THE UNIVERSITY OF NEW SOUTH WALES

SCHOOL OF ELECTRICAL ENGINEERING AND
COMPUTER SCIENCE AND ENGINEERING

The Élan Am386SC300 Portable Computer

John Zaitseff (2120715)

Bachelor of Engineering (Computer Engineering)

October 1995

Supervisor: A/Prof. Branko Celler
Assessor: Dr. Tim Hesketh

Table des matières

Table des figures

Liste des tableaux

Introdution

Nous faisons tous de l'optimisation, dans notre vie quotidienne, nous cherchons à optimiser notre temps de travail, nos espaces de rangement, où encore le trajet que nous aurons à parcourir pour nous rendre quelque part, etc. Nous cherchons tous une meilleure solution aux problèmes qui jalonnent notre existence. De manière générale, l'optimisation va donc consister à trouver cette meilleure solution .

L'optimisation est devenue une discipline incontournable du monde moderne dans lequel nous vivons, car celui-ci est sujet à une compétition internationale excessive et croissante. Dés lors, il devient nécessaire, voir vital pour les entreprises comme pour les gouvernements, de maximiser ou de minimiser toutes sortes de choses ; par exemple maximiser les profits tout en minimisant les pertes, améliorer si possible de façon optimale certaines processus de fabrication ou les fonctionnalités, de certaines processus objets ou produits

L'optimisation va consisté à rechercher dans le domaine initiale une solution qui maximise ou minimise une fonction objectif, pour un domaine continu et discret, on distingue classiquement deux type d'optimisation :

- **L'optimisation locale** : cherche une solution qui est la meilleure localement, cette solution est appelée un optimum local.
- **l'optimisation globale** : cherche quant à elle la meilleure solution du domaine en entier, c'est à dire que dans tout le domaine. il n'existe aucune solution qui lui soit meilleur tout en respectant les contraintes. Cette solution est appelée globale.

L'optimum global est aussi une solution locale. En revanche il est bien plus épineux de trouver l'optimum global, car aprés avoir trouvé cet optimum. Est la différence entre la solution globale et une solution locale est bien souvent significative mais l'intérêt ne pas que complétif. L'optimum global dans nombreux problème est la solution mathématique.

Le premier chapitre de notre étude commence par rappel des notions élémentaires de optimisation globale ainsi que des notations et définitions qui en découlent, afin d'établir clairement les bases de la programmation non linéaire sans contraintes.

Le deuxième chapitre est basé sur l'analyse d'intervalle et différent notations utilisées dans ce rapport. Après avoir présenté l'arithmétique d'intervalle ainsi que les diverses notations et définitions utilisées. Nous présenterons les théorèmes fondamentaux permettant de créer diverses fonctions d'inclusion, qui conduiront à l'encadrement de l'optimum global, ensuite, nous détaillons le principe de l'algorithme du séparation et d'évaluation qui est plus connu sous le nom anglais de **Branch and Bound** par intervalle, en mettant en évidence son caractère modulaire. Quelques techniques d'accélération précédemment étudiées seront également rappelles.

Le troisième chapitre est consacré à la présentation de la méthode d'encadrement de l'optimum global. Cette méthode, utilise l'arithmétique d'intervalle, nous allons d'abord faire la construction des hyperplans d'appui en chacun des sommets de pavé considéré. Ensuite on a présenté les méthodes différentiables de plusieurs variables. Et des tests numériques sur le problème d'optimisation sans contrainte valideront ces différentes méthodes intégrées dans un algorithme de type **Branch and Bound.**

Le dernier chapitre est consacré à une application informatique de la méthode de simplexe droit permettant de calcul le minimum globale d'une fonction différentiable à plusieurs variables

Chapitre 1

Quelques généralités

Ce premier chapitre a pour but de rappeler quelques définitions et notions de base dont nous servons dans les chapitres suivants :

1.1 Fonction à plusieurs variables

On appelle fonction numérique à plusieurs variables toutes application f définie sur un ensemble E de $\mathbb{R}^n$ et a valeurs dans $\mathbb{R}$:

tel que :

$$f : E \ \subset \mathbb{R}^n \rightarrow \mathbb{R}$$

$$(x_1, ..., x_n) \rightarrow f(x_1, ..., x_n)$$

le gradient :

soit $\quad f : E \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$

on définit le gradient de f par :

$$\nabla_i \ f(x) = \left(\frac{\partial f(x)}{x_1}, ..., \frac{\partial f(x)}{x_n} \right)^T$$

la matrice hessienne :

soit $f(x)$ une fonction de n variable, la matrice hessienne de $f(x)$,noté $Hf(x)$, est donnée par

$$H_i(x) = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & ... & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & ... & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & ... & \frac{\partial^2 f(x)}{\partial x_2^2} & ... & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & ... & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & ... & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} \end{pmatrix}$$

1.2 Optimisation globale

Optimiser : rendre optimal, donner à quelque chose les meilleures conditions d'utilisation, de fonctionnement ou de rendement au regard de certaines circonstances. (Déf. du LAROUSSE).

Le problème que l'on étudie ici est celui de la recherche du minimum d'une fonction réelle f de n variables réelles $x_1, x_2, ..., x_n$.chacune de ces variables pouvant prendre n'importe quelle valeur de $-\infty$ à $+\infty$. De tels problème apparaissent fréquemment dans les applications.

Soit $f : X \to \mathbb{R}$, pour tout $x \in X$, $x = (x_1, x_2, ..., x_n)^T$,on associe la valeur réelle :

$$f(x) = f(x_1, x_2, ..., x_n)$$

On cherche a résoudre :

$$(\mathbf{P}) \qquad \begin{cases} \min\limits_{x \epsilon E} f(x) \\ x \epsilon X \end{cases} \tag{1,1}$$

où E est un espace de Banach, X un sous-ensemble de E.

TERMINOLOGIE :

• f est appelée fonction coût, objectif ou encore critère

•X est appelé ensemble des contraintes ou domaine admissible du problème (P).

• Le problème (P) est dit réalisable si : $X \neq \varnothing$;

• Si l'espace d'arrivée de f est $\mathbb{R}$,est de dimension supérieure ou égale à 2, on parle d'optimisation multi-critères, mais ce n'est pas l'objet de ce mémoire.

•Un cas fréquent en optimisation est celui où X est définie par des égalités et des inégalités :

$$X = \{f(x) \epsilon \mathbb{R}^n : h(x) = 0; g(x) \leq 0\}$$

avec :

•$h : \mathbb{R}^n \to \mathbb{R}^p$ est supposée continue. "$h(x) = 0$" représente p contraintes d'égalités :

$$h_i(x) = 0 \ ; \ i = 1, ..., p$$

• $g : \mathbb{R}^n \to \mathbb{R}^{\ p}$ est elle aussi supposée continue et "$g(x) \leq 0$" représente q contraintes d'inégalités :

$$g_i(x) \leq 0; \ \ i = 1, ..., q$$

Dans ce cas, on dit qu'il s'agit d'un problème d'optimisation à n variables de décision, p contraintes d'égalités et q contraintes d'inégalités. Résoudre le problème (P) revient généralement à en chercher des solutions locales (faute de mieux!) au sens de la définition suivante :

Définition 1(Minimum local/Minimum global)

*$x_0 \epsilon E$ est un minimum local de f sur $X \subset E$ si et seulement si :

$$x_0 \epsilon X \text{ et } \exists\, V_{x_0} \text{ un voisinage de } x_0 \text{ tq} : \forall\, x \in V x_0 \cap X \,;\, f(x) \geq f(x_0) \tag{1.2}$$

*$x_0 \epsilon E$ est minimum global de f sur X si et seulement si :

$$x_0 \epsilon X \text{ et } \forall\, x \epsilon\, X \,;\, f(x) \geq f(x_0) \tag{1.3}$$

Les notions de maximum local et global peuvent être définies de façon tout à fait similaire. En fait, on démontre facilement que les problèmes (avec ou sans contraintes) :

$$\underset{x}{min} f(x) \quad \text{et} \quad \underset{x}{max}\, x - f(x)$$

sont équivalents dans le sens où ils ont même ensemble de solutions et :

$$\underset{x}{min} f(x) = - \ \underset{x}{max}(-f(x)) \text{ , ou encore, } max f(x) = -\underset{x}{min}(-f(x))$$

Ainsi la recherche d'un maximum pouvant se ramener à la recherche d'un minimum, nous porterons une attention plus particulière à la recherche du minimum.

1.3 Convexité et optimisation :

Les problèmes dont les données sont convexes, constituent une classe importante en optimisation, car fréquemment rencontrés dans les applications et à la base de nombreuses méthodes développées pour des problèmes plus généraux.

1.3.1 Eléments d'analyse convexe :

Définition2 (Ensemble convexe) :

Soit $C \subset \mathbb{R}^n$. L'ensemble C est convexe si et seulement si :

$$\forall (x; y) \in C^2 \text{ , } \lambda \in [0, 1], x + (1 - \lambda) y \ \in C$$

c.'est-a-dire, si x et y sont deux éléments de C alors le segment qui relie x à y est inclus dans C

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

Définition 3(Fonction convexe/strictement convexe)

Soit $C \subset \mathbb{R}^n$ convexe et $f : C \to \mathbb{R}$,f est convexe ssi :

$$\forall (x;y) \in C^2 \text{ , } \forall \lambda \in [0;1], f(\lambda x + (1-\lambda)y) \leq f(x) + (1-\lambda)f(y) \tag{1,4}$$

f est strictement convexe ssi :

$$\forall (x;y) \in C^2 \text{ , } x \neq y \text{ , } \forall \lambda \in]0;1[\text{ , } f(\lambda x + (1-\lambda)y) < f(x) + (1-\lambda)f(y) \tag{1,5}$$

Théorème 1 :

Soit $C \in \mathbb{R}^n$ convexe et $f : C \to \mathbb{R}$ différentiable. La fonction f est convexe ssi :

$$\forall (x;y) \in C^2, f(y) \geq f(x) + \langle \bigtriangledown f(x) \text{ , } y - x \rangle \tag{1.6}$$

ou de façon équivalente, ssi :

$$\forall (x;y) \in C^2 \text{ ; } \langle \nabla f(y) - \bigtriangledown f(x) \text{ , } y - x \rangle \geq 0 \tag{1.7}$$

Preuve :

Soit $(x;y) \in C^2$. Par convexité de f , on a donc pour tout $t \in]0,1[$:

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y) = f(x) + t(f(y) + f(x))$$

soit :

$\dfrac{f(x + t(y-x)) - f(x)}{t} \leq f(y) - f(x)$: En passant à la limite pour $t \to 0^+$, il suit (1.6).

Réciproquement, on applique (1.6) à $tx + (1-t)y$ et x , puis à $tx + (1-t)y$ et y, d'où :

$$f(x) \geq f(tx + (1-t)y) + (1-t) \langle \bigtriangledown f(tx + (1-t)y), y - x \rangle$$

$$f(x) \geq f(tx + (1-t)y) - t \langle \bigtriangledown f(tx + (1-t)y), y - x \rangle$$

En combinant ces deux inégalités, on obtient : $tf(x) + (1-t)f(y) \geq f(tx + (1-t)y)$; et donc la convexité de f. En échangeant les rôles de x et y dans (1.6), puis en sommant les deux inégalités obtenues, on démontre sans problème que (1.6) implique (1.7). Pour montrer la réciproque, on introduit :

$$\varphi : t \in [0;1] \to tf(x) + (1-t)f(y) - f(tx + (1-t)y);$$

φ est positive sur $[0;1]$.

Remarque 1.1 :

Si ∇f est strictement monotone i.e. si les inégalités (1.6) et (1.7) sont strictes pour $x \neq y$, alors f est strictement convexe. Si de plus, la fonctionnelle f est deux fois différentiable, on a alors une caractérisation d'ordre deux de la convexité via la Hessienne.

Théorème 2 :

Soit $f \ : \mathbb{R}^n \to \mathbb{R}$ une fonctionnelle de classe C^2.

•Si la hessienne $H[f](x)$ de f est une matrice symétrique définie positive pour tout $x \in \mathbb{R}^n$, alors f est strictement convexe.

•Si $H[f](x)$ est une matrice symétrique semi définie positive pour tout $x \in \mathbb{R}^n$, alors f est convexe.

Cas des fonctions convexes (condition nécessaire et suffisante d'optimalité globale) :

Dans le cas d'une fonction convexe f définie sur $\mathbb{R}^n$, la condition nécessaire et suffisante pour que x^n soit un minimum global de f et que 0 définit un sous-gradient de f en x^*. Pour une fonction continument différentiable, on obtient donc :

Théorème 3 :

si f est un fonction convexe continument différentiable, une condition nécessaire et suffisante pour que x^* soit un optimum global de f sur $\mathbb{R}^n$ est que : $\nabla f(x) = 0$. Autrement dit, dans le cas convexe, la stationnarité à elle seul constitue une condition nécessaire et suffisante d'optimalité globale.

