REPUBLIQUE ALGERIENNE DEMOCRATIQUE et POPULAIRE.
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique.

UNIVERSITE MOULOUD MAMMERI, TIZI-OUZOU
Faculté des Sciences
Département de Mathématiques

MÉMOIRE DE MASTER
en
MATHÉMATIQUES

Spécialité

Recherche Opérationnelle : Méthodes et modèles de décision

Thème

L'amélioration du DCA dans la recherche d'un Optimum Global d'une Fonction non convexe

Présenté par

OUZOUIR Noredine
BOUZID Mahrez

Devant le jury d'examen composé de :

M. djamal Hamadouch ;	Professeur ;	président
Mme Fadila Leslous ;	M. Assist A ;	Promoteur
M. Mohand Ouanes ;	Professeur ;	Examinateur

Soutenu le 28 \ 06 \ 2016.

Remerciements

Nous remercions avant tout ALLAH le tout puissant de nous avoir donné la force et le courage de réaliser ce travail.

Nous tenons à remercier notre promotrice, Mme LESLOUS, qui nous a encadrés pendant notre processus,et nous tenons à lui exprimer notre profonde reconnaissance pour son très bon encadrement, sa patience, sa gentillesse, sa disponibilité, nous la remercions vivement pour le temps qu'elle a consacrée à notre mémoire, sa confiance, de nous proposés un thème très intéressant.

Sans oublier d'exprimer notre reconnaissance et gratitude à tous nos enseignants durant notre cursus au département Mathématiques.

Nous tenons à remercier très sincèrement l'ensemble des membres du jury qui nous font le grand honneur d'avoir accepté de juger notre travail.

Dédicace

Je dédier ce modeste travail

- A mes très chers parents.
- A mes fréres : Mouhamed, Sofiane et sa petite famille surtout Lounes.
- A mes soeurs : Nassima, Souad et son mari Zahir.
- A toute la famille.
- A tous mes enseignants.
- A tous mes amies et plus précisément à :
  Farid, adnane, abd allah, Djamal, Noureddine, Mohamed, Lhachmi, Rafik, abd alrazaq .

O.Noredine

Dédicace

Je dédie ce travail

- A mes chers parents et grands parents
- A ma Sœur Meriam et son conjoint et particulièrement mes neveux Malek et Samy.
- A mes deux frères Naguib et Menad. Et a mes cousins.
- A mes tantes (Dahbia,Zhor,Malika,Samia et Nadia) et leur familles. Et a toute ma famille de plus proche et le plus loin ?
- A tous mes amis et particulièrement Rabah, Achour, Smail, Karim, Hakim et Nacer.
- A tous enseignants et enseignantes de département de Mathématique.

B.Mahrez

Table des matières

Chapitre 1

Introduction générale

1.1 La programmation DC et DCA

L'optimisation est une branche importante des mathématiques.En général, elle consiste à déterminer les meilleurs éléments d'un ensemble au sens d'un critère donné. On peut rencontrer des problèmes d'optimisation dans divers domaines, par exemple l'economie, la gestion de portefeuille, la théorie des jeux, la gestion de production, la télécommunication,... pour résoudre ces problèmes, l'optimisation offre un cadre algorithmique très riche. on peut distinguer deux branches de l'optimisation déterministe : la programmation convexe et la programmation non convexe. La résolution d'un programme convexe permet d'établir des caractérisations (sous forme de conditions nécessaires et suffisantes) de solutions optimales et ainsi de construire des méthodes itératives convergeant vers des solutions optimales.

L'optimisation non convexe connaît une explosion spectaculaire depuis une quinzaine d'années car dans les milieux industriels, on a commencé à remplacer les modèles convexes par des modèles non convexes plus complexes mais plus fiables qui présentent mieux la nature des problèmes étudiés. Durant ces dernières années, la recherche en optimisation non convexe a largement bénéficié des efforts des chercheurs et s'est enrichie de nouvelles approches. Nous pouvons distinguer deux approches différentes mais complémentaires en programmation non convexe :

1. Approches continues globales dont les nouveaux outils algorithmique sont inspirés par les techniques combinatoires de la Recherche Opérationnelle. Elles consistent à localiser les solutions optimales à l'aide des méthodes d'approximation, des techniques de coupe, des méthodes de décomposition, de séparation et évaluation. Elles ont connu de très nombreux développements importants depuis plus d'un quart de siècle, à travers les travaux de Hoang Tuy (reconnu comme le pionnier), R. Horst [19], Pham Dinh Tao [15], NIU-Yi-Shuaiy [1]...

L'inconvénient majeur des méthodes globales est leur cout exhorbitant en temps de calcul qui les empêche de traiter, en pratique, des problèmes d'optimisation non convexe réels de très grande taille. Et tout le monde convient qu'il faut développer les approches locales performantes et économiques capables de résoudre ces problèmes à très grande dimension. On rejoint ainsi la communauté de l'Optimisation Non Linéaire avec des outils théoriques et algorithmique spécifiques : la Programmation DC et DCA. Dans ce contexte, les problèmes d'optimisation combinatoire sont reformés grâce à des techniques de pénalité exacte en Programmation DC et traités par DCA.

2. Approches locales et globales d'analyse convexe qui sont basées sur l'analyse et l'optimisation convexe. Ici la programmation DC (Difference of Convex functions) et DCA (DC Algorithms) jouent le rôle central car la plupart des problèmes d'optimisation non convexe sont formulés/reformulés sous la forme des programmes DC. La programmation DC et DCA constituent l'épine dorsale de la programmation non convexe et de l'optimisation globale. Ils sont introduits par Pham Dinh Tao en 1985 [30, 31] à l'état préliminaire et développes intensivement à travers de nombreux travaux communs de Le Thi Hoai An et Pham Dinh Tao depuis 1994 [27, 28, 29, 32] pour devenir maintenant classiques et de plus en plus utilisés par des chercheurs et praticiens de par le monde, dans différents domaines des sciences appliquées. Ces outils théoriques et algorithmique constituent une

extension de l'analyse et l'optimisation convexes, assez large pour couvrir la quasi-totalité des problèmes d'optimisation non convexe de la vie courante.

La programmation DC inclut les problémes de la programmation convexe et aussi presque tous les problèmes d'optimisation non convexes. La forme standard d'un programme DC est :

$$(P_{dc}) \qquad \alpha = \inf\{f(x) = g(x) - h(x) : x \in \mathbb{R}\}$$

où g, h : $\mathrm{R}^n \to \mathbb{R}\bigcup\{+\infty\}$ sont convexes semi-continues inférieurement et propres. Une telle fonction f est appelée fonction DC et g, h sont des composantes DC de f.

Pour une utilisation optimale de DCA, il est crucial d'apporter une réponse à ces deux questions :

1. Comment trouver une décomposition DC pour la fonction objectif f bien adaptée à la structure spécifique du probléme traité ?
2. Quelle stratégie d'initialisation utiliser pour DCA ?

En pratique, DCA est moins coùteux et ainsi capable de traiter les programmes non convexes de grande dimention.

Nous détaillerons dans cet partie les bases théoriques et algorithmique de la programmation DC et DCA. Nous commencerons par quelques rappels d'analyse convexe, ensuite nous présenterons la programmation DC sur lesquelles est basé l'algorithme DCA. Nous terminerons par des applications de DCA.

1.2 L'amélioration du DCA (MDCA)

dans cette partie nous présentons un Algorithme pour résoudre un probléme non convexe DC sur un intervalle [a,b] de R, on utilise le DCA (Difference of convexe Algorithm) et le minimum de la fonction moyenne de deux approximation de f à partir de a et b.

Cette stratégie a l'avantage de fournir en général un minimum qui sera situé dans la zone d'attraction du minimum global cherché.

Comme le DCA est très sensible au choix du point de dèpart, nous proposons donc de ne pas choisir de point de départ.

On applique le DCA à partire de ce minimum trouvé, on arrive au minimum global cherché.

Chapitre 2

Les techniques Principales de la programmation DC et DCA

2.1 La programmation DC

2.1.1 Introduction

Les fonctions DC possèdent de nombreuses propriétés importantes qui ont été établies à partir des années 50 par Alexandroff (1949) [22], Hartman (1959)[23], une des principales propriétés est leur stabilité relative aux opérations fréquemment utilisées en optimisation. Cependant, il faut attendre le milieu des années 80 pour que la classe des fonctions DC soit introduite en optimisation, élargissant ainsi la classification des problèmes d'optimisation avec l'apparition de la programmation DC. On distingue deux grandes approches DC :

1. L'approche combinatoire (cette terminologie est due au fait que les nouveaux outils introduits ont été inspirés par les concepts de l'optimisation combinatoire) en optimisation globale continue, et

2. L'approche de l'analyse convexe en optimisation non convexe.

2.1.2 Eléments d'analyse convexe

cette section est consacrée à un rapide rappel d'analyse convexe.qui fondent la base de la programmation DC. Pour plus de détails en analyse convexe, on pourra se reporter, par exemple aux ouvrage de R.T. Rockafellar [33], Hiriart-Urruty [34].

soit X l'espace euclidienne dans $\mathbb{R}^n$ muni du produit scalaire usuel noté $\langle .,. \rangle$ et la norme euclidienne associée $\|x\| = \langle x, x \rangle^{\frac{1}{2}}$ et Y l'espace vectoriel dual de X relatif au produit scalaire, que l'on peut identifier $\acute{a}$ X. Lanalyse convexe moderne permet aux fonctions de prendre les valeurs $\pm\infty$.

On notera $\overline{\mathbb{R}} = \mathbb{R} \bigcup \{\pm\infty\}$ muni d'une structure algébrique déduite de celle de $\mathbb{R}$ avec la $(+\infty) - (+\infty) = +\infty$.

Définition 2.1. (Ensemble convexe) un ensemble $C \subset X$ est dit convexe lorsque, pour tout x et y de C, le segment $[x, y]$ est tout entier contenu dans C, c.à.d,

$$\forall \lambda \in [0,1], \ \lambda x + (1 - \lambda) y \in C.$$

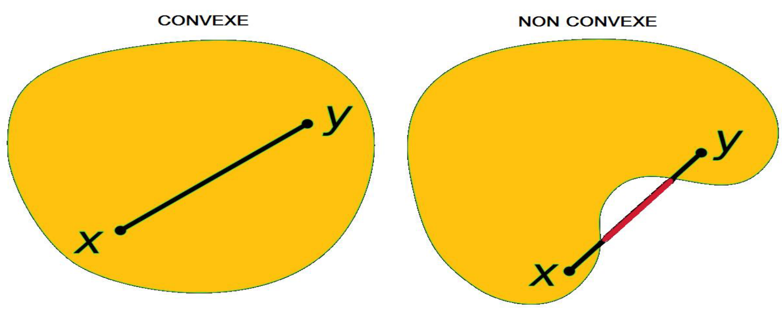

Fig.2-1 Ensemble Convexe et Non convexe

Définition 2.2. (envloppe convexe) soit $S \subset X$. L'envloppe convexe de S, notée co(S), est l'intersection de tous les sous-ensembles convexes de X qui contiennent S. elle est en effet le plus petit ensemble convexe de X qui contient S.

Sachant que X est un espace vectoriel de dimension finie, on peut en déduire facilement que co(S) est l'ensemble des points $x \in X$ qui peuvent s'écrire sous forme de combinaison convexe finie d'éléments de S, c.à.d.

$$\text{co(S)} = \{ \ x = \sum_{i=1}^{m} \lambda_i x^i : x^i \in S, \ \lambda_i \geq 0, \ \forall i = 1, ..., m; \ \sum_{i=1}^{m} \lambda_i = 1\}.$$

Théorème 2.1. ***(Théorème de carathéodory)*** *Dans un espace vectoriel de dimension finie n, co(S) est l'ensemble des borycentres à coefficients positifs ou nuls de familles de n+1 points de S.*

Définition 2.3. (Variété affine) Soit $C \in X$ un ensemble convexe non vide. La variété affine engendrée par C est notée $aff(C)$ et se définit par :

$$aff(C) = \{ \ \sum_{i} \lambda_i x^i : x^i, \ \sum_{i} \lambda_i{=}1, \ \lambda_i \in \mathbb{R}\}$$

où seules les sommes finies sont prises en compte.

