(L_j^k\right)}{U_j^k - L_j^k}\\
&= \frac{\left(-\left(U_j^k\right)^2 - \left(-\left(L_j^k\right)^2\right)\right)}{U_j^k - L_j^k}\\
&= \frac{\left(L_j^k\right)^2 - \left(U_j^k\right)^2}{U_j^k - L_j^k}\\
&= \frac{\left(L_j^k - U_j^k\right)\left(L_j^k + U_j^k\right)}{U_j^k - L_j^k}\\
&= -\left(L_j^k + U_j^k\right)
\end{aligned}$$

d'autre part on a :

$$\begin{aligned}
\eta_j &= h_j\left(L_j^k\right) - \delta_j L_j^k \\
&= -\left(L_j^k\right)^2 - \left(-\left(L_j^k + U_j^k\right)\right) L_j^k \\
&= -\left(L_j^k\right)^2 + \left(L_j^k\right)^2 - L_j^k U_j^k \\
&= -L_j^k U_j^k
\end{aligned}$$

on peut alors écrire :

$$\begin{aligned}
\varphi_{S_k}^j(x) &= \left(\delta_j x_j + \eta_j\right) \\
&= -\left(L_j^k + U_j^k\right) x_j - L_j^k U_j^k
\end{aligned}$$

donc :

$$\begin{aligned}
\varphi_{S_k}(x) &= \sum_{j=1}^{n} \varphi_{S_k}^j(x) \\
&= \sum_{j=1}^{n} \left(-\left(L_j^k + U_j^k\right) x_j - L_j^k U_j^k\right) \\
&= -\left(L^k + U^k\right)^T x - \left(L^k\right)^T U^k
\end{aligned}$$

alors, la preuve est complète. ■

Remarque 5.3 : La fonction linéaire $\varphi_{S_k}(x)$ est la bonne borne inférieure du la fonction $h(x)$ sur le rectangle $\left[L^k, U^k\right]$.

D'après les deux relations *(5.10)* et *(5.11),* on trouve que la fonction linéaire approximative peut prendre deux formes qui seront données dans le théorème suivant :

Théorème 5.1 :[6] *Les deux formes linéaires approximées inférieurement la fonction non convexe optimiser $f(x)$ sur le rectangle choisie $S^k = \left[L^k, U^k\right]$ sont définies par :*

$$L_{S^k}(x) = \left(a_{S^k}\right)^T x + \left(b_{S^k}\right) \qquad (5.16)$$

où :

$$
\begin{aligned}
a_{S^k} &= d + 2\,(Q + \theta I)\, L^k - \theta\left(L^k + U^k\right) \\
b_{S^k} &= -\left(L^k\right)^T (Q + \theta I)\, L^k + \theta\left(L^k\right) U^k
\end{aligned}
\tag{5.17}
$$

et :

$$
U_{S^k}(x) = (\overline{a}_{S^k})^T x + \left(\overline{b}_{S^k}\right) \tag{5.18}
$$

où :

$$
\begin{aligned}
\overline{a}_{S^k} &= d + 2\,(Q + \theta I)\, U^k - \theta\left(L^k + U^k\right) \\
\overline{b}_{S^k} &= -\left(U^k\right)^T (Q + \theta I)\, U^k + \theta\left(L^k\right) U^k
\end{aligned}
\tag{5.19}
$$

Preuve : D'après la relation *(5.10),* l'enveloppe convexe de cette forme est donnée par :

$$
L_{S^k}(x) = (a_{S^k})^T x + (b_{S^k})
$$

avec :

$$
\begin{aligned}
a_{S^k} &= \frac{f\left(U^k\right)-f\left(L^k\right)}{U^k-L^k} \\
&= \frac{d^T U^k-\theta \sum\limits_{i=1}^{n}\left(U_i^k\right)^2+2\left(L^k\right)^T(Q+\theta I)\,U^k-\left(L^k\right)^T(Q+\theta I)\,L^k}{U^k-L^k} \\
&\quad -\frac{d^T L^k-\theta \sum\limits_{i=1}^{n}\left(L_i^k\right)^2+2\left(L^k\right)^T(Q+\theta I)\,L^k-\left(L^k\right)^T(Q+\theta I)\,L^k}{U^k-L^k} \\
&= \frac{d^T U^k-\theta \sum\limits_{i=1}^{n}\left(U_i^k\right)^2+2\left(L^k\right)^T(Q+\theta I)\,U^k-d^T L^k+\theta \sum\limits_{i=1}^{n}\left(L_i^k\right)^2}{U^k-L^k} \\
&\quad -\frac{2\left(L^k\right)^T(Q+\theta I)\,L^k}{U^k-L^k} \\
&= \frac{d^T\left(U^k-L^k\right)-\theta \sum\limits_{i=1}^{n}\left(\left(U_i^k\right)^2-\left(L_i^k\right)^2\right)+2\left(L^k\right)^T(Q+\theta I)\left(U^k-L^k\right)}{U^k-L^k} \\
&= \frac{d^T\left(U^k-L^k\right)-\theta \sum\limits_{i=1}^{n}\left(U_i^k-L_i^k\right)\left(U_i^k+L_i^k\right)+2\left(L^k\right)^T(Q+\theta I)\left(U^k-L^k\right)}{U^k-L^k} \\
&= d-\theta \sum_{i=1}^{n}\left(U_i^k+L_i^k\right)+2\left(L^k\right)^T(Q+\theta I) \\
&= d-\theta\left(U^k+L^k\right)+2\left(L^k\right)^T(Q+\theta I)
\end{aligned}
$$

et :

$$
\begin{aligned}
b_{S^k} &= f\left(L^k\right)-a_{S^k}L^k \\
&= \left(d^T L^k-\theta \sum_{i=1}^{n}\left(L_i^k\right)^2+2\left(L^k\right)^T(Q+\theta I)\,L^k-\left(L^k\right)^T(Q+\theta I)\,L^k\right) \\
&\quad -\left(d-\theta\left(U^k+L^k\right)+2\left(L^k\right)^T(Q+\theta I)\right)L^k \\
&= d^T L^k-\theta \sum_{i=1}^{n}\left(L_i^k\right)^2+\left(L^k\right)^T(Q+\theta I)\,L^k-d^T L^k+\theta\left(U^k+L^k\right)^T L^k- \\
&\quad 2\left(L^k\right)^T(Q+\theta I)\,L^k \\
&= -\left(L^k\right)^T(Q+\theta I)\,L^k-\theta \sum_{i=1}^{n}\left(L_i^k\right)^2+\theta\left(U^k\right)^T L^k+\theta\left(L^k\right)^T L^k \\
&= -\left(L^k\right)^T(Q+\theta I)\,L^k+\theta\left(U^k\right)^T L^k
\end{aligned}
$$

donc, la relation (5.16) est satisfaite ;

d'autre part, pour la deuxième forme approximative $U_{S_k}(x)$ on utilise la forme (5.11), et la même définition de l'enveloppe convexe qu'été utilisée préçidament dans la définition de la borne inférieure $L_{S^k}(x)$ de $f(x)$ sur le rectangle $S^k = \left[L^k, U^k\right]$.

5.1.2 La relation effective entre la fonction optimiser $f(x)$ et leur estimateur linéaire :

On peut trouver la relation entre la fonction optimiser et les deux bornes inférieures $L_{S^k}(x)$ et $U_{S_k}(x)$, et on peut aussi estimer la différence ou bien la distance entre elles, ça sera bien démontrer par les deux théorèmes suivants.