Théorème 4(Condition suffisante d'optimalité globale) :

Soient $C \subset \mathbb{R}^n$ un convexe et $f : C \to \mathbb{R}$ une fonctionnelle. Soit x^* un point de minimum local de$f$.

$i.$ Si f est convexe, alors x^* est un point de minimum global de f.

$ii.$ Si f est strictement convexe, alors x^* est l'unique point de minimum global de f.

1.3.2 Optimisation sans contrainte

Nous nous intéressons ici à la conception de méthodes numériques pour la recherche des point $x \in \mathbb{R}^n$ qui réalisent le minimum d'une fonction $f : \mathbb{R}^n \to \mathbb{R}$:

$$(P) \quad \min_{x \in \mathbb{R}^n} f(x)$$

où f est supposée au moins différentiable. On parle d'optimisation sans contrainte.

Conditions d'optimalité :

Théorème 5 (Conditions nécessaires d'optimalité locale)

Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction différentiable.

· Si $x^* \in \mathbb{R}^n$ réalise un minimum local (resp. maximum local) de f, alors :

$$\nabla f(x^*) = 0 \qquad \text{(CN d'optimalité du 1er ordre)}$$

·Si, de plus, f est deux fois différentiable dans un voisinage ouvert de x^*, alors :

$H[f](x^*)$ semi définie positive (CN d'optimalité du 2nd ordre)
(resp. $H[f](x^*)$ semi définie négative)

Théorème 6 (Condition Suffisante d'optimalité locale) :

Soit O un ouvert de $\mathbb{R}^n$. Soit $f : \mathbb{R}^n \to \mathbb{R}$ une application supposée de classe C^2 sur O.

·Si $\overline{x} \in O$ vérifie : $\nabla f(\overline{x}) = 0$ *et* $H[f](\overline{x})$ symétrique, définie positive (resp. définie négative), Alors $\overline{x}$ est un point de minimum local (resp. maximum local) de f.

Définition 4(Points critiques) :

Soit $f : \mathbb{R}^n \to \mathbb{R}$ une application différentiable. Tout point $x \in \mathbb{R}^n$ vérifiant :

$$\nabla f(x) = 0$$

est appelé point critique (ou point stationnaire) de f.

1.4 Boite

on désigne par boite, un produit cartésien d'intervalles, en pratique, c'est un vecteur d'intervalles qui définit un espace de recherche dans le quel se trouvent les valeurs des inconnues.

1.5 Fonction affine

une fonction affine est une fonction de la variable réelle dont la représentation graphique est une droite, c'est une fonction polynôme de degré inférieur où égale à un, elle est définie par :

$$\begin{array}{l} f : \mathbb{R}^n \to \mathbb{R} \\ x \to f(x) = ax + b \ , \qquad \text{avec } a, b \ \epsilon \mathbb{R}. \end{array}$$

1.6 Cône polyédrique

1.6.1 polyédrique :

Un ensemble convexe $p \subset \mathbb{R}^n$ est un polyédre, s'il est intersection d'une famille finie où infinie de demi-espaces fermés. En d'autre termes, un polyèdre est un ensemble de solutions d'un système fini d'inégalité linéaire de la forme :

$$\left\langle a^i, x \right\rangle \leq b \ , \ i = 1, ..., n$$

où A est une matrice $m \times n$, $b \in \mathbb{R}^m$, m et n sont deux entiers positifs.

1.6.2 Cône :

un cône est un ensemble de points tels que :

$$\begin{array}{c} x \in C \Rightarrow \lambda x \in C, \quad \text{pour tout} \ \ \lambda \succeq 0. \\ x_1 \in C \ , \ x_2 \in C \Rightarrow x_1 + x_2 \ \in C \end{array}$$

un cône est dite polyedrique si il a la forme :

$$C = \{ x \setminus \langle a_j, x \rangle \preceq 0, j = 1, ..., r \}.$$

1.7 Théorème de valeur moyenne

En analyse réelle, le théorème de la moyenne est un résultat classique concernant l'intégration des fonctions continues d'une variable réelles, selon lequel la moyenne d'une fonction continue sur un segment se réalise comme valeur de la fonction. Pour toute fonction F à valeurs réelles, définie et continue sur un segment $[a, b]$, avec $a \prec b$, il existe un réel c compris entre a et b, vérifiant :

$$f(c) = \frac{1}{b\ -\ c} \int_a^b f(x)dx$$

1.8 la formule de Taylor

Soit I un intervalle de $\mathbb{R}$ et $a \epsilon I$

E : un espace vectoriel normé de dimension finie.

f : une fonction définie de I dans E qui soit dérivable en a jusqu'à l'ordre n (un entier naturel). Alors, pour tous x dans I, ona :

$$f(x) = f(a) + \frac{f^{\ (1)}(a)}{1!}(x-a) + \frac{f^{\ (2)}(a)}{2!}(x-a)^2 + ... + \frac{f^{\ (n)}(a)}{n!}(x-a)^n + O(x-a)^n$$

Chapitre 2

Analyse d'intervalle

L'analyse d'intervalle est née dans les années 1960 avec l'avènement des premiers ordinateurs. En effet. Au début de l'invention des calculs par ordinateur. Les formats pour coder les nombre flottants n'étaient pas s'accumuler et aboutir à des résultats aberrants. Le premier livre sur ce domaine est celui de Moore 1966 [1]. Ce n'est pas historiquement le premier travail sur ce sujet, mais c'est à la suite de ce livre que plus de 1000 articles fussent écrits, ce qui donna un formidable élan à l'analyse d'intervalles.

L'idée de l'analyse d'intervalle est de représente tous les nombres réelles par deux nombres flottants qui l'encadrent. L'analyse d'intervalle a rapidement été utilisée en optimisation, pour la résolution de systèmes linéaires et non-linéaires [2]. Grâce à celle-ci de nombreux algorithmes d'optimisation globale ont été développés , l'analyse d'intervalle va être utilisée ici pour encadrer de façon précise l'optimum global, et ainsi permettre de déterminer l'optimum global ainsi que tous les optimums d'un problème d'optimisation avec ou sans contraintes.

2.1 Arithmétique d'intervalle standard :

Un intervalle X est caractérisé par sa borne inférieure x^l (*lower bound*) et sa borne supérieure x^u (*upper bound*) avec

$$X = \left[x^l \ , \ x^u\right] = x \in \mathbb{R}/ \ x^l \leq x^u$$

Notations et définitions :

$\nsucc$: l'ensemble des intervalles compact réels.

$$\nsucc = X = \{[x^l, x^u]/x^l, x^u \in \mathbb{R} \ \text{ et } \ x^l \leq x^u\}$$

définition1 (le milieu d'un intervalle) : C'est la fonction de $\nsucc$ dans $\mathbb{R}$ par :

$$mid(X) = \frac{x^l \ + \ x^u}{2}$$

Par extension aux formes vectorielles, nous obtenons la fonction vectorielle définie de $\nsucc^n$ dans $\mathbb{R}^n$ par :

$$mid(X) = (mid(X_1), \, mid(X_2), ..., \, mid(X_n))^T \qquad \forall \ X = (x_1, ..., x_n) \in \nsucc^n$$

définition2 (largeur d'un intervalle) : C'est la fonction de $\nsucc$ dans $\mathbb{R}$ par :

$$w(X) = x^u - x^l \quad , \quad X \,\epsilon \nsucc .$$

Par extension aux formes vectorielles, nous obtenons la fonction vectorielle définie de $\nsucc^n$ dans $\mathbb{R}^n$ par :

$$\vec{w}(X) = (w(x_1), ..., w(x_n))^T \qquad\qquad \forall\ X \in \nsucc^n .$$

définition3(rayon de l'intervale) : C'est la fonction de $\nsucc$ dans $\mathbb{R}$ par :

$$rad\,(X) = \frac{x^u \ - \ x^l}{2}$$

définition4 :

On appelle " boite de dimension n " tout vecteur d'intervalles $X = (x_1, x_2, ..., x_n)$. Autrement dit chacune des composantes est un intervalle compact réel. L'ensemble des boites de dimension n est noté $\nsucc^n$. La taille d'une boite de $\nsucc^n$ est définie par :

$$\forall X \in \nsucc^n : w(X) = \max_{i=1,...,n} w(X_i)$$

2.2 calculs :

Définition5 :

Soit X et Y deux intervalles appartenant à $\nsucc$, les opération usuels de l'arithmétique d'intervalles sont de telle sorte que l'intervalle résultat soit le plus petit intervalle contenant tout les points images des éléments des intervalles de départ. Ainsi, la forme générale des opérateurs unaires devient :

$$X \ op \ Y = \{x \ op \ y \ / \ x \in X \ et \ y \in Y\} \text{ avec } op \in \{+, -, \times, \div\}$$

2.2.1 Opérations arithmétiques :

Quand on applique la définition précédente aux opérations , $+, =, -, \times, ^2, \div$ ou $\sqrt{}$, on obtient les formules suivantes, plus utilisables en pratique que la définition abstraite.

$$\begin{cases} [x^l \ , \ x^u] \ + \ [y^l \ , \ y^u] = [x^l + y^l, x^u + y^u] \\ [x^l \ , \ x^u] \ - \ [y^l \ , \ y^u] = [x^l - y^u, x^u - y^l] \\ [x^l \ , \ x^u] \times \ [y^l \ , \ y^u] = \begin{bmatrix} \min(x^l \times \ y^l, \ x^l \ \times y^u, x^u \ \times y^l, x^u \times \ y^u) \ , \\ \max(x^l \times \ y^l, \ x^l \ \times y^u, x^u \ \times y^l, x^u \times \ y^u) \end{bmatrix} \end{cases}$$

$$\begin{cases} \left[x^l\ ,\ x^u\right]^2 = \left[\min\left((x^l)^2\ ,\ (x^u)^2\right), \max\left((x^l)^2\ ,\ (x^u)^2\right)\right] & \text{si } 0 \notin \left[x^l\ ,\ x^u\right] \\ \left[0\ , \max\ \left((x^l)^2\ ,\ (x^u)^2\right)\right] & \text{sinon} \end{cases}$$

$$\begin{cases} 1/\left[y^l\ ,\ y^u\right] = \left[\min\left(1/\ y^l\ , 1/\ y^u\right), \max(1/\ y^l\ , 1/\ y^u)\right] & \text{si } 0\ \notin \left[y^l\ ,\ y^u\right] \\ \left[x^l\ ,\ x^u\right]\ /\ \left[y^l\ ,\ y^u\right] = \left[x^l\ ,\ x^u\right] \times (1/\ y^l\ , 1/\ y^u) & \text{si } 0 \notin \left[y^l\ ,\ y^u\right] \end{cases}$$

Fonction usuelles :

Soit X un intervalle de $\nsucc$, nous pouvons définir aussi :

1. **puissance** :

$$X^n = \begin{cases} [1,1] & n = 1 \\ \left[(x^l)^n\ , (\ x^u\ )^n\right] & x^l \geq 0 \text{ ou si } 0\ \epsilon\ X \text{ et n impair} \\ \left[(x^u)^n\ , (\ x^l\ )^n\right] & x^l \leq 0 \\ \left[0, \max((x^l)^n\ , (\ x^u\ )^n)\right] & \text{si } 0\ \epsilon\ X \text{ et n pair} \end{cases}$$

Ceci pour tout n appartenant à $\mathbb{N}$.

2. **Logarithme** :

Le logarithme étant une fonction strictement croissante sur l'intervalle $[0, +\infty[$, nous obtenons :

$$\log(x) = (\log(x^l), \log(x^u))\ , \text{ avec } X \subset [0, +\infty[\,.$$

3. **Racine carrée** :

Laracine carée étant une fonction strictement croissante sur l'itervalle $]0, +\infty[$, nous obtenouns :

$$\sqrt{x} = \left[\sqrt{x^l}\ ,\ \sqrt{x^u}\right] \qquad \text{avec } X \subset [0, +\infty[$$

4. **Exponentielle :**

L'exponentielle étant une fonction strictement croissante sur $\mathbb{R}$ on obtient :

$$\forall A \in \nsucc : \exp(A) = \{\exp x/x \in A\} = [\exp(a^L),\ \exp(a^U)].$$

5. **valeur absolue :**

$$| A |= \begin{cases} \left[0, \max\left\{| x^l |, | x^u |\right\}\right] & 0 \in X \\ \left[x^l \ , \ x^u\right] & x^l \geq 0 \\ \left[| x^l | \ , \ | x^u |\right] & x^u \ \preceq 0 \end{cases}$$

2.3 Quelque propriétés de l'arithmitique d'intervalle :

1. On peut d'ores et déjà constater que les opérations définies ci-dessus ne présentent pas les propriétés algébriques de leurs contreparties ponctuelles. Tout d'abord, la soustraction n'est pas la réciproque de l'addition. Par exemple,

si $x = [2, 4]$, $x - x = [2, 4] - [2, 4] = [-2, 2] \neq 0$. même s'il le contient. En effet,

$x - x = \{x - y \mid x \in x, y \in x \} \supset \{ x - x \mid x \in x\} = \{0\}$.

et l'inclusion est stricte.

