

UNIVERSITE FERHAT ABBAS SETIF 1

THESE

**Présentée à la Faculté des Sciences
Département de Mathématiques**

Pour l'obtention du diplôme de

Doctorat en-sciences

Option: Mathématiques Appliquées

Par

GUETTAL Djaouida

Thème

**Efficacité et fiabilité des méthodes utilisant l'approche
'Branch-and-Bound' pour l'optimisation globale non convexe.**

Soutenue le 26/06/2014, devant le jury composé de:

Mr. M. ACHACHE	Prof. U. Ferhat Abbas sétif 1	Président
Mr. A. ZIADI	Prof. U. Ferhat Abbas sétif 1	Rapporteur
Mr. L. DJEFFAL	M. C A. U. Hadj Lakhdar de Batna	Examineur
Mr. A. MERZOUGUI	M. C A. U. Mohamed Boudiaf de Msila	Examineur

Remerciements

Au terme de ce travail, je tiens à exprimer mes remerciements envers toutes les personnes qui ont contribué au bon déroulement de cette thèse.

*Je remercie particulièrement, mon encadreur **Abdelkader Ziadi**, Professeur à l'université Ferhat Abbas Sétif 1, pour m'avoir encadré et orienté durant mon travail.*

*Je remercie vivement Monsieur **Achache Mohamed**, Professeur à l'université Ferhat Abbas Sétif 1, pour l'honneur qu'il me fait en acceptant de présider le Jury de cette thèse.*

*J'adresse mes vifs remerciements à monsieur **Merzougui Abdelkrim**, Maitre de conférences à l'université Mohamed Boudiaf de M'sila et monsieur **Djeffal Lakhdar**, Maitre de conférences à l'université Hadj Lakhdar de Batna, d'avoir bien voulu faire parti de ce jury, et d'avoir pris le temps de lire et juger ce travail.*

Mes remerciement s'adressent aussi à Rahal Mohamed, pour l'aide et le soutien qu'il m'a accordé au long de cette thèse.

Je ne peux pas oublier de remercier mes parents, ainsi que tous les autres membres de ma famille, pour leurs soutiens et leurs encouragements.

Table des matières

Introduction	4
1 Optimisation locale et Optimisation globale	9
1.1 Introduction	9
1.2 Généralités sur l'optimisation	10
1.3 Classification des problèmes et des approches	11
1.3.1 Problèmes avec ou sans contraintes	11
1.3.2 Approches déterministes et approches probabilistes	12
1.4 Théorèmes généraux d'existence et d'unicité d'extremum global	14
1.5 Quelques algorithmes efficaces d'optimisation locale	14
1.5.1 Algorithme général des méthodes de descentes	15
1.5.2 Méthodes de Quasi-Newton	15
1.6 Algorithmes déterministes d'optimisation globale unidimensionnelle	16
1.6.1 Méthodes d'optimisation globale basées sur des techniques de recherche locale	16
1.7 Méthodes de recouvrement	18
1.7.1 Principe général	19
1.7.2 Algorithme de Piyavskii-Shubert unidimensionnel ($n = 1$)	21
1.7.3 Différentes extensions de la méthode de Piyavskii-Shubert ($n \geq 1$)	25
1.7.4 Méthodes utilisant des fonctions minorantes de classe C^2	25
1.7.5 La méthode DIRECT unidimensionnelle	27
1.7.6 Structure générale de la méthode Branch-and-Bound	32
1.8 Conclusion	33

2	Optimisation globale sans contraintes des fonctions lipschitziennes à plusieurs variables	35
2.1	Introduction	35
2.2	La méthode DIRECT multidimensionnelle	36
2.2.1	Division dans les dimensions supérieures	37
2.2.2	Hyperrectangles potentiellement optimaux	38
2.3	La méthode Branch-and-Bound multidimensionnelle et l'arithmétique par intervalles	39
2.3.1	Calcul par intervalles	40
2.3.2	Forces et faiblesses de l'arithmétique par intervalles	43
2.3.3	Un algorithme prototype de Branch-and-Bound par intervalles	45
2.4	Algorithme de Piyavskii généralisé	47
2.4.1	Description de l'algorithme	47
2.4.2	Convergence de l'algorithme	49
2.5	La méthode de la transformation réductrice Aliénor	50
2.5.1	Présentation de la méthode	50
2.5.2	Application de la méthode Aliénor à l'optimisation globale	55
2.6	La méthode hybride Branch-and-Bound-Piyavskii	57
2.6.1	Description de la méthode	57
2.6.2	Convergence de la méthode	61
2.7	La méthode hybride Aliénor-Piyavskii	63
2.7.1	Présentation de la méthode	63
2.7.2	Convergence de la méthode	65
2.8	Applications numériques	66
2.8.1	Etude comparative	69
3	Problèmes d'optimisation globale lipschitziennes avec contraintes	71
3.1	Introduction	71
3.2	La méthode de la transformation réductrice et l'optimisation globale avec contraintes	73
3.3	Construction des courbes α -denses dans un compact de \mathbb{R}^n	74
3.3.1	Construction des courbes α -denses à partir de certaines classes de fonctions	74

3.3.2	Construction des courbes α -denses à partir des fonctions périodiques . . .	83
3.3.3	Génération concrète des courbes α -denses dans des compacts de \mathbb{R}^n . . .	85
3.3.4	Exemple de courbes α -dense	88
3.4	Transformation à une dimension de quelques problèmes d'optimisation globale avec contraintes par Alienor	90
3.4.1	Quelques problèmes tests	90
3.4.2	Résultats numériques	93
3.5	Conclusion	94
	Conclusion générale	95
	Bibliographie	98

Introduction

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. Les algorithmes d'optimisation ont été développés dans pratiquement plusieurs problèmes économiques et industriels dans le but de maximiser ou minimiser toutes sortes de choses ; par exemple, maximiser les profits tout en minimisant les pertes, améliorer si possible de façon optimale certains processus de fabrication ou les fonctionnalités de certains objets ou produits. Ainsi, l'optimisation entre en jeu dans beaucoup de domaines scientifiques : conception de moteurs électriques, changement d'orbite d'un satellite, météo, biologie-mathématique, génie des procédés chimiques...

Le thème de l'optimisation globale étudie la question de trouver les extrema (locaux et globaux), d'une fonction d'une ou de plusieurs variables. Il n'y a pas si longtemps, mettre au point des méthodes numériques permettant de déterminer la solution d'un problème d'optimisation non linéaire, non convexe et non différentiable pouvant répondre à un ensemble de contraintes, elles aussi non linéaires et non convexes, paraissait très difficile. Bien que la théorie classique de l'optimisation ne peut pas être appliquée directement dans les problèmes d'optimisation globale, les outils traditionnels tels que l'analyse convexe, sont largement utilisés dans la construction des méthodes d'optimisation globale. Cette approche constitue une partie importante de l'optimisation globale déterministe. Par exemple, un remarquable progrès a été accompli dans la construction des algorithmes de minimisation des fonctions concaves dans des régions convexes, et aussi la minimisation des fonctions différence de deux fonctions convexes.

Ces dernières années, de nombreux travaux ont été réalisés. On peut les classer en deux grandes familles :

Les approches déterministes, comme leur nom l'indique, nous offrent la certitude d'obtenir l'optimum global et ne laissent aucune place au hasard et conduiront pour un contexte initial donné à une même solution finale. Pour ces méthodes, l'exploration de l'espace de solutions se fait grâce à des procédures de recherche qui sont élaborées à partir de la constante de Lipschitz ou celle de Hölder, des dérivées ou d'autres informations locales et globales concernant la fonction objectif. La théorie mathématique de l'optimisation globale est assez développée et possède de nombreuses applications importantes. Récemment, une série de travaux de Floudas, Horst et

Tuy [19], Strongin et Sergeyev [46], Pardalos [20] et Thoai [21], etc., sur l'optimisation globale a été publiée.

Dans le cas où la fonction objectif est donnée sous forme de "boite noire" le problème d'optimisation est particulièrement difficile. Les modèles déterministes ne traitent pas adéquatement les informations disponibles sur la fonction objectif. Les approches stochastiques peuvent souvent faire face à ce genre de problèmes plus facilement et plus efficacement que les algorithmes déterministes. Les algorithmes stochastiques explorent l'espace de solutions grâce en partie à des procédures de transitions aléatoires. Ainsi, plusieurs exécutions successives de ces algorithmes, pourront conduire à des résultats différents (pour un même point initial). L'avantage des méthodes stochastiques est leur simplicité et leur pertinence pour les problèmes où les évaluations de la fonction objectif sont corrompues par un bruit aléatoires, ainsi que leur robustesse à l'égard de la croissance de la dimension. L'inconvénient majeur de ces méthodes est qu'elles peuvent diverger, et passer plusieurs fois à côté de la solution. L'obtention de l'optimum global n'est pas garantie, il est seulement repéré avec une probabilité proche de 1.