Théorème 5.2 :[6] *Soit la matrice $(Q + \theta I)$ être une matrice semi définie positive, alors on a :*

$$\begin{aligned} L_{S^k}(x) &\leq f(x) ; \forall x \in \left(\chi_f\right) \cap S^k \\ U_{S_k}(x) &\leq f(x) ; \forall x \in \left(\chi_f\right) \cap S^k \end{aligned} \tag{5.20}$$

(i,e : les deux fonctions linéaires $L_{S^k}(x)$ et $U_{S_k}(x)$ sont les deux bornes inférieurs de la fonction non convexe optimiser sur le rectangle choisie $S^k = \left[L^k, U^k\right]$).

Preuve : On a déjà préçidament définie la fonction optimiser avec l'utilisation des deux bornes L^k et U^k, alors, d'après la relation *(5.10)* on a :

$$\begin{aligned} f(x) &= \left(x - L^k\right)^T (Q + \theta I) \left(x - L^k\right) + d^T x - \theta \sum_{i=1}^{n} x_i^2 + 2\left(L^k\right)^T (Q + \theta I) x \\ &\quad - \left(L^k\right)^T (Q + \theta I) L^k \end{aligned}$$

d'autre part on a :

$$\varphi_{S_k}(x) = \sum_{j=1}^{n} \left(-\left(U_j^k + L_j^k\right) x_j + U_j^k L_j^k\right) = -\left(L^k + U^k\right)^T x + \left(L^k\right)^T U^k$$

est la borne inférieure de la fonction :

$$h(x) = -\|x\|^2 = -\sum_{j=1}^{n} x_j^2$$

on peut alors écrire que :

$$\begin{aligned} f(x) \geq & \left(x - L^k\right)^T (Q + \theta I)\left(x - L^k\right) + d^T x - \theta\left(L^k + U^k\right)^T x + \theta\left(L^k\right)^T U^k \\ & + 2\left(L^k\right)^T (Q + \theta I) x - \left(L^k\right)^T (Q + \theta I) L^k \end{aligned}$$

donc, d'après la définition de l'estimateur linéaire $L_{S^k}(x)$ on peut écrire :

$$f(x) \geq \left(x - L^k\right)^T (Q + \theta I)\left(x - L^k\right) + L_{S^k}(x)$$

mais par hypothèse, la matrice $(Q + \theta I)$ est semi définie positive alors, elle satisfait la relation suivante :

$$\langle (Q + \theta I) x, x \rangle \geq 0; \forall x \in \mathbb{R}^n$$

on peut alors écrire :

$$\left(x - L^k\right)^T (Q + \theta I)\left(x - L^k\right) \geq 0$$

s'implique que :

$$f(x) \geq L_{S^k}(x); \forall x \in S^k = \left[L^k, U^k\right]$$

la même chose pour la deuxième borne $U_{S_k}(x)$ où on utilise la forme *(5.11)* de la fonction optimiser, et on continue sous les même étapes préçédentes, alors, la preuve est complète.

■

Théorème 5.3 :[6] *Soit la matrice $(Q + \theta I)$ être semi définie positive, et soit $\rho(Q + \theta I)$ être la plus grande valeur propre de la matrice $(Q + \theta I)$, alors, on a :*

$$\max\left\{\left|f(x) - L_{S^k}(x)\right| : x \in S^k = \left[L^k, U^k\right]\right\} \leq \left(\rho(Q + \theta I) + \theta\right)\left\|L^k - U^k\right\|^2 \qquad (5.21)$$

et :

$$\max\left\{\left|f(x)-U_{S_k}(x)\right| : x\in S^k=\left[L^k,U^k\right]\right\}\leq\left(\rho\left(Q+\theta I\right)+\theta\right)\left\|L^k-U^k\right\|^2 \qquad (5.22)$$

Preuve : Prendre la forme *(5.10)* de la fonction optimiser, et la relation *(5.16)* alors, on a :

$$\begin{aligned}
f(x)-L_{S^k}(x) &= \left(x-L^k\right)^T\left(Q+\theta I\right)\left(x-L^k\right)+d^Tx-\theta\sum_{i=1}^{n}x_i^2+2\left(L^k\right)^T\left(Q+\theta I\right)x \\
&\quad -\left(L^k\right)^T\left(Q+\theta I\right)L^k-d^Tx-2\left(L^k\right)^T\left(Q+\theta I\right)x+\theta\left(L^k+U^k\right)^Tx \\
&\quad +\left(L^k\right)^T\left(Q+\theta I\right)L^k-\theta\left(L^k\right)^TU^k \\
&= \left(x-L^k\right)^T\left(Q+\theta I\right)\left(x-L^k\right)-\theta\left(\sum_{i=1}^{n}x_i^2-\left(L^k+U^k\right)^Tx+\left(L^k\right)^TU^k\right) \\
&= \left(x-L^k\right)^T\left(Q+\theta I\right)\left(x-L^k\right)-\theta\left(\|x\|^2+\varphi_{S_k}(x)\right) \\
&= \left(x-L^k\right)^T\left(Q+\theta I\right)\left(x-L^k\right)+\theta\left(-\|x\|^2-\varphi_{S_k}(x)\right)
\end{aligned}$$

d'un autre côté, on peut écrire :

$$\begin{aligned}
\left(-\|x\|^2-\varphi_{S_k}(x)\right) &= -x^Tx+\left(L^k+U^k\right)^Tx-\left(L^k\right)^TU^k \\
&= -x^Tx+\left(L^k\right)^Tx+\left(U^k\right)^Tx-\left(L^k\right)^TU^k \\
&= -x^T\left(x-L^k\right)+\left(U^k\right)^T\left(x-L^k\right) \\
&= \left(U^k-x\right)^T\left(x-L^k\right)
\end{aligned}$$

alors :

$$\begin{aligned}
\left\|f(x)-L_{S^k}(x)\right\|_\infty &= \max\left\{\left|f(x)-L_{S^k}(x)\right| : x\in S^k=\left[L^k,U^k\right]\right\} \\
&= \left\|\left(x-L^k\right)^T\left(Q+\theta I\right)\left(x-L^k\right)+\theta\left(U^k-x\right)^T\left(x-L^k\right)\right\|_\infty
\end{aligned}$$

d'autre part, on a :

$$\begin{aligned}
\left|f(x) - L_{S^k}(x)\right| &= \left|\left(x - L^k\right)^T (Q + \theta I)\left(x - L^k\right) + \theta\left(U^k - x\right)^T \left(x - L^k\right)\right| \\
&\leq \rho(Q + \theta I)\left\|U^k - L^k\right\|^2 + \theta\left\|\left(U^k - x\right)^T\left(x - L^k\right)\right\| \\
&\leq \rho(Q + \theta I)\left\|U^k - L^k\right\|^2 + \theta\left\|U^k - L^k\right\|^2 \\
&= \left(\rho(Q + \theta I) + \theta\right)\left\|U^k - L^k\right\|^2
\end{aligned}$$

donc :

$$\max\left\{\left|f(x) - L_{S^k}(x)\right| : x \in S^k = \left[L^k, U^k\right]\right\} \leq \left(\rho(Q + \theta I) + \theta\right)\left\|U^k - L^k\right\|^2$$

la même chose pour la deuxième borne $U_{S_k}(x)$ où on utilise la forme *(5.11)* de la fonction optimiser, et on continue sous les même étapes préçédentes, alors, la preuve est complète. ■

5.1.3 Construction du problème linéaire (LBP) :

D'après le théorème (**5.2**) on a :

$$\begin{aligned}
L_{S^k}(x) &\leq f(x)\,;\forall x \in \left(\chi_f\right) \cap S^k \\
U_{S_k}(x) &\leq f(x)\,;\forall x \in \left(\chi_f\right) \cap S^k
\end{aligned}$$

et à cause de ça on peut écrire :

$$f(x) \geq \max\left\{L_{S^k}(x), U_{S_k}(x)\,;\forall x \in \left(\chi_f\right) \cap S^k\right\} = \gamma(x) \tag{5.23}$$

où la fonction $\gamma(x)$ est présentée comme *une borne inférieure* pour la fonction optimiser, on peut alors construire le problème linéaire suivant :

$$\left\{\begin{array}{l}
\min\max\left\{L_{S^k}(x), U_{S_k}(x)\right\} \\
\qquad Ax \leq b \\
\qquad x \in S^k
\end{array}\right. \tag{LBP}$$

on peut aussi le construire sous une autre forme plus simple comme suit :

$$\begin{cases} \min t \\ L_{S^k}(x) \leq t \\ U_{S_k}(x) \leq t \\ Ax \leq b \\ x \in S^k \end{cases}$$

d'après la construction de le problème linéaire **(LBP)**, on trouve une résultat très importante donnée dans le corollaire suivant :

Corollaire 5.1 :[**6**]

(1) La valeur optimale du **(LBP)** est une borne inférieure de la valeur optimale globale du problème original **(PQN)** sur le rectangle S^k.