2. De la même façon, la division n'est pas la réciproque de la multiplication : si $x = [2, 3]$, l'intervalle

$x \ / \ x = [2, 3] \ / \ [2, 3] = [2/3, 3/2]$ n'est pas égal à 1 même s'il le contient.

De plus, la multiplication d'un intervalle par lui-même n'est pas égal à l'élévation au carré : si $x = [-3, 2]$,

$x \times x = [-3, 2] \times [-3, 2] = [-6, 9]$,alors que $x^2 = \{x^2 \mid x \in x\} = [0, 9]$.

Enfin, la multiplication n'est pas distributive par rapport à l'addition : si $x = [-2, 3]$, $y = [1, 4]$ et $z = [-2, 1]$,

$x \times (y + z) = [-2, 3] \times ([1, 4] + [-2, 1]) = [-2, 3] \times [-1, 5] = [-10, 15]$

$x \times y + x \times z = [-2, 3] \times [1, 4] + [-2, 3] \times [-2, 1] = [-8, 12] + [-6, 4] = [-14, 16]$

Comme l'illustre cet exemple, la multiplication est sous-distributive par rapport à l'addition, c'est-à-dire que

$x \times (y + z) x \times y + x \times z$.

3. pour tout (x,y) $\in X \times Y$, $x \ op \ y \in X \ op \ Y$

4. principe d'inclusion si $X \subset C$ et $Y \subseteq D$ alors

$$X \ \mathbf{op} \ Y \quad \subseteq \quad C \ \mathbf{op} \ D$$

5. les nombres réels n'étant pas tous représentables en machines, le résultat exact de l'opération x op y ne pourra généralement pas être obtenu. Ce résultat sera en cadré par l'intervalle X **op** Y .

op : étant l'une des quatre opérations élémentaire : $+, -, \times, \div$.

6. la relation d'ordre ($\prec$) est une relation d'ordre partiel sur :

$$\not\succ : \ x < y \ \textit{si et seulement si} \ x^u \preceq y^l.$$

7. les opérateurs sur les ensembles gardent la même définition pour les intervalles. Ainsi l'union et l'intersection de deux intervalles donnent :

$$X \cup Y = \left[\min\left(x^l, y^l\right), \max\left(x^u, y^u\right)\right]$$

$$X \cap Y = \begin{cases} \left[\max(x^l, y^l), \min(x^u, y^u)\right] & \text{sinon} \\ 0 \qquad x^u < y^l \qquad \text{ou} \qquad y^u < x^l \end{cases}$$

2.4 Evaluation d'une expression

Puisque l'on sait calculer le résultat d'une opération arithmétique ou d'une fonction élémentaire quand les variables prennent pour valeur des intervalles, on sait également calculer le résultat d'une expression mêlant opérations arithmétiques ou algébriques et fonctions élémentaires sur des intervalles. Voici simplement quelques exemples pour illustrer ce propos.

L'expression polynomiale :

$f : x \to x^2 - 2x + 1, \qquad I = [-1, 3]$

$I^2 - 2.I + 1 = [-1, 3]^2 - 2[-1, 3] + 1 = [0, 9] + [-6, 2] + 1 = [-5, 12]$

et en écrivant $\quad x^2 - 2x + 1 = x(x - 2) + 1$

$I(I\ -2) + 1 = [-1, 3][-3, 1] + 1 = [-9, 3] + 1 = [-8, 4]$

alors que $\quad F(I) = f(I) = [0, 4]$.

2.5 Arithmétique Arrondie

Nous avons vu dans les paragraphes précédents comment l'arithmétique d'intervalle nous permettait de calculer des bornes précises de l'encadrement des opérations élémentaires $+$, $-$,$\times$, $\div$, en supposant que le calcul effectué sur les bornes soit d'une grande précision. ce pendant même au niveau des opérations élémentaires le calcul peut manquer de précision.

Exemple :

$[0.123\ , 0.456] + [0.0116\ , 0.0214] = [0.1346\ , 0.46774]$

qui peut être arroundie par $[0.135\ , 0.47\ ]$ avec une précision de 3 décimales.

2.6 Fonction d'inclusion :

Soit $D \subseteq \mathbb{R}$ et $f\ : D \to \mathbb{R}$

Définition 1 (image directe d'une application) :

Soit $F\,(Y) = \{\ f(y) \setminus\ y\ \in\ Y\}$, pour tout Y appartenant à $\ngtr(D)$, Où $\ngtr(D)$ est l'ensemble des intervalles compacts de D.
$f(Y)$ est appelée image direct de f sur Y.

Définition 2 (**Fonction d'inclusion**) :

Une fonction $F : \ngtr(D) \to \ngtr$ est appelé fonction d'inclusion pour f si seulement si :
$f(y) \subseteq F(Y)$, est applée fonction d'inclusion pour tout $Y \in \ngtr(D)$, où $f(Y)$ représent l'image directe de f sur Y , est une fonction retournant un minorant et majorant de f sur un intervalle donne Y .

2.7 Extension naturelle et propiétés :

En pratique l'arithmétique d'intervalles est utilisée pour calculer des minorants et des majorants de fonction $\mathbb{R}^n$ dans $\mathbb{R}$. Pour cela, introduisons d'abord quelques définitions. Il existe des définitions plus précises d'une fonction explicite (facturable fonctions en anglais) et d'un problème explicite (facturable program en anglais). Mais, par souci de concision, nous avons préféré utiliser des définition plus intuitives.

Définition :

Soit $f :\ \mathbb{R}^n \to \mathbb{R}$, f possède une expression explicite si l'expression analytique de f est connue de façon explicite, c'est-à-dire qu'elle peut s'écrire en n'utilisant que des variables, des fonctions et des opérateurs élémentaire tels que $+, -, \sqrt{}, \times, \div, \log, \exp, abs, cos,\ sin\ ,\ arcos$. On dit que f est une fonction explicite.

Définition :

Soit $f\ :\ X\ \subset\ \mathbb{R}^n\ \to\ \mathbb{R}$, une fonction d'inclusion de f est une fonction retournant un minorant et un majorant de f sur un intervalle donné. On la note, de manière générale.

$$F : X \cap \not\succ^n \rightarrow \not\succ. \quad \text{Ainsi} \ : \qquad \forall z \in Z \subseteq X \ , \ f(z) \in F(z)$$

Définition :

Soit $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$. une extension naturelle aux intervalles d'une fonction est la réécriture de f en remplaçant toutes les occurrences d'une variable par intervalle correspondant et les opérateurs classiques par leur équivalent en arithmétique d'intervalles. Elle est noté $EN_f\,(x)$.

Toutes les fonctions explicites possèdent une extension naturelle. Mais, l'évaluation de celle-ci n'est pas unique. En effet, certaines fonctions pouvent s'écrire sous forme développée ou factorisée. Il existe donc une extension naturelle pour chaque façon d'écrire la fonction.

Proposition :

L'extention naturelle aux intervalles d'une fonction est une fonction d'inclusion.

Définition :

Soit $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$,l'image directe de f sur X, notée $f(x)$, est l'intervalle de largeur minimale telque :

$$\forall x \in X \ , \ f(x) \in f(X).$$

Théorème :

Soit $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$, si EN_f l'extention naturelle de f ne posséde qu'une seule occurence de chaque variable, alors l'image directe de f sur X est exactement $EN_f\,(X)$.

$$\forall Z \subseteq X,\ EN_f\,(Z) = f(Z)$$

Remarque :

Toutes les extention naturelles aux intervalles de differentes expressions d'une même fonction ne sont pas equivalentes [5].

exemple :

Considérons la fonctions $f : X_1 \times X_2 = [0,2] \times [-1,4] \rightarrow \mathbb{R}$ définie par

$$f(x_1, x_2) = x_1^3 - 2x_1x_2^2 + x_2$$

alors l'extention naturelle de f est :

$$F(Y) = F(Y_1, Y_2) = Y_1^3 - 2Y_1Y_2^2 + Y_2 \qquad \forall Y = (Y_1,\ Y_2) \subseteq (X_1,\ X_2)$$

$$\begin{aligned} F(A) &= [0,2]^3 - 2[0,2][-1,4]^2 + [-1,4] \\ &= [-65, 12]. \end{aligned}$$

2.8 Algorithme de Branch and Bound par intervalles :

De nos jour, les algorithmes de Branch and Bound ont pris une place non négligeable dans le monde de l'optimisation globale. Les raisons de cet engouement sont simples, c'est un principe à la fois basique et générale sur le quel peuvent venir se greffe de nombreuses idées pour améliorer la convergence. Dans la suit de notre étude, nous ne nous intéresserons qu'au Branch and Bound basé sur l'analyse d'intervalles, autrement appelé les algorithmes de Branch and Bound par intervalle.

Dans les algorithmes de Branch and Bound l'analyse d'intervalle est utilisée pour calculer des bornes inférieure et supérieure de l'image directe d'une fonction sur un pavé : ie l'image directe de f sur un pavé X est définie par l'intervalle $\left[\min_{x\in X} f(x), \max_{x\in X} f(x)\right]$ et elle est notée $f(X)$. Cet algorithme permet de résoudre des problèmes d'optimisation globale sans contraintes, ces problèmes peuvent s'écrire sous la forme (BP) avec f peut être non-convexe la seul condition est que la fonction objectif possède une expression explicite, pour pouvoir utiliser l'arithmétique d'intervalle, le domaine X est un sous-ensembles convexe de $\mathbb{R}^n$.

$$(\mathbf{BP}) \qquad \begin{cases} \min\limits_{x \epsilon E} f(x) \\ x \epsilon X \end{cases}$$

L'algorithme schématise la structure générale des Branch and Bound par intervalles, si une solution réalisable existe, l'algorithme retourne dans $\widetilde{f}$ un majorant de la valeur du minimum global et dans $\widetilde{z}$ une solution telle que $f(\widetilde{z}) = \widetilde{f}$.

Algorithme de Branch and Bound par Intervalle : **IBBA**

· Soit $X =$ le domaine initial dans le quel le minimum global est recherché.
  X : un pavé dans $\mathbb{R}^n$ en général.
  $\widetilde{f} = +\infty$: désigne le majorant courant du minimum global.
  $L = \{(X, -\infty)\}$: l'initialisation de la solution de données pour stoker les éléments.
· **Répéter**
· extraire de L, l'élément $V = (Z, f_z)$ dont la borne inférieure est la plus petite
· découper V en deux sous-pavé $V_1 = (Z_1, f_{z_1})$, $V_2 = (Z_2, f_{z_2})$
· **Pour** $j = 1$ à 2 **Faire**
  ·procédure d'accélération : monotonie, convexite,...
  ·mise a jour f_{z_j} : un minorant de la fonction objectif sur Z_j
·**Si** $\widetilde{f} \geq f_{z_j}$ **alors**
  ·Insertion (Z_j, f_{z_j}) dans L
  ·$\widetilde{f} = \min(\widetilde{f}, f(midZ_j))$
  ·**Si** $\widetilde{f}$ a été modifie **alors**
    ·$\widetilde{Z} = mid(Z_j)$
    ·Retirer de L tous les éléments (Z, f_z) telque $\widetilde{f} < f_z$
  ·**FinSi**
·**Finsi**
·**Finpour**
·**Jusqu'à** $\widetilde{f} - \min\limits_{(Z,f_z)\in L} f_z < \varepsilon$ oú lorsque $L = \varnothing$

Chapitre 3

Algorithmes d'encadrement de l'optimum global d'une fonction différentiable

l'algorithme étudié ici est un algorithme de type **Branch and Bound** par intervalle. Nous allons utiliser le développements de **Taylor** du premier ordre et le théorème de la valeur moyenne pour construire des hyperplans d'appui de la fonction sur le pavé. En calculant l'intersection de ces hyperplans nous déterminerons un minorant de cette fonction sur ce pavé. Ce minorant est souvent meilleur que ceux engendrés par les diverses fonctions d'inclusion issues l'arithmétique d'intervalle. Nous verrons ensuite des méthodes basées sur la construction d' hyperplans d'appui affines ainsi que le principe des algorithmes permettant de trouver ce minorant des fonctions différentiables de plusieurs variables. Nous testerons leur efficacité au sein de l'algorithme **Branch and Bound,** sur des exemples de minimisation sans contrainte des fonctions à plusieurs variables.