Définition 2.4. (Intérieur et intérieur relatif d'un convexe) Soit $C \in X$ un ensemble convexe non vide. L'intérieur du convexe C est noté $intC$ et se définit par :

$$intC = \{x \in C : \exists r > 0, \ B(x, r) \subset C\},$$

ou $B(x, r)$ désigne la boule euclidienne de rayon r centrée en x. L'intérieur relatif de C est noté ir(C) et se définit par

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

où $aff(C)$ désigne la variété affine de C.

Définition 2.5. (Domaine d'une fonction) Soit $f : X \to]-\infty, +\infty]$ Le domaine de la fonction f est l'ensemble, noté $domf$, défini par

$$domf = \{x \in X : f(x) < +\infty\}.$$

On dit que f est propre si $domf \neq \varnothing$ (i.e. elle n'est pas identiquement égale à l'infini).

Définition 2.6. [2] **(Fonction coercive)** Une fonction $f : X \to]-\infty, +\infty]$ est dite coercive si elle vérifie :

$$\lim_{\|x\| \to \infty} f(x) = +\infty$$

.

Définition 2.7. (Fonction propre) La fonction f est dite propre si elle ne prend jamais la valeur $-\infty$ et si elle n'est pas identiquement égale à $+\infty$.

Définition 2.8. (Fonction convexe) une fonction $f : X \to]-\infty, +\infty]$ est dite convexe si elle vérifie :

$$\forall(\ (x, y) \in X^2\ )\ (\ \forall\lambda \in]0, 1[\ )\ f(\ \lambda x + (1-\lambda)y\ ) \leq \lambda f(x) + (1-\lambda)f(y). \tag{3.1}$$

si l'inégalité est stricte pour tout $\lambda \in]0, 1[$ et tout x et y dans C avec $x \neq y$, alors f est dite strictement convexe.

Cette définition est aussi valable pour $f : X \to \mathbb{R} \cup \{+\infty\}$. il est clair que si $f : X \to \mathbb{R} \cup \{+\infty\}$ est convexe alors $dom(f)$ est convexe et la restriction $F(f)$ est une fonction convexe à valeurs dans $\mathbb{R}$ sur dom(F).

Remarque 2.1.

- Soit un ensemble $S \subset X$. On appelle fonction indicatrice de S, notée $\chi_s(x)$, définie par $\chi_s(x) = 0$ si $x \in S$; et $+\infty$ dans le cas contraire. La fonction $\chi_s(x)$ est convexe si et seulement si S est un ensemble convexe.

- Soit $C \subset X$ un ensemble convexe. La fonction $f : C \to \mathbb{R}$ est convexe si et seulement si la fonction $F + \chi_s(x)$ est convexe sur X à valeurs dans $\mathbb{R} \cup \{+\infty\}$.

Définition 2.9. (Épigraphe) L'épigraphe d'une fonction $F : X \to]-\infty, +\infty]$ est noté Epi(f) et se définit par :

$$Epi(f) = \{(x, \alpha) \in X \times \mathbb{R} : f(x) \leq \alpha\}.$$

La fonction f est convexe si et seulement si Epi(f) est un sous-ensemble convexe de $X \times \mathbb{R}$.

Définition 2.10. (Fonction fortement convexe) La fonction $f : C \to \mathbb{R}$ est dite fortement convexe de module ρ sur l'ensemble convexe C (autrement dit $\rho - convexe$) lorsqu'il existe $\rho > 0$ tel que pour tout x et y de C et tout $\lambda \in]0, 1[$, on ait :

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) - \lambda(1-\lambda)\frac{\rho}{2}\|x - y\|^2.$$

en d'autres termes la fonction $f - \frac{\rho}{2}\|.\|^2$ est convexe sur C. Le module de forte convexité de f sur C, noté $\rho(f, C)$, est défini par :

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

ainsi la fonction f est fortement convexe sur C si et seulement si $\rho(f, C) > 0$. on peut aussi remarquer que toute fonction fortement convexe est strictement convexe et toute fonction strictement convexe est convexe.

Définition 2.11. (Sous-gradient) soit une fonction convexe propre $f : X \rightarrow]-\infty, +\infty]$. un vecteur y dans Y est appelé sous-gradient de f au point $x^0 \in dom(f)$ si pour tout $x \in X$ on a :

$$f(x) \geq f(x^0) + \langle y, x - x^0 \rangle.$$

Définition 2.12. (Sous-différentiel) L'ensemble de tous les sous-gradients de f au point $x^0(f)$ est le sous-différentiel de f au point x^0, noté $\partial f(x^0)$.

Le domaine du sous-différentiel de f est $dom(\partial f) = \{x : \partial f(x) \neq \varnothing\}$.

Théorème 2.2. [2]

$*$ *$\partial f(x)$ est une partie convexe fermée de Y.*

$*$ *ir(dom(f)) $\subset dom(\partial f) \subset dom(f)$.*

Définition 2.13. (ε-sous-gradient) Soit ε un réel positif. Un vecteur $y \in Y$ est appelé ε-sous-gradient de f au point x^0 si :

$$\mathrm{f(x)} \geq f(x^0) + \langle y, x - x^0 \rangle - \varepsilon, \forall x \in X.$$

Le ε-sous-différentiel de f au point x^0, noté $\partial_\varepsilon f(x^0)$, est l'ensemble de tous les ε-sous-gradient de f au point x^0.

Définition 2.14. (Fonction s.c.i.) La fonction f est dite semi-continue inférieurement (s.c.i) au point $x \in X$ si

$$\lim_{y \rightarrow x} \inf f(y) = f(x).$$

On note $\Gamma_0(X)$ l'ensemble des fonction convexes s.c.i. est propres sur X.

Proposition 2.1. [1]

- $f \in \Gamma_0(X) \Longleftrightarrow f^* \in \Gamma_0(Y),\ et\ (f^*)^* = f.$

- *Si* $f \in \Gamma_0(X)$ *est différentiable en* x *si et seulement si* $\partial f(x) = \{\nabla f(x)\}$.

- *Si* $f \in \Gamma_0(X)$ *alors*

$$y \in \partial f(x) \Longleftrightarrow x \in \partial f^*(y).$$

- $y \in \partial f(x) \Longleftrightarrow f(x) + f^*(y) = \langle x, y\rangle.$

- $x^0 \in$ *arg min* $\{f(x) : x \in X\} \Longleftrightarrow 0 \in \partial f(x^0).$

Définition 2.15. (Fonction conjuguée) La fonction conjuguée de $f \in \Gamma_0(X)$, notée $f^* : Y \to \mathbb{R} \cup \{+\infty\}$, est définie par :

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

Définition 2.16. (Polyédre convexe) Une partie convexe C est dit convexe polyédrale si :

$$\mathrm{C}=\bigcap_{i=1}^{m}\{x \in \mathbb{R}^n / \langle a_i, x\rangle - b_i \leq 0, a_i \in \mathbb{R}^n, b_i \in \mathbb{R}\}$$

Définition 2.17. (Fonction polyédrale)
Une fonction $f : \mathbb{R}^n \to]-\infty, +\infty]$ est dite convexe polyédrale si son épigraphe est un polyédre de $\mathbb{R}^{n+1}$

Remarque 2.2. fonction polyédrale $\Rightarrow$ fonction convexe.

Remarque 2.3. Toute fonction polyédrale est convexe propre et s.c.i.

Proposition 2.2. [1] *Soit $f \in \Gamma_0(X)$. Alors f est une fonction convexe polyédrale si et seulement si dom(f) est un polyédre convexe et*

$$f(x)=\sup\{\langle a_i, x\rangle - b_i : i = 1, ..., l\} + \chi_{dom(f)}(x)$$

Proposition 2.3. [1] *Les fonctions convexes polyédrales possèdent les popriétés intéressantes suivantes :*

- *Si f_1 et f_2 sont convexes polyédrales, alors $f_1 + f_2, \max\{f_1, f_2\}$ sont convexes polyédrales.*

- *Si f est convexe polyédrale alors f^* l'est aussi et $dom(\partial f) = dom(f)$. De plus, si f est finie partout alors*

$$dom(f^*) = co\{a_i : i = 1, ..., l\}$$

$$f^*(y) = \min\{\sum_{i=1}^{l} \lambda_i a_i : y = \sum_{i=1}^{l} \lambda_i b_i, \sum_{i=1}^{l} \lambda_i = 1, \lambda_i \geq 0, i = 1, ..., l\}.$$

- *Si f est polyédrale alors $\partial f(x)$ est une partie convexe polyédrale non vide $\forall x \in dom(f)$.*

- *soient $f_1, f_2, ..., f_l$ des fonctions convexes polyédrales sur X telles que les ensembles convexes $dom(f_i), i = 1, ..., l$ aient un point commun. Alors*

$$\partial(f_1 + f_2 .... + f_l)(x) = \partial f_1(x) + \partial f_2(x) + ... + \partial f_l(x), \forall x.$$

2.1.3 Classe des fonctions DC

Dans cette section, nous allons introduire la définition des fonctions DC et des propriétés essentielles de l'ensemble des fonctions DC notamment sa stabilité par rapport aux opérations usuelles en optimisation.

Définition 2.18. (Fonction DC) Soit C un sous-ensemble convexe fermé non vide de X. Une fonction $f : C \rightarrow]-\infty, +\infty]$ est dite DC sur C si elle peut s'écrire sous la forme

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

où g et h sont deux fonctions convexes sur C. On dit que $g - h$ est une décomposition DC de f.

Remarque 2.4. Si $f = g - h$ est une fonction DC sur C alors pour toute fonction convexe finie ϕ sur C, $f = (g + \phi) - (h + \phi)$ est aussi une décomposition DC de f. Ainsi une fonction DC possède une infinité de décompositions DC.

Définition 2.19. (Ensemble des fonctions DC) Soit Conv(C) l'ensemble des fonctions convexes sur C. L'ensemble des fonctions DC sur C, noté DC(C), est l'espace vectoriel engendré par Conv(C) définit comme :

$$DC(C) = Conv(C_g) - Conv(C_h)$$

L'espace vectoriel DC(C) est une classe de fonctions assez large. Il contient toutes les fonctions convexes sur C, les fonctions concaves sur C, ainsi que beaucoup des fonctions ni convexes ni concaves sur C. Il contient la quasi-totalité des fonctions rencontrées dans les problèmes concrets en optimisation non convexe .

Les fonctions DC possèdent beaucoup de propriétés importantes qui ont été établies à partir des années 50 par Alexandroff(1949) [22] et hartman(1959) [23] etc. La classe des fonctions DC est stable par rapport aux opérations usuelles utilisées en optimisation.