(2) La solution optimale du **(LBP)** notée par x^k est une solution réalisable pour le problème original (**PQN**).

Preuve :

(1) posons ϖ être la valeur optimale associer au problème **(LBP)**, et $f(x^*)$ être la valeur optimale globale de **(PQN)**, donc immédiatement on a :

$$\begin{aligned} \varpi &= \min\max\{L_{S^k}(x), U_{S_k}(x)\} \qquad (5.24) \\ f(x^*) &= \left\{\min f(x); x \in (\chi_f) \cap S^k\right\} \end{aligned}$$

donc, il y a deux possibilités :

(a) si on a :

$$\varpi = \min L_{S^k}(x) = L_{S^k}(\widetilde{x}) \leq f(x); \forall x$$

pour $x = x^*$ on a :

$$\varpi = L_{S^k}(\widetilde{x}) \leq f(x^*)$$

(b) Si on a :

$$\varpi = \min U_{S_k}(x) = U_{S_k}(\widehat{x}) \leq f(x); \forall x$$

s'implique que :

$$\varpi = U_{S_k}(\widehat{x}) \leq f(x^*)$$

on peut alors écrire que :

$$\varpi \leq f(x^*)$$

c'est à dire que la relation (1) est satisfaite. $\natural$

(2) si $\widetilde{x}$ est la solution optimale de **(LBP)** alors, nécessairement elle satisfait les contraints suivantes :

$$A\widetilde{x} \leq b \text{ et } \widetilde{x} \in S^k$$

donc, elle satisfait les contraintes du **(PQN)**, s'implique que $\widetilde{x}$ appartient à l'ensemble des solutions réalisables du **(PQN)**, alors, la relation (2) est satisfaite. $\blacksquare$

5.2 Les techniques de partition et réduction sur un rectangle :

Dans cette section on va donner une nouvelle méthode de partition est appelé *bisection méthod* ou *the two-partition méthode*, et d'un autre côté, on va expliquer les techniques de réduction (élimination) sur un rectangle dans $\mathbb{R}^n$, prendre par exemple :

$$S^k = \left\{x \in \mathbb{R}^n; L^k \leq x \leq U^k\right\}$$

un rectangle dans $\mathbb{R}^n$ et un point x^k telle que $x^k \in S^k$.

5.2.1 Méthode de partition sur un rectangle :

La méthode est expliquée par l'algorithme suivant :

Algorithme(1) :

initialisation :

itération k : trouver le point w_j telle que :

$$w_j = \max\left\{\left(x_i^k - L_i^k\right)\left(U_i^k - x_i^k\right); i = 1, 2, ..., n\right\}$$

Si $w_j \neq 0$ **alors** on subdiviser le rectangle S^k à deux subrectangles :

$$\begin{aligned} S_1^K &= \left\{x \in \mathbb{R}^n : L_j^k \leq x_j \leq w_j; L_i^k \leq x_i \leq U_i^k; i \neq j\right\} \\ S_2^K &= \left\{x \in \mathbb{R}^n : w_j \leq x_j \leq U_j^k; L_i^k \leq x_i \leq U_i^k; i \neq j\right\} \end{aligned}$$

Sinon

on subdiviser le rectangle S^k à deux subrectangles où l'indice j est déterminée comme étant l'indice de la plus

longue arrête de $S^k = \left[L^k, U^k\right]$ avec : $w_j = \frac{L_j^k + U_j^k}{2}$ le point médial de l'arrête $\left[L_j^k, U_j^k\right]$.

Finsi ;

k⟵k+1

Fin.

5.2.2 Description de techniques du réduction (élimination) :

Dans cette section on va expliquer le principe des techniques du réduction rectangulaire dans le but d'accélérer la convergence de l'algorithme proposée (i,e l'algorithme d'optimisation global).

Cette technique est résumée dans l'algorithme suivant qu'est appelée *linearity-based range réduction algorithme* pour réduce et delete (éliminer) le rectangle S^k.

Cette élimination est dans le but d'obtenir un nouveau rectangle, noté aussi S^k, qui n'est pas plus large que le rectangle avant la réduction, d'autre part, l'algorithme suivant nous donne une nouvelle méthode qui peut éliminer quelles que subrectangles et quelles que contraintes linéaires.

Algorithme(2) :

initialisation : $I'_k := \{1, 2, ..., m\}$, $P_k := P$

itération k :

Pour $i = 1, 2, ..., m$ **faire ;**

$rU_I = \sum_{j=1}^{n} \max \left\{ a_{ij} L_j^k, a_{ij} U_j^k \right\}$

$rL_i = \sum_{j=1}^{n} \min \left\{ a_{ij} L_j^k, a_{ij} U_j^k \right\}$

Si $rL_i > 0$ **alors ; stop.**

le problème **(PQN)** n'est pas réalisable sur le rectangle S^k et dans se cas $S^k = \phi$.

Sinon

Si $rU_I < b_i$ **alors ;**

la contrainte est *redundant* et :

$I'_k := I'_k - \{i\}$

$P_k = P_k - \left\{ x \in \mathbb{R}^n : (a_i)^T x \leq b_i \right\}$

Sinon

Pour $j = 1, 2, ..., n$ **faire ;**

Si $a_{ij} > 0$ **alors ;**

$U_j^k := \min \left\{ U_j^k, \frac{b_i - rL_i + \min\left\{ a_{ij} L_j^k, a_{ij} U_j^k \right\}}{a_{ij}} \right\}$

Sinon

$L_j^k := \max \left\{ L_j^k, \frac{b_i - rU_i + \max\left\{ a_{ij} L_j^k, a_{ij} U_j^k \right\}}{a_{ij}} \right\}$

FinSi ;

FinPour ;

FinSi ;

FinSi ;

FinPour ;

k⟵k+1

Fin.

Remarque 5.3 : Sachons que, tous les contraintes linéaires de le problème **(PQN)** sont exprémées comme :

$$\sum_{j=1}^{n} a_{ij}x_j \leq b_i : i = 1, 2, .., m \tag{5.25}$$

d'autre part, le rectangle S^k définie préçidament être ajouter comme une contrainte pour le problème **(PQN)**.

Les minimums et maximums de chaque $(a_{ij}x_j)$ sur l'intervalle $\left[L_j^k, U_j^k\right]$ sont atteints en les points extrêmes de se intervalle, c'est à dire en les points L_j^k, U_j^k.

5.3 L'algorithme rectangulaire de séparation et réduction (élimination) :

D'après tous les lignes précédentes on peut présenter un algorithme rectangulaire pour résoudre les problèmes quadratiques non convexes, qu'est appelée *l'algorithme rectangulaire de séparation et réduction* ou *Branch and Bound.*

Dans cette algorithme, et pour chaque rectangle S^k, on calcule la valeur optimale $\alpha\left(S^k\right)$ par la résolution de problème min-max qui noté par **(LBP)**, telle que cette valeur présente une borne inférieure de la fonction optimiser $f(x)$ sur S^k, on peut alors écrire :

$$\alpha\left(S^k\right) \leq f(x) : \forall x \in \left(\chi_f\right) \cap S^k$$

donc :

$$\alpha\left(S^k\right) \leq f\left(x^*\right)$$

où $f(x^*)$ est la valeur optimale globale de **(PQN)**.