Le problème de minimisation s'écrit :

$$\begin{cases} \min\limits_{x \in X} f(x) \\ x \in \mathbb{R}^n \end{cases}$$

où f est une fonction différentiables de n-variables.

Le principe de cette méthode a déjà été étudiée dans le cas de fonction polynômiale d'une seule variables, par **Jaumard, Hansen,** et **Xiong** , par **Alefed** .et par **Visweswaran** et **Floudas** .**Mladineo** donne, pour les fonctions qui satisfont la conditions **Lipchitz** un procédé basé sur la construction des cônes avec des bases sphériques, òu un système linéaire doit être résolu.

Notre méthode permet de construire un cône polyédrique en résolvant un système linéaire très simple.

3.1 construction des hyperplans :

Dans une première partie nous construirons des hyperplans minorant de la fonction f sur différents pavés de $\not\succ^n$. Nous en déduirons une minoration de la fonction f sur le pavé considéré.

Nous allons étudier dans cette partie des méthodes permettant d'améliorer l'encadrement du minimum global.

D'après la forme de Taylor d'une fonction f , nous obtenons quelque fonctions de l'inclusion de f dans la boite X [1] ,cette fonction d'inclusion est appelée **la forme Taylor :**

$$\forall (x, y) \in X^2, f(y) \in f(x) + (X - x)^T g(X) \qquad (2,1)$$

où

$g(X) = \left(\dfrac{\partial f}{\partial x_1}, ..., \dfrac{\partial f}{\partial x_n} \right)$, est un vecteur de gradient de f, g sera calculée par une méthode de différentiation automatique [2] [10] .

soit $X_i = \left[x_i^L, x_i^U\right]$ et $G_i(x) = \left[g_i^L, g_i^U\right]$ un encadrement du gradient de f sur X , pour $i = 1, 2, ..., n$, donc :

$$\forall\ x \in X, \frac{\partial f}{\partial x_i}(x) \in G_i(x)$$

soit $S = (s_1, ..., s_n)$ les coordonées d'un sommet S de la boite X , et $g^S(X) = \left(g_1^S(X), ..., g_n^S(X)\right)^T$.

pour tout k , on définit le vecteur $g_k^S(X)$ par :

$$g_k^S(X) = \frac{s_k - x_k^U}{x_k^L - x_k^U} g_k^L(X) + \frac{s_k - x_k^L}{x_k^U - x_k^L} g_k^U(X)\ \forall k \in \{1, 2, ...\} \qquad (2,2)$$

On obtient les inégalité suivantes :

$$\left(y_i - x_i^L\right) g_i^L(X) \le \left(y_i - x_i^L\right) \alpha \le \left(y_i - x_i^L\right) g_i^U(X)$$

et

$$\left(y_i - x_i^U\right) g_i^U(X) \le \left(y_i - x_i^U\right) \alpha \le \left(y_i - x_i^U\right) g_i^L(X)$$

Pour tout $i = 1, ..., n$ et tout $\alpha \in G_i(X)$, nous pouvons déduire pour chaque sommet S de la boîte X , un hyperplan d'appui d'une fonction f sur X qui est définit comme suit :

$$f(y) \ge f(s) + (y - s)^T g^S(X)\ , \forall y \in X \qquad (2,3)$$

Donc , nous pouvons obtenir 2^n hyperplans de la fonction f pour la boîte X . chacun de ces hyperplans est un hyperplan secondaire maximal au sommet S pour la boîte X .

On suppose par la suite :

$$g_i^L(X) \cdot g_i^U(X) < 0$$

pour tout i ; cela est obtenu par l'application d'une épreuve de monotonie à f sur X [10] , ceci nous conduit à la construction de l'hyper plan $\not\succ_k$ sous la forme :

$$u_k(x) = f(s_k) + (x - s_k)^T g^{S_k}(X) \qquad (2,4)$$

Proposition1 :

Pour tout y_i appartenant à X_i, les deux minorations sont obtenues aux deux points x_i^L et x_i^U :

$$\bullet \forall x \in (y_i - x_i^L) g_i(X) \ , \ x \geq (y_i - x_i^L) g_i^L(X),$$

$$\bullet \forall x \in (y_i - x_i^U) g_i(X) \ , \ \ x \geq (y_i - x_i^U) g_i^U(X).$$

Ceci bien sûr pour tout i appartenant à $\{1, ..., n\}$

Démonstration :

Pour tout $i = 1, ..., n$, et pour tout $y_i \in X_i$, nous avons par définition de l'opérateur $\times$ de l'arithmétique d'intervalle [5] :

$$\begin{aligned}(y_i - x_i^L) g_i(X) = [\ & \min\{(y_i - x_i^L) g_i^L(X), (y_i - x_i^L) g_i^U(X)\}, \\ & \max\{(y_i - x_i^L) g_i^L(X), (y_i - x_i^L) g_i^U(X)\}] \ ,\end{aligned}$$

or $\left(y_i - x_i^L\right) \geq 0$, donc :

$$\begin{aligned}(y_i - x_i^L) g_i(X) &= (y_i - x_i^L)[\min \left\{g_i^L(X), g_i^U(X)\right\}, \{g_i^L(X), g_i^U(X)\}] \\ &= \left[(y_i - x_i^L) g_i^L(X), (y_i - x_i^L) g_i^U(X)\right].\end{aligned}$$

Nous démontrerions de la même façon que

$$(y_i - x_i^U) g_i(X) = \left[(y_i - x_i^U) g_i^U(X), (y_i - x_i^U) g_i^L(X)\right],$$

parce que $y_i - x_i^U \leq 0$.

3.2 Fonctions minorantes et hyperplans d'appui :

Définition1 :(Fonction minorants sur un pavé)

Soit f une fonction définie de $X \subset \mathbb{R}^n$ dans $\mathbb{R}$, une fonction g définie de X dans $\mathbb{R}$ **minore** f sur le pavé X , si pour tout $x \in X$, $f(x) \geq g(x)$.

Définition2 :(Hyperplan d'appui en un point du pavé)

Un hyperplan $\not>$ défini sur $\mathbb{R}^n$, de fonction affine associée u , est **un hyperplan d'appui** en un point x du pavé X , si et seulement si :

$$\forall y \in X \ , \ f(y) \geq u(y) \text{ et } f(x) = u(x)$$

E^+ sera le demi-espace associé à l'hyperplan d'appui $\not>$, il contient par construction le graphe de f sur le pavé X.

Théorème1 :

En chacun des 2^n sommets de X , un hyperplan d'appui peut étre construit :

$$\forall x \in X, f(x) \geq f(x^S) + f(x - x^S)^T g^S(X) \qquad (2,5)$$

où x^S est un sommet de l'hyperplan X et suivant que $x_i^S = x_i^L$, ou $x_i^S = x_i^U$, on prendra $g_i^S(X) = g_i^L(x)$, ou $g_i^S(X) = g_i^L(x)$, ceci pour tout i appartenant à $\{1, ..., n\}$.

Démonstration :
Par contraction , la fonction $f(x^S) + (x - x^S)^T g(X)$ est une fonction d'inclusion de f sur le pavé X.

$$\forall x \in X \ , \ f(x) \in f(x^S) + (x - x^S)^T g(X),$$

or d'après la proposition 1 , comme x^S est un sommet du pavé,

$$(x_i - x_i^S) g_i(X) \geq (x_i - x_i^S) g_i^S(X) \ , \ \forall i \in \{1, ..., n\}$$

suivant que $x_i^S = x_i^L$, ou $x_i^S = x_i^U$.

Nous obtenons donc , $(x - x^S)^T g(X) \geq (x - x^S)^T g^S(X)$.

Proposition 2 :

Si 0 appartient à $g_i(X)$, ceci pour tout $i \in \{1, ..., n\}$, alors tous les hyperplans construits seront des hyperplans d'appui en chacun des différents sommets.

Propriété 1 :
Chaque fonction de l'affine u_k satisfaites les relations :

$$\begin{aligned} u_k(x^S) &= f(x^S) \\ u_k(x) &\leq f(x) \\ u_k(x^S) &\geq u_k(x) \end{aligned}$$

Déffinition 3 :

Un ensemble de $n+1$ sommets est dit admissible si le sommet c du cône polyédrique défini par $n+1$ hyperplans, construit comme au dessus , satisfaites, $z_c \leq f(x)$, $\forall x \in X$.

3.3 Procédé d'encadrement de l'optimum global pour les fonctions différentiables de n variables :

Dans cette partie , une méthode d'encadrement pour les fonctions différentiables de plusieurs variables est présentée. Elle est basée sur la mise en évidence d'une condition suffisante de minoration. Cette condition nous permettra de choisir l'un des meilleurs simplexes. Cela sera accompli par l'intersection de $n+1$ hyperplans choisi parmi les 2^n hyperplans que nous pouvons construire avec les propriété donnés ci-dessus soit un "point bas", qui sera un minorant de la fonction f sur un pavé considéré. Si cette condition n'est pas vérifiée, alors plusieurs stratégies sont envisageables : intersection avec la frontière du cylindre infini de base le pavé, relaxation des gradients, et élimination de variables. Enfin une méthode recherchant les $n+1$ hyperplans dont l'intersection donne un minorant de f sur le pavé sera exposée.[7], [8].

3.3.1 Principe du procédé d'encadrement :

Soit X un pavé de $\mathbb{R}^n$. Nous noterons $\not\succ_k$ les 2^n hyperplans d'appui construits grâce au théorème 1, u_k les fonctions affines associées et E_k^+ les demi-espaces positifs associés ($k \in \{1, ..., n\}$), où

$$E_k^+ = \{(x, z) \in \mathbb{R}^n \times \mathbb{R},\, z \geq u_k\,(x)\,,\ \forall x \in \mathbb{R}^n\} \qquad (3,1)$$

Pour simplifier, nous supposerons que ces hyperplans d'appui vérifient la propriété1 ; ce qui revient à éliminer les cas de monotonie dans certaines directions[2], [11](ceci sera la première étape de ces algorithmes).

Il en résulte que le graphe $G = \{(x, f(x)); x \in X\}$ de la fonction vérifie :

$$G \subset \bigcap_{k \in K} E_k^+,\, \forall K \subseteq \{1, ..., 2^n\}$$

Propriété2 :

Si $\bigcap_{k \in K} E_k^+$ est un cône polyédrique avec sommet C qui contient le graphique de la fonction f alors , notons $\{C^*\} = \bigcap_{k \in K} \not\succ_k$, et $C^* = (x^*, z^*)$, avec $|K| = n+1$.

Proposition 3 :

Tous les points C issus de l'intersection de $(n+1)$ hyperplans parmi les 2^n hyperplans d'appui, ne sont pas tous des minorants de f sur le pavé X .

Exemple1

Nous n'effectuerons pas une preuve complète de cette décomposition ; pour élucider le problème nous prenderons un exemple à 2 variables.
Considérons le pavé $X = [-5, 5] \times [-15, 10]$, on définit la fonction f par :

$$\begin{aligned} f : X \subset \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x_1, x_2) &\longrightarrow f(x_1, x_2) = x_1^2 - 2x_1x_2 + 3x_1 - 5x_2 \end{aligned}$$

Le gradient de f sur le pavé X est :

$$G(X) = \begin{pmatrix} [-27, 43] \\ [-15, 5] \end{pmatrix}$$

Soit $s_1 = (-5, -15)^T$, $s_2 = (5, -15)^T$, $s_3 = (5, 10)^T$

On note que $(0, 0)$ est dans $G(X)$, donc on est dans un cas non trivial. Considérons les trois hyperplans coorspondant au sommet s_1, s_2, s_3 parmi les quatre possibles :

$u_1(x_1, x_2) = f(-5, -15) - 27(x_1 + 5) - 15(x_2 + 15)$, plan d'apui de f en $(-5, -15)$, noté $\nsucc_1$,vérifiant la propriété 1.

$u_2(x_1, x_2) = f(5, -15) + 43(x_1 - 5) - 15(x_2 + 15)$, plan d'apui de f en $(5, -15)$, noté $\nsucc_2$,vérifiant la propriété 1.

$u_3(x_1, x_2) = f(-5, 10) - 27(x_1 + 5) + 5(x_2 - 10)$, plan d'apui de f en $(-5, 10)$, noté $\nsucc_3$,vérifiant la propriété 1.

Les valeurs de la fonction en chacun des sommets du pavé X sont :

$f(-5, -15) = -65$, $f(5, -15) = 265$, $f(-5, 10) = 60$, $f(5, 10) = -110$.

$\nsucc_1 \cap \nsucc_2$ donne une droite dont sa composante par rapport à la première variable est déterminée par $x_1^* = -3,57$.