De nombreux algorithmes, où l'aléatoire et l'argument statistique sont impliqués ont été proposés heuristiquement. Ces algorithmes sont basés sur des analogies avec les processus naturels. Les exemples les plus connus de tels algorithmes sont l'optimisation évolutive, le recuit simulé, l'algorithme génétique et la recherche tabou. Les algorithmes heuristiques d'optimisation globale sont très populaires dans les applications. Le recuit simulé, étant méthode importante, a été intensivement étudié par plusieurs chercheurs.

Dans cette thèse, on s'intéresse à quelques méthodes déterministes multidimensionnelles. Nous nous sommes concentrés sur la résolution de problèmes explicites, c'est-à-dire des problèmes dans lesquels toutes les expressions des équations de la fonction objectif et des contraintes sont connues de façon explicite.

Le premier chapitre de notre étude commence par un rappel de quelques algorithmes d'optimisation locale, ensuite nous présenterons quelques méthodes d'optimisation globale basées sur l'utilisation des techniques de recherche locale (multistart, tenneling, clustering,...). Notre attention sera portée sur les méthodes de recouvrement qui ont la réputation d'être efficaces en dimension 1. Parmi ces méthodes, il y a celles qui sont basées sur la construction des fonctions minorantes de la fonction objectif sur le domaine faisable (Piyavskii-Shubert, Direct,

Brent,...), et d'autres sur les techniques de partition et d'élimination (**Branch-and-Bound**.

Dans le second chapitre nous étudierons la généralisation de certaines méthodes de recouvrement présentées dans le premier chapitre, au cas des problèmes d'optimisation globale des fonctions lipschitziennes à plusieurs variables, définies sur un pavé $P = \prod_{i=1}^n [a_i, b_i]$ de \mathbb{R}^n et à valeurs dans \mathbb{R} . Dans ce chapitre nous donnerons une nouvelle approche en matière d'optimisation globale déterministe basée sur la technique de réduction de la dimension du problème et qui est connue sous le nom de méthode de la transformation réductrice **Alienor**. La méthode **Alienor** a été élaborée par Cherruault et coll. [7], [8], [9], [56]. L'idée consiste à approcher la fonction objectif à n variables ($n \geq 2$) définie sur le pavé P , par une fonction d'une seule variable $\theta \geq 0$, en densifiant le pavé P de manière aussi précise que l'on veut, à l'aide d'une courbe dite α -dense, assez régulière. Ceci permet de ramener le problème d'optimisation multidimensionnel à un problème d'optimisation unidimensionnel, pour pouvoir utiliser les méthodes unidimensionnelles connues par leur efficacité et performance. Dans cette partie nous présentons d'abord la méthode **Alienor** de base [6] et nous étudions le couplage de cette méthode avec l'algorithme unidimensionnel de **Piyavskii-Shubert**. Ensuite, nous étudions le couplage de la méthode de **Piyavskii** multidimensionnelle avec celle de **Branch-and-Bound** (qui procède à une minimisation unidimensionnelle sur la diagonale principale du pavé P). Nous verrons que les méthodes mixtes offrent des perspectives intéressantes pour diminuer le temps de calcul. Enfin, nous terminerons ce chapitre par une série d'applications numériques sur des algorithmes mixtes tirés de la littérature, suivie d'une étude comparative sur les deux algorithmes mixtes appliqués aux problèmes posés.

Dans le troisième chapitre nous étudions quelques problèmes d'optimisation globale multidimensionnelle avec contraintes [18], [19], [36] de la forme suivante :

$$\left\{ \begin{array}{l} \min f(x), \quad x \in \mathbb{R}^n \\ \text{sujet à } g_i(x) \leq 0, \\ \text{pour } i = 1, \dots, m \end{array} \right. \quad (\mathbf{P})$$

où f et $g_i(x)$ sont des fonctions lipschitziennes.

L'existence de multiples minima locaux de la fonction objectif à plusieurs variables rend les problèmes d'optimisation globale avec contraintes très difficile à résoudre.

Dans ce chapitre nous allons étudier l'application de la méthode de la transformation réductrice **Alienor** pour résoudre quelques classes de problèmes d'optimisation avec contraintes. Dans la première partie nous présentons des résultats originaux [15] sur la construction des courbes α -denses dans des compacts de \mathbb{R}^n à partir d'une certaines classes de fonctions et à partir des fonctions périodiques. Nous donnons des conditions suffisantes sur les coordonnées de la paramétrisation des courbes α -denses pour obtenir de nouvelles familles de courbes d'expression analytique plus simple. Ceci permet d'améliorer le temps de calcul dans la méthode **Alienor** pour résoudre le problème **(P)** d'optimisation globale avec contraintes.

Nous proposons quelques exemples de courbes α -denses puis voir comment ramener le problème avec contraintes **(P)** à un problème d'optimisation globale unidimensionnel de la forme

$$\min_{\theta \in [0, \theta_{\max}]} f^*(\theta), \quad (\mathbf{P}^*)$$

Avec $f^*(\theta) = f(h(\theta))$, où $h(\theta)$ est la courbe α -dense.

Enfin, nous appliquons l'algorithme unidimensionnel de **Piyavskii-Shubert** pour résoudre le problème **(P*)** et nous terminerons notre travail par une série d'applications numériques suivi d'une conclusion générale sur les différentes approches et méthodes utilisées en optimisation globale.

Notations

\mathbb{R}	espace des nombres réels
\mathbb{R}^n	espace euclidien de dimension n
$\ \cdot\ $	la norme euclidienne dans \mathbb{R}^n
$Ent(\cdot)$	partie entière d'un nombre réel
$\langle \cdot, \cdot \rangle$	produit scalaire
$f(\cdot)$	fonction objectif
x^*	$x^* \in \arg \min f(x^*)$
$\arg \min f(D)$	l'ensemble des minimiseurs globaux de f sur D
$\text{globmin } f(D)$	le minimum global de f sur l'ensemble D
$\mu(\cdot)$	la mesure de Lebesgue sur \mathbb{R}
$\mathbf{d}(x,y)$	la distance entre x et y
$V(x,r)$	la boule (dans un espace métrique) de centre x et de rayon r
C^2	l'espace des fonctions 2 fois continûment différentiables
∇f	le gradient de la fonction f
$\nabla^2 f$	le hessien de la fonction f
$V(x^*, r)$	une boule ouverte de centre x^* et de rayon r
$C \subset \mathbb{R}^n$	un ensemble fermé et borné
$\overline{\text{co}(C)}$	la fermeture de l'enveloppe convexe de C

Chapitre 1

Optimisation locale et Optimisation globale

1.1 Introduction

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction continue.

La programmation non linéaire ou optimisation (locale) non linéaire consiste en un ensemble de méthodes pour trouver un optimum local (minimum ou maximum). Mais comme maximiser f revient à minimiser $-f$, nous considérons donc tout problème d'optimisation comme un problème de minimisation c'est-à-dire :

$$\text{Pr b.}_{loc} \left\{ \begin{array}{l} \text{Trouver un point } x_{loc} \text{ de } \mathbb{R}^n \text{ tel qu'il existe un voisinage } V \text{ de } x_{loc} \text{ avec :} \\ f(x_{loc}) \leq f(x), \quad \forall x \in V \end{array} \right.$$

Cependant en pratique, il existe, en général, plusieurs minimiseurs locaux avec des valeurs de fonction différentes. Ainsi le problème de l'optimisation globale se veut que :

$$\text{Pr b.}_{glb} \left\{ \begin{array}{l} \text{Trouver un point } x_{glb} \text{ de } \mathbb{R}^n \text{ tel que :} \\ f(x_{glb}) \leq f(x), \quad \forall x \in \mathbb{R}^n \end{array} \right.$$

Nous nous intéressons dans cette thèse à la résolution numérique de ce problème.