Proposition 2.4. [1] *Si* $f = g - h,\ f_i = g_i - h_i,\ i = 1, \ldots.., m$ *sont des fonctions dans DC(C) alors*

- *Une combinaison linéaire finie des fonction DC est une fonction DC, c.à.d*

$$\sum_{i=1}^{m} \lambda_i f_i \in DC(C), \forall \lambda_i \in \mathbb{R}$$

- $\max\limits_{i=1,\ldots,m}(f_i) \in DC(C)$ *et* $\min\limits_{i=1,\ldots,m}(f_i) \in DC(C)$ *car*

$$\max_{i=1,\ldots,m} f_i = \max_{i=1,\ldots,m}[\ g_i\ + \sum_{j=1,j\neq i}^{m} h_j\ ] - \sum_{j=1}^{m} h_j.$$

$$\min_{i=1,\ldots,m} f_i = \sum_{j=1}^{m} g_i\ -\ \max_{i=1,\ldots,m}[\ h_i + \sum_{j=1,j\neq i}^{m} g_j\ ].$$

- $|f| \in DC(C),$ *car*

$$|f| = 2\max(g, h) - (g + h).$$

- $f^+, f^- \in DC(C)$ *où* $f^+(x) = \max(0, f(x))$ et $f^-(x) = \min(0, f(x))$ car

$$f^+ = \max(g, h) - h$$

$$f^- = g - \max(g, h)$$

- $\prod\limits_{i=1}^{m} f_i \in DC(C)$

Définition 2.20. (Fonction localement DC) Supposons que l'ensemble convexe C est ouvert. Une fonction $f : C \to \mathbb{R} \cup \{+\infty\}$ est dite localement DC sur C si pour tous les points $x^0 \in C$, il existe une boule de centre x^0 et de rayon ε, notée $B(x^0, \varepsilon)$, telle que f est une fonction DC sur $B(x^0, \varepsilon)$, $\forall \varepsilon \geq 0$.

Soit C un ouvert convexe de X. On note $C^{1.1}(C)$ le sous espace vectoriel de $C^1(C)$ des fonctions dans le gradient est localement lipschitzien sur C. $C^{1.1}(C)$ est contenu dans le cône convexe des fonctions $LC^2(C)$, dites fonctions de sous$-C^2$, qui sont localement enveloppe supérieur d'une famille de fonctions de classe $C^2(C)$. Le théorème suivant montre la richesse de la classe des fonctions DC :

Théorème 2.3. [1]

- *Une fonction sous $- C^2$ sur C est une fonction localement DC sur C*

- *Une fonction localement DC sur C est une fonction DC sur C . finalement on a :*

$$C^2(C) \cup C^{1.1}(C) \cup C^1(C) \cup Conv(C) \cup LC^2(C) \subset DC(C).$$

Si C est en plus compact, alors DC(C) est un sous-espace vectoriel dense dans l'ensemble des fonctions continues sur C muni de la norme de la convergence uniforme sur C.

2.1.4 Programmation DC

Définition 2.21. (Programme DC) On appelle programme DC tout problème d'optimisation de la forme

$$(\mathrm{P}_{dc}) \quad \inf\{f(x) = g(x) - h(x) : x \in \mathbb{R}^n\}$$

où g et h sont convexes. C'est un programme sans contraintes. En effet, beaucoup de problèmes d'optimisation non convexe sous contraintes (convexes et non convexes) peuvent être transformés en (P_{dc}) sous cetaines conditions techniques. Les problèmes suivants sont bien connus en optimisation non convexe.

(1) $\sup\{f(x) : x \in C\}, f$ et C sont convexes

(2) $inf\{f_0(x) = g_0(x) - h_0(x) : x \in C\},\ g_0,\ h_0$ et C sont convexes.

(3) $inf\{f_0(x) = g_0(x) - h_0(x) : x \in C,\ g_i(x) - h_i(x) \leq 0, i = 1...m\}$ $g_i,\ h_i,\ i = 0...m$ et C sont convexes.

La raison est bien simple : La majorité des problèmes d'optimisation de la vie réelle sont de nature non convexes. De plus dans les modèles mathématiques industriels les approches convexes montrent leurs limites et sont de plus en plus remplacées par des approches non convexes qui sont plus réalistes.

Le programme (1) est équivalent à (P_{dc}) dans la mésure *où*, d'une part, on peut transformer le programme (1) en (P_{dc}) avec $g = X_C$ (la fonction indicatrice de C définie par $X_C(x) = 0$ si $x \in C; +\infty$ dans le cas contraire) et $h = -f$; d'autre part, (P_{dc}) peut se réécrire sous la forme du programme (1) avec une variable additionnelle t telle que

$$\sup\{h(x) - t : g(x) - t \leq 0\}.$$

Selon la même technique, le programme (2) (un programme DC sous contraintes convexe) et aussi équivalent à (P_{dc}) avec $g = g_0 + X_C$ et $h = h_0$

Dans le programme (3), les contraintes $g_i(x) - h_i(x) \leq 0,\ i = 1, ..., m$ sont des contraintes non convexes appelées contraintes DC, qui sont équivalentes à une seule contrainte DC $\hat{g} - \hat{h} \leq 0$ du fait que $\{\ g_i(x) - h_i(x) \leq 0,\ i = 1...m\} \equiv \{\max_{i=1,....,m}(g_i(x) - h_i(x)) \leq 0\}$ (voir la proposition (2.4)). Le programme suivant est donc un programme équivalent au programme (3)

$$\inf\{f_0(x) = g_0(x) - h_0(x) : x \in C,\ \hat{g}(x) - \hat{h}(x) \leq 0\}.$$

Gr$\hat{a}$ce au théorème de pénalité exacte relative $\grave{a}$ la contrainte DC $\hat{g}(x) - \hat{h}(x) \leq 0$, on peut ramener ce dernier programme à la forme (P_{dc})

Ainsi, la résolution de (P_{dc}) entra$\hat{\imath}$ne celle des autres. Le programme (P_{dc}) est par conséquent un problème essentiel en programmation DC. La structure spéciale de (P_{dc}) a permis d'importants développements, tant sur le plan théorique que sur le plan algorithmique.

2.1.5 Dualité en programmation DC

Dans ce paragraphe, nous allons introduire brièvement le concept de dualité en programmation DC. Bien qu'établie depuis longtemps en analyse convexe, la notion de dualité dans le cas non convexe a été proposée récemment avec les travaux dans les cas Quasi-convexe et Anti-convexe et introduite par Toland dans le cadre de la programmation DC comme une généralisation des travaux sur la maximisation convexe de Pham Dinh Tao [24]. Cette théorie de dualité DC est très riche et on pourra se référer aux travaux de Let Thi Hoai An [27] pour plus de détails.

Considérons le programme DC

$$(\mathrm{P}_{dc}) \qquad \alpha = \inf\{f(x) = g(x) - h(x) : x \in X\}$$

où g, $h \in \Gamma_0(X)$. On adopte la convention $(+\infty)-(-\infty) = (+\infty)$ comme en optimisation convexe. puisque $h \in \Gamma_0(X)$, on a donc $h^{**} = h$ et

$$\mathrm{h(x)}=(\mathrm{h}^*)^*(x) = \sup\{\langle x, y\rangle - h^*(y) : y \in Y\}.$$

Avec cette relation, on trouve que

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

Avec la convention $(+\infty) - (+\infty) = (+\infty)$, nous obtenons le problème dual de (P_{dc}), noté (D_{dc}) :

$$(\mathrm{D}_{dc}) \quad \alpha = \inf\{h^*(y) - g^*(y) : y \in Y\}.$$

Comme g^* et h^* appartiennent à $\Gamma_0(Y)$, le problème (D_{dc}) est un programme DC. De plus, (P_{dc}) et (D_{dc}) ont la même valeur optimale α. En autre, il existe une parfaite symétrie entre (P_{dc}) et (D_{dc}), le dual de (D_{dc}) est exactement (P_{dc}). Les résultats suivants donnent quelques propriétés concernant les solutions des problèmes (P_{dc}) et (D_{dc}).

Théorème 2.4. [2]

- *x^* est une solution optimale globale du problème (P_{dc}) si et seulement si*

$$\alpha = (g-h)(x^*) \leq (h^* - g^*)(y), \forall y \in Y$$

- *y^* est une solution optimale globale du problème (D_{dc}) si et seulement si*

$$\alpha = (h^* - g^*)(y^*) \leq (g-h)(x), \forall x \in X.$$

Le théorème ci-dessous montre que la résolution du problème primal (P_{dc}) implique la résolution du problème dual (D_{dc}). On observe la parfaite symétrie entre les problèmes primal et dual.

Théorème 2.5. [2] *Soient $g,\ h \in \Gamma_0(X)$. Alors*

1.

$$\inf\{g(x) - h(x) : x \in dom(g)\} = \inf\{h^*(y) - g^*(y) : y \in dom(h^*)\}\ .$$

2. Si y^0 est un minimum de $(h^* - g^*)$ alors chaque $x^0 \in \partial g^*(y^0)$ est un minimum de $(g-h)$.

3. Si x^0 est un minimum de $(g-h)$ alors chaque $y^0 \in \partial h(x^0)$ est un minimum de $(h^* - g^*)$.

Ce dernier théorème montre que la résolution de l'un des deux problèmes (P_{dc}) et (D_{dc}) implique celle de l'autre.

Théorème 2.6. *soient $g,\ h \in \Gamma_0(X)$ et h(x)=ng(mx), alors les fonctions conjuguées h^* et g^* satisfont*

$$h^*(y) = ng^*(\frac{y}{mn}),$$

où $m \neq 0$ et $n > 0$.

Preuve du théorème. *par définition de la fonction conjuguée, on a*

$$\text{h}^*(y) = sup\{\langle x, y\rangle - h(x) : x \in X\} = sup\{\langle x, y\rangle - ng(mx) : x \in X\}.$$

Sachant que $m \neq 0, n > 0$, alors

$\sup\{\langle x, y\rangle - ng(mx) : x \in X\} = nsup\{\langle mx, \frac{y}{mn}\rangle - g(mx) : x \in X\}$
$= ng^*(\frac{y}{mn})$.

D'où $\text{h}^*(y) = ng^*(\frac{y}{mn})$.

Lemme 2.1. soient $g,\ h \in \Gamma_0(X)$ et $h(x) = ng(\frac{x}{n}),\ n > 1$. Le problème primal est défini par :

$$\text{(P)} \qquad \inf\{g(x) - h(x) : x \in X\}$$

est son problème dual associé :

$$\text{(D)} \qquad \inf\{h^*(y) - g^*(y) = (n-1)g^*(y) : y \in Y\}$$

est un programme convexe.

Exemple 2.1. soient $x \in \mathbb{R},\ n > 1,\ g(x) = e^x$ et $h(x) = ne^{\frac{x}{n}}$. Le problème primal

$$(\mathrm{P}) \qquad \inf\{g(x) - h(x) = e^x - ne^{\frac{x}{n}} : x \in \mathbb{R}\}.$$

est un programme DC non convexe. Son problème dual (D) :

$$\begin{aligned} &\inf\{h^*(y) - g^*(y) = (n-1)g^*(y) : y \in \mathbb{R}\} \\ =\ &\inf\{sup\{\langle x, y\rangle - h(x)\} - sup\{\langle x, y\rangle - g(x)\} : x, y \in \mathbb{R}\} \\ =\ &(n-1)sup\{\langle x, y\rangle - g(x) : x, y \in \mathbb{R}\} \\ =\ &(n-1)sup\{\langle x, y\rangle - e^x : x, y \in \mathbb{R}\} \end{aligned}$$

est un programme convexe.

2.1.6 Conditions d'optimalité en programmation DC

Soient P et D les ensembles de solutions des problèmes (P_{dc}) et (D_{dc}). Un point est dit point critique du problème (P_{dc}) si $\partial g(x^*) \cap \partial h(x^*) \neq \varnothing$ (un point de KKT généralisé).

Il est bien connu en analyse convexe qu'une condition nécessaire et suffisante pour que x minimise une fonction convexe f est

$$0 \in \partial f(x).$$

Cette condition bien que nécessaire et suffisante dans le cas convexe ne l'est plus dans le cas non convexe. En programmation DC, une condition nécessaire et suffisante d'optimalité globale utilisant le $\varepsilon - sous$ différentiel est donnée par le théorème suivant.

Théorème 2.7. [6] ***(Condition d'optimalité globale)*** *Soient $g, h \in \Gamma_0(X)$ et $f = g - h$. Alors x^* est un minimum global de $g - h$ sur X si et seulement si*

$$\partial_\varepsilon h(x^*) \subset \partial_\varepsilon g(x^*), \forall \varepsilon > 0.$$

Cette caractérisation est une traduction directe du théorème (2.5) et par conséquent invérifiable en pratique.