$\nsucc_1 \cap \nsucc_3$ donne une droite dont sa composante par rapport à la seconde variable est déterminée par $x_2^* = -15$.

D'où

$z^* = u_1(x^*) = u_2(x^*) = u_3(x^*) = -103,6$. Or $f(5, 10) = -110$.

et donc z^* n'est pas un minorant de f sur X .

3.4 Recherche d'un simplexe admissible :

Nous somme conduit à envisager deux types de simplexes admissibles, dans chacun des deux cas la solution (x^*, z^*) du système est obtenue de façon élémentaire.

On pose donc le problème suivant :

$$(P) \qquad \begin{cases} \quad Minimiser \quad z \\ slc \quad (x,z) \in E_k^+, \forall k \in K \\ \quad x \in X \end{cases} \qquad (4,1)$$

$x \in X$ est assuré parce que la propriété 1 est supposée vérifiée.

Soit e_k la $k^{ième}$ variable d'écart $(e_k \geq 0)$, nous obtenons :

$$z = f(x^{S_k}) + e_k + (x - x^{S_k})^T g^{S_k}(X), \ \forall k \in K \qquad (4,2)$$

Nous pouvons écrire dans l'une des équations :

$$z = z^* + \sum_{k \in} c_k e_k$$

Si C^* est la solution optimale du programme linéaire $(4,1)$, alors (x^*, z^*) satisfait la propriété 2 ; le z^* est un minorant de f sur la boîte X et $x^* \in X$, pour $e_k = 0$, si $c_k \geq 0, \forall k \in K$; l'inégalité $z^* \leq f^*$ (minimum global) est alors assurée.
Donc, la forme standard de problème (4,1) :

$$\begin{cases} \quad Minimize \quad z \\ slc \quad z = e_k + f(x^{S_k}) + (x - x^{S_k})^T g^{S_k}(X), \\ \quad 0 \leq e_k, k = 0,1,2,...;n, \\ \quad (x,z) \in E_k^+, \forall \ k \in K \end{cases} \qquad (4,3)$$

définition 4 :(sommets admissibles)

On dit qu'un ensemble de $n+1$ sommets de X est **admissible** si le sommet C^* du cône simplicial, défini en chacun de ces sommets par un hyperplan d'appui construit grâce au théorème 1, minore f sur un pavé X .Le simplexe non-dégénéré engendré par ces $n+1$ sommets est dit admissible.

3.4.1 Simplexe droit :

La première étude que nous avons effectuée porte sur les simplexes droits ; c'est à dire, on prend un sommet de référence s_0 et tous ces n voisins où les n sommets adjacents d'une boite X (changement d'une seule composante du vecteur représentant le sommet de référence) ; par exemple , pour trois variables soit (x_1^L, x_2^L, x_3^L) le sommet de référence, alors les trois sommets voisins (adjacents) définissant le simplexe droit sont :(x_1^U, x_2^L, x_3^L) , (x_1^L, x_2^U, x_3^L)et (x_1^L, x_2^L, x_3^U), fig.4.1.

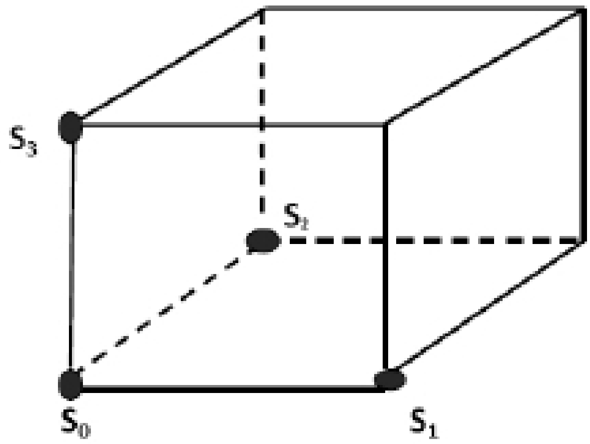

Fig 4,1-Simplexe Droit

L'intérêt d'un tel simplexe est de fournir un système linéaire à matrice diagonale d'où le calcul simple de x^* et z^*.Cependant pour que ce simplexe droit soit admissible, il doit vérifier la condition suffisante d'optimalité donnée par le théorème 2.

Théorème2

soit s_0 un sommet d'une boite X dans $\mathbb{R}^n$, appartir duquel est construit le simplexe droit , ayant pour sommets adjacents $s_{i,i\in\{1,\ldots,n\}}$,qui sont diffère de s_0.

Pour les simplexes droits , les coefficients c_k sont tous positifs par construction sauf un. D'où l'obtention de la **condition suffisante d'optimalité** suivante :

$$\sum_{i=1}^{n} \frac{g_i^{S_0}(X)}{g_i^{S_i}(X) - g_i^{S_0}(X)} \geq -1 \tag{4,4}$$

Où

$$g_i^{S_k}(X) = g_i^L(X) \qquad \text{si} \quad (S_k)_i = x_i^L$$

et

$$g_i^{S_k}(X) = g_i^U(X) \qquad \text{si} \quad (S_k)_i = x_i^U$$

Démonstration :

Notons S_0 sommet à partir duquel est construit le simplexe droit , $S_1, ..., S_n$ les n sommets adjacents de S_0 tel que :

$$x_j^{S_0} = x_j^{S_i}, \forall j \in \{1, ..., n\} \setminus \{i\}$$

et $x_i^{S_0}$ est opposé à $x_j^{S_i}$.

L'intersection de l'équation issue de $(4,2)$:

$$z = f(x^{S_0}) + e_0 + (x - x^{S_0})^T g^{S_0}(X) \qquad (4,5)$$

avec l'une des n autres équations donne : (pour tout $k \in \{1, ..., n\}$)

$$f(x^{S_0}) - f(x^{S_k}) + e_0 - e_k - x_k^{S_0} g_k^{S_0}(X) + x_k^{S_k} g_k^{S_k}(X) + x_k(g_k^{S_0}(X) - g_k^{S_k}(X)) = 0$$

$$\begin{aligned}
\Rightarrow x_k &= \frac{f(x^{S_0}) - f(x^{S_k}) + e_0 - e_k - x_k^{S_0} g_k^{S_0}(X) + x_k^{S_k} g_k^{S_k}(X)}{g_k^{S_k}(X) - g_k^{S_0}(X)} \\
&= x_k^* + \frac{e_0 - e_k}{g_k^{S_k}(X) - g_k^{S_0}(X)}
\end{aligned}$$

En reprenant l'équation $(4,5)$, nous obtenons :

$$z = f\left(x^{S_0}\right) + e_0 + (x - x^{S_0})^T g^{S_0}(X)$$

$$\Leftrightarrow z = z^* + \sum_{k=1}^{n} \frac{e_0 - e_k}{g_k^{S_k}(X) - g_k^{S_0}(X)} g_k^{S_0}(X) + e_0$$

$$\Leftrightarrow z = z^* + \left(1 + \sum_{k=1}^{n} \frac{g_k^{S_0}(X)}{g_k^{S_k}(X) - g_k^{S_0}(X)}\right) e_0 - \sum_{k=1}^{n} \frac{g_k^{S_0}(X)}{g_k^{S_k}(X) - g_k^{S_0}(X)} e_k$$

Donc nous obtenons les équations suivantes :

$$x^* = \frac{f(x^{S_0}) - f(x^{S_k})}{g_k^{S_k}(X) - g_k^{S_0}(X)} + \frac{x_k^{S_k} g_k^{S_k}(X) - x_k^{S_0} g_k^{S_0}(X)}{g_k^{S_k}(X) - g_k^{S_0}(X)} \qquad (4,6)$$

$$z^* = f\left(x^{S_0}\right) + e_0 + (x^* - x^{S_0})^T g^{S_0}(X) \qquad (4,7)$$

Etude du signe des $c_k = \dfrac{g_k^{S_0}(X)}{g_k^{S_k}(X) - g_k^{S_0}(X)}$

- $g_k^{S_k}(X) - g_k^{S_0}(X) < 0 \Rightarrow g_k^{S_k}(X) \leq 0, g_k^{S_0}(X) \geq 0$ et $g_k^{S_k}(X) \neq g_k^{S_0}(X)$ d'où $c_k \geq 0$,

- $g_k^{S_k}(X) - g_k^{S_0}(X) > 0 \Rightarrow g_k^{S_k}(X) \geq 0, g_k^{S_0}(X) \leq 0$ et $g_k^{S_k}(X) \neq g_k^{S_0}(X)$ d'où $c_k \geq 0$.

Tous ces coefficient c_k sont positifs.

Il ne nous reste plus qu'à vérifier :

$$1 + \sum_{k=1}^{n} \frac{g_k^{S_0}(X)}{g_k^{S_k}(X) - g_k^{S_0}(X)} \geq 0$$

cette inéquation devient notre condition suffisante de minoration.

Théorème3 :

Une condition suffisante d'existence d'un simplexe droit permettant de vérifier la condition suffisante de minoration exposée au théorème 2, est donnée par :

$$1 - \sum_{k \in K \setminus \{0\}} \frac{\left|g_k^{S_0}(X)\right|}{W\left(G_i\left(X\right)\right)} \geq 0 \qquad (4,8)$$

Où $\quad w\left(G_i\left(X\right)\right) = g_i^U(X) - g_i^L(X)$.

Proposition 4 :

Une condition suffisante d'existence d'un simplexe droit permettant de vérifier la condition suffisante de minoration exposée au théorème 2 est donnée par :

$$\sum_{k=1}^{n} \frac{\min\left\{g_i^U(X), \left|g_i^L(X)\right|\right\}}{w\left(G_i\left(X\right)\right)} \leq 1 \qquad (4,9)$$

Le sommet engendré par ce choix sera le sommet à partir duquel sera construit le simplexe droit. Si cette inéquation est en plus vérifiée, le simplexe droit est admissible. Sinon nous ne pouvons pas trouver de simplexe droit admissible dans ce cas.

Démonstration :

La démonstration est immédiate. Si cette condition suffisante n'est pas vérifiée, nous ne pouvons pas savoir si l'un des 2^n simplexes droits est admissible.

Lemme 1 :

Pour une fonction de 2 variables la condition suffisante d'existence (proposition 4) est toujours vérifiée.

Démonstration :

Par définition de $w\left(G_i\left(X\right)\right)$, nous avons :

$$\frac{g_i^U(X) - g_i^L(X)}{w(G_i(X))} = 1$$

Comme les hyperplans d'appui construits satisfont la propriété 1 : $g_i^U(X) \geq 0$, $g_i^L(X) \leq 0$ et $g_i^U(X) \neq g_i^L(X)$ d'où l'égalité suivante :

$$\frac{g_i^U(X)}{w(G_i(X))} + \frac{g_i^L(X)}{w(G_i(X))} = 1$$

d'où

$$\frac{\min\left\{g_i^U(X) - \left|g_i^L(X)\right|\right\}}{w(G_i(X))} \leq \frac{1}{2}$$

Donc la somme de ces deux termes est inférieure ou égale à 1, ce qui implique la condition suffisante de minoration est bien vérifiée.

3.4.2 Algorithme d'encadrement de l'optimum utilisant les simplexes droits :

Cet algorithme permet de trouver un **minorant** de f (fonction différentiable) sur un pavé quelconque. Cette méthode sera incluse dans un algorithme d'optimisation de type **Branch and Bound** , pour améliorer les bornes inférieures fournies par les fonctions d'inclusion standards.

Algorithme

!Vérification de la propriété 1 de la condition suffisante d'existance
Calcule du gradient de f : notons le g
$\sum \rightarrow 0$
Pour tout i allant de1 à n faire
Si $g_i^L(X) \geq 0$ Alors
$x^S \leftarrow x_i^L$
Le problème ne dépend plus de cette variabble : insérer i dans un liste L
Si*non* Si $g_i^U(X) \leq 0$ Alors
$x^S \leftarrow x_i^U$
Le problème ne dépend plus de cette variabble : insérer i dans un liste L
Si*non*
Si $\left|g_i^L(X)\right| > g_i^U(X)$ Alors
$g_i^S \leftarrow g_i^U(X)$
$x_i^S \leftarrow x_i^U$
Si*non*
$g_i^S \leftarrow g_i^L(X)$
$x_i^S \leftarrow x_i^L$
Fin Si
$\sum \leftarrow \sum + \dfrac{g_i^S(X)}{w(G_i(X))}$
Fin Si
Fin Pour
Si $\sum \leq 1$ Alors
Calcul de $C^* = (x^*, z^*)$ à partir du simplexe droit construit autour de x^S.
Si*non*
3 possibilités :
– Intersection avec la frontière du cylindre infini de base le pavé X : $X \times \mathbb{R}$,
– Relaxation des gradients pour satisfaire la condition suffisante de minoration,
– Elimination de variables pour satisfaire la condition suffisante de minoration.
Fin Si
Retourne le minorant de f ainsi obtenu.