Pour des raisons numériques on suppose l'existence d'un compact C contenant au moins une

solution x_{glb} , et le problème global devient celui que nous allons considérer dans toute la suite, soit :

$$(\mathbf{P}) \left\{ \begin{array}{l} \text{Trouver } x_{glb} \text{ de } C \text{ tel que :} \\ f(x_{glb}) \leq f(x), \quad \forall x \in C \end{array} \right.$$

1.2 Généralités sur l'optimisation

Dans toute la thèse, on note par (x_1, x_2, \dots, x_n) un vecteur de \mathbb{R}^n , $(x^k)_{k \in \mathbb{N}}$ une suite de vecteurs de \mathbb{R}^n et par $\|\cdot\|$ la norme euclidienne dans \mathbb{R}^n . Considérons le problème général d'optimisation :

$$\min_{(x_1, x_2, \dots, x_n) \in C} f(x_1, x_2, \dots, x_n) \quad (\mathbf{P})$$

Pour lequel on doit trouver un point $x^* = (x_1^*, x_2^*, \dots, x_n^*) \in C$, réalisant le minimum de f , ce que nous écrirons :

$$\min_{(x_1, x_2, \dots, x_n) \in C} f(x_1, x_2, \dots, x_n) = f(x_1^*, x_2^*, \dots, x_n^*)$$

où $f : C \rightarrow \mathbb{R}$ une fonction à n variables réelles x_1, x_2, \dots, x_n , continue sur C , appelée fonction objectif et C une partie non-vide de l'espace \mathbb{R}^n munie de la norme euclidienne, appelée l'ensemble faisable.

Définition 1.1 (minimum local)

$$\left\{ \begin{array}{l} \text{On dira que } x_{loc} \text{ de } \mathbb{R}^n \text{ est un minimiseur local de } f \text{ si} \\ \exists V \text{ un voisinage de } x_{loc} \text{ tel que } f(x_{loc}) \leq f(x), \quad \forall x \in V. \end{array} \right.$$

Rappelons qu'un voisinage $V(x^*, r)$ d'un point x^* de \mathbb{R}^n , défini comme une boule ouverte par :

$$V(x^*, r) = \{x \in \mathbb{R}^n : \|x - x^*\| < r\}$$

de centre x^* et de rayon r .

Définition 1.2 (minimum global)

$$\left\{ \begin{array}{l} \text{On dira que } x_{glb} \text{ de } \mathbb{R}^n \text{ est un minimiseur global de } f \text{ si} \\ f(x_{glb}) \leq f(x), \quad \forall x \in \mathbb{R}^n. \end{array} \right.$$

Il est clair qu'un minimum global est un minimum local, mais la réciproque n'est pas en général vraie.

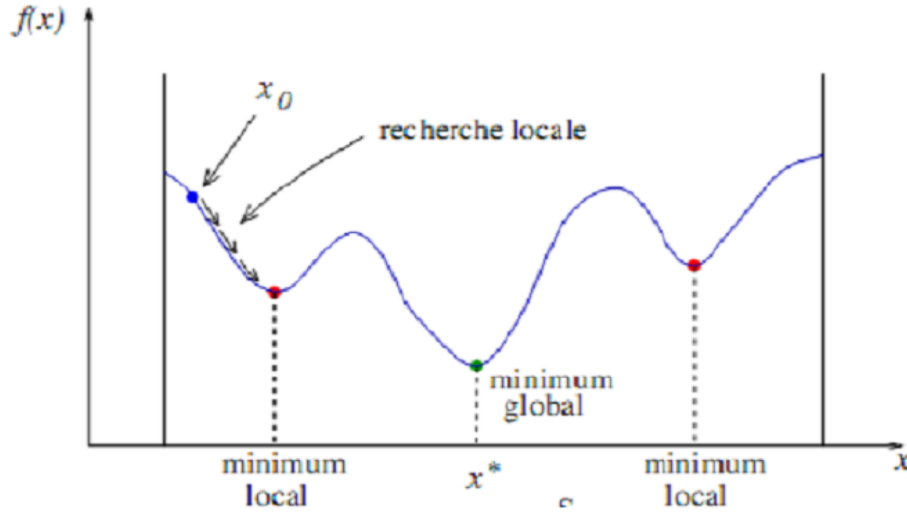


Fig.1 Minima locaux et minimum global d'une fonction à une variable

1.3 Classification des problèmes et des approches

1.3.1 Problèmes avec ou sans contraintes

Une classification naturelle des problèmes d'optimisation globale serait de distinguer les problèmes sans contraintes (P_{sc}) et ceux avec contraintes (P_c). Les problèmes sans contraintes, ce sont ceux dont l'ensemble faisable est $C = \mathbb{R}^n$. Cependant, sont aussi considérés essentiellement sans contraintes les problèmes dont les contraintes spécifient seulement la région d'intérêt pour l'exploration globale mais ne sont pas des contraintes importantes; comme par exemple le cas où C est un pavé de \mathbb{R}^n , i.e., $C = \prod_{i=1}^n [a_i, b_i]$ avec $(a_1, a_2, \dots, a_n) \in \mathbb{R}^n$ et $(b_1, b_2, \dots, b_n) \in \mathbb{R}^n$. Si $C \neq \mathbb{R}^n$, le problème (**P**) est dit problème de minimisation avec contraintes. Dans le cas $C = \mathbb{R}^n$, et pour que le problème (**P**) admette un minimum global, il faut imposer une hypothèse

supplémentaire sur f dite condition de croissance à l'infini, connue sous le nom de la coercivité, c'est-à-dire $f(x) \xrightarrow{\|x\| \rightarrow +\infty} +\infty$ où $\|x\|$ désigne la norme euclidienne de x dans \mathbb{R}^n . Notons que cette hypothèse implique l'existence d'un ensemble robuste compact $C' \subset \mathbb{R}^n$, tel que C' contient tous les minimiseurs globaux de f , et la valeur de $f(x)$ quand x est sur la frontière de C' est plus grande que la valeur de $f(x)$ pour tout point x à l'intérieur de C' . Rappelons qu'un sous ensemble fermé $C' \subset \mathbb{R}^n$ est dit robuste s'il est la fermeture d'un ensemble ouvert $U \neq \emptyset$ de \mathbb{R}^n , c'est-à-dire $C' = \overline{U}$.

Numériquement, il est plus aisé de résoudre les problèmes sans contraintes et cela, quelle que soit la nature de la méthode utilisée. En effet, le domaine (des variables) de recherche n'étant pas restreint, on a moins de soucis à se faire que dans les cas des problèmes avec contraintes où il faut tout mettre en œuvre pour s'assurer à chaque itération que l'on reste dans le domaine faisable ou sinon s'imposer un moyen d'y revenir à la limite du domaine.

1.3.2 Approches déterministes et approches probabilistes

Différentes classifications des méthodes d'optimisation globale sont trouvées dans la littérature. Une classification, basée sur [19], est présentée dans cette section. Elle n'a pas pour objectif d'être complète, ni de montrer tous les détails des méthodes, mais plutôt de mettre en évidence les caractéristiques de chaque classe.

Approches déterministes

Dans ce type de méthodes, l'aléatoire n'intervient pas, c'est-à-dire que pour résoudre un problème, l'algorithme se comportera toujours de la même façon et donnera toujours la même réponse. Ces algorithmes peuvent se classer en fonction du type de problèmes pouvant être résolu les programmes linéaires, les problèmes convexes, quadratiques, polynomiaux ou plus généraux.

Ces techniques ont généralement l'avantage de ne pas nécessiter de point de départ. Mais avant tout elles fournissent une réponse déterminante sur la qualité des solutions trouvées : l'optimum est-il local ou global? Quel est le degré de certitude? etc. Cette précision a une importance significative, car il est souvent beaucoup moins coûteux de trouver une solution que de prouver qu'il s'agit bien de l'optimum global.

Pour ces méthodes, l'exploration de l'espace des solutions se fait grâce à des procédures de recherche qui sont élaborées à partir de la constante de Lipschitz ou celle de Hölder, des dérivées ou d'autres informations locales et globales concernant la fonction objectif. La théorie mathématique de l'optimisation globale est assez développée et possède de nombreuses applications importantes. Récemment, une série de travaux de Floudas, Horst et Tuy [19], Strongin et Sergeyev [46], Pardalos [20] et Thoai [21], etc. sur l'optimisation globale a été publiée.

Dans le cas où la fonction objectif est donnée sous forme de "boite noire" le problème d'optimisation est particulièrement difficile. Les modèles déterministes ne traitent pas adéquatement les informations disponibles sur la fonction objectif.

Approches stochastiques

Les approches stochastiques peuvent souvent faire face à ce genre de problèmes plus facilement et plus efficacement que les algorithmes déterministes. Les algorithmes stochastiques explorent l'espace des solutions grâce en partie à des procédures de transitions aléatoires. Ainsi, plusieurs exécutions successives de ces algorithmes, pourront conduire à des résultats différents (pour un même point initial). L'avantage des méthodes stochastiques est leur simplicité et leur pertinence pour les problèmes où les évaluations de la fonction objectif sont corrompues par un bruit aléatoire, ainsi que leur robustesse à l'égard de la croissance de la dimension. L'inconvénient majeur de ces méthodes est qu'elles peuvent diverger, et passer plusieurs fois à côté de la solution. L'obtention de l'optimum global n'est pas garantie, il est seulement repéré avec une probabilité proche de 1.