Remarque 2.5.

- Si $f \in \Gamma_0(X)$, on peut écrire $g = f$ et $h = 0$. Dans ce cas, l'optimalité globale de la programmation DC est identique à celle de la programmation convexe : $0 \in \partial f(x^*)$, du fait que

$$\partial_\varepsilon h(x^*) = \partial h(x^*) = \{0\}, \forall \varepsilon > 0.$$

- D'une manière plus générale, considérons les décompositions DC de $f \in \Gamma_0(X)$ de la forme $f = g - h$ avec $g = f + h$ et $h \in \Gamma_0(X)$ finie partout sur x. Dans ce cas, la condition d'optimalité $0 \in \partial f(x^*)$ équivaut à $\partial h(x^*) \subset \partial g(x^*)$.

Définition 2.22. Soient g et h deux fonctions de $\Gamma_0(X)$. Un point $x^* \in dom(g) \cap dom(h)$ est un minimum local de $g - h$ sur X si et seulement si

$$g(x) - h(x) \geq g(x^*) - h(x^*), \forall x \in V_{x^*}.$$

(V_{x^*} désigne un voisinage de x^*).

Théorème 2.8. [6] ***(Condition nécessaire d'Optimalité locale)***

si x^* est un minimum local de $g - h$, alors

$$\partial h(x^*) \subset \partial g(x^*)$$

Cette derniére condition étant suffisante dès que h est polyédrale.

Corollaire 2.1.
Si $h \in \Gamma_0$ (x) est polyédrale alors une condition nécessaire et suffisante pour que $x^* \in X$ soit un minimum local de $g - h$ est

$$\partial h(x^*) \subset int(\partial g(x^*)).$$

Théorème 2.9. [6] ***(Condition suffisante D'Optimalité local)*** *:*
si x^* admet un voisinage V tel que

$$\partial h(x) \cap \partial g(x) \neq \varnothing, \forall x \in V \cap dom(g),$$

alors x^* est un minimum local de $f = g - h$.

Corollaire 2.2.
si $x^* \in int(dom(h))$ vérifie

$$\partial h(x^*) \subset int(\partial g(x^*));$$

alors x^* est un minimum local strict de $g - h$.

Théorème 2.10. [6]

1. *$\partial h(x) \subset \partial g(x), \forall x \in P$ et $\partial g^*(y) \subset \partial h^*(y), \forall y \in D$.*

2. *Transport de minima globaux :*

$$\bigcup\{\partial h(x)/x \in P\} \subseteq D \subset dom(h^*)$$

La première inclusion se transforme en égalité si g^* est sous-différentiable sur D, d'autre façon si $D \subset ir(dom(g^*))$ ou si g^* est sous-différentiable sur $dom(h^*)$ et on écrit : $D \subset (dom(\partial g^*) \cap dom(\partial h^*))$.

$$\bigcup\{\partial g^*(y)/y \in D\} \subseteq P \subset dom(g)$$

La première inclusion se transforme en égalité si h est sous-différentiable sur P, d'autre façon si $P \subset ir(dom(h))$ ou si h est sous-différentiable sur $dom(g)$ et on écrit : $P \subset (dom(\partial g \cap dom \partial h))$.

3. *Transport de minima locaux : Soit $x^* \in dom(\partial h)$ un minimum local de $f = g - h$. Et soit $y^* \in \partial h(x^*)$ et V_{x^*} un voisinage de x^* tel que $g(x) - h(x) \geq g(x^*) - h(x^*), \forall x \in V_{x^*} \cap dom(g)$ Si*

$$x^* \in int(dom(g^*)) \text{ et } \partial g^*(y^*) \subset V_{x^*},$$

alors y^* est un minimum local de $h^* - g^*$.

2.2 DCA (DC Algorithm)

DCA (DC Algorithm) est une nouvelle méthode itérative d'optimisation locale basé sur l'optimalité locale et la dualité en programmation DC (non différentiable). Cette approche est complètement différente des méthodes classiques de sous-gradient en optimisation convexe. Dans les DCA, la construction algorithmique cherche à exploiter la structure DC du problème. Elle nécessite, en premier lieu, de disposer d'une représentation DC de la fonction à minimiser, i.e. $f = g - h$ (g, h convexe), car toutes les opérations s'effectueront uniquement sur les composantes convexes. Ainsi, la séquence des directions de descente est obtenue en calculant une suite de sous-gradient non directement à partir de la fonction f, mais des composantes convexes des problèmes primal et dual.

DCA a été introduit par Pham Dinh Tao [24, 25, 26] a l'état préliminaire en 1985 et puis intensivement dévloppé par Pham Dinh Tao, le Thi Hoai An et al. Depuis 1994. Cette approche permet de construire deux suites $\{x^k\}$ et $\{y^k\}$ (candidats supposés pour des solutions optimales des programme DC primal et dual respectivement) telles que leurs limites (x^* et y^*) soient des points KKT généralisés de ces programmes, et telles que les suites $\{(g-h)(x^k)\}$ *et* $\{(g^*-h^*)(y^k)\}$ soient décroissantes et tendent vers la même limite $\beta = \{(g-h)(x^*)\} = \{(g^*-h^*)(y^*)\}$

Proposition 2.5. [1]

1. Les suites $\{g(x^k) - h(x^k)\}$ et $\{h^*(x^k) - g^*(x^k)\}$ décroissent et tendent vers la même limite β qui est supérieure ou égale à la valeur optimale globale α.

2. Si $(g-h)(x^{k+1}) = (g-h)(x^k)$ 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 $h^* - g^*$).

3. Si la valeur optimale du problème (P_{dc}) est finie et si les suites $\{x^k\}$ et $\{y^k\}$ sont bornées, alors toute valeur d'adhérence x^* de la suite $\{x^k\}$ (resp. y^* de la suite $\{y^k\}$) est un point critique de $g-h$ *(resp.* h^*-g^**)*

pour construire les deux suites $\{x^k\}$ et $\{y^k\}$, on définit deux programme convexe (P_k) et (D_k) comme :

$$(D_k) \qquad y^k \in \arg\min\{h^*(y) - [g^*(y^{k-1}) + \langle y - y^{k-1},\ x^k\rangle] : y \in Y\}$$

$$(P_k) \qquad x^{k+1} \in \arg\min\{g(x) - [h(x^k) + \langle x - x^k,\ y^k\rangle] : x \in X\}$$

On a ensuite le schéma simple suivant pour décrire DCA : Alors le point x^{k+1} (resp. y^k) est une solution optimale du programme (P_k) (resp. (D_k)). On peut facilement comprendre que (P_k) (resp. (D_k)) est obtenu en remplaçant h (resp. g^*) de (P_{dc}) (resp. (D_{dc})) par sa minorante affine $h_k(x) = h(x^*) + \langle x - x^k,\ y^k\rangle$ au voisinage de x^k avec $y^k \in \partial h(x^k)$ (resp. $g_k^*(y) = g_k^*(y^{k-1}) + \langle y - y^{k-1},\ x^k\rangle$ au voisinage de y^{k-1} avec $x^k \in \partial g^*(y^{k-1})$). Par conséquent, le problème (P_k) (resp. (D_k)) est un problème de la borne supérieure du programme DC (P_{dc}) (resp. (D_{dc})).

On a ensuite le schéma simple suivant pour décrire DCA :

$$\begin{array}{ccc} x^k & \longrightarrow & y^k \in \partial h(x^k) \\ & \swarrow & \\ x^{k+1} \in \partial g^*(y^k) & \longrightarrow & y^{k+1} \in \partial h(x^{k+1}). \end{array}$$

DCA est donc un schéma de point fixe $x^{k+1} \in (\partial g^* \circ \partial h)(x^k)$.

Grâce à la proposition (2.5), DCA s'arrête si au moins l'une des suites $\{(g-h)(x^k)\}, \{(h^*-g^*)(y^k)\}, \{x^k\}, \{y^k\}$ converge. En pratique, nous utilisons souvent les conditions d'arrêt suivantes :

- $|(g-h)(x^{k+1}) - (g-h)(x^k)| \leq \varepsilon.$
- $\|x^{k+1} - x^k\| \leq \varepsilon$

pour obtenir une solution $\varepsilon-$optimale. On peut dès lors décrire DCA

Algorithme DCA
Étape ***0 :*** x^0 donné*, k=0*
Étape ***1 :*** on calcule $y^k \in \partial h(x^k)$
Étape ***2 :*** on détermine $x^{k+1} \in \partial g^*(y^k)$
Étape ***3 :*** Si les conditions d'arrêt sont vérifiées alors on termine DCA ; sinon k=k+1 et on répète *Étape* ***1 :***

2.2.1 Existence et bornitude des suites générées par DCA

Pour un programme DC, si DCA peut effectivement construire les deux suite $\{x^k\}$ et $\{y^k\}$ à partir d'un point initial arbitraire $x^0 \in X$, alors les deux suites sont dites bien définies.

Lemme 2.2. [1] **(Existence des suites)** Les relations suivantes sont équivalentes :

1. les suites $(x^k)_k$ et $(y^k)_k$ sont bien définies.
2. $\mathrm{dom}(\partial g) \subset dom(\partial h)$ et $dom(\partial h^*) \subset dom(\partial g^*)$

La proposition suivante établit les conditions de bornitude pour les suites générées par DCA.

Lemme 2.3. [1] **(Bornitude des suites)** si $g - h$ est coercive[2], alors

1. la suite $\{x^k\}$ est bornée.
2. si $\{x^k\} \subset int(dom(h))$ alors la suite $\{y^k\}$ est aussi bornée

si $g^* - h^*$ est coercive, alors

1. la suite $\{y^k\}$ est bornée.
2. si $\{y^k\} \subset int(dom(g^*))$ alors la suite $\{x^k\}$ est aussi bornée

La convergence de DCA est donnée par le théorème suivant :

Théorème 2.11. [1] ***(Convergence de DCA)*** *: On suppose que les suites $\{x^k\}$ et $\{y^k\}$ sont bien définies*

1. *Les suites $\{g(x^k) - h(x^k)\}$ et $\{h^*(y^k) - g^*(y^k)\}$ sont décroissantes et*
   - *$g(x^{k+1}) - h(x^{k+1}) = g(x^k) - h(x^k)$ si et seulement si $y^k \in \partial g(x^k) \cap \partial h(x^k)$, $y^k \in \partial g(x^{k+1}) \cap \partial h(x^{k+1})$ et $[\rho(g) + \rho(h)]\ \|x^{k+1} - x^k\| = 0$. De plus, si g et h sont strictement convexes sur X, alors $x^k = x^{k+1}$. Dans ce cas, DCA termine en un nombre fini d'itérations. x^k et x^{k+1} sont des points critiques de la fonction $g - h$.*
   ($\rho(g) = \rho(g, X)$ est le module de forte convexité de g sur X).
   - *$h^*(y^{k+1}) - g^*(y^{k+1}) = h^*(y^k) - g^*(y^k)$ si et seulement si $x^{k+1} \in \partial g^*(y^k) \cap \partial h^*(y^k), x^{k+1} \in \partial g^*(y^{k+1}) \cap \partial h^*(y^{k+1})$ et $[\rho(g^*, D) + \rho(h^*, D)]\ \|y^{k+1} - y^k\| = 0$. De plus, si g^* et h^* sont strictement convexes sur Y, alors $y^k = y^{k+1}$. Dans ce cas, DCA termine en un nombre fini d'itérations. y^k et y^{k+1} sont des points critiques de la fonction $h^* - g^*$.*
2. *Si $\rho(g,\ C) + \rho(h,\ C) > 0$ (resp. $\rho(g^*,\ D) + \rho(h^*,\ D) > 0$), alors la suite*
   $\{\|x^{k+1} - x^k\|^2\}$ (resp. $\{\|y^{k+1} - y^k\|^2\}$) converge.
3. *Si la valeur optimale du problème P_{dc} est finie et si les suites $\{x^k\}$ et $\{y^k\}$ sont bornées, alors toute valeur d'adhérence x^* (resp. y^*) de $\{x^k\}$*
   (resp. $\{y^k\}$) est un point critique de $g - h$ (resp. $h^* - g^*$)

2.2.2 Programmation DC polyédrale

Dans cette partie, on présente une classe de problèmes DC qui se rencontre fréquemment dans la pratique et possède des propriétés importantes, tant sur le plan théorique que sur le plan algorithmique. Cette classe de problèmes DC s'appelle DC polyédrale.