D'autre part, on note par x^k la solution du **(LBP)** associée à la valeur optimale $\alpha\left(S^k\right)$.

La borne supérieure de la valeur optimale globale de **(PQN)** sur le rectangle S^k est définie par les solutions x^k du **(LBP)** où $x^k \in \left(\chi_f\right) \cap S^k$.

Remarque 5.4 :[6] La procédure rectangulaire *Linearity-based range réduction* sera utilisée dans l'algorithme suivante pour accélérer leur convergence.

L'algorithme est donnée par :

Algorithme(3) : *Séparation et réduction.*

initialisation : déterminée le rectangle initial S° telle que $\left(\chi_f\right) \subset S^\circ$ et posons

$$\text{LBP}_{S^\circ} := S^\circ \cap \left(\chi_f\right);$$

itération k :

Si $\text{LBP}_{S^\circ} \neq \phi$ **alors ;**

résoudre le problème linéaire (LBP) si k=0.

soient x° être la solution optimale du (LBP) et $\alpha\left(x^\circ\right)$ être leur valeur optimale, et soient :

$H := \{S^\circ\}$ (l'ensemble des subrectangles de le rectangle initial S°) ;

$\alpha_\circ := \min\left\{\alpha\left(S^\circ\right)\right\}, \beta_\circ := f\left(x^\circ\right)$ (la borne supérieure de $f\left(x^*\right)$) ;

k :=0

Pour Stop=false **faire ;**

Si $\alpha_k = \beta_k$ **alors ;**

Stop=true (x^k est une solution optimale globale du (PQN))

Sinon

on divise le rectangle S^k à deux subrectangles $\left\{S_j^k : j = 1, 2\right\}$ par la méthode choisie.

Pour $j = 1, 2$ **faire ;**

appliquée les techniques de réduction (delete) sur les deux subrectangles S_j^k, le rectangle qui reste sera notée aussi par S_j^k.

Si $S_j^k \neq \phi$ **alors ;**

$(\text{LBP})_{S_j^k} := \left\{x \in \mathbb{R}^n : x \in S_j^k \cap \left(\chi_f\right)\right\}$, puis résoudre le problème linéaire (LBP) où $S^k := S_j^k$.

notée par x^{k_j} la solution optimale et $\alpha\left(S_j^k\right)$ être la valeur optimale associer,

$H := H \cup \left\{S_j^k\right\}$;

$\beta_{k+1} := \min\left\{f\left(x^k\right), f\left(x^{k_j}\right)\right\}$;

$x^k := \arg\min \beta_{k+1}$;

FinSi ;

FinPour ;

$H := H - \{S^k\}$;

$\alpha_{k+1} := \min\limits_{S\in H} \{\alpha(S)\}$; on choisie un rectangle $S^{k+1} \in H$

telle que $\alpha_{k+1} = \alpha(S^{k+1})$;

k⟵*k+1 ;*

FinSi ;

FinPour ;

FinSi ;

Fin.

5.4 La convergence de l'algorithme proposée :

Cette algorithme est convergente par rapport les algorithmes proposées dans plusieurs références, la différence restreint dans l'application de la partition, le calcule de la borne inférieure et la stratégie de la réduction (élimination) rectangulaire.

On peut démontrer brièvement cette convergence dans les lignes qui suivent.

Théorème 5.4 :[6] *La solution optimale globale du problème* (PQN) *peut être obtenir si l'algorithme de Séparation et interpolation terminée dans un nombre finie de pas.*

Preuve : Soit x^k la solution obtienne lorsque l'algorithme termine dans l'étape k, alors, on peut écrire :

$$\alpha_k = \min_{S\in H} \alpha(S) = \alpha(S^k) = \beta_k = f(x^k) \tag{5.26}$$

telle que :

$$\alpha_k = \alpha(S^k) = \min\max\{L_{S^k}(x), U_{S_k}(x)\}$$

c'est à dire que :

$$\alpha(S^k) = f(x^k) \leq f(x) : \forall x \in S^k \tag{5.27}$$

donc, nécessairement (x^k) est une solution optimale globale de **(PQN).** ■

Si l'algorithme ne termine pas dans un nombre finie de pas alors, on a le théorème suivante où la solution globale est existe mais se n'est pas une solution unique.

Théorème 5.5 :[6, 21] *Si l'algorithme est génère un nombre infinie des suites* $\left\{x^k\right\}_{k\in\mathbb{N}}$ *alors, tout point d'accumulation* x^* *de* $\left\{x^k\right\}_{k\in\mathbb{N}}$ *présente une solution optimale globale pour le problème originale* (PQN).

Preuve : posons x^* être le point d'accumulation pour la suite $\left\{x^k\right\}_{k\in\mathbb{N}}$, et soit $\left\{x_q^k\right\}_{k\in\mathbb{N}}$ une sou suite de la suite $\left\{x^k\right\}_{k\in\mathbb{N}}$, et on la note aussi par $\left\{x^{k_q}\right\}_{k\in\mathbb{N}}$.

d'autre part, dans l'algorithme presque préçédent, la suite $\{\alpha_k\}_{k\in\mathbb{N}}$ qu'est appellée *une suite inférieure* est strictement croissante, la même chose pour la suite $\{\beta_k\}_{k\in\mathbb{N}}$ qu'est appellée *une suite supérieure,* cette suite est strictement décroissante où :

$$\alpha_k = L_{S^k}\left(x^k\right) \text{ et } \beta_k = f\left(x^k\right)$$

et à cause de ça, on peut dire que les deux suites $\{\alpha_k\}_{k\in\mathbb{N}}$ et $\{\beta_k\}_{k\in\mathbb{N}}$ sont convergente sur S^k, et dans se cas on a :

$$\lim_{k\longrightarrow\infty}\beta_k = \lim_{q\longrightarrow\infty}\beta_{k_q} = \lim_{q\longrightarrow\infty}\left(f\left(x^{k_q}\right)\right) = f\left(x^*\right)$$

au plus de ça, et son perde de généralité, posons x^{k_q} être la solution optimale du problème linéaire (LBP) sur le rectangle S^{k_q}, telle que :

$$S^{k_{q+1}} \subset S^{k_q}, q = 1, 2, .., n$$

et puisque la partition rectangulaire est *exhaustive,* c'est à dire que :

$$\lim_{q\longrightarrow\infty} S^{k_q} = x^*$$

d'un autre côté, d'après le théorème (5.3) on a :

$$0 \leq \beta_{k_q} - \alpha_{k_q} = f\left(x^{k_q}\right) - L_{S^{k_q}}\left(x^{k_q}\right) \leq \left(\rho\left(Q+\theta I\right)+\theta\right)\left\|U^{k_q} - L^{k_q}\right\|^2$$

alors, pour k tend vers à ∞, on trouve :

$$\lim_{q\longrightarrow\infty}\left\|U^{k_q}-L^{k_q}\right\|^2=\lim_{k\longrightarrow\infty}\left\|U^k-L^k\right\|^2=0$$

s'implique que :

$$0\leq\lim_{q\longrightarrow\infty}\left(\beta_{k_q}-\alpha_{k_q}\right)=\lim_{q\longrightarrow\infty}\left(f\left(x^{k_q}\right)-L_{S^{k_q}}\left(x^{k_q}\right)\right)\leq 0$$

c'est à dire que :