De nombreux algorithmes, où l'aléatoire et l'argument statistique sont impliqués ont été proposés heuristiquement. Ces algorithmes sont basés sur des analogies avec les processus naturels. Les exemples les plus connus de tels algorithmes sont l'optimisation évolutive, le recuit simulé, l'algorithme génétique et la recherche tabou. Les algorithmes heuristiques d'optimisation globale sont très populaires dans les applications. Le recuit simulé, comme étant méthode importante, a été intensivement étudié par plusieurs chercheurs.

1.4 Théorèmes généraux d'existence et d'unicité d'extremum global

Théorème 1.1 (Weirstrass) *Si C est un ensemble compact et f une fonction continue alors l'ensemble des solutions du problème **(P)** est non vide (en d'autres termes, f atteint son minimum global dans C).*

Le second résultat est moins général car il considère le cas particulier $C = \mathbb{R}^n$.

Théorème 1.2 *Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction continue sur \mathbb{R}^n , si $\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$, alors le problème **(P)** admet une solution globale x^* .*

Démonstration. Puisque $\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$, alors par définition, $\forall T > 0, \exists M > 0$, tel que : $\|x\| > M \Rightarrow f(x) > T$. Soit $x_0 \in C$. On pose $T = f(x_0)$, alors $\exists M > 0$ tel que on a $\|x\| > M \Rightarrow f(x) > f(x_0)$, donc $\exists M > 0, f(x) \leq f(x_0) \Rightarrow \|x\| \leq M$. Puisque x^* est caractérisé par $f(x^*) \leq f(x), \forall x \in \mathbb{R}^n$, on a donc forcément $\|x^*\| \leq M$. Donc x^* est solution du problème :

$$\min_{\|x\| \leq M} f(x),$$

et le théorème de Weirstrass s'applique, la boule $\{x \in \mathbb{R}^n : \|x\| \leq M\}$ étant compacte. C'est-à-dire x^* est un minimiseur global de f .

L'unicité résulte en général des propriétés de convexité (de f et de C).

1.5 Quelques algorithmes efficaces d'optimisation locale

Bien que l'optimisation locale ne soit pas le thème essentiel de notre travail, nous présentons les grandes lignes. En effet comme nous le verrons après, beaucoup d'algorithmes d'optimisation globale utilisent des algorithmes d'optimisation locale comme sous-procédures. Donc l'efficacité de ces derniers affecte celle des premiers. Il est utile donc de montrer qu'il y a dans le domaine de l'optimisation locale des algorithmes efficaces. Citons les algorithmes des initialisations multiples, l'algorithme de Tunnelier (cf. la section 1.5).

Nous nous intéressons ici à l'aspect numérique de la résolution du Prb._{loc} . Pour ce faire, on construit une suite itérative $x_1, x_2, \dots, x_n, \dots$ de façon à trouver un point x_{loc} solution de Prb._{loc} .

1.5.1 Algorithme général des méthodes de descentes

On dira que la direction d_k est une direction de descente si $f(x_k + t_k \cdot d_k) < f(x_k)$ pour un pas t_k donné. Les méthodes de descentes sont celles où les t_k et d_k sont supposés être ceux qui donnent une direction de descente et un pas optimal. Ce qui donne :

Initialisation de x_0, t_0, d_0 (point initial, pas, direction)

Tant que Test d'arrêt faux faire

calcul de la direction (de descente) d_k

calcul du pas (optimal) t_k

$$x_{k+1} = x_k + t_k \cdot d_k$$

Fin (de tant que)

$(x_k, f(x_k))$ est solution

t_k et d_k sont appelés respectivement pas et direction d'avancée. Pour les calculer, l'algorithme peut utiliser des informations sur les valeurs du gradient (voir du Hessien) de f au point x_k courant.

1.5.2 Méthodes de Quasi-Newton

Ces méthodes sont celles qui sont les plus utilisées en optimisation locale différentiable.

On note par $g_k = \nabla f(x_k)$ le vecteur gradient de f en x_k

$$H_k = H(x_k) \quad \text{une matrice } N \times N \text{ symétrique définie positive}$$

Algorithme

Initialisation $\epsilon, \delta, x_0, H_0$:

si $|g_k| < \epsilon$ stop,

sinon poser $\delta_k = -H_k * g_k$;

trouver (en partant de $\lambda = 1$) λ_k qui minimise la fonction de $\lambda : f(x_k + \lambda \cdot d_k)$;

poser alors $V_k = \lambda_k d_k$

$$x_{k+1} = x_k + V_k$$

calculer g_{k+1} ;

$$\text{faire } U_k = g_{k+1} - g_k$$

réactualiser H_{k+1} (en utilisant H_k, V_k et U_k voir ci-dessous) :

calcul de H_k dans le cas général

$H_{k+1} = H_k + B_k$ de façon que :

- L'équation de Quasi-Newton, $H_{k+1}U_k = V_k$, soit satisfaite pour tout k ;
- H_k soit symétrique définie positive pour tout k ;
- B_k soit minimale en un certain sens.

Ce qui laisse beaucoup de choix possibles.

1.6 Algorithmes déterministes d'optimisation globale unidimensionnelle

En optimisation locale, il existe des algorithmes efficaces pour résoudre pratiquement tous types de problèmes : des procédures de recherche directe, indirecte, avec ou sans calcul de gradients, etc...

Cependant, il arrive qu'un utilisateur, en exécutant plusieurs fois de suite avec des points initiaux différents, tel ou tel procédure, se rende compte qu'il trouve plusieurs minimums dont les valeurs diffèrent de façon significative. Il devient alors, nécessaire de faire appel à d'autres outils pour aider à déterminer quel minimum est le plus petit. C'est ce genre de problèmes que l'optimisation globale se propose de résoudre.

1.6.1 Méthodes d'optimisation globale basées sur des techniques de recherche locale

Si la fonction à minimiser possède plusieurs minima locaux, l'application d'une méthode d'optimisation locale nous donne un minimum qui n'est pas généralement global. Un effort très important a été fait, basé sur le développement de la théorie d'optimisation locale, dans le but d'utiliser des techniques de la recherche locale dans la construction d'une stratégie d'optimisation globale [47].

La méthode des initialisations multiples (multistart)

Cette méthode consiste à faire des recherches multiples de minima locaux en partant de différents points initiaux. Ces points sont généralement choisis parmi les éléments d'un maillage

uniforme. L'avantage est que ces méthodes sont généralement faciles à programmer et la solution trouvée est au moins un optimum local. La difficulté de cette méthode est essentiellement due au fait que, pour être sûr de trouver le minimum global avec une précision voulue, on doit prendre un nombre de points initiaux beaucoup plus grand que le nombre de minimiseurs locaux, or ce dernier est presque toujours inconnu. Si les minimiseurs locaux sont éloignés on doit envisager des modifications de `multistart`.

La méthode de tunnel (tunneling)

L'algorithme tunneling est l'une des méthodes de pénalisation. La méthode de percement du tunnel (*tunneling*), présentée par Levy et coll. [27], se décompose en deux phases :

- Dans la première phase on utilise une technique de descente locale permettant de trouver un minimiseur local x^* de la fonction objectif f .
- Dans la deuxième phase (la phase de tunneling) on construit une fonction auxiliaire $T(x)$ appelée fonction de tunneling (appelée encore fonction pénalité ou fonction Filled). Cette fonction admet un maximum (peut être local) au point x^* possédant des dérivées premières continues (sauf, peut être en x^*) et dépend de f , de x^* et d'un nombre fini de paramètres choisis automatiquement par l'algorithme.

On cherche alors un minimiseur y^* de $T(x)$ dans l'ensemble : $\mathbb{E}(x^*) = \{x \in D, f(x) \leq f(x^*)\} \setminus \{x^*\}$, en appliquant une minimisation locale qui commence en un point x_0 proche du point record x^* . Ensuite, on retourne à la minimisation de f en appliquant une descente qui commence cette fois-ci au point y^* ; le minimiseur local obtenu est un nouveau point record. L'itération précédente peut être répétée. On arrête la recherche lorsque le même point $y^* \in \mathbb{E}(x^*)$ est obtenu plusieurs fois par les itérations précédentes.

Méthodes de regroupement (clustering)

Ces méthodes estiment la position du minimum global en trouvant les minima locaux. Elles sont composées de plusieurs étapes. Premièrement l'espace de recherche est échantillonné. Ensuite, au moyen d'analyse de proximité (clustering techniques), les points sont regroupés par régions prometteuses. Finalement des recherches locales sont exécutées et celles qui convergent vers des régions déjà explorées sont abandonnées. Les analyses de proximité divisent l'espace de

recherche en régions, et utilisent, entre autres informations, les valeurs de la fonction ou de la matrice Hessienne des points échantillonnés. Un désavantage de ces méthodes est qu'elles sont peu performantes pour les fonctions ayant de nombreux minima. Dans ce cas, en effet, un grand nombre d'échantillons est nécessaire.