Définition 2.23. (programmation DC polyédrale) Un programme DC est dit polyédrale si l'une des composantes convexes (g et h) de la fonction $f = g - h$ est une fonction convexe polyédrale.

On peut montrer que, pour résoudre un programme DC polyédrale, DCA a une convergence finie.

Notons que h_k est la minorante affine de la fonction convexe h au voisinage de x^k. On construit h^k, fonction convexe polyédrale engendrée par $\{h_k\}$.

$$\mathrm{h}_k(x) = h(x^k) + \langle x - x^k, y^k\rangle = \langle x, y^k\rangle - h^*(y^k), \forall x \in X$$

$h^k(x) = sup\{h_i(x) : i = 0, ..., k\} = sup\{\langle x, y^i\rangle - h^*(y^i) : i = 0, ..., k\}$,
$\forall x \in X$,
où les suites $\{x^k\}$ et $\{y^k\}$ sont obtenues via la procédure DCA.
on définit de manière analogue la fonction duale g_k^* (resp. $(g^*)^k$)

$$g_k^*(y) = g^*(y^{k-1}) + \langle y - y^{k-1}, x^k\rangle = \langle y, x^k\rangle - g(x^k), \forall y \in Y.$$

$$(g^*)^k(y) = sup\{g_i^*(y) : i = 0, ..., k\} = sup\{\langle y, x^i\rangle - g(x^i) : \text{i=0,....,k}\}, \forall y \in Y$$

Sachant qu'une fonction convexe propre s.c.i. est caractérisée comme le suprémum de ses minorantes affines, il s'avère donc plus judicieux d'utiliser h^k (resp. $(g^*)^k$) comme sous estimée de la fonction convexe h (resp. g^*) sur X (resp. Y), plutôt que la minorante affine h_k (resp. $(g^*)_k$). On peut ensuite obtenir les programmes suivants :

$$(\mathrm{P}^k) \qquad \inf\{g(x) - h^k(x) : x \in X\}$$

$$(\mathrm{D}^k) \qquad \inf\{h^*(y) - (g^*)^k(y) : y \in Y\}$$

Les deux problèmes (P^k) et (D^k) sont des problèmes DC non convexes. On peut voir la différence avec les sous-problèmes (P^k) et (D^k) qui sont convexes.

Théorème 2.12.

(1) *$g(x^{k+1}) - h(x^{k+1}) = h^*(y^k) - g^*(y^k)$ si et seulement si $h^k(x^{k+1}) = h(x^{k+1})$.*

(2) *$h^*(y^k) - g^*(y^k) = g(x^k) - h(x^k)$ si et seulement si $g^*(y^k) = (g^*)^k(y^k)$.*

(3) *$g(x^{k+1}) - h(x^{k+1}) = g(x^k) - h(x^k)$ si et seulement si $h^k(x^{k+1}) = h(x^{k+1})$ et $g^*(y^k) = (g^*)^k(y^k)$*

par conséquent, si $g(x^{k+1}) - h(x^{k+1}) = g(x^k) - h(x^k) = h^*(y^k) - g^*(y^k)$

alors les propositions suivantes sont vraies :

- *x^{k+1} (resp. y^k) est une solution optimale du problème (P^k) (resp (D^k)).*
- *si h et h^k coincident en une solution de (P_{dc}) ou/et g^* et $(g^*)^k$ coincident en une solution de (D_{dc}), alors x^{k+1} (resp y^k) est également une solution de (P_{dc}) (resp. (D_{dc})).*

Supposons que la valeur optimale du problème (P_{dc}) est finie et que la suite $\{x_k\}_k$ générée par DCA est bornée, alors pour toute valeur d'adhérence x^∞ on a

$$\mathrm{g}(x^\infty) - h_\infty(x^\infty) = inf\{g(x^{i+1}) - h_i(x^{i+1}),\ i = 0, ..., \infty\}$$

où h_∞ est la minorante affine de h au voisinage du point x^∞ définie par :

$$\mathrm{h}_\infty = h(x^\infty) + \langle x - x^\infty,\ y^\infty\rangle = \langle x, y^\infty\rangle - h^*(y^\infty), \forall x \in X,$$

avec $y^\infty \in \partial h(x^\infty)$ une valeur d'adhérence de $\{y^k\}_k$. Le point x^∞est donc une solution du programme DC

$$(\mathrm{P}^\infty) \qquad inf\{g(x) - h^\infty(x) : x \in X\}$$

De manière analogue, le point y^∞ est une solution du programme DC

$$(\mathrm{D}^\infty) \qquad inf\{h^*(y) - (g^*)^\infty(y) : y \in Y\}$$

on a le théorème suivant :

Théorème 2.13. *Si la valeur optimale du problème (P_{dc}) (resp. (D_{dc})). est finie et la suite $\{x^k\}_k$(resp. $\{y^k\}_k$) est bornée, alors toute valeur d'adhérence x^∞ (resp. y^∞) est une solution de problème (P^∞) (resp. (D^∞)). De plus, les valeur optimales sont égales, c.à.d*

$$g(x^\infty) - h^\infty(x^\infty) = h^*(y^\infty) - (g^*)^\infty(y^\infty)$$

Si l'un des condition suivantes est vérifiée :

- *Les fonctions h et h^∞ coincident en une solution optimale de (P_{dc})*
- *Les fonctions g^* et $(g^*)^\infty$coincident en une solution optimale de (D_{dc})*

alors x^∞ et y^∞ sont également des solutions optimales de (P_{dc}) et (D_{dc}) respectivement.

2.2.3 Interprétation Géométrique de DCA

Notons d'abord que DCA ne travaille qu'avec les composantes DC g et h. A la $k^{ième}$ itérations de DCA, on remplace la composante h par sa minorante affine

$$h_k(x) = h(x^k) + \langle x - x^k,\ y^k \rangle$$

au voisinage de x^k. Comme h est une fonction convexe, on a alors

$$h(x) \geq h_k(x); \forall x \in X.$$

donc on aura par suite :

$$f^k(x) := g(x) - [h_k(x) = h(x^k) + \langle x - x^k,\ y^k \rangle] \geq g(x) - h(x);\ \forall x \in X.$$

Donc f^k est une fonction majorante de la fonction f et de plus lorsque g et h sont différentiables sur X ; alors on constate que :

$$f^k(x^k) = f(x^k)\ et\ \nabla f^k(x^k) = \nabla f(x^k)$$

Proposition 2.6. *soit g, h deux fonctions convexes et différentiables sur X. Notons*

$$f^k(x) := g(x) - [h_k(x) = h(x^k) + \langle x - x^k,\ y^k \rangle].$$

On a

1. $f^k(x) \geq f(x),\ \forall x \in X.$
2. $f^k(x^k) = f(x^k)$
3. $\nabla f^k(x^k) = \nabla f(x^k).$

A l'aide de cette proposition, on peut obtenir une simple interprétation géométrique de DCA. En effet, la surface de f^k peut être imaginée comme un "bol" se plaçant au-dessus de la surface de f ; de plus, les deux surfaces se touchent au point $(x^k,\ f(x^k))$ (voire la figure.2.2)

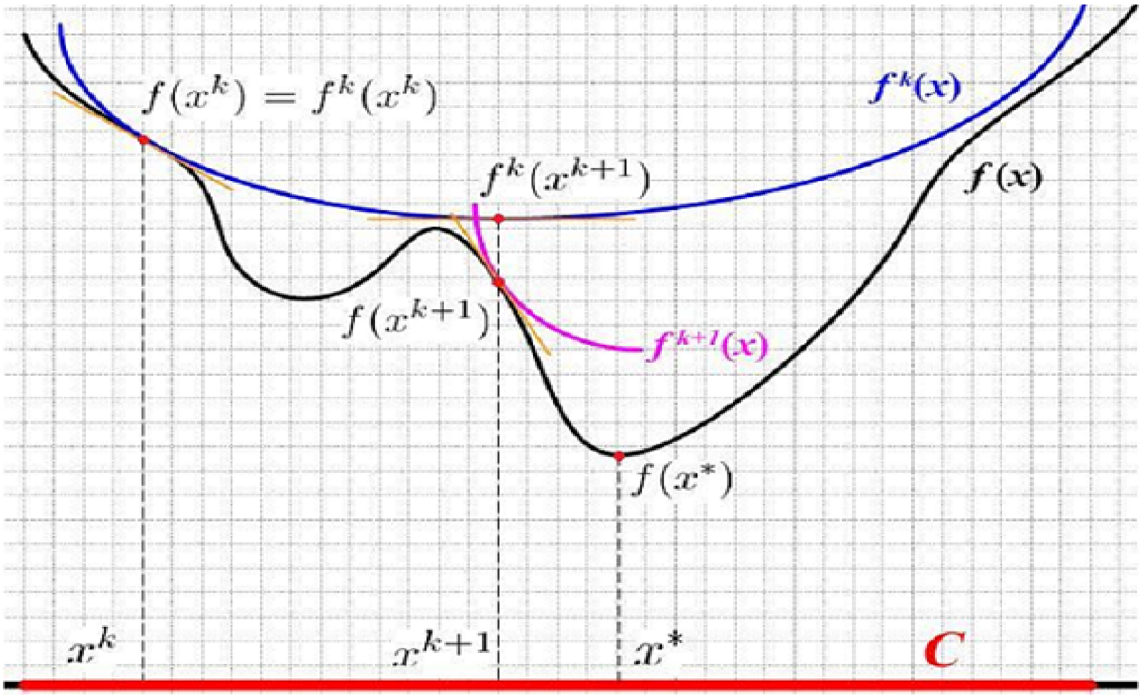

Fig.2.2-Interprétation Géométrique de DCA

On peut voir sur la figure 2.2 que $f^k(x) \geq f(x)$, $\forall\ x \in \mathcal{C}$ et $f^k(x^k) = f(x^k)$. DCA construit d'abord un suprémum $f(x^k)$ de la fonction f(x) au point x^k, et ensuite minimise cette fonction $f^k(x)$ sur $\mathcal{C}$ afin d'obtenir le point x^{k+1}, et ainsi de suite pour obtenir le point x^{k+2}.....On peut également retrouver une explication explicite de beaucoup de propriétés importantes de DCA sur la figure (e.g. décroissance, bornitude, convergence, etc.).

Par exemple, théoriquement, DCA construit une suite $\{x^k\}$ telle que la suite $\{f(x^k)\}$ est décroissante. Ceci peut $\hat{e}$tre facilement vérifié sur cette figure car x^{k+1} est un minimum de f^k (donc $f^k(x^k) \geq f^k(x^{k+1})$ et $f(x^{k+1}) \leq f^k(x^{k+1})$, ainsi $f^k(x^k) = f(x^k)$) . Finalement, on a $f(x^k) \geq f^k(x^{k+1}) \geq f(x^{k+1})$. Ceci montre que la suite $\{f(x^k)\}$ est décroissante.

Pour la bornitude et la convergence de DCA, sachant que $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ est bornée inférieurement sur $\mathbb{R}^n$, la suite $\{f(x^k)\}$ est aussi bornée inférieurement. On sait qu'une suite $\{f(x^k)\}$ décroissante et bornée inférieurement est convergente.