$$\lim_{q\longrightarrow\infty}\left(\beta_{k_q}-\alpha_{k_q}\right)=\lim_{q\longrightarrow\infty}\left(f\left(x^{k_q}\right)-L_{S^{k_q}}\left(x^{k_q}\right)\right)=0$$

on peut alors écrire que :

$$\begin{aligned}
\left(\lim_{q\longrightarrow\infty}\left(\beta_{k_q}-\alpha_{k_q}\right)=0\right) &\Longrightarrow \left(\lim_{q\longrightarrow\infty}\beta_{k_q}=\lim_{q\longrightarrow\infty}\alpha_{k_q}\right)\\
&\Longrightarrow \left(\lim_{q\longrightarrow\infty}\alpha_{k_q}=\lim_{q\longrightarrow\infty}\left(\beta_{k_q}-\left(\beta_{k_q}-\alpha_{k_q}\right)\right)\right)\\
&\Longrightarrow \left(\lim_{q\longrightarrow\infty}\alpha_{k_q}=\lim_{q\longrightarrow\infty}\beta_{k_q}-\lim_{q\longrightarrow\infty}\left(\beta_{k_q}-\alpha_{k_q}\right)\right)\\
&\Longrightarrow \left(\lim_{q\longrightarrow\infty}\alpha_{k_q}=\lim_{q\longrightarrow\infty}\beta_{k_q}\right)\\
&\Longrightarrow \left(\lim_{k\longrightarrow\infty}\alpha_k=\lim_{k\longrightarrow\infty}\beta_k\right)\\
&\Longrightarrow \left(\lim_{k\longrightarrow\infty}\alpha_k=\lim_{k\longrightarrow\infty}\left(f\left(x^k\right)\right)=f\left(x^*\right)\right)
\end{aligned}$$

alors, x^* est une solution optimale globale pour le problème **(PQN)**, c'est à dire que chaque point d'accumulation pour les suites générées par l'algorithme proposée est une solution optimale globale pour le problème **(PQN).** ■

5.5 Exemple :

Soit le problème donné par :

$$\begin{cases} \min f(x) = \left(-\sum_{i=1}^{n} \left(x_i^2 + \frac{1}{2} x_i \right) \right) \\ -1.0 \leq x_i \leq 1.0; i = 1, 2, ..., n \end{cases}$$

5.6 Subdivisions rectangulaires normales :

Soit le rectangle S° définie par :

$$S^\circ = \{ x \in \mathbb{R}^n ; L_i^\circ \leq x_i \leq U_i^\circ : i = 1, 2, ..., n \}$$

un rectangle dans $\mathbb{R}^n$, et soit $\Psi_{S^\circ}(x)$ le sous estimateur linéaire de la fonction optimiser sur le rectangle S° définie par :

$$\Psi_{S^\circ}(x) = \max \{ L_{S^\circ}(x), U_{S^\circ}(x) \} \tag{5.28}$$

d'autre part, notons par x^k et $\alpha\left(S^k\right)$ la solution optimale et la valeur optimale respectivement du le problème linéaire **(LBP)**.

Soit le processus de subdivision rectangulaire, qui génère une famille de rectangles qui peuvent être représenté par un arbre dont la racine est S°.

Définition 5.3 :[6, 21] Un noeud est le successeur d'un autre si et seulement s'il représente un élément de la partition du rectangle correspondant au dernier noeud.

On peut poser un chemin d'arbre de subdivision qui nous donne une suite de subrectangles $\left\{S^k\right\}_{k \in \mathbb{N}}$ comme suit :

$$\begin{tikzcd}
 & & & S^\circ \arrow[dll] \arrow[drr] & & & \\
 & S_{+2} \arrow[dl] \arrow[dr] & & & & S_{+1} \arrow[dl] \arrow[dr] & \\
S_{+22} & & S_{+21} & & S_{+12} & & S_{+11}
\end{tikzcd}$$

$$\swarrow\searrow \quad \swarrow\searrow \quad \swarrow\searrow \quad \swarrow\searrow$$

$$\vdots \quad \vdots \; \vdots \quad \vdots \; \vdots \quad \vdots \; \vdots \quad \vdots$$

$$\curvearrowleft\curvearrowright \; \curvearrowleft\curvearrowright \; \curvearrowleft\curvearrowright \; \curvearrowleft\curvearrowright \; \curvearrowleft\curvearrowright \; \curvearrowleft\curvearrowright \; \curvearrowleft\curvearrowright \; \curvearrowleft\curvearrowright$$

$\boxed{\textit{Arbre de subdivision rectangulaire}}$

On a besoin des définitions suivantes :

Définition 5.4 :[**6, 21**] Une suite de rectangle $\left\{S^k\right\}_{k\in\mathbb{N}}$ est dite normale si :

$$\lim_{k\longrightarrow\infty}\left|f\left(x^k\right)-\Psi_{S^k}\left(x^k\right)\right|=0 \tag{5.29}$$

Définition 5.5 :[**6, 21**]Une subdivision rectangulaire est dite normale si toute suite infinie de rectangles qu'elle génère est normale.

5.6.1 Construction de subdivisions rectangulaires normales :

On va présenter quatre types très importants de subdivision rectangulaire normale qui sont utiles.

Supposons que le rectangle :

$$S^k=\left\{x\in\mathbb{R}^n; L_i^k\leq x_i\leq U_i^k : i=1,2,...,n\right\}$$

est sélectionnée, on a notée plus haut par S_{+1} et S_{+2} les deux subrectangles génèrent par un *bissection méthode* entre les bissections méthodes suivantes, telle que, il y a plusieurs choies de l'indice s et le point h_s qui'est notée plus haut par w_j.

Les bissections méthodes sont :

La bissection exhaustive :

De manière générale, l'indice s est déterminée comme étant l'indice de la plus longue arrête de S^k, c'est à dire telle que :

$$\left(U_s^k-L_s^k\right)^2=\max\left\{\left(U_i^k-L_i^k\right)^2 : i=1,2,...,n\right\} \tag{5.30}$$

avec :

$$h_s = w_j = \frac{U_s^k + L_s^k}{2} \tag{5.31}$$

le point midiale de l'arrête $\left[U_s^k, L_s^k\right]$, et dans se cas, S^k est subdivisée en deux subrectangles :

$$\begin{aligned} S_{+1}^k &= \left\{x \in S^k : L_s^k \leq x_s \leq h_s\right\} \\ S_{+2}^k &= \left\{x \in S^k : h_s \leq x_s \leq U_s^k\right\} \end{aligned} \tag{5.32}$$

ω-subdivision :

Pour un rectangle sélectionné S^k, et soient $L_{S^k}(x)$, $U_{S^k}(x)$ les deux bornes inférieures de la fonction optimiser sur S^k, on peut alors prendre la borne inférieure $\Psi_{S^k}(x)$ de la fonction optimiser définie par :

$$\Psi_{S^k}(x) = \max\left\{L_{S^k}(x), U_{S^k}(x)\right\} \tag{5.33}$$

donc, on a :

$$f(x) - \Psi_{S^k}(x) \geq 0 : \forall x \in S^k$$

et dans se cas, l'indice s est déterminée par le quelle :

$$s \in \arg\max_i \left\{f\left(x_i^k\right) - \Psi_{S^k}\left(x_i^k\right)\right\} \tag{5.34}$$

puis, on subdivisée le rectangle S^k en deux subrectangles :

$$\begin{aligned} S_{+1} &= \left\{x \in \mathbb{R}^n : L_s^k \leq x_s \leq h_s : L_i^k \leq x_i \leq U_i^k; i = 1, 2, ..., n\right\} \\ S_{+2} &= \left\{x \in \mathbb{R}^n : h_s \leq x_s \leq U_s^k : L_i^k \leq x_i \leq U_i^k; i = 1, 2, ..., n\right\} \end{aligned} \tag{5.35}$$

où :

$$h_s = x_s^k \tag{5.36}$$

avec x^k est la solution du problème approximative linéaire **(LBP)**, et dans se cas, l'indice s nous donne la plus grande différence entre $f\left(x_i^k\right)$ et $\Psi_{S^k}\left(x_i^k\right)$.