Transition vers l'optimum global

Comme la théorie de l'optimisation locale est largement plus développée que celle de l'optimisation globale, certains chercheurs ont tenté de faire quelques modifications des méthodes d'optimisation locale de sorte que ces méthodes permettent de passer d'un minimiseur local à un autre. L'idée de ces méthodes consiste à construire une trajectoire qui, dans le cas idéal, passe par tous les minima (locaux ou globaux). Ceci permet d'obtenir le minimum global. Parmi ces méthodes, celle basée sur la résolution d'équations différentielles ordinaires (comme celle de **Branin**).

En conclusion on peut affirmer que les algorithmes d'optimisation globale, basés sur l'utilisation des méthodes locales et associés à des heuristiques, sont souvent peu efficaces car le minimum absolu n'est obtenu que dans des cas très particuliers. De plus, il n'existe aucun résultat général concernant la convergence vers un minimum global.

1.7 Méthodes de recouvrement

Ce sont des méthodes avec garantie de convergence. Les cas les plus simples de ces méthodes sont celles qui consistent en la détermination de sous-régions ne peuvent contenir le minimum global, et leur élimination pour la suite. C'est comme une procédure de localisation qui se termine par une sous-région de taille suffisamment réduite pour donner la solution avec la précision souhaitée.

Rappelons le problème général :

$$f^* = \min_{x \in C} f(x), \quad (\mathbf{P})$$

où C est un compact de \mathbb{R}^n , et $f : C \rightarrow \mathbb{R}$ une fonction continue sur C .

Théoriquement, résoudre **(P)** reviendrait à déterminer exactement le plus petit de tous les minimums locaux (ou un de ses arguments x^* au cas où il y aurait plusieurs minimiseurs

globaux). Cependant, du point de vue numérique, une solution ne pourrait être obtenue qu'avec une certaine précision donnée. Dès lors, le problème **(P)** sera considéré résolu dès qu'on aura atteint un élément, pour un seuil $\epsilon > 0$ donné, de l'un au moins des ensembles $C_\epsilon(x^*), C_\epsilon(f_{x^*})$ suivants (où x^* est un minimiseur global) :

$$\begin{aligned} C_\epsilon(x^*) &= \{x \in C \text{ tel que } \|x - x^*\| < \epsilon\} \\ C_\epsilon(f_{x^*}) &= \{x \in C \text{ tel que } |f(x) - f(x^*)| < \epsilon\} \end{aligned}$$

1.7.1 Principe général

Donnons d'abord l'idée clé des méthodes par recouvrement (on va par la suite spécifier la description). Supposons que la fonction f est évaluée aux points x_1, x_2, \dots, x_k . Nous appelons record d'ordre k , la valeur suivante :

$$R_k = \min \{f(x_1), f(x_2), \dots, f(x_k)\}; \quad (1.1)$$

et nous appelons point record d'ordre k tout point $x_i, i = 1, \dots, k$ vérifiant : $f(x_i) = R_k$.

Définissons l'ensemble suivant :

$$E_k = \{x \in \mathbb{R}^n : R_k - \epsilon \leq f(x)\}, \quad (1.2)$$

évidemment,

$$R_k - \epsilon \leq \min_{x \in E_k} f(x)$$

alors :

$$\min_{x \in C} f(x) = \min_{x \in C - E_k} f(x), \text{ si } C \neq E_k.$$

Donc les points de E_k ne représentent pas d'intérêt dans la recherche du minimum global et par conséquent il peut être omis de l'ensemble faisable. On continue la recherche du minimum global seulement sur l'ensemble $C - E_k$. En pratique si :

$$C \subset E_k \quad (1.3)$$

alors le problème initial est résolu et le point record x_k est considéré comme solution approximée (R_k est le minimum global de f), avec la garantie que $x_k \in C$.

Maintenant, si f est évaluée en un nouveau point x_{k+1} , alors un autre record R_{k+1} est obtenu, et par conséquent un ensemble plus grand E_{k+1} peut être omis. La suite réelle (R_k) est décroissante, par contre la suite d'ensembles $\{E_k\}$ est croissante au sens de l'inclusion, i.e.,

$$R_{k+1} \leq R_k \quad , \quad E_k \subset E_{k+1}. \quad (1.4)$$

Ainsi le problème de minimisation est ramené à la construction d'une suite de points x_1, x_2, \dots, x_k vérifiant (1.3).

Généralement, les fonctions rencontrées dans la plupart des problèmes pratiques, ont un taux de variation borné (hypothèse moins restrictive que la dérivabilité), c'est pourquoi l'hypothèse selon laquelle la fonction objectif f est lipschitzienne est très commune en analyse numérique. Les méthodes de recouvrement ont été construites pour exploiter théoriquement cette propriété. Dans ce cas, l'évaluation de f aux seuls points d'un maillage suffisamment dense du domaine C garantit la détection du minimum global avec une précision donnée. On les appelle méthodes de recouvrement [47] parce qu'ils génèrent une famille de boules $(B(x_i, r_i))_{1 \leq i \leq l}$ dont la réunion couvre C .

Définition 1.3 Une fonction $f : C \subset \mathbb{R}^n \rightarrow \mathbb{R}$ est dite lipschitzienne sur C , s'il existe une constante $L = L(f, C) > 0$ telle que

$$|f(x) - f(y)| \leq L \|x - y\|, \quad \forall x, y \in C. \quad (1.5)$$

Evidemment, si f est lipschitzienne de constante L sur C , alors elle est de même pour toute constante $L' > L$. De plus de l'inégalité (1.5), on déduit que toute fonction lipschitzienne est continue sur C .

Dans la plupart des cas on utilise la norme euclidienne. Cependant dans plusieurs applications on a considéré l'inégalité (1.5) par rapport à d'autres normes.

Théorème 1.3 Soit $C \subset \mathbb{R}^n$ un ensemble fermé et borné, et soit f une fonction continûment différentiable sur un ensemble convexe ouvert C_0 contenant C . Alors f est lipschitzienne sur C .

Preuve Soit $x^0 \in C$ un point arbitraire. Alors le développement de Taylor à l'ordre 1 au voisinage de x^0 est donné par :

$$f(x) = f(x^0) + g(x^r)^T(x - x^0), \quad x \in C, \quad (1.6)$$

où g désigne le gradient de f , et x^r un point convenablement choisi de l'intervalle $[x^0, x]$. Maintenant, en appliquant l'inégalité de Cauchy Schwarz à l'équation (1.6), on obtient

$$|f(x) - f(x^0)| = |g(x^r)^T(x - x^0)| \leq \|g(x^r)\| \|x - x^0\|.$$

Désignons par $\overline{co(C)}$ la fermeture de l'enveloppe convexe de C . Alors, $\overline{co(C)} = co(C) \subseteq C_0$, et donc $\overline{co(C)}$ est compact et la norme du gradient g est bornée sur $\overline{co(C)}$. Alors $\max_{x \in \overline{co(C)}} \|g(x)\|$ est une constante de Lipschitz appropriée, et f donc lipschitzienne.

Algorithmes de recouvrement itératif. Dans les algorithmes de recouvrement itératif, les éléments x_1, x_2, \dots, x_k du maillage sont plus denses dans la région où la fonction prend des petites valeurs et moins denses ailleurs. En effet, à l'étape i de l'algorithme, la sélection du point x_i ($1 \leq i \leq n$) dépend des points choisis précédemment x_1, x_2, \dots, x_{i-1} et de leurs images $f(x_1), f(x_2), \dots, f(x_{i-1})$. La construction du maillage dépend de la fonction f , de l'ensemble faisable C et de la métrique choisie. En plus de la constante de Lipschitz, on fait appel parfois aux dérivées premières et secondes de f lorsqu'elles existent.