Lorsque DCA converge vers un point x^*, ce point doit être un point KKT généralisé. Cela est aussi aisé à comprendre, à l'aide de la figure. Si on lance DCA à partir d'un point KKT généralisé de f (e.g. x^* sur la figure), alors DCA s'arrête immédiatement à ce point car il est aussi un minimum de f^* (la fonction majorante convexe définie au point x^* par $f^* = g(x) - [h(x^*) + \langle x - x^*,\ y^*\rangle], y^* \in \partial h(x^*)$).

Gr$\hat{a}$ce à la figure, on peut constater que DCA a la faculté de sauter certains voisinages de minima locaux. Par exemple, DCA saute un minimum local entre x^k et x^{k+1} et parvient à un voisinage de la solution globale x^k. Ce phénomène est très intéréssant et important pour l'optimisation globale. Bien que l'on ne puisse pas toujours garantir que ce phénomène se produit, on peut comprendre que la performance de DCA est susceptible de dépendre de la decomposition DC et de la position du point initial. Par conséquent, pour appliquer DCA, il faut toujours penser aux deux questions suivantes :

1. comment choisir un point initial x^0 ?

2. comment choisir les composantes *DC*

Actuellement, sur le plan théorique, il n'y a pas d'approche déterministe pour répondre à ces deux questions. Elles restent encore ouvertes. En pratique, on choisit des décompositions DC adaptées à la structure spécifique du problème DC donné. On préfère souvent une décomposition de la structure plus simple et moins coûteuse pour la construction, telle que les suites x^k et y^k soient facilement calculables. On cherche habituellement à ce que :

• La fonction h permette de calculer facilement ∂h.

• Le programme convexe $min\{g(x) - \langle x,\ y^k\rangle : x \in X\}$ soit facile à résoudre, cela est notamment important pour les problèmes de grande dimensions.

• Le point initial x^0 soit aussi proche que possible d'une solution optimale globale.

2.2.4 DCA et l'algorithme de Gradient Projeté

Considérons le problème d'optimisation suivant :

$$(P) \qquad \alpha = inf\{f(x) : x \in c\}$$

où f :$\mathbb{R}^n \to \mathbb{R}$ est une fonction de classe $C^2(\mathbb{R}^n)$, et $\mathcal{C}$ un convexe fermé de $\mathbb{R}^n$. Soit une fonction quadratique convexe $g_\lambda(x) = \frac{\lambda}{2}\|x\|^2$ avec $\lambda > 0$. On peut choisir une décomposition DC $f(x) = g_\lambda(x) - [g_\lambda(x) - f(x)]$, ou $g_\lambda(x) - f(x)$ et $g_\lambda(x)$ sont deux fonctions convexe sur C il est facile de vérifier qu'un tel nombre λ existe si et seulement si $\rho(\nabla^2 f(x)) < +\infty, \forall x \in C$, ou, $\rho(\nabla^2 f(x))$ est le rayon spectrale de la matrice hessienne $\nabla^2 f(x)$. On peut donc choisir

$$\lambda = min\{\lambda : \lambda > 0, \lambda \geq \rho(\nabla^2 f(x)), x \in C\}$$

Dans ce cas, on dit que g_λ est plus convexe que f sur C.

Par conséquent, le problème (p) est équivalent au problème DC

$$(\bar{P}) \qquad \alpha = inf\{f(x) = (g + Xc)(x) - h(x) : x \in \mathbb{R}^n\}$$

où $g(x) = \frac{\lambda}{2}\|x\|^2$, Xc est la fonction indicatrice de C, $h(x) = g(x) - f(x)$. $(\bar{P})$ est un programme DC car $(g + Xc)$ et h sont deux fonctions convexe sur $\mathbb{R}^n$. Sachant que g et f sont deux fonctions de classe C^2 alors

$$\partial h(x) = \nabla h(x) = \nabla g(x) - \nabla f(x) = \lambda x - \nabla f(x).$$

DCA pour résoudre ce problème particulier $(\bar{P})$, peut être décrit par le schéma de point fixe suivant :

$$x^{k+1} \in argmin\{(g + Xc)(x) - \langle x, \nabla h(x^k)\rangle : x \in \mathbb{R}^n\}.$$

ceci équivaut à

$$x^{k+1} = P_c(x^k - \tfrac{\nabla f(x^k)}{\lambda}).$$

où P_{dc} désigne l'opérateur de projection orthogonale sur C. On reconnaît l'algorithme de gradient projeté avec le pas $\frac{1}{\lambda}$.

Chapitre 3

La méthode MDCA

3.1 Introduction

La fonction fminbnd de MATLAB est une méthode standard de résolution de minimisation d'une fonction réelle définie sur un intervalle fermé borné [a, b]$\subset \mathbb{R}$.

Elle nous fournit donc que des minima locaux, non nécessairement globaux si la fonction n'est pas unimodale.

Dans cet partie, nous proposons une méthode alternative basée sur la décomposition de la fonction à minimiser on une différence de fonction convexes (DC) et l'application de l'algorithme DCA.

Lalgorithme DCA fournira généralement lui aussi un minimum local non nécessairement global (ou même un point critique).

3.2 Position du Problème

Soit le programme DC suivant :

$$(P_{dc}) \Longleftrightarrow min\{f(x) = g(x) - h(x) : x \in [a, b]\}$$

$f : R^n \to \mathbb{R}$ non convexe
$g : \mathbb{R}^n \to \mathbb{R}$ convexe
$h : \mathbb{R}^n \to \mathbb{R}$ convexe

On cherche à résoudre le programme P_{dc} en appliquant le DCA sur le minimum de la moyenne des deux approximation de f à partir de a et b (MDCA) .

3.2.1 Le principe de la Méthode (MDC)

Le DCA est trés sensible au choix du point de départ.
Dans le cas d'une minimisation d'une fonction réelle définie sur [a, b], le minimum trouvé en partant de a sera en général différent de celui trouvé en partant de b.

Nous proposons de ne pas choisir de point de départ. A la place nous cherchons à minimiser la moyenne de deux approximation de f à partir de a et de b (MDC) soit :

$$\min \tfrac{1}{2}(f_k(x,a) + f_k(x,b))$$

avec :

$$f_k(x,a) = g(x) - h'(a)(x-a) - h(a)$$
$$f_k(x,b) = g(x) - h'(b)(x-b) - h(b)$$

Cette stratégie a l'avantage de fournir en général un minimum qui sera situé dans la zone d'attraction du minimum global cherché comme le montre bien l'exemple ci dessous :

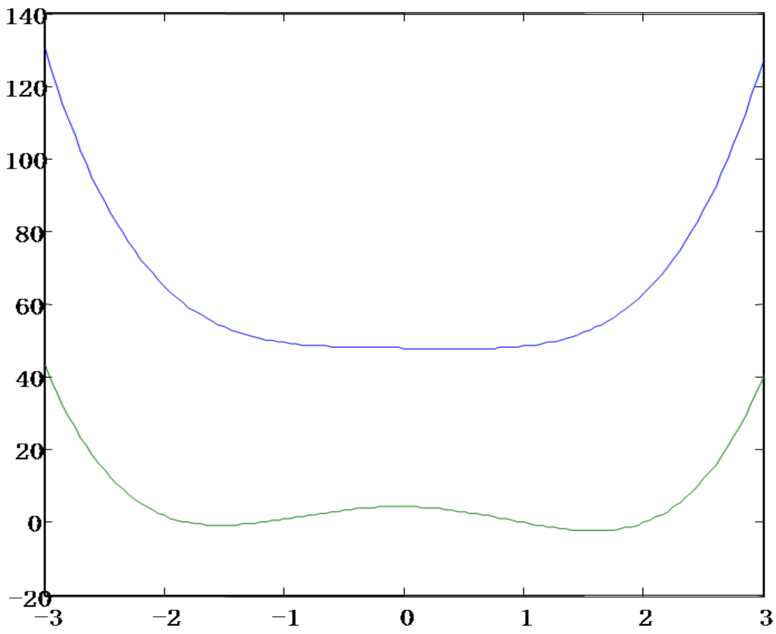

Fig.1-principe de la Méthode

3.2.2 Le principe de DCA

Notons que DCA ne fonctionne q'uavec les composantes DC g et h.
A la K-ième itération de DCA, on remplace h par sa minorante affine $\mathrm{h}_k(x) = h(x^k) + \langle x - x^k,\ y^k\rangle$ au voisinage de x^k.

Sachant que h est une fonction convexe, on a donc h(x)$\geq h_k(x)$,$\forall x \in X$. Par suite, $g(x) - [h(x^k) + \langle x - x^k,\ y^k\rangle] \geq g(x) - h(x), \forall x \in X$. C'est à dire, $g(x) - [h(x^k) + \langle x - x^k,\ y^k\rangle]$ est une fonction majorante de la fonction $f(x)$.

En effet, la surface de f^k peut être imaginée comme un bol se plaçant au dessus de la surface de f ; de plus, les deux surfaces se touchent au point $(x^k,\ f(x^k))$.

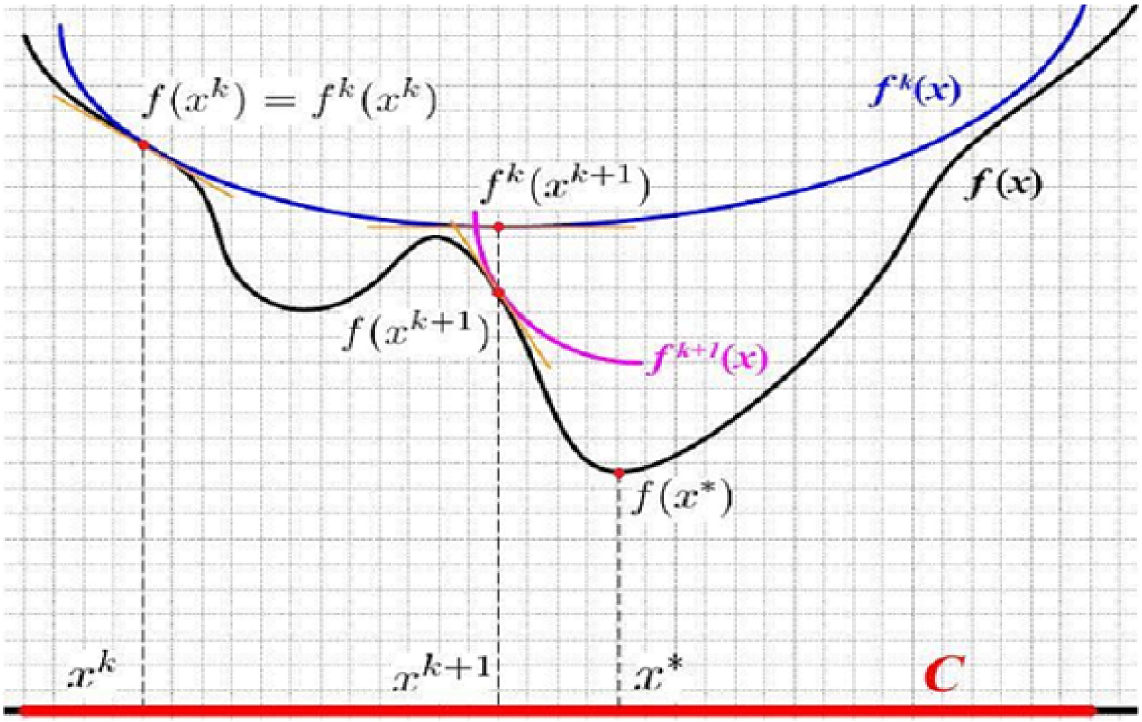

Fig.2-Principe du DCA

Proposition 1 : [1]

Soit g, h deux fonction convexes et différentiables sur X. Notons $f^k(x) = g(x) - [h(x^k) + \langle x - x^k,\ y^k \rangle]$. *On a :*

- $f^k(x) \geq f(x), \ \forall x \in X.$

- $f^k(x^k) = f(x^k).$

- $\bigtriangledown f^k(x^k) = \bigtriangledown f(x^k).$

Preuve :

Sachant que h est convexe, $h(x) \geq h_k(x),\ \forall x \in X$. *Par conséquent,*

$f^k(x)$=$g(x)$-$[h(x^k) + \langle x - x^k,\ y^k \rangle] \geq g(x) - h(x) = f(x),\ \forall x \in X,$ *c.à.d.*

$f^k(x) \geq f(x),\ \forall x \in X.$

$f^k(x)$=$g(x^k) - [h(x^k) + \langle x^k - x^k,\ y^k \rangle] = g(x^k) - h(x^k) + 0 = f(x^k).$

$\nabla f^k(x) = \nabla g(x) - y^k.$

Comme g est différentiable, on a $y^k = \nabla h(x^k)$. *C'est pourqoi*

$\nabla f^k(x) = \nabla g(x) - \nabla h(x^k)$. *Finalement, on a*

$\nabla f^k(x^k) = \nabla g(x^k) - \nabla h(x^k) = \nabla f(x^k).$

Proposition 2 :

1. Les suites $\{g(x^k) - h(x^k)\}$ *et* $\{h^*(y^k) - g^*(y^k)\}$ *déccroissentes et tendent vers la même limite* β *qui est supérieur ou égale à la valeur optimale globale* α.