Subdivision adaptive :

L'indice s est déterminée comme suit :

$$\left|\upsilon_s^k - x_s^k\right| := \max_i \left|\upsilon_i^k - x_i^k\right| \tag{5.37}$$

où x^k est la solution optimale du **(LBP)** et :

$$\upsilon_i^k \in \arg\min_i \left\{\Psi_{S^k}^i\left(L^k\right), \Psi_{S^k}^i\left(U^k\right)\right\} \tag{5.38}$$

qui n'est rien d'autre que :

$$\upsilon^k \in \arg\min_{y \in S^k} \left\{\Psi_{S^k}(y)\right\} \tag{5.39}$$

et dans se cas, le point h_s est définie par :

$$h_s = \frac{\left(\upsilon_s^k + x_s^k\right)}{2} \tag{5.40}$$

Largest distance bisection :

Cette méthode est proposée par Xue et Xu [11], et elle est détailler dans l'algorithme suivante :

Algorithme(4) : *LDB*

initialisation : déterminée le rectangle initial S° telle que $\left(\chi_f\right) \subset S^\circ$, posons $S^k := S^\circ$ pour $k \in \{1, 2, ...\}$;

itération k :

étape1 calculée $\delta_i = \frac{f_i\left(U_i^k\right) - f_i\left(L_i^k\right)}{U_i^k - L_i^k}$ pour $i = 1, 2, ..., n$

définie la distance entre $\Psi_{S^k}^i(x_i)$ et $f_i(x_i)$ pour $x_i \in \left[L_i^k, U_i^k\right]$ par :

$d_i(x_i) = f_i(x_i) - \Psi_{S^k}^i(x_i) = f_i(x_i) - \delta_i x_i - \eta_i$;

étape2 maximisée la distance $d_i(x_i)$ pour trouver la solution h_i^k,

posons $d_i\left(h_i^k\right)$ être le maximum ;

étape3 déterminée l'indice s telle que :

$d_s\left(h_s^k\right) = \max\left\{d_i\left(h_i^k\right) : i = 1, 2, ..., n\right\}$;

étape4 déterminée le point h_s^k telle que $h_s^k = x_s^k$ ou bien, h_s^k est le point déterminée dans *l'étape2*.

ANNEXES :

Annexe I ***Application particulière de la méthode de la transformation duale canonique*** **:**

Entre les applications très importantes de la transformation duale canonique et les méthodes paramétriques, il y a la programmation quadratique sur un sphère, alors, on considère le problème suivant :

$$\begin{cases} \min f(x) = \frac{1}{2}x^T Qx - d^T x \\ x^T x \leq 2\mu \end{cases}$$

qui présente un problème quadratique avec un seul contrainte quadratique,

si on utilise la norme $\|.\|_2$, on peut alors présenter le problème presque préçédent comme suit :

$$\begin{cases} \min f(x) = \frac{1}{2}x^T Qx - d^T x \\ \frac{1}{2}\|x\|_2^2 \leq \mu \end{cases}$$

il est défficile de trouver la solution exacte de se type de problèmes, à cause de ça, plusieurs méthodes étaient proposées pour trouver la solution approximée de ce problème, d'autre part, entre ses méthodes il y a *la transformation duale canonique.*

Remarquons que la contrainte d'inégalité $Bx \leq b$ est absente, alors, avec l'utilisation de cette méthode ce type de problèmes peut être résoud complètement.

En effet, la forme duale canonique associer à ce problème est donnée par :

$$\begin{cases} \text{Ext}\, f^d(\rho^*) = \frac{-1}{2}d^T (Q + \rho^* I)^{-1} d - \mu\rho^* \\ \rho^* \geq 0, \det(Q + \rho^* I) \neq 0 \end{cases}$$

qui présente une maximisation concave avec un seul degré de libertie.

Le théorème suivant nous donne un ensemble complet des solutions.

Théorème 1 :[5] *Soit la matrice symétrique Q admet $p \leq n$ valeurs propres distinctes, avec l'indice $id \leq p$ telle que :*

$$\lambda_1 < \lambda_2 < \ldots < \lambda_{id} < 0 \leq \lambda_{id+1} < \ldots < \lambda_p$$

alors, on a :

(1) *Pour un vecteur donné* $d \in \mathbb{R}^n$ *et un paramètre suffisamment large* $\mu > 0$, *le problème dual canonique* (P_μ^d) *admet au plus de* $(2id+1)$ *points de* (KKT) *notées par :*

$$\left\{\overline{\rho_i^*}; i = 1, 2, ..., 2id+1\right\}$$

et satisfais la relation suivante :

$$\overline{\rho_1^*} > -\lambda_1 > \overline{\rho_2^*} \geq \overline{\rho_3^*} > -\lambda_2 > \overline{\rho_4^*} \geq \overline{\rho_5^*} > -\lambda_3 > ... > -\lambda_{id} > \overline{\rho_{2id}^*} \geq \overline{\rho_{2id+1}^*} > 0$$

(2) *Pour chaque* $\left\{\overline{\rho_i^*}; i = 1, 2, ..., 2id+1\right\}$ *le vecteur définie par :*

$$\overline{x_i} = \left(Q + \overline{\rho_i^*} I\right)^{-1} d$$

est un point de (KKT) *pour le problème paramétrique* (P_μ), *et on a aussi :*

$$\left\{f\left(\overline{x_i}\right) = f^d\left(\overline{\rho_i^*}\right) : i = 1, 2, ..., 2id+1\right\}$$

(3) *Si l'indice* $id > 0$ *alors, le problème dual canonique* (P_μ^d) *admet au plus de* $(2id+1)$ *points de* (KKT) *sur la bornitude de sphère définie par :*

$$\left\{\frac{1}{2}\left|x_i\right|^2 = \mu : i = 1, 2, ..., 2id+1\right\}$$

et dans se cas, le point $\overline{x_i}$ *définie préçidament est un minimum global de* (P_μ).

Preuve :

par hypothèse on a :

$$\lambda_1 < \lambda_2 < ... < \lambda_{id} < 0 \leq \lambda_{id+1} < ... < \lambda_p$$

les valeurs propres de la matrice Q, alors,

(1) Puisque la matrice Q est symétrique alors, il existe une matrice orthogonale R telle que :

$$R^T = R^{-1} \text{ et } Q = R^T D R$$

avec D une matrice diagonale définie par :

$$D = \lambda_i \delta_{ij}$$

donc, pour un vecteur $d \in \mathbb{R}^n$ on peut supposer que :

$$\left.\begin{array}{c} g = Rd = (g_i)_i \\ \Psi(\rho^*) = \frac{1}{2} d^T (Q + \rho^* I)^{-2} d \end{array}\right\}$$

Sachons que :

$$\Psi(\rho^*) = \frac{1}{2} \sum_{i=1}^{p} g_i^2 (\lambda_i + \rho^*)^{-2}$$

en effet, on a :

$$\begin{aligned}
\Psi(\rho^*) &= \frac{1}{2} d^T (Q + \rho^* I)^{-2} d \\
&= \frac{1}{2} d^T \left(R^T D R + \rho^* I\right)^{-2} d \\
&= \frac{1}{2} d^T R^T (D + \rho^* I)^{-2} (Rd) \\
&= \frac{1}{2} (Rd)^T (D + \rho^* I)^{-2} (Rd) \\
&= \frac{1}{2} g^T (D + \rho^* I)^{-2} g \\
&= \frac{1}{2} \sum_{i=1}^{p} g_i^2 (\lambda_i + \rho^*)^{-2}
\end{aligned}$$