Le calcul du C^* à partir du simplexe droit de sommet S est relativement aisé,et dans notre cas revient à resoudre un système linéaire diagonale.

Algorithme :

$f^S \to f(x^S)$
$z^* \to f^S$
Pour i allant de 1 à n Faire
Si i n'appartient pas à la liste L Alors
$x^E \to x^S$
Si $x_i^S = x_i^L$ *Alors*
$x_i^E \to x_i^U$
$g^E \to g_i^U$
Si *non*
$x_i^E \to x_i^L$
$g^E \to g_i^L$
Fin Si
!l'hp issu de x^E ne diffère de l'hp issu de x^S que par la constante et le coefficient directeur de x_i
$$x_i^* \leftarrow \frac{f^S - f(x^E) - x_i^S g_i^S(X) + x_i^E g^E}{g^E - g_i^S(X)}$$
$$z^* \leftarrow z^* + \left(x_i^* - x_i^S\right) g_i^S$$
Fin Si
FinPour

Nous verrons dans un prochain paragraphe comment améliorer le calcul de gradients de f par des méthodes de pente, nous en déduirons une nouvelle façon de calculer les coordonnées du point C^*

Méthode de l'intersection avec la frontière du cylindre infini de base de pavé
Cette méthode permet de trouver rapidement une borne inférieure de f sur un pavé X. Si on n'a pas pu vérifier la condition suffisante de minoration alors on effectuera l'intersection des divers hyperplans que nous avons pu construire avec le cylindre infini de base de pavé $X : X \times \mathbb{R}$.

Propriété 3 :

$(n - n_l) + 1$ hyperplan d'appui, ont été construits. La propriété suivante est vérifiée, puisque l'on a éliminé les cas par monotonie :

$$\forall k \in \{1, ..., n\} \setminus L, g_k^L(X) < 0 \text{ et } g_k^U > 0$$

L'intersection de deux de ces hyperplans d'appui ainsi construits, est une fonction minorants de f sur le pavé considéré.

En notant $\not>_0$, l'hyperplan d'appui en S_0 sommet du simplexe droit autour duquel celui-ci est construit.$\not>_k$,$\forall k \in \{1, ..., n\} \setminus L$ sont les hyperplans d'appui des autres sommets du simplexes, tel que :

$$\not>_0 \cap \not>_k \qquad \text{donne} \quad x_k^*$$

La borne inférieure de f sur le pavé est donc :

$$\max_0 \not> \cap \not>_k \cap \mathit{Frantière\ du\ sylindre\ infini}$$

Algorithme

Pour $i\ = 1$ à n Faire
Si i n'appartient pas à L Alors
Détermination direct de x_i^* en effectuant $\not\succ_0 \cap \not\succ_i$
$$z_i^* = f(x^{S_0}) + (x_i^* - x_i^{S_0})g_i^{S_0}(X) - \sum_{k\in\{1,\ldots,n\}\setminus(L\cup\{i\})} w(X_k)|g_k^{S_0}(X)|$$
Fi Si
Fin Pour
Borne inférieure : $\max_{k\in\{1,\ldots,n\}\setminus L}\{z_i^*\}$

Méthode de relaxation des gradients Cette méthode permet d'accroitre l'écart entre les deux bornes du gradient de f sur X de façon à satisfaire la condition suffisante de minoration, et ainsi de pouvoir appliquer le calcul de C^* grâce à l'algorithme 3.4.2.

Algorithme :

Pour $i = 1$ à n Faire
Si i n'appartient pas a la liste L Alors
Si $\dfrac{g_i^S(X)}{d_i} > \dfrac{1}{n - n_l}$ Alors
Si $g_i^S(X) = g_i^L(X)$ Alors
$g_i^U \to (1 - (n - n_l))g_i^L(X)$
Si *non* $g_i^S(X) = g_i^U(X)$
$g_i^L \to (1 - (n - n_l))g_i^U(X)$
Fin Si
Fin Si
FinPour
!on est sur que $\sum \leq 1$

Où n_l est le nombre d'élément de liste L

Nous appliquerons l'algorithme 3.2.4 permettant de calculer C^* avec les nouveaux g_i^L et g_i^U relaxés.

Méthode d'élimination de variables : Cette méthode permet d'éliminer certaines variables du problème, de façon à vérifier la condition suffisante de minoration ; donc l'intersection des hyperplans construits sur le simplexe droit restant sera une **borne inférieure** de f sur le pavé considéré.

Algorithme :

$z^* \leftarrow f(x^S)$
Pour i allant de 1 à n Faire
Si i n'appartient pas à L Alors
Si $\sum > 1$ Alors
chercher le $\dfrac{\left|g_j^S\right|}{d_j}$ maximum dans la liste $\left\{\dfrac{\left|g_1^S\right|}{d_1}, ..., \dfrac{\left|g_n^S\right|}{d_n}\right\}$ sans les k appartenant à L
$z^* \rightarrow z^* - w(X_j)|g_j^S(X)|$
Supprimer $\dfrac{\left|g_j^S\right|}{d_j}$ de la liste
$\sum \rightarrow \sum - \dfrac{\left|g_j^S\right|}{d_j}$
Si *non*
Sortir de la boucle pour $\sum \leq 1$
Fin Si
FinPour
$! \sum \leq 1$
Tant que la liste de $\dfrac{\left|g_j^S\right|}{d_j}$ n'est pas vide Faire
Extraire le premier élément de cette liste, notons le : $\dfrac{\left|g_p^S\right|}{d_p}$
Déterminer l'intersection entre $\not\succ_0$ et $\not\succ_p$ ce qui donne x_p^*
$z^* \rightarrow z^* + (x_p^* - x_p^S)g_p^S(X)$
Fin Tant que

3.4.3 Simplexe relative à deux sommets opposés :

On associe au pavé X , de façon naturelle à partir de ses sommets et de ses arêtes, un graphe orienté symétrique .Le résultat suivant montre alors que , étant donné un sommet quelconque S et son opposé $\overline{S}$ sur le pavé, on peut toujours trouver un chemin de longueur n :$SS_{k_1}S_{k_2}...S_{k_{n-1}}S_{k_n}$,tel que l'ensemble de ces $n+1$ sommets soit admissible, voir figure $(4-2)$.Soit $x^S = (x_1^S, x_2^S, ..., x_n^S)$ un sommet de X.

On note :

$$K_i = -\frac{\left|g_i^S(X)\right|}{g_i^U(X) - g_i^L(X)} \;, i = 1, ..., n \qquad (4,10)$$

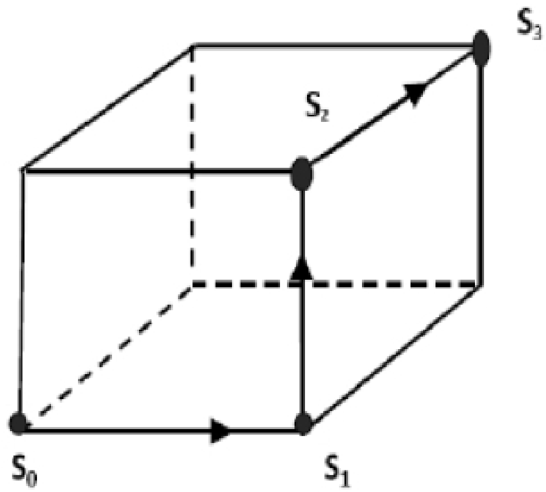

Fig. 3.1 — Recherche d'un simplexes admissible-

Théorème 4 :

Soit S un sommet arbitraire de la boite X, $\overline{S}$ son contraire .

La suite des nombres réels K_i étant classée dans l'ordre croissant

$$-1 < K_{k_1} \leq K_{k_e} \leq ... \leq K_{k_n} < 0 \qquad (4,11)$$

On obtient que l'ensemble des sommets obtenus en modifiant la coordonnée K_1 de S pour obtenir S_{k_1}, puis la coordonnée de K_2 de S_{k_1} et ainsi de suite jusqu'à la coordonnée K_n de $S_{k_{n-1}}$ est admissible. Quand le chemin de S à $\overline{S}$ est définie pour $j = 0, 1, 2, ..., n$ par :

$$\begin{cases} S_i^{k_{j+1}} := S_i^{k_j} \ , i \in \{1, ..., n \setminus k_{j+1}\} \\ S_{k_{j+1}}^{k_{j+1}} := x_{k_{j+1}}^L + x_{k_{j+1}}^U - S_{k_{j+1}}^{k_j} \end{cases} \qquad (4,12)$$

Où $S_{S_0} = S$ et $S_{k_n} = \overline{S}.$

Ce chemin est unique pour les inégalités strictes.

Théorème 5 :

Le problème $(4,1)$ associé aux hyperplans $\not\succ_{K_0}, \not\succ_{k_1}, ..., \not\succ_{k_{n-1}}, \not\succ_{k_n}$, respectivement construits en $S, S_{k_1}, ..., S_{k_{n-1}}, \overline{S}$ admet une solution optimale c^* si et seulement si $(4,11)$ est vérifiée et l'on a l'encadrement recherché

$$z^* \leq f_{\min} \mid X \leq \min \left\{ f\left(x^*\right), f\left(x^S\right), f\left(x^{\bar{S}}\right), f\left(x^{S_j}\right) ; \ j = 1, ..., n-1 \right\} \qquad (4,13)$$

avec

$$x_{k_j}^* = \frac{f\left(x^{S_{k_j}}\right) - f\left(x^{S_{k_{j+1}}}\right) + g_{k_j}^{S_{k_{j+1}}}(X)\, x^{S_{k_{j+1}}} - g_{k_j}^{S_{k_j}}(X)\, x^{S_{k_j}}}{g_{k_j}^{S_{k_{j+1}}}(X) - g_{k_j}^{S_{k_j}}(X)}, \; j = 0, 1, ..., n \qquad (4,14)$$

avec $S_{k_0} = S$, $S_{k_n} = \overline{S}$, et enfin :

$$z^* = f(x^S) + \sum_{l=1}^{n} g_l^S(X)\,(x_l^* - S_l) \qquad (4,15)$$

Démonstration :

On pose

$$d_k = w\left(G_i(X)\right) = g_i^U(X) - g_i^L(X)$$

$$\begin{aligned} U_k &= g_i^U(X) \\ L_k &= g_i^L(X) \end{aligned}$$

D'où

$$d_k = U_k - L_k$$

Le principe de démonstration est le même que pour les simplexes droits, seule change la forme du système d'équations et donc l'expression des couts marginaux qui en découle. Les équations associées à un chemin de longueur n d'un sommet S_0 à son opposé S_n sur le pavé après une renumérotation convenable.

$$\left\{\begin{array}{ll} (\not>_0, S_0) & z = L_1x_1 + L_2x_2 + ... + L_nx_n + c_0 \\ (\not>_1, S_1) & z = U_1x_1 + L_2x_2 + ... + L_nx_n + c_1 \\ (\not>_2, S_2) & z = U_1x_1 + U_2x_2 + ... + L_nx_n + c_2 \\ \vdots & \\ (\not>_n, S_n) & z = U_1x_1 + U_2x_2 + ... + U_nx_n + c_n \end{array}\right.$$

D'où

$$x_{k+1} = \frac{c_k - c_{k+1}}{U_k - L_k}, \qquad k = 1, 2, ..., n$$

D'où

$$z = z^* + e_0 + \sum_{k=1}^{n} L_k \left(\frac{e_{k-1} - e_k}{U_k - L_k}\right)$$

$$z = z^* + \sum_{k=1}^{n} \left(\frac{L_{k+1}}{d_{k+1}} - \frac{L_k}{d_k} \right) e_k + e_0 \left(1 + \frac{L_1}{d_1} \right) + \left(- \frac{L_n}{d_n} \right) e_n$$

D'où les condition d'optimalité de c^* :

$$\begin{aligned} cm_0 &= 1 + \frac{L_1}{d_1} \geq 0 \\ cm_n &= -\frac{L_n}{d_n} \geq 0 \\ cm_k &= \frac{L_{k+1}}{d_{k+1}} - \frac{L_k}{d_k} \geq 0, \quad k = 1, 2, ..., n-1 \end{aligned}$$

Les deux premières conditions sont satisfaites par $g_i^L \leq 0 \leq g_i^U$, et $g_i^L \neq g_i^U$ pour $i = 1, ..., n$.Les suivantes, si et seulement si :

$$\frac{L_1}{d_1} \leq \frac{L_2}{d_2} \leq ... \leq \frac{L_n}{d_n}$$

et alors le chemin $S_1, S_2, ..., S_n$ est admissible.