2. Si $(g-h)(x^{k+1}) = (g-h)(x^k)$ *l'algorithme s'arrête à l'itération K+1, et le point* x^k *(resp* y^k*) est un point critique de* $g - h$ *(res* $h^* - g^*$*).*

3. Si la valeur optimale du problème (P) est finie et si les suites $\{x^k\}$ *et* $\{y^k\}$ *sont bornées. alors toute valeur d'adhérence* x^* *de la suite* $\{x^k\}$ *(resp.* y^* *de la suite* $\{y^k\}$*) est un point critique de* $g - h$ *(res* $h^* - g^*$*).*

Remarque :

Grace à la proposition 2, DCA s'arrête si au moins l'un des suites $\{(g - h)(x^k)\}$, $\{(\mathrm{h}^* - g^*)(y^k)\}, \{x^k\}, \{y^k\}$ converge. En pratique nous utilisons souvent les conditions d'arrêt suivantes :

- $|(g-h)(x^{k+1}) - (g-h)(x^k)| \leq \epsilon.$
- $\| x^{k+1} - x^k \| \leq \epsilon.$

3.2.3 propriétés de DCA

1- Le DCA construit une suite $\{\mathrm{x}^k\}$ telle que la suite $\{f(x^k)\}$ est décroissante. Ceci peut être ficilement vérifié sur cette figure car x^{k+1} est un minimum de f^k (donc $f^k(x^k) \geq f^k(x^{k+1})$ et $f(x^{k+1}) \leq f^k(x^{k+1})$, ainsi $f^k(x^k) = f(x^k)$. Finalement, on a $f(x^k) \geq f^k(x^{k+1}) \geq f(x^{k+1})$. Ceci montre que la suite $\{f(x^{k+1})\}$ est décroissante.

2- Pour la bornitude et la convergence de DCA, sachant que f : $\mathrm{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$ est bornée infrieurement sur R^n, la suite $\{f(x^k\}$ est aussi bornée inférieurement. On sait q'une suite $\{f(x^k)\}$ décroissante et bornée inférieurement est convergente.

3- Lorsque DCA converge vers un point x^*, ce point doit être un point KKT généralisé. On peut le comprendre, à laide de la figure.2. Si on lance DCA à partir d'un point de KKT généralisé de f(x^* sur la figure), alors DCA s'arrête immédiatement à ce point car il est aussi un minimum de f* (la fonction majorante convexe définie au point x^* par

$$f(x^*) = g(x) - [h(x^*) + \langle x - x^*, y^* \rangle], y^* \in \partial h(x^*)).$$

4- On peut constater, grace à la figure que DCA a la faculté de sauter certains voisinages de minima locaux. Par exemple, DCA saute un minimum local entre x^k et x^{K+1} et parvient à un voisinage de la solution globale. Bien que l'on ne puisse pas toujours garantir que ce phénomène se produie, on peut comprendre que la performance de DCA est susceptible de dépendre de la decomposition DC et de la position du point initial.

3.3 DCA (DC algorithm)

Étape 0 : x^0 donné, $K = 0$.

Étape 1 : On calcule $\mathrm{y}^K \in \partial h(x^K)$.

Étape 2 : On détermine $\mathrm{x}^{K+1} \in \partial g^*(y^K)$.

Étape 3 : Si les conditions d'arrêt sont vérifiées alors on termine DCA ; Si non K=K+1 et on répète Étape 1.

3.4 Application de DCA :

3.4.1 Exemple 1 :

$$(\mathrm{P}) \Longleftrightarrow \begin{cases} f(x) &= (x-1)(x-2)(x+1)(x+2) \longrightarrow Min \\ x &\in [-3, +3] \end{cases}$$

On applique le DCA à partir du point 3 et du point -3, aprés on cherche le minimum des deux minimum pour trouver le minimum global.

$\mathrm{g(x)}=\mathrm{x}^4 + 4,\ h(x) = 5x^2,\ h'(x) = 10x,\ y^k = \nabla h(x^k)$

$Si\ \mathrm{x}^0 = 3,\ K = 0,\ y^0 = 30$

1^{er} itération :

$\mathrm{x}^1 = 1.96,\ y^1 = 19.6$

2^{eme} itération :

$\mathrm{x}^2 = 1.7,\ y^2 = 17$

3^{eme} itération :

$\mathrm{x}^3 = 1.58,\ y^3 = 15.8$

Comme $|x^3 - x^2| = |1.58 - 1.7| < \epsilon$

Alors x=1.58 est la solution optimal du problème (minimum global) avec :

f(1,58)=-2.25

$Si\ \mathrm{x}^0 = -3,\ k = 0,\ y^0 = -30$

1^{er} itération :

$\mathrm{x}^1 = -1.96,\ y^1 = -19.6$

2^{eme} itération :

$\mathrm{x}^2 = -1.7,\ y^2 = -17$

3^{eme} itération :

$\mathrm{x}^3 = -1.58,\ y^3 = -15.8$

Comme $|x^3 - x^2| = |-1.58 + 1.7| < \epsilon$

Alors x=-1.58 est la solution optimal du problème (minimum global) avec :

f(-1.58)=-2.25

Remarque :

Comme f(-1.58)=f(1.58), appliqué le DCA à partir de -3 ou à partir de 3 on arrive au minimum global car la fonction f est symétrique.
La fonction standard fminbnd de MATLAB donne la même solution x=1.58 (minimum global)

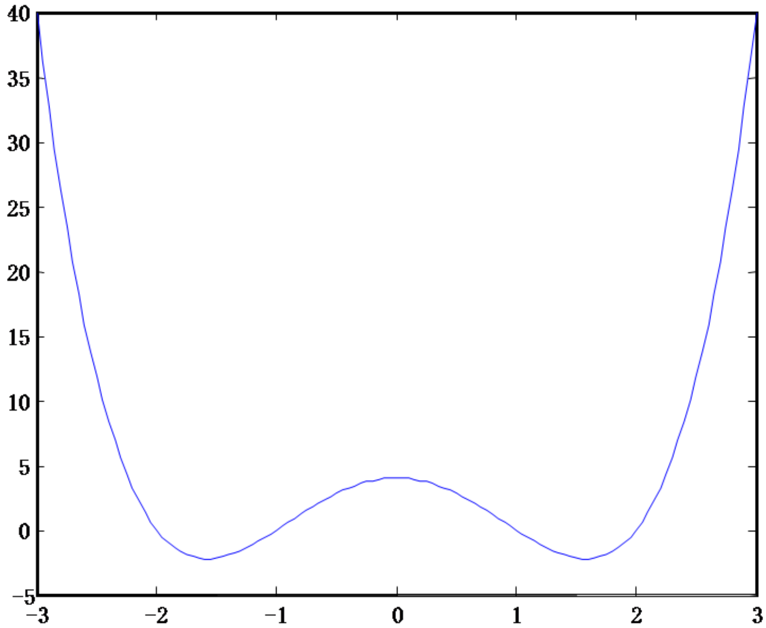

Fig.3-La Solution Graphique de l'exemple 1 (DCA)

3.4.2 Exemple 2 :

$$(\mathrm{P}) \Longleftrightarrow \begin{cases} f(x) & = & (x-1.2)(x-1.8)(x+1)(x+2) \longrightarrow Min \\ x & \in & [-3,+3] \end{cases}$$

On applique le DCA à partir du point 3 et du point -3, aprés on cherche le minimum des deux minimum pour trouver le minimum global.

g(x)=$x^4 + 0.48x + 4.32$, $h(x) = 4.84x^2$, $h'(x) = 9.68x$, $y^k = \nabla h(x^k)$

Si $x^0 = 3$, $k = 0$

3^{eme} itération de DCA donne :

$x^3 = +1.58$ (minimum local pas global)

Si $x^0 = -3$, $k = 0$

3^{eme} itération de DCA donne :

$x^3 = -1.58$ (minimum global).

La fonction fminbnd de MATLAB donne la solution x=1.58 (minimum local pas global).

Remarque :

Dans cet exemple on a deux minimums : Un est global, et un autre local. Alors on est obligé de prendre le minimum des deux qui est un minimum global.
La solution de l'exemple 2 avec le DCA est : x=-1.58 (minimum global).
La solution de l'exemple 2 avec la fonction standard fminbnd de MATLAB est : x=+1.58 (minimum local).

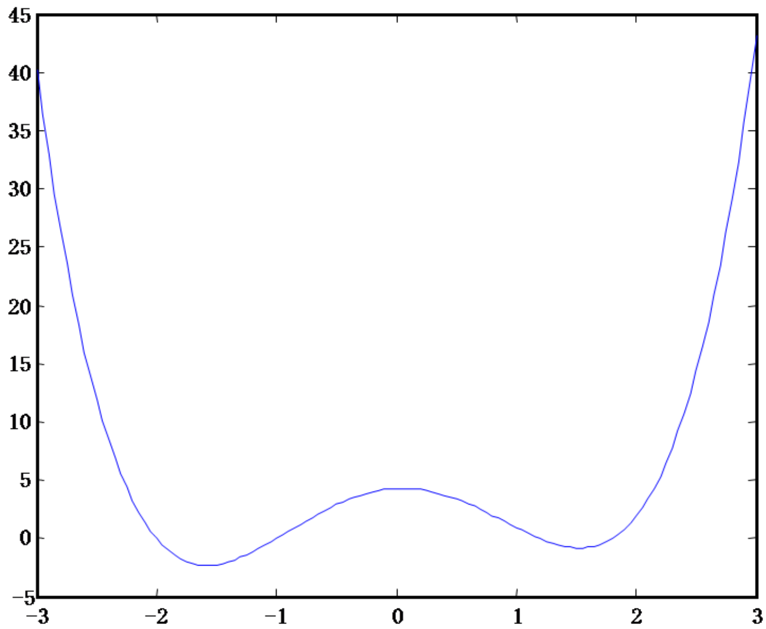

Fig.4-La Solution Graphique de l'exemple 2 (DCA)

3.4.3 Exemple 3 :

$$(\mathrm{P}) \Longleftrightarrow \begin{cases} f(x) = (x-1)(x-2)(x+1.8)(x+1.2) \longrightarrow Min \\ x \in [-3,+3] \end{cases}$$

$g(x) = x^4 + 0.48x + 4.32,\ h(x) = 4.84x^2,\ h'(x) = 9.68x$

$y^k = \nabla h(x^k)$

Si $x^0 = +3$

La 3^{eme} itération de DCA nous donne la solution x=+1.58 (minimum global).

Si $x^0 = -3$.