1.7.2 Algorithme de Piyavskii-Shubert unidimensionnel ($n = 1$)

Le premier algorithme proposé pour résoudre les problèmes d'optimisation lipschitziennes à une seule variable ($n = 1$) a été indépendamment donné par Piyavskii et Shubert. La méthode de Piyavskii-Shubert [38, 44], consiste à construire des recouvrements de plus en plus raffinés de la fonction objectif sur le domaine d'optimisation $C \subset \mathbb{R}^n$. Cette méthode nécessite la connaissance de la constante de Lipschitz L . Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction lipschitzienne de constante $L > 0$ sur $[a, b]$, c'est-à-dire

$$|f(x) - f(y)| \leq L|x - y|, \quad \forall x, y \in [a, b] \quad (1.7)$$

La méthode de Piyavskii minimise la fonction lipschitzienne f définie sur l'intervalle $[a, b]$, en construisant une fonction minorante de la fonction objectif f sur l'intervalle $[a, b]$. L'idée est de poser $y = a$ et $y = b$ dans l'inégalité (1.7) pour obtenir les deux inégalités suivantes :

$$\begin{aligned} f(x) &\geq f(a) - L(x - a), \forall x \in [a, b], \\ f(x) &\geq f(b) + L(x - b), \forall x \in [a, b]. \end{aligned}$$

Utilisant ces deux inégalités, on peut construire une fonction minorante linéaire telle que

$$\tilde{f}(x) \leq f(x), \forall x \in [a, b].$$

\tilde{f} est donnée par

$$\tilde{f} = \begin{cases} f(a) - L(x - a), & x \in [a, x(a, b)] \\ f(b) + L(x - b), & x \in [x(a, b), b], \end{cases}$$

où

$$x(a, b) = [f(a) - f(b)] / (2L) + [a + b] / 2, \tag{1.8}$$

et la valeur minimale de \tilde{f} est donnée par

$$B(a, b) = [f(a) + f(b)] / 2 - L(b - a) / 2.$$

Le lemme suivant montre que la fonction \tilde{f} est toujours définie. Pour ce faire, nous devons seulement nous assurer que $x(a, b) \in [a, b]$.

Lemme 1.1 *Soient f une fonction lipschitzienne sur $[a, b]$ de constante de Lipschitz L et $x(a, b)$ donné par l'équation (1.8). Alors $x(a, b) \in [a, b]$.*

Preuve Nous allons prouver que $a \leq x(a, b)$.

$$\begin{aligned}
x(a, b) &= \frac{f(a) - f(b)}{2L} + \frac{a + b}{2} \\
&= -\frac{f(b) - f(a)}{2L} + \frac{L(b - a)}{2L} + a \\
&= \frac{L(b - a) - (f(b) - f(a))}{2L} + a \\
&= \frac{1}{2L} \underbrace{(L(b - a) - (f(b) - f(a)))}_{\geq 0} + a \\
&\quad \begin{array}{l} f \text{ lipschitzienne} \\ \geq a. \end{array}
\end{aligned}$$

De même on peut montrer que $x(a, b) \leq b$. Ce qui achève la preuve.

La figure 2 montre un exemple d'une fonction f et sa fonction minorante \tilde{f} , où f est donnée par

$$f(x) = \sin\left(\left(x - \frac{1}{2}\right)4\pi\right) + 6x^2 + 2, \forall x \in [0, 1].$$

Donc

$$f'(x) = 4\pi \cos\left(\left(x - \frac{1}{2}\right)4\pi\right) + 12x, \forall x \in [0, 1].$$

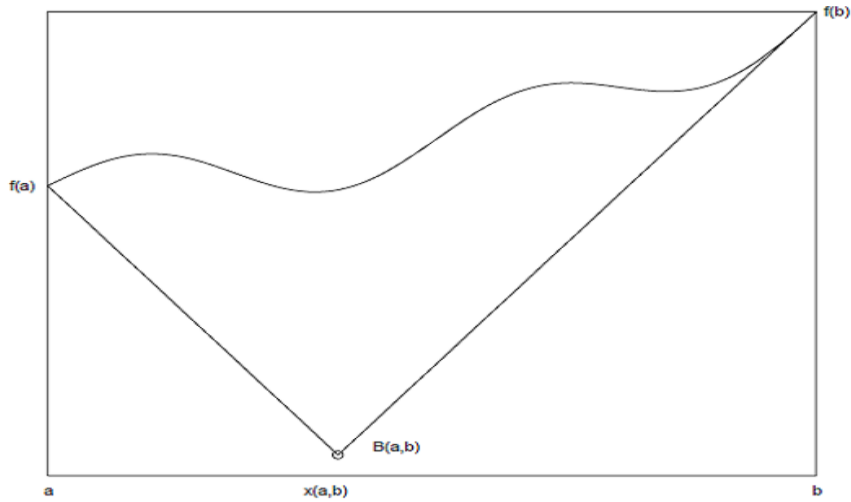


Fig.2 Exemple d'une fonction lipschitzienne avec $x(a, b)$ et $B(a, b)$

En utilisant le théorème 1.3, on obtient la constante de Lipschitz L

$$L = \sup_{x \in [0,1]} |f'(x)| = 4\pi + 12.$$

L'idée de l'algorithme de Piyavskii est de diviser l'intervalle initial en deux intervalles $I_1 = [a, x(a, b)]$ et $I_2 = [x(a, b), b]$. Ensuite, nous évaluons f en $x(a, b)$ et calculons les nouvelles valeurs $x_1 = x(a, x(a, b)), x_2 = x(x(a, b), b)$ et $B_1 = B(a, x(a, b)), B_2 = B(x(a, b), b)$ pour chacune de ces deux intervalles. B_1 et B_2 sont des bornes inférieures de la fonction f respectivement sur ces deux intervalles. Nous divisons ensuite l'intervalle correspondant à la plus petite borne inférieure dans l'étape suivante.

Dans l'algorithme de Piyavskii qui suit, les entrées sont les bornes inf. et bornes sup. de l'intervalle $[a, b]$, la constante de Lipschitz L et la précision $\epsilon > 0$. L'algorithme retourne x_{opt} et $f_{opt} = f(x_{opt})$, qui sont des solutions du problème, i.e., $f_{opt} \leq f_* + \epsilon$ où $f_* = \min_{x \in [a,b]} f(x)$.

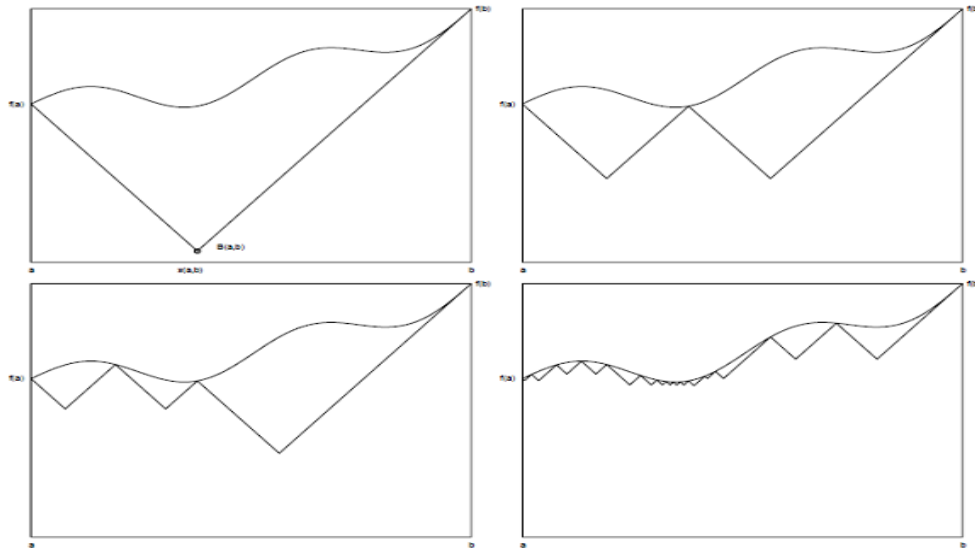


Fig.3 Exemple des itérations de l'algorithme de Piyavskii

Algorithme de Piyavskii unidimensionnel (a, b, L, ϵ)

$k = 1, sample = 1, l_{sample} = a, u_{sample} = b$

Calculer $B_1 = B(a, b), x_1 = x(a, b)$

Soit $f_{opt} = \min \{f(a), f(b)\}$, $x_{opt} = \arg \min \{f(a), f(b)\}$, $B_{opt} = B_1$.

tant que $f_{opt} - B_{opt} > \epsilon$ **faire**

choisissez un nouvel intervalle d'échantillonnage, cet intervalle a *sample* pour indice.

$l_{k+1} = x_k$, $u_{k+1} = u_{sample}$, $u_{sample} = x_k$, $f_k = f(x_k)$

calculer $B_{k+1} = B(l_{k+1}, u_{k+1})$, $x_{k+1} = x(l_{k+1}, u_{k+1})$,

$B_{sample} = B(l_{sample}, u_{sample})$, $x_{sample} = x(l_{sample}, u_{sample})$

mise à jour $f_{opt} = \min_{1 \leq i \leq k+1} \{f(a), f(u_i)\}$, $x_{opt} = \arg \min_{1 \leq i \leq k+1} \{f(a), f(u_i)\}$ et $B_{opt} = \min_{1 \leq i \leq k+1} B_i$

$k = k + 1$

fin tant que

1.7.3 Différentes extensions de la méthode de Piyavskii-Shubert ($n \geq 1$)

Beaucoup d'extensions de la méthode de Piyavskii ont été faites par plusieurs auteurs. Citons dans le cas unidimensionnel l'application de cette méthode aux fonctions non nécessairement lipschitziennes, aux fonctions seulement continues sur un intervalle fermé de \mathbb{R} [44]. La dernière approche est basée sur le résultat suivant :

Une fonction réelle f définie sur un intervalle $[a, b]$ est uniformément continue si et seulement si, pour tout $\epsilon > 0$, il existe une constante $K(\epsilon) > 0$ telle que

$$\forall x, y \in [a, b], \quad |f(x) - f(y)| \leq K(\epsilon) |x - y| + \epsilon.$$

L'extension au cas des fonctions höldériennes a été développée par plusieurs auteurs [40], [26], [28].