Pour trouver un chemin admissible d'un sommet S à son opposé $\overline{S}$ il suffit donc de se ramener à cette situation, ce qui est toujours possible.

$$\frac{L_k}{d_k} = \frac{g_k^L}{g_k^U - g_k^L}, \quad \text{si } L_k = g_k^L$$

et

$$\frac{L_k}{d_k} = -\frac{g_k^U}{g_k^U - g_k^L}, \quad \text{si } L_k = g_k^U$$

Alors si et sommet initial est

$$x^S = \left(x_1^S, x_2^S, ..., x_j^S, ..., x_n^S \right)$$

On considère la suite

$$K_j = \frac{g_j^L}{d_j}, \quad \text{si } \; x_j^s = x_j^L$$

et

$$K_j = -\frac{g_j^U}{d_j}, \quad \text{si } x_j^s = x_j^U$$

Que l'on classe dans l'ordre croissant

$$-1 < K_{k_1} \leq K_{k_e} \leq ... \leq K_{k_n} < 0$$

On obtient donc un chemin $SS_{i_1}S_{i_2}...\overline{S}$ admissible en partant du sommet S en modifiant successivement la coordonnée i_1 de S qui donne le sommet S_{i_1} , puis la coordonnée i_2 de S_{i_1} pour obtunir S_{i_2}...jusqu'à la coordonnée i_n de $S_{i_{n-1}}$ qui donne $\overline{S}$.

La coordonnée est modifiée de la façon suivante si $K_{i_p} = \dfrac{g_{i_p}^L}{d_{i_p}}$ alors $x_{i_p}^{S_{i_p}} = x_{i_p}^U$ et $x_{i_p}^U$ sinon.

Le cône simplicial défini à partir des hyperplans d'appui associé aux sommets $S, S_{i_1}, S_{i_2}, ..., \overline{S}$ et construits grâce au théorème 1, convient ; son sommet $c^* = (x^*, z^*)$ vérifier les conditions d'optimalité de $(4,1)$.

Le chemin est unique si les inégalités sont strictes.

Notons que le chemin inverse de $\overline{S}$ à S est obtenu à partir de l'ordre

$$-1 < K_{k_1} \leq K_{k_e} \leq ... \leq K_{k_n} < 0$$

Avec $\overline{K} = -K_j - 1$
Application :

Nous appliquons cette méthode à l'exemple 1 :
Soit $S = (5, -15)^T$ est le sommet initial, $g^S = (43, -15)$ on obtient :

$K_1 = -\dfrac{43}{70}$, $K_2 = -\dfrac{15}{20}$

lequel donne

$$-1 \leq K_2 \leq K_1 < 0$$

Alors un chemin admissible de $S = (5, -15)^T$ à $\overline{S} = (-5, 10)^T$ est obtenu par modifier en premier le x_2 de la coordonnée qui mène au sommet $(5, 10)$ et deuxièmes le x_1 de la coordonnée au quel mène $(-5, 10)$ le sommet terminal. On obtient avec l'hyperplan correspondant $x^* = \left(\dfrac{8}{7}, -5\right)$ et $z^* = -\dfrac{2211}{7} \cong -315,9$.

Nous voyons maintenant pour quoi le simplexe (S_1, S_2, S_3) n'était pas admissible dans cet exemple ; généralement pour une fonction du d'une seule variable sur une boîte X ; deux simplexes droits sont admissibles ; lequel ; dans ce cas, coïncide avec les bornes.

3.4.4 Algorithme de recherche d'un simplexe admissible :

Ce algorithme permet de trouver un **minorant** de f (fonction différentiable) sur un pavé quelconque. Il est bien sûr basé de théorème 5, que nous venons d'exposer. Ceci garantit l'obtention d'un minorant de f sur le pavé X. Pour simplifier , nous présenterons l'algorithme recherchant un chemin optimal entre les deux sommet opposés suivants : $x^L = \left(x_1^L, x_2^L, ..., x_n^L\right)$

et $x^U = \left(x_1^U, x_2^U, ..., x_n^U\right)$, cette méthode sera ensuite incluse dans un algorithme d'optimisation de type **Branch and Bound**, pour améliorer les bornes inférieures fournies par les fonctions d'inclusions standards.

Algorithme

!Vérification de la propriété 1 de la condition suffisante d'existance
Calcule du gradient de f : notons le g
$\sum \to 0$
Pour i =1 à n *faire*
  Si $g_i^U(X) \geq 0$ *Alors*
    $x^S \leftarrow x_i^L$
    Le problème ne dépend plus de cette variabble : insérer i dans un liste L
  Si *non* Si $g_i^U(X) \leq 0$ *Alors*
    $x^S \leftarrow x_i^U$
    Le problème ne dépend plus de cette variabble : insérer i dans un liste L
      Si *non*
        $g_i^S \leftarrow g_i^L(X)$
        $x_i^S \leftarrow x_i^L$
        Insérer $K_i = \dfrac{g_i^S}{d_i}$ dans la liste LK
  Fin Si
FinPour
Trier LK par ordre croissant
! Calcul de $c^* = (x^*, z^*)$ à partir du simplexe admissible.
$z^* \leftarrow f(x^S)$
Tant que LK par ordre croissant
Extraire le premier élément de LK triée, et son indice i
$$x_i^* = \frac{f(x^S) - f(x_1^S, ..., x_i^U, ...x_n^S) - x_i^L g_i^L(X) + x_i^U g_i^U(X)}{d_i}$$
  $x_i^S \leftarrow x_i^U$
  $z^* \leftarrow z^* + (x_i^* - x_i^L) g_i^L(X)$
Fintantque
Retourner le **minorant** z^* de f ainsi obtenu.

3.5 Utilisation de ces méthodes de minoration dans les algorithmes de type Branch an Bound

Les méthodes que nous avons exposées dans les paragraphes suivants s'intègrerant aisément dans les algorithmes d'optimisation globale de type **Branch and Broun**. Aprés avoir exposé l'algorithme de Banch and Bound utilisant ces méthodes d'encadrement pour la résolution de problèmes continus sans contraintes, nous démontrerons de celui-ci.

3.5.1 Algorithme de type Branch and Bound

La recherche d'une borne inférieure de f sur un pavé considéré, s'effectuera en utilisant l'une des trois méthodes précédemment citées. La borne supérieure de l'optimum global, pourra être aussi améliorée, parce que ces méthodes de recherche de bornes inférieures utilisent plusieurs évaluation de f en différents points x^{S_0} et les autres sommets du simplexes droits $x^{S_1}, ..., x^{S_n}$ (en éliminant les sommets non utilisés($\in L$). La borne supérieure sera donc :

$$\widetilde{f} = \min_{i \in \{0,1,\dots,n\} \setminus L} f(x^{S_i})$$

si x^* a été calculé au cours de l'algorithm, on prendra comme borne supérieure :

$$\widetilde{f} = \min\{\widetilde{f}, f(x^*)\}$$

Algorithme :

Encadrement de l'optimum sur $X \Rightarrow (\widetilde{y}, \widetilde{f})$
Insérer $(X, \widetilde{y})$ dans L
Tant Que non Test de Terminaison Faire
    Extraire le premier élément de L :$(Y, \widetilde{y})$
    Partager Y par rapport à sa plus large arête :v
    $V^{(1)} \to Y$ et $V^{(2)} \to Y$
    $V_v^{(1)} \to [y^L, \min(Y)]$ et $V_v^{(2)} \to [\min(Y), y^U]$
    Pour i allant de 1 à 2 Faire
        Encadrer l'optimum global de f sur un pavé $V^i \Rightarrow (z_m, z_M)$
        Si $z_m > \widetilde{f}$ Alors
            Passer au i suivant
        Si *non*
            Si $z_m < \widetilde{f}$ Alors
                $\widetilde{f} \to z_M$
                Eliminer tous les couples $(Z, \widetilde{z})$ de la liste L telque $z > \widetilde{f}$
            Fin Si
            Insérer $(V^{(i)}, z_m)$ dans L dans l'ordre croissant des z_m
        Fin Si
    Fin Pour
Fin Tant Que

3.5.2 Convergence

Consérons la fonction d'inclusion de **Taylor** à l'ordre 1, et notons la T1.

$$\forall\ \overline{x}\ \in X,\ T_1 = f(\overline{x}) + (X + \overline{x})^T g(X)$$

ou g est le gradient de f .

Montrons que $\forall\ Y \subseteq X$, $T_1(Y) \subseteq T_1(X)$, en considérant des développements de Taylor sur les sommets de X et de Y, et en utilisant les propriétés d'inclusion de l'arithmétique d'intervalle.

$$
\begin{array}{rcl}
\text{Soit } Y & \subseteq & X, \\
T_1(Y) & = & f(y^E) + (Y - y^E)^T g(Y), \\
\text{or } f(y^E) & = & f(x^S) + (y^E - x^S)^T g(\varepsilon), \ \text{avec } (y^E, x^x, \varepsilon) \in \mathbb{R}^3, \\
\text{d'où } T_1(Y) & = & f(x^S) + (Y - y^E)^T g(Y) + (y^E - x^S)^T g(\varepsilon), \\
& \subseteq & f(x^S) + (Y - y^E)^T g(Y) + (y^E - x^S)^T g(X), \\
& \subseteq & f(x^S) + (Y - y^E)^T g(X) + (y^E - x^S)^T g(X), \\
& \subseteq & f(x^S) + (Y - y^E)^T g(X) \\
& \subseteq & f(x^S) + (X - y^E)^T g(X) = T^1,
\end{array}
$$

De plus, la fonction d'inclusion T_1 vérifie la propriété de contraction nécessaire à son utilisation dans les algorithmes de **Branch and Bround :**

Quand $\quad W(X) \to 0,\ X - x^S \to 0,$ donc $W(T_1(X)) = w(f(x^S) + (X - x^S)^T g(X)) = 0$

Nous allons maintenant montrer que toutes les bornes inférieure déterminées par l'une de ces trois méthodes est bien supérieure à la borne inférieure déterminée par la fonction d'inclusion de **Taylor** à l'ordre 1.

• **Méthode de l'intersection avec la frontière :**

$$
b_i = \max_{i \in \{1,\ldots,n\}} \ f(x^S) + (x_i^* - x_i^S) g_i^S(x) - \sum_{\substack{j=1 \\ j \neq i}}^{n} w(X_j) \left| g_j^S(X) \right|
$$

Or comme T_1 est calculée en un sommet de X,

$$
T_1^L(X) = f(x^S) - \sum_{j=1}^{n} w(X_j) \left| g_j^S(X) \right|
$$

Donc $\ b_i \geq T_1^L(X)$.

• Méthode d'élimination de variables :

$$
b_i = \ f(x^S) + \sum_{j \in J} (x_i^* - x_i^S) g_i^S(x) - \sum_{\substack{j=1 \\ j \neq i}}^{n} w(X_j) \left| g_j^S(X) \right|
$$

Or comme $T1$ est calculée en un sommet de X,

$$
T_1^L = \ f(x^S) - \sum_{j=1}^{n} w(X_j) \left| g_j^S(X) \right|
$$

Donc $b_i \geq T_1^L(X)$.

• Méthode de relaxation :

Dans sa forme actuelle, nous ne pouvons démontrer la convergence de cette méthode, mais dans tous les tests numériques que nous avons effectués, cette méthode est plus rapide que les deux précédentes.

• Méthode de recherche d'un chemin admissible :

$$b_i = \ f(x^S) + \sum_{j \in J} (x_i^* - x_i^L) g_i^L(x)$$

Donc forcément dans ce cas, $b_i \geq T_1^L(X)$

Les bornes inférieures calculées par ces différentes méthodes sont donc supérieures au bornes inférieures fournies par T_1. Comme la fonction d'inclusion T_1 vérifie les propriétés de convergence de l'algorithme de **Branch and Bound,** nos différentes méthodes convergeront, au pire, de façon équivalente à l'utilisation de T_1.

Pour assurer une convergence à chaque itération, nous pourrons utiliser la nouvelle borne inférieure calculée par l'une des trois méthodes envisagées, que si celle-ci est améliorée.

Chapitre 4

Résultats numériques

Nous aboutissons maintenant à l'étape finale à savoir l'élaboration d'un logiciel aussi convivial que possible, sans perdre de vue son aspect pratique et encore moins sa performance, muni d'une interface claire et accessible, facilitant ainsi son utilisation.

Avant de procéder à la présentation du logiciel, une description de l'environnement de programmation utilisé s'avère nécessaire, nous citons aussi que nous avons utilisé l'outil de programmation : Matlab7

Définition :

Matlab est un environnement de programmation scientifique et de génération de graphisme. Il peut travailler en deux modes différents : immédiate ou diffère(interpréteur ou compilateur).