La 3^{eme} itération de DCA nous donne la solution x=-1.58 (minimum local).

La fonction standard fminbnd de MATLAB nous donne la solution x=-1.58 (minimum local).

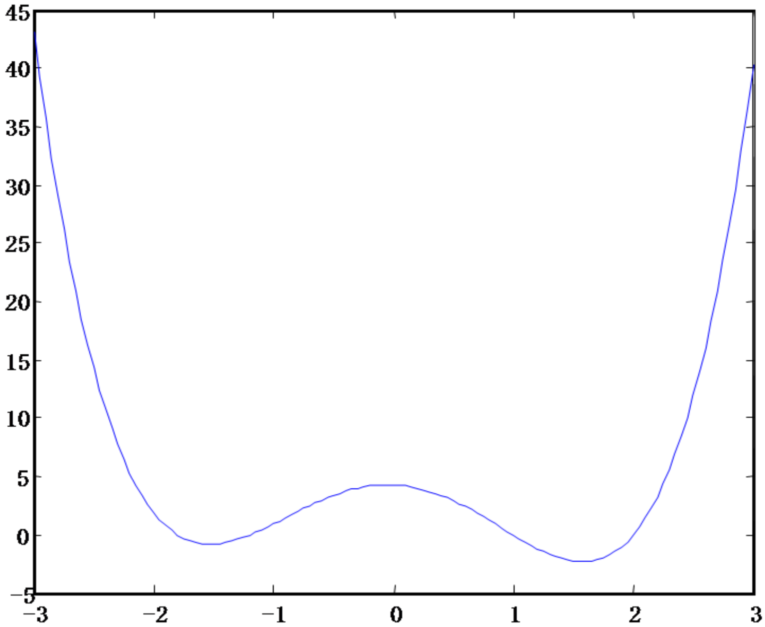

Fig.5-La Solution Graphique de l'exemple 3 (DCA)

Remarque :

Le DCA est très sensible au point de départ. Alors on propose dans l'exemple 4 de résoudre le problème de l'exemple 3 avec la Méthode proposée. Et avoir la solution global avec une seule itération de DCA.

3.5 Algorithme de la Méthode (MDCA)

*É*tape 0 : $x^0 = min\frac{1}{2}(f_k(x,a) + f_k(x,b)),\ k = 0.$

*É*tape 1 : Application du DCA à partire de x^0.

3.6 Application de l'algorithme (MDCA)

3.6.1 Exemple 4 :

$$(P)\Longleftrightarrow\begin{cases} f(x) &= (x-1)(x-2)(x+1.8)(x+1.2) \longrightarrow Min \\ x &\in [-3,+3] \end{cases}$$

$g(x) = x^4 + 0.48x + 4.32, h(x) = 4.84x^2, h'(x) = 9.68x$

$y^k = \nabla h(x^k)$

Nous cherchons à minimiser la moyenne des deux approximation de f à partire de -3 et de +3, $(f_k(x,-3), f_k(x,+3))$ avec :

$f_k(x,-3) = g(x) - h'(-3)(x+3) - h(-3)$

$f_k(x,+3) = g(x) - h'(+3)(x-3) - h(+3)$

Donc :

$f_k(x,-3) = x^4 + 28.56x + 47.88$

$f_k(x,+3) = x^4 - 29.52x + 47.88$

Etape :0

Résoudre le problème convexe (P') suivant :

$$(P')\Longleftrightarrow min\frac{1}{2}[f_k(x,-3) + f_k(x,+3)]$$

x=0.48 est la solution du problème (p')

Etape :1

Application de DCA à partire de x=0.48

$\mathrm{x}^0 = 0.48,\ k = 0,\ y^0 = 4.64$

1^{er} itération

x^1 est la solution du problème convexe :

$\min\{\mathrm{x}^4 - 16.11x + 5.44\}$

x=1.58 est la solution optimale (minimum globale) du problème (P) Alors que la fonction standard de MATLAB nous donne un minimum locale (x=-1.58)

Remarque :

Avec la méthode proposé on choisie pas de point de départ, à la place nous cherchons à minimiser la moyenne des deux approximation de f à partir de 3 et de -3 qui nous fournira un minimum qui sera situé dans la zone d'attraction du minimum global.

On applique le DCA à partir de ce minimum. qui nous donne le minimum globale cherché.

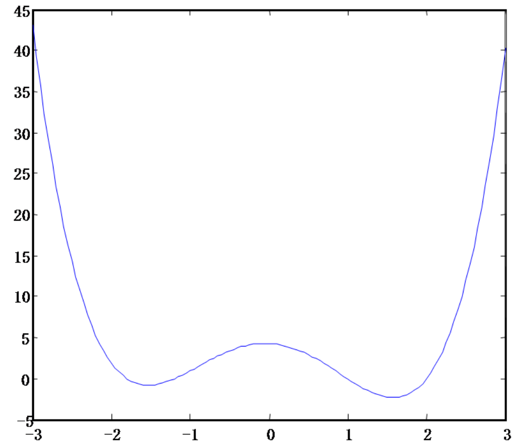

Fig.6-La Solution de l'exemple 4 (DCA)

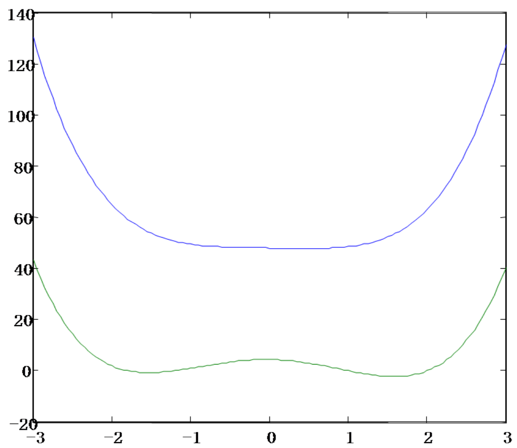

Fig.7-La Solution de l'exemple 4 (MDCA)

Chapitre 4

conclusion et perspectives

Nous nous sommes intéressé dans ce mémoire à une classe particulière de problèmes d'optimisation, à savoir : les problèmes non convexe (DC).
la programmation DC et DCA qui est une approche déterministe pour résoudre des problèmes non convexes. Cette approche utilise des outils d'analyse convexe et exploite la structure DC de ces problèmes.

La stratégie de minimisation de la moyenne suivie de l'application standard de DCA a conduit à l'obtention du minimum global de la fonction f alors que la fonction standard de MATLAB fminbnd trouve un minimum local pas global.

Il reste maintenant à téster d'autres exemples pour mieux évaluer la pértinence de cette stratégie qui renforce l'importance de DCA dans la résolution des problèmes non convexes.

Bibliographie

[1] NIU-Yi-Shuaiy : Programmation DC et DCA en Optimisation Combinatoire et Optimisation Polynomiale via les Techniques de SDP, France, Rouen 2010. thèse.

[2] Mamadou THIAO : Approches de la programmation DC et DCA en data mining, France, Rouen 2006. thèse.

[3] MOEINI Mahdi : La programmation DC et DCA pour l'optimisation de portefeuille, Paul Verlaine-Metz 2008. thèse.

[4] Gasmi Bouthaina : Contribution a l'étude des méthodes de résolution des problèmes d'optimisation , Algérie, Batna, 2007. thèse.

[5] Rozenn Texier-Picard : convexité et applications, Université Rennes1. thèse.

[6] PHAM Viet Nga ; Programmation DC et DCA pour l'optimisation non convexe/optimisation globale en variables mixtes entières. Codes et Applications. France,Rouen,2013. thèse.

[7] ALABBOUD HASSAN : La programmation semi-définie combinée et comparée avec d'autres probèmes d'optimisation, Université du Havre ,2007.

[8] Pierre Maréchal : éléments d'analyse convexe, Université Paul Sabatier.

[9] George B.Dantzing : Applications et prolongement de la programmation linéaire ;Ed Dunod-paris,(2000).

[10] Isabelle-Vialle : Mathématiques Appliquées à L'éconmie ; Ed Math sciences,Springer.

[11] jacques Teghem : programmation linéaire ; Springer(2005).

[12] Phan Dinh,Tao.,Lethi, H.A. : Convex analysis approach to D.C. programming : Theory, Algorithms and Applications. Acta Mathematica vietnamica 22(1), 287-367 (1997).

[13] Pham Din, T., Le Thi, H.A. : The DC programming and DCA Revisited With DC Models of Real World Nonconvex Optimization Problems. Annals of Operations Rechearch 133, 23-46 (2005).

[14] Phan Dinh, T.,Niu, Y.S. : DC Programming for Mixed Integer Program. Technical Report, LMI INSA-Rouen(2008).

[15] Fédéric Bonnans : Optimisation continue, Ed Hermann,(2005).

[16] Horst R. : A general class of branch-and-bound methods in global optimization with some new approaches for concave minimization. J. Optim. Theory Appl. 51, no. 2, 271-291, (1986).

[17] Horst R. : Deterministic global optimization with partition sets whose feasibility is not known : application to concave minimization, reverse convex constraints, DC-programming, and Lipschitzian optimization. J. Optim. Theory Appl. 58 ,no. 1, 11-37, (1988).

[18] Abedlkarim Keraghel : Élements d'analyse convexe dans R^n, Théorie fondamentale et exercises, Université de Sétif, avril, 2001.

[19] Dominique Az : élement d'analyse convexe et variationnelle, Ellipses, Paris, 1997.

[20] Aleksandroff A.D. : On surfaces represented as the difference of convex functions. Izvestiya Akad. Nauk Kazah. SSR. 60, Ser. Mat. Meh. 3, (1949).

[21] Hartman P. : On functions representable as a difference of convex functions. Pacific J. Math. 9, 707-713, (1959).

[22] Pham Dinh T. : Algorithmes de calcul d'une forme quadratique sur la boule unité de la norme maximum, Numer. Math. 45, 377-440, (1985).

[23] Pham Dinh T. : Algorithms for solving a class of non convex optimization problems. Methods of subgradients. Fermat days 85. Mathematics for Optimization, J.B. Hiriart Urruty (ed.), Elsevier Science Publishers, B.V. North-Holland, (1986).

[24] Pham Dinh T. : Duality in d.c. (difference of convex functions) optimization. Subgradient methods. Trends in Mathematical Optimization, K.H. Hoffmann et al. (ed.), International Series of Numer Math. 84, Birkhauser, (1988).

[25] H. A. Le Thi. Analyse numérique des algorithmes de l'optimisation D.C. : Approches locale et globale. Codes et simulations numériques en grande dimension. Applications. PhD thesis, Université de Rouen, France, Décembre 1994.

[26] H. A. Le Thi and T. Pham Dinh. Large-Scale Molecular Optimization from Distance Matrices by a D.C. Optimization Approach. SIAM Journal on Optimization, 14(1) :77-14-2003.

[27] T. Pham Dinh and H. A. Le Thi. Stabilité de la dualité lagrangienne en optimisation DC (différence de deux fonctions convexes). C. R. Acad, Paris, t. 318, Série I :379-384, 1994.

[28] T. Pham Dinh. Algorithmes de calcul d'une forme quadratique sur la boule unité de la norme maximum. Numerische Mathematik, 45 :377-440, 1985.

[29] T. Pham Dinh. Duality in DC (différence of convex functions) optimization. Subgradient methods. Trends in Mathematical Optimization, International Série of Numer Math, Birkh$\ddot{a}$auser, 84 :277-293, 1988.

[30] T. Pham Dinh and H. A. Le Thi. A DC optimazation algorithm for solving the trust region subproblem. SIAM Journal of Optimization, 8(2) :476-505, May 1998.

[31] Rockafellar R.T. : Convex Analysis. Princeton University Press, N.J., (1970).

[32] Hiriart-Urruty J.B., Lemarechal C. : Convex Analysis and Minimization Algorithms, Springer, Berlin, (1993).