Rappelons qu'une fonction f est höldérienne sur $[a, b]$ si,

$$\exists h > 0 \text{ et } 0 < \alpha < 1 \text{ tel que,} \quad |f(x) - f(y)| \leq h |x - y|^\alpha, \forall x, y \in [a, b].$$

Des extensions intéressantes de la méthode de Piyavskii dans le cas multidimensionnel où C est un pavé $\prod_{i=1}^n [a_i, b_i]$ ou un compact quelconque de \mathbb{R}^n seront détaillées dans les chapitres 2 et 3.

1.7.4 Méthodes utilisant des fonctions minorantes de classe C^2

C'est une méthode unidimensionnelle élaborée par Brent [5]. Elle s'applique aux fonctions deux fois continûment différentiables dont la dérivée seconde est bornée. Autrement dit, il existe une

constante $M > 0$ telle que : $|f''(x)| \leq M$, pour tout $x \in [a, b]$. L'idée consiste à construire une suite croissante de fonctions $(\varphi_k)_{k \in \mathbb{N}}$, paraboliques par morceaux sur $[a, b]$ telle que :

$$\varphi_k(x) \leq \varphi_{k+1}(x) \leq f(x), \quad \forall x \in [a, b], \forall k \in \mathbb{N}.$$

Rappelons qu'une fonction F est parabolique par morceaux sur $[a, b]$, s'il existe une famille finie d'intervalles $(I_i)_{1 \leq i \leq k}$ disjoints deux à deux tels que : $[a, b] = \bigcup_{1 \leq i \leq k} I_i$ et $\forall i = 1, \dots, k$, la restriction de F à l'intervalle I_i est une fonction parabolique. Ceci permet de générer une suite de points $x_k^* = \arg \min_{x \in [a, b]} \varphi_k$, qui converge vers une solution optimale du problème **(P)**.

La méthode de Brent s'appuie sur le résultat suivant :

Théorème 1.4 Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction deux fois continûment dérivable et vérifie $|f''(x)| \leq M, \forall x \in [a, b]$. Alors $\forall x_1, x_2 \in [a, b]$ vérifiant $x_1 \leq x_2$, la parabole $\varphi(x)$ définie par :

$$\varphi''(x) = M, \quad \forall x \in [a, b] \quad \text{et} \quad \begin{cases} \varphi(x_1) = f(x_1) \\ \text{et} \quad \varphi(x_2) = f(x_2) \end{cases}$$

est un sous estimateur de f sur $[x_1, x_2]$.

Démonstration. Posons $g(x) = \varphi(x) - f(x)$ pour $x \in [a, b]$, alors

$g''(x) = \varphi''(x) - f''(x) = M - f''(x) \geq 0$, donc g est convexe et on a : $g(x_1) = g(x_2) = 0$ d'où nécessairement $g(x) \leq 0, \forall x \in [x_1, x_2]$, et donc $\varphi(x) \leq f(x), \forall x \in [x_1, x_2]$.

Remarque 1. Si la fonction f est évaluée aux points $a = x_0 \leq x_1 \leq \dots \leq x_k \leq x_{k+1} = b$ de $[a, b]$, alors on peut construire à partir de chaque couple (x_j, x_{j+1}) de points successifs une parabole notée $\varphi_{[x_j, x_{j+1}]}$ définie sur l'intervalle $[x_j, x_{j+1}]$ pour $j = 0, \dots, k$ et vérifiant :

$$\begin{cases} \varphi_{[x_j, x_{j+1}]}(x_j) = f(x_j) \\ \varphi_{[x_j, x_{j+1}]}(x_{j+1}) = f(x_{j+1}) \\ \text{et} \quad \varphi_{[x_j, x_{j+1}]}''(x) = M \text{ pour } x \in [x_j, x_{j+1}]. \end{cases}$$

On obtient ainsi une suite de fonctions paraboliques par morceaux dont le terme général sera noté par :

$$F_k(x) = \left[\varphi_{[a, x_1]}, \varphi_{[x_1, x_2]}, \dots, \varphi_{[x_{k-1}, x_k]} \right].$$

Maintenant, cherchons une parabole $\varphi(x)$ vérifiant $\varphi''(x) = M$ et passant par les deux points $(x_1, f(x_1)), (x_2, f(x_2)), x_1 \neq x_2$. Posons $\varphi(x) = \frac{M}{2}x^2 + \alpha x + \beta$ on a

$$\begin{cases} f(x_1) = \frac{M}{2}x_1^2 + \alpha x_1 + \beta \\ f(x_2) = \frac{M}{2}x_2^2 + \alpha x_2 + \beta. \end{cases}$$

Ce système a pour solution : $\alpha = \frac{f(x_1)-f(x_2)}{x_1-x_2} - \frac{M}{2}(x_1+x_2)$ et $\beta = \frac{x_1f(x_2)-x_2f(x_1)}{x_1-x_2} + \frac{M}{2}x_1x_2$.

Le minimum de φ est nécessairement réalisé au point x_0 et vérifie $\varphi'(x_0) = 0$ d'où

$$x_0 = -\frac{\alpha}{M} = \frac{f(x_2)-f(x_1)}{M(x_1-x_2)} + \frac{x_1+x_2}{2}.$$

Algorithme de Brent

1. Initialisation.

Poser $k = 1$, $x_1 = \frac{f(a)-f(b)}{M(b-a)} + \frac{a+b}{2}$, $x_\epsilon = \arg \min \{f(a), f(x_1), f(b)\}$, $f_\epsilon = f(x_\epsilon)$.

$$F_\epsilon = \varphi_{[a,b]}(x_1), F_1 = \left[\varphi_{[a,x_1]}, \varphi_{[x_1,b]} \right].$$

2. Etape $k = 2, 3, \dots$

Si $f_\epsilon - F_\epsilon \leq \epsilon$, alors arrêter.

Sinon, déterminer $x_{k+1} = \arg \min F_k([a, b])$,

si $f(x_{k+1}) \leq f_\epsilon$, alors poser $f_\epsilon = f(x_{k+1})$, $x_\epsilon = x_{k+1}$.

Poser $F_{k+1}(x) = \left[\varphi_{[a,x_1]}, \varphi_{[x_1,x_2]}, \dots, \varphi_{[x_k,b]} \right]$ et $F_\epsilon = \min F_{k+1}([a, b])$.

Poser $k = k + 1$ aller à 2.

La méthode de **Brent** génère un nombre de points d'évaluations plus petit que des méthodes d'Evtushenko et de Piyavskii, mais elle est applicable seulement à la classe des fonctions deux fois dérivables et de dérivées secondes bornées alors que les autres algorithmes s'appliquent à une classe plus grande (fonctions lipschitziennes, höldériennes,...). En plus, son inconvénient est qu'elle ne possède pas de généralisation au cas multidimensionnel, et perd de son efficacité si la fonction objectif est assez plate au voisinage du minimum global.

1.7.5 La méthode DIRECT unidimensionnelle

L'algorithme **DIRECT** a porté certaines remèdes aux difficultés rencontrées dans la méthode de Piyavskii, particulièrement dans le cas multidimensionnel. Cette méthode a été développée par D. R. Jones, C. D. Perttunen et B. E. Stuckman [24] en 1993. **DIRECT** tire son nom de la Division

des Rectangles. Pour expliquer cet algorithme, nous avons d'abord supposé que la constante de Lipschitz L est connue, nous laissons tomber cette exigence plus tard. Revenons d'abord à l'inégalité (1.7) :

Soit $c = (a + b)/2$ et posons $y = c$ dans (1.9). Alors $\forall x \in [a, b]$

$$x \in [a, c] : f(c) + L(x - c) \leq f(x) \leq f(c) - L(x - c)$$

$$x \in [c, b] : f(c) - L(x - c) \leq f(x) \leq f(c) + L(x - c).$$

Ces deux inégalités définissent une région dans laquelle le graphe de la fonction est contenue. De plus, nous obtenons une borne inférieure $D(a, b)$ de f sur $[a, b]$ par la substitution de $y = a$ ou $y = b$ dans les inégalités précédentes.

$$D(a, b) = f(c) - L(b - a)/2.$$

L'algorithme DIRECT divise l'intervalle d'origine en trois sous intervalles de longueurs égales et évalue la fonction f au point milieu de chaque sous intervalle. Par conséquent l'une des principales différences entre l'algorithme DIRECT et l'algorithme de Piyavskii est que l'algorithme de Piyavskii divise l'intervalle d'origine en deux intervalles de longueurs différentes.