D'autre part, la fonction qui définit presque préçédent est une fonction strictement convexe avec des valeurs réels dans chaque intervalle $]-\lambda_{i+1}, -\lambda_i[$, donc, pour un paramètre $\mu > 0$ suffisamment large, l'équation algébrique suivant :

$$\Psi(\rho^*) = \frac{1}{2} \sum_{i=1}^{p} g_i^2 (\lambda_i + \rho^*)^{-2} = \mu$$

admet au plus $(2p)$ solutions notons par $\left\{\overline{\rho_i^*}\right\}_i$ satisfais les deux relations :

$$\left\{-\lambda_{j+1} < \overline{\rho_{2j+1}^*} \leq \overline{\rho_{2j}^*} < -\lambda_j : j = 1, 2, ..., p-1\right\}$$

et

$$\overline{\rho_1^*} > -\lambda_1 \text{ et } \overline{\rho_{2p}^*} < -\lambda_p$$

d'un autre cotée, la matrice Q admet id valeurs propres négatives alors, l'équation algébrique :

$$\Psi\left(\rho^*\right) = \mu$$

admet au plus $(2id+1)$ racines strictement positifs notées par :

$$\left\{\overline{\rho_i^*} > 0 : i = 1, 2, ..., 2id+1\right\}$$

Et avec la condition de *(KKT)* donnée par :

$$\left(\overline{\rho_i^*}\right)^T \left(\frac{1}{2}\left|x_i\right|^2 - \mu\right) = 0$$

on trouve que le problème primal paramétrique (P_μ) admet au plus $(2id+1)$ points de *(KKT)* notées par $\overline{x_i}$ sur l'ensemble :

$$\left\{\frac{1}{2}\left|x_i\right|^2 = \mu\right\}$$

Donc, la relation une est satisfaite. $\natural$

(2) Pour $d \in \mathbb{R}^n$ un vecteur donné et un paramètre positif $\mu > 0$, on dit que le scalaire $\overline{\rho^*}$ est un point de *(KKT)* si et seulement si :

$$0 \leq \overline{\rho^*} \perp \frac{1}{2} d^T \left(Q + \overline{\rho^*} I\right)^{-2} d - \mu \leq 0$$

Prend par exemple le point :

$$\overline{x_i} = \left(Q + \overline{\rho^*} I\right)^{-1} d$$

alors, la condition presque préçédente devient :

$$0 \leq \overline{\rho^*} \perp \frac{1}{2}\overline{x}^T\overline{x} - \mu \leq 0$$

En effet, on a :

$$0 \leq \overline{\rho^*} \perp \frac{1}{2}d^T \left(Q + \overline{\rho^*}I\right)^{-1} \left(Q + \overline{\rho^*}I\right)^{-1} d - \mu \leq 0$$

qui implique que :

$$0 \leq \overline{\rho^*} \perp \left(\left(Q + \overline{\rho^*}I\right)^{-1} d\right)^T \left(Q + \overline{\rho^*}I\right)^{-1} d - \mu \leq 0$$

alors :

$$0 \leq \overline{\rho^*} \perp \frac{1}{2}\overline{x}^T\overline{x} - \mu \leq 0$$

qui présente la condition de *(KKT)* pour le problème original **(P)**, donc, le point $\overline{x_i}$ qui définie précidament est un point de *(KKT)* pour **(P)**, alors, on a trouvée que pour chaque point de *(KKT)* $\overline{\rho^*}$ pour le problème dual paramétrique (P_μ^d) le vecteur qui définie par :

$$\overline{x_i} = \left(Q + \overline{\rho^*}I\right)^{-1} d$$

est un point de *(KKT)* pour le problème primal paramétrique (P_μ), d'autre part, on a :

$$\begin{aligned}
f\left(\overline{x}\right) &= \frac{1}{2}\overline{x}^T Q\overline{x} - d^T\overline{x} \\
&= \frac{1}{2}\left(\left(Q+\overline{\rho^*}I\right)^{-1}d\right)^T Q\left(\left(Q+\overline{\rho^*}I\right)^{-1}d\right) - d^T\left(\left(Q+\overline{\rho^*}I\right)^{-1}d\right) \\
&= \frac{1}{2}d^T\left(Q+\overline{\rho^*}I\right)^{-1}Q\left(Q+\overline{\rho^*}I\right)^{-1}d - d^T\left(Q+\overline{\rho^*}I\right)^{-1}d \\
&= -d^T\left(Q+\overline{\rho^*}I\right)^{-1}d + \frac{1}{2}d^T\left(Q+\overline{\rho^*}I\right)^{-1}\left(Q+\overline{\rho^*}I-\overline{\rho^*}I\right)\left(Q+\overline{\rho^*}I\right)^{-1}d \\
&= -d^T\left(Q+\overline{\rho^*}I\right)^{-1}d + \frac{1}{2}d^T\left(Q+\overline{\rho^*}I\right)^{-1}\left(Q+\overline{\rho^*}I\right)\left(Q+\overline{\rho^*}I\right)^{-1}d \\
&\quad -\frac{1}{2}d^T\left(Q+\overline{\rho^*}I\right)^{-1}\left(\overline{\rho^*}I\right)\left(Q+\overline{\rho^*}I\right)^{-1}d \\
&= -d^T\left(Q+\overline{\rho^*}I\right)^{-1}d + \frac{1}{2}d^T\left(Q+\overline{\rho^*}I\right)^{-1}d - \frac{1}{2}\overline{\rho^*}\overline{x}^T\overline{x} \\
&= -\frac{1}{2}d^T\left(Q+\overline{\rho^*}I\right)^{-1}d - \frac{1}{2}\overline{\rho^*}\overline{x}^T\overline{x}
\end{aligned}$$

mais on a la condition d'optimalité suivante :

$$\frac{1}{2}\overline{\rho^*}\overline{x}^T\overline{x} = \overline{\rho^*}\mu$$

On peut alors écrire :

$$\begin{aligned}
f\left(\overline{x}\right) &= -\frac{1}{2}d^T\left(Q+\overline{\rho^*}I\right)^{-1}d - \frac{1}{2}\overline{\rho^*}\overline{x}^T\overline{x} \\
&= -\frac{1}{2}d^T\left(Q+\overline{\rho^*}I\right)^{-1}d - \overline{\rho^*}\mu \\
&= f^d\left(\overline{\rho^*}\right)
\end{aligned}$$

alors, la deuxième relation est satisfaite. ♮

(3) Puisque la fonction $\Psi\left(\rho^*\right)$ est strictement convexe, les points de *(KKT)* :

$$\{\overline{x_i} : i = 1, 2, ..., 2id + 1\}$$

sont des minimums globaux pour le problème primal paramétrique (P_μ). ■

Annexe II ***Preuve de lemme3.4*** **:**

(1) La première relation présente le théorème classique de min-max *saddle Lagrangian.*

$\natural$

(2) Puisque la fonction lagrangien est concave en x et v alors, si le vecteur $(\overline{x}, \overline{v}) \in (X_f) \times \mathbb{R}_+^m$ est un point de *(KKT)* alors, on a :

$$\max_{x \in X_f} L(x, \overline{v}) = L(\overline{x}, \overline{v}) = \max_{v \in \mathbb{R}_+^m} L(\overline{x}, v) \tag{A}$$

d'autre part, pour x un point de (X_f) on peut écrire :

$$L(x, \overline{v}) = \max_{v \in \mathbb{R}_+^m} L(x, v) = \max_{v \geq 0} L(x, v) = f^d(v) \tag{B}$$

et pour chaque $v \geq 0$ donné, on a :

$$L(\overline{x}, v) = \max_{x \in \mathbb{R}^n} L(x, v) = \max_{x \in X_f} L(x, v) = f^d(v) \tag{C}$$