4.1 Description du langage de programmation Matlab

MATLAB est un logiciel commercial de calcul interactif. Il permet de réaliser des simulations numériques

basées sur des algorithmes d'analyse numérique. Il peut donc être utilisé pour la résolution approchée d'équations différentielles, d'équations aux dérivées partielles ou de systémes linéaires, etc...

La connaissance de ce logiciel est en soi indispensable parce qu'il est de plus en plus utilisé dans l'industrie et les banques pour développer des prototypes de logiciels et tester de nouveaux algorithmes.

La barre de titre :

Comme toutes les fenêtres Windows, la fenêtre MATLAB est surmontée par une barre de titre, contenant à sa gauche une icône et à sa droite les trois boutons « mise en icône », « minimisation/maximisation » et « fermeture ».

La barre de Menu :

La barre de menu en contient 5 (en général)

- File (Fichier) permet d'obtenir l'éditeur de programmes ;
- Edit (Edition) permet de couper / coller dans la ligne de commande et autres ;
- Debug permet l'exécution d'un programme et autres ;

- Window (Fenêtre) permet le passage aux différentes fenêtres du logiciel ;
- Help (Aide) accède au menu d'aide

La barre d'outils :

La barre d'outils en contient 9, qui sont souvent des raccourcis de fonctions contenues dans les menus. De gauche à droite (entre autre) :

- Ouvrir un nouveau fichier dans l'éditeur ;
- Rappeler un ancien fichier dans l'éditeur ;
- Couper ;
- Copier ;
- Coller ;
- Annuler ;
- appeler l'aide.

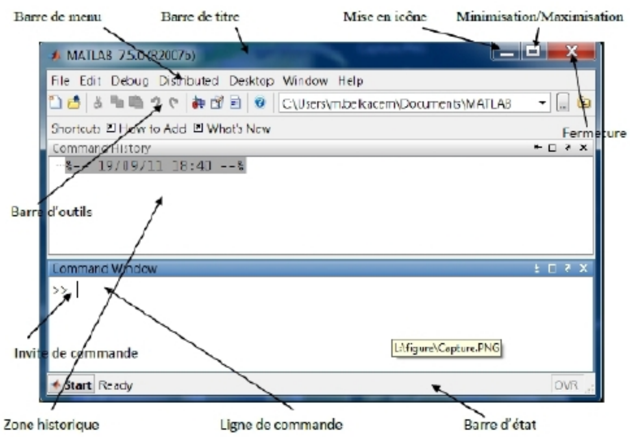

FIG 4.1 : La fentre principale du logiciel MATLAB

La fenêtre de commande :

Elles se divisent en deux zones.

La zone historique, qui ne peut être modifiée, mais dont on peut copier des parties ;

La zone de commande éditable.

La zone de commande permet (comme le nom l'indique) de taper une commande qui sera accepté à l'aide de la touche <return> ou <entrée>■.

L'éditeur de MATLAB : (Script)

Ecrire des scripts, c'est regrouper la suite des instructions que l'on aurait pu taper directement sur la ligne de commande, en un même fichier : « m-fichier ».

Le nom d'un fichier sauvé doit commencer par une lettre et ne contenir que des lettres des chiffres et le souligné (surtout pas de caractères spéciaux(comme *+@ etc.) ni de blanc).

Il doit toujours posséder l'extension « .m », faute de quoi MATLAB ne le reconnaitra pas comme exécutable. S'il s'agit d'une fonction, le nom du fichier (sans extension) doit être identique au nom de la fonction.

Exemple d'utilisation :

Pour cela on ouvre la fenêtre d'édition de fichier (M-File) afin d'écrire le programme comme le montre la figure suivante :

Soit le pavé $X = [-5, 5] \times [-15, 10]$, et soit la fonction réelle de deux variable f définit par :
$f : X \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$

$(x_1, x_2) \longrightarrow f(x_1, x_2) = x_1^2 - 2x_1x_2 + 3x_1 - 5x_2$

Le gradient de f est le suivent :

$g(x_1, x_2) = [g_1(x_1, x_2), g_2(x_1, x_2)] = (2x_1 - 2x_2 + 3\ ,\ -2x_1 - 5)$

On utilise l'analyse d'intervalle pour calculé le gradient de f sur le pavé X.

$$\begin{aligned}
g_1 &= 2x_1 - 2x_2 + 3 \\
&= 2 \times [-5, 5] - 2 \times [-15, 10] + 3 \ = [-10, 10] - [-30, 20] + 3 \\
&= [-10 - 20, 10 - (-30)] + 3 = [-30, 40] + 3 \\
&= [-30 + 3, 40 + 3] \ = [-27, 43] \\
g_2 &= -2x_1 - 5 \\
&= -2 \times [-5, 5] - 5 = [-10, 10] - 5 \\
&= [-10 - 5,\ 10 - 5] \\
&= [-15, 5]
\end{aligned}$$

Donc le gradient de f sur le pavé X est :

$$G(X) = \begin{pmatrix} [-27, 43] \\ [-15, 5] \end{pmatrix}$$

On note que $(0, 0)$ est bien dans $G(X)$, donc on est dans un cas trivial.

Soient S_1, S_2, S_3, S_4 les quatre sommet de pavé X, telque $S_1 = (-5, -15)^T$, $S_2 = (5, -15)^T$, $S_3 = (-5, 10)^T$

$S_4 = (5, 10)^T$,
* soient les gradient correspondant aux sommets :

$g_1 = (-27, -15)^T$, $g_2 = (43, -15)^T$, $g_3 = (-27, 5)^T$, $g_4 = (43, -15)^T$.

* les valeurs de la fonction en chacune des sommets du pavé X sont :

$f(-5, -15) = -65,\ f(5, -15) = 265,\ f(-5, 10) = 60,\ f(5, 10) = -110.$

*Considérons les plans d'appui de f en chaque sommet :

$$u_1(x_1, x_2) = f(-5, -15) - 27(x_1 + 5) - 15(x_2 + 15) = -27x_1 - 15x_2 - 425$$

$$u_2(x_1, x_2) = f(5, -15) + 43(x_1 - 5) - 15(x_2 + 15) = 43x_1 - 15x_2 - 175$$

$$u_3(x_1, x_2) = f(-5, 10) - 27(x_1 + 5) + 5(x_2 - 10) = -27x_1 + 5x_2 - 125$$

$$u_4(x_1, x_2) = f(5, 10) + 43(x_1 - 5) + 5(x_2 - 10) = 43x_1 + 5x_2 - 375$$

Donc on calcul le point $C^* = (x^*, z^*)$, à partir du simplexe droit S. On utilisant l'algorithme de simplexe droit on obtient le programme suivent :

$$\begin{aligned}
X &= [-55; -1510];\\
G &= [-2743; -155];\\
somme &= 0;\\
L &= 1;\\
U &= 2;\\
for\ i &= 1:2,\\
if\ \ G(i, L) &> = 0,\\
XS(i) &= X(i, L);\\
L(i) &= i;\\
& \quad else\\
if\ G(i, U) &< = 0,\\
XS(i) &= X(i, U);\\
L(i) &= i;\\
& \quad else\\
if\ abs(G(i, L)) &> G(i, U)\\
gs(i) &= G(i, U);\\
XS(i) &= X(i, U);\\
& \quad else\\
gs(i) &= G(i, L);\\
XS(i) &= X(i, L);\\
& \quad end\\
somme &= somme + gs(i)/(G(i, U) - G(i, L));\\
& \quad end\\
& \quad end\\
& \quad end\\
& \quad somme\\
if\ \ \ somme &< 1,\\
& \quad XS\\
XS1 &= [XS(1)X(2, 1)]
\end{aligned}$$

$$
\begin{aligned}
XS2 &= [X(1,2)XS(2)] \\
& \quad gs \\
g &= [-27-15;-275;435] \\
f &= [-6560-110] \\
for \quad i &= 1:3, \\
for \quad j &= 1:2, \\
if \quad i &= =1, \\
S(i,j) &= XS1(j); \\
& \quad else \\
if \quad i &= =2, \\
S(i,j) &= XS(j); \\
& \quad else \\
S(i,j) &= XS2(j); \\
& \quad end \\
& \quad end \\
& \quad end \\
& \quad end \\
& \quad S \\
for \quad i &= 1:3, \\
sb &= 0; \\
for \quad j &= 1:2, \\
sb &= sb + S(i,j) * g(i,j); \\
& \quad end \\
b(i) &= sb - f(i); \\
& \quad end \\
& \quad b \\
Sol(2) &= (b(1)-b(2))./(g(1,2)-g(2,2)); \\
Sol(1) &= (b(1)-b(3)-((g(1,2)-g(3,2)))*Sol(2))./(g(1,1)-g(3,1)); \\
& \quad sol \\
& \quad end
\end{aligned}
$$

$$
\begin{aligned}
ZMIN &= (Sol(1))\hat{}2 - 2*(Sol(1)*Sol(2)) + 3*Sol(1) - 5*Sol(2) \\
ZMIN1 &= -27*Sol(1) - 15*Sol(2) - 425 \\
ZMIN2 &= 43*Sol(1) - 15*Sol(2) - 175 \\
Zmin3 &= -27*Sol(1) + 5*Sol(2) - 375
\end{aligned}
$$

et les résultat obtenent par cet programme est donné par :

```
somme =

  -0.1357

XS =

   -5    10

XS1 =

   -5   -15

XS2 =

    5    10

gs =

  -27     5

g =

  -27   -15
  -27     5
   43     5
```

```
   -65    60  -110

S =

    -5   -15
    -5    10
     5    10

b =

   425   125   375

Sol =

    3.5714  -15.0000

ZMIN =

  205.6122

>>
```

Conclusion

Nous nous somme intéressé dans notre travail à résoudre des problèmes d'optimisation globales basée sur l'analyse d'intervalle.

L'algorithme **B&B** par intervalles s'est avéré assez efficace pour les dimensions inférieures à quatre si l'expression de la fonction à optimiser n'est pas trop compliquée. Il est également efficace pour les dimension supérieures à 4 dont les expressions de la fonction objectif et de ses dérivées premières et secondes sont simples. Cependant, ils nécessitent des calculs auxiliaires énormes. le test de monotonie améliore le temps d'exécution si les expression des dérivées ne sont pas coûteuses. Enfin, notons que les algorithmes B&B ont l'avantage de converger avec certitude vers le minimum global. En particulier, ceux basés sur l'analyse des intervalles possèdent des techniques plus efficaces pour extraire les bornes inférieures de la fonction objectif. Ils peuvent être améliorés s'ils sont enrichis par des tests d'accélération ou couplés avec d'autres méthodes d'optimisation globale ou même d'optimisation locale ces dernières permettent d'améliorer la recherche du minimum global.

Nous avons introduit deux méthodes au calcul des bornes inférieurs d'une fonction, considérée sur un pavé, afin d'améliorer la résolution de problème de minimisation a l'aide d'algorithme de type **Branch and Bround** par intervalles. Ces deux méthodes simples d'encadrement du minimum d'une fonction différentiable de n variables sur un pavé : La première est conditionnelle mai on peut toujours choisir le sommet le plus favorable, l'autre est inconditionnelle et permet de partir d'un sommet quelconque, de plus ils ont des avantages qui souvent du problème traité. En général, les méthodes dites du "simplexe droit" ou du "simplexe admissible".

D'un point de vue méthodologie, de nombreux développements sont en cours, telque :

- recherche de nouvelles conditions suffisantes de minoration, sur simplexe ,
- recherches d'un " meilleur" simplexe.

Bibliographie

[1] R.E.Moore.Interval Analysis.Prentice.Hall Inc Englewood Cliffs, N.J.,1966.

[2] E.Hansen. Global optimisation using Interval Analysis. Marcel Dskker, INC.270 Madison Avenue, New York 1OO16, 1992.

[3] J.P.Colosetti. Application de l'Arithmétique d'Intervalle. aux problèmes d'Optimisatio Globale. Technical report, INTP-ENSEEIHT, 2 rue Camichel 31000 Toulouse, 1985.

[4] K.Ichida and Y.Fuju. An Intervalle Arithmetic Méthode for Global Optimization. Computing,23-85-97, 1979

[5] Ramon E. Moore. Intervalle Analysis. Prentice Hall series in Automatic computation, Prentice Hall Internationnal, INC, London, 1966.

[6] J.L. LAGOUANELLE, F. MESSINE Algorithme d'encadrement de l'optimum global d'une fonction differentiable, Compte Rendu à l'Académie des sciences, Numerical Analysis, T.326, Serie I,(198).

[7] RatscheK.H., Rokne.J., New computer Méthodes for Global Optimization, Ellis Horwood timited Market Cross House, Cooper Street, chichester, West Sussec, P019EB, England, 1988.