D'après tous sa, on peut écrire :

$$\begin{aligned} \max_{x \in X_f} L(x, \overline{v}) &= {}^B \max_{x \in X_f} \max_{v \in \mathbb{R}_+^m} L(x, v) =^A L(\overline{x}, \overline{v}) \\ &= {}^C \max_{v \in \mathbb{R}_+^m} \max_{x \in X_f} L(x, v) \end{aligned}$$

alors, la deuxième relation est satisfaite. $\natural$

(3) D'après la relation (B) on a :

$$\min_{x \in X_f} L(x, \overline{v}) =^B \min_{x \in X_f} \max_{v \in \mathbb{R}_+^m} L(x, v) =^A L(\overline{x}, \overline{v})$$

D'un autre côté on a :

$$\min_{v \geq 0} L(\overline{x}, v) =^C \min_{v \geq 0} \max_{x \in \mathbb{R}^n} L(x, v) =^A L(\overline{x}, \overline{v})$$

Donc, d'après les deux relations presque précédentes, la troisième relation est satisfaite. $\blacksquare$

Conclusion :

Dans ce mémoire on s'est intéressé à l'étude de quatre méthodes pour résoudre les problèmes de la programmation quadratiques non convexes.

- **La première méthode DCA** est basée sur la théorie de la programmation **DC** (différence de deux fonctions convexes). Elle est de type *primal dual*.

- **La deuxième méthode** est appelée *la transformation duale canonique* ***(TDC)***. Dans cette méthode on ajoute une condition très importante est appelé *la condition de normalité* (région de confiance) pour assurer l'existence de la solution globale. On transforme d'abord le problème initial à un problème paramétrique équivalent, puis, on définie le problème dual canonique sous forme d'un système algébrique facile à résoudre.

- **La troisième méthode** est de type Branch and Bound. On utilise la structure propre de la forme quadratique pour définir une autre forme séparable de type convexe. Et on a étudie l'erreur de cette approximation.

- **La quatrième méthode** est de type *Branch and Bound.* La différence entre cette méthode et la troisième méthode réside dans la stratégie d'évaluation et de séparation. Dans cette méthode on présente une nouvelle forme linéaire approximante inférieurement la fonction optimiser $f(x)$ sur les n-rectangles, et on explicite les techniques de partition et de réduction sur un rectangle choisie.

Nous pensons qu'un effort consédirable doit être porté sur la mise en oveure de ces méthodes sur des exemples de grand taille et proposition des variantes à ces méthodes.

Bibliographie

[1] **Abdelkrim Keraghel**, *Élements d'analyse convexe dans* $\mathbb{R}^n$, *Théorie fondamentale et exercises*, Université de Sétif, avril, 2001.

[2] **Dominique Azé**, *élement d'analyse convexe et variationnelle, Ellipses, Paris, **1997**.*

[3] **L. Ekeland, R. Temam**, *Analyse convexe et Problèmes Variationnels*, Dunod, Paris, **1974**.

[4] **François Bertrand AKOA**, *Approches de points intérieurs et de la programmation DC en optimisation non convexe.* Thèse de Doctorat, Laboratoire de mathématiques de l'institut national des sciences appliquées de Rouen.**2005**.

[5] **David Yang GAO**. *Canonical duality theory and solutions to constrained non convex quadratic programming.* Dedicated to professor Ivar Ekeland on the occasion of his 60th birthday. Journal of global optimisation, 29 pp 377-399. Mathematics subject classifications 90C, 49N. Kluwer academic publishers, **2004**.

[6] **Yuelin Gao, Honggang Xue, Peiping Shen**, *A new rectangle Branch and reduce approch for solving non convex quadratic programming problèms,* Applied mathematics and computation 168 (**2005**), 1409-1418.

[7] **Reiner Horst. Panous. M. Pardalos and Nguyen V. Thoai**, *introduction to global optimisation.* Kluwer academic publishers. Dord Echt/ Boston/ London. Volume 3.**(1995)**.

[8] **J.B. Hiriart-Urruty**, *Generalized differntiability, duality and optimization for problems dealing with differences of convexs functions,* Lecture notes in Economics en Math. Systems, 256 (**1985**), pp 37-70.

[9] **J.B. Hiriart-Urruty**, *From convex optimization to non convex optimization.* Part I : *Necessary and sufficient conditions for global optimality, Non smooth optimization and related topics,* Ettore Majorana international Sciences, Series 43, Plenum Press.

[10] **Le Thi Hoai An, Pham Dinh Tao**, *The DC programming and DCA revisited with dc models of real world non convex optimization problems.* Annals of operations research 133, pp 23-46 (**2005**).

[11] **Le Thi Hoai An**, *Analyse numérique des algorithmes de l'optimisation DC, Approches locale et globale,* Thèse de doctorat, Université de Rouen, Décembre **1994**.

[12] **Le Thi Hoai An, Pham Dinh Tao**, *DCprogramming : Theory, Algorithms and applications.* (**2002)**

[13] **Panos. M. Pardalos**, *Global optimisation algorithms for linearly constrained indéfinite quadratic problems,* Computers math. applic. Vol. 21, NO 6/7, pp 87-97 (**1991**).

[14] **Panos. M. Pardalos**, **H. Edivin Romeijn, Hoang Tuy**. *Recent developments and trends in global optimisation.* Journal of computational and applied mathémtics 124 (**2000**). PP 209-228

[15] **Pham Dinh Tao and Le Thi Hoai An,** *DCA for solving the Trust-region subproblem.* SIAM J. optimisation, pp 476-505, **1998**.Vol. 8m. N02.

[16] **Pham Dinh Tao and Le Thi Hoai An,** *Convex analysis approach to DC programming : Theory, Algorithms and Applications,* Acta Mathamatica Vietnamica, dedicated to professor Hoang Tuy on the occasion of his 70th birthday, Vol.22, NO 1 (**1997**).

[17] **P.Hansen, B.Jaumard, G.Savard**, *New branch-and-bound rules for linear bilevel programming, SIAM J. Sci. Stat. Comp.* 13 (**1992**) 1194-1217.

[18] **R.T.Rockafillar**, *Convex Analysis*, Princeton Mathematic Ser.28.Princeton Univerity Press, Princeton, New Jersey, **1993**.

[19] **D. Rachid Benacer**, *contribution a l'etude des algorithmes de l'optimisation non convexe et non differentiable. Thèse de doctorat,* univercité scientifique, Technologique et médicale de Grenoble, **1986**

[20] **M. Selmani,** *Méthodes de résolution de problèmes avec contraintes. Thèse de doctorat,* Institut de mathématique appliquées de Grenoble. Sept-**1985**.

[21] **Xue Honggang, Xu Chengxian**, *A Branch and Bound algorithme for solving a class of DC-Programming.* Applied mathematics and computation 165 (**2005**). pp 291-302.

Résumé :

Ce travail est principalement consacré à l'association des méthodes de recherche globale pour résoudre les problèmes d'optimisation quadratiques en général non convexes.

Nous avons utilisées quatre méthodes différentes :

1. L'Algorithme DC (différence de deux fonctions convexes).
2. La méthode de la transformation duale canonique (TDC).
3. La méthode de la Séparation et interpolation.
4. méthode de séparation et évaluation qui' est appelée aussi Branch and Bound.

Mots clés :

Programmation DC, Optimisation globale, Programmation quadratique non convexe, Méthodes de Branch and bound, Tactique de réduction rectangulaire.

Abstract :

This work is mainly devoted to global research method four the quadratic optimization problems.

We are chose for different methods witch are :

1. The DC Algorithm (different of tow convex function).
2. The canonical dual transformation method (CDT(.
3. The Branch and Bound method.
4. The new Branch and Reduce method.

Key words :

DC-programming, Global optimization, Non convex quadratic programming, Branch and bound method, Rectangle reducing tactics.