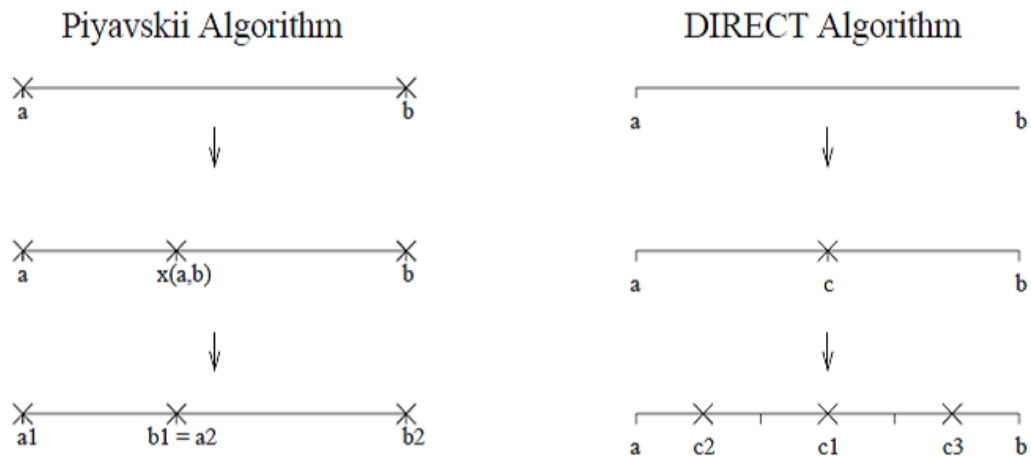


Fig. 4 La stratégie de division pour Piyavskii et DIRECT

Dans l'algorithme **DIRECT** on doit savoir la façon dont nous décidons quel intervalle prochain sera divisé, cela signifie que nous donnons quelques informations lorsque nous divisons l'intervalle. Dans **DIRECT**, nous gagnons parce que nous évaluons la fonction seulement sur le milieu de l'intervalle. L'avantage de cette idée peut facilement être étendu à des dimensions plus élevées sans augmenter le nombre nécessaire d'évaluations de la fonction lors des divisions à des hyperrectangles. Cela, est dû au fait que dans les dimensions supérieures la fonction sera évaluée seulement aux centres des hyperrectangles. Si nous étendons l'idée d'évaluer la fonction aux extrémités des intervalles à des dimensions supérieures, nous devons évaluer la fonction à tous les sommets. Cela augmente le nombre d'évaluations de la fonction exponentiellement avec la dimension du problème. Par ailleurs, la stratégie de **DIRECT** ne dépend pas de la constante de Lipschitz, puisque les points où la fonction est évaluée sont toujours les milieux. Cela va devenir important dans la suite.

Intervalles potentiellement optimaux

Il reste à décrire comment **DIRECT** choisit les intervalles à diviser. Une caractéristique importante de **DIRECT** est qu'elle fait cette décision sans avoir besoin de connaître la constante de Lipschitz L . Pour expliquer cette méthode, nous avons d'abord supposé que L est connue.

Soit $D(a_i, b_i)$ une borne inférieure de f dans chaque intervalle $[a_i, b_i]$. Nous allons maintenant choisir l'intervalle avec la plus petite borne $D(a_i, b_i)$ et diviser l'intervalle dans la prochaine itération. Nous faisons cela puisque la possibilité d'amélioration la valeur de f est plus importante pour cet intervalle.

Comme on a indiqué auparavant, dans de nombreuses applications la constante de Lipschitz L n'est pas connue. Par conséquent, nous essayons d'obtenir une estimation de la constante de Lipschitz basée sur des données connues dans une étape particulière. En particulier, nous regardons les points pour lesquels il existe une certaine constante $\widetilde{K}_i > 0$ telle que l'intervalle correspondant serait choisi si $L = \widetilde{K}_i$. Cela signifie que nous pourrions déposer une ligne de pente \widetilde{K}_i à travers le point $(\frac{b_i - a_i}{2}, f(c_i))$ tel que tous les autres points créés se situent au dessus de cette ligne. Supposons que plusieurs intervalles ont la même longueur, donc nous aurons besoin seulement de considérer les points qui représentent l'un de ces intervalles avec la plus petite valeur de la fonction en leur milieu. Maintenant, tous les points pour lesquels un tel \widetilde{K}_i existe promettrait la meilleure amélioration de la fonction si la constante de Lipschitz serait

\tilde{K}_i .

Nous appelons les intervalles pour lesquels un tel \tilde{K}_i existe : *intervalles potentiellement optimaux*.

La définition formelle d'un intervalle potentiellement optimal est comme suit :

Définition 1.4 Soit $\epsilon > 0$ une constante positive et soit f_{\min} la meilleure valeur courante de la fonction. Un intervalle j est dit être potentiellement optimal s'il existe un taux de variation \tilde{K} tel que

$$f(c_j) - \tilde{K}(b_j - a_j)/2 \leq f(c_i) - \tilde{K}(b_i - a_i)/2, \quad \forall i \quad (1.9)$$

$$\text{et} \quad f(c_j) - \tilde{K}(b_j - a_j)/2 \leq f_{\min} - \epsilon |f_{\min}|. \quad (1.10)$$

Dans cette définition, l'inégalité (1.9) nous indique seulement la décision de choisir des intervalles qui promettent la meilleure amélioration de la valeur de la fonction si f serait lipschitzienne de constante de Lipschitz \tilde{K} .

L'inégalité (1.10) assure que nous avons la possibilité d'une descente significative de f dans l'intervalle considéré.

Algorithme DIRECT unidimensionnel ($a, b, f, \epsilon, n_i, n_{fe}$)

$m = 1, c_1 = (a + b)/2$

Evaluer $f(c_1), f_{\min} = f(c_1), k = 0$

Tant que $k < n_i$ et $m < n_{fe}$ **faire**

 Identifier l'ensemble \mathcal{P} des intervalles potentiellement optimaux

Tant que $\mathcal{P} \neq \emptyset$ **faire**

 Prendre $j \in \mathcal{P}$

 Echantillonner des nouveaux points (c_{m+1}, c_{m+2}) , mettre à jour les extrémités

 Evaluer $f(c_{m+1}), f(c_{m+2})$, mettre à jour f_{\min}

 Poser $m = m + 2, \mathcal{P} = \mathcal{P} \setminus \{j\}$

Fin tant que

$k = k + 1$

Fin tant que

m : compteur pour le nombre d'évaluations de la fonction réalisées par l'algorithme.

k : compteur pour le nombre d'itérations.

Jones et al. arrêtent l'algorithme après un certain nombre d'itérations n_i . Mais comme ils ont remarqué que si la constante de Lipschitz est connue, les critères d'arrêt de l'algorithme de Piyavskii peuvent être utilisés. Une autre façon évidente d'arrêter l'algorithme est une fois que le nombre d'évaluations atteint un nombre prédéfini n_{fe} .

D'après la définition 1.4, il n'est pas clair comment effectivement identifier les intervalles potentiellement optimaux. Par conséquent, nous donnons le lemme suivant qui montre un moyen simple d'identifier les intervalles potentiellement optimaux.

Lemme 1.2 Soit $\epsilon > 0$ une constante positive et soit f_{\min} une meilleure valeur courante de la fonction. Soit I l'ensemble des indices de tous les intervalles, $d_i = (b_i - a_i)/2$ et $j \in I$ soit donné. Soit :

$$\begin{aligned} I_1 &= \{i \in I : d_i < d_j\}, \\ I_2 &= \{i \in I : d_i > d_j\}, \\ I_3 &= \{i \in I : d_i = d_j\}. \end{aligned}$$

Un intervalle j est potentiellement optimal si

$$f(c_j) \leq f(c_i), \forall i \in I_3, \quad (1.11)$$

il existe $\tilde{K} > 0$ tel que

$$\max_{i \in I_1} \frac{f(c_j) - f(c_i)}{d_j - d_i} \leq \tilde{K} \leq \min_{i \in I_2} \frac{f(c_i) - f(c_j)}{d_i - d_j}, \quad (1.12)$$

et

$$\epsilon \leq \frac{f_{\min} - f(c_j)}{|f_{\min}|} + \frac{d_j}{|f_{\min}|} \min_{i \in I_2} \frac{f(c_i) - f(c_j)}{d_i - d_j}, \text{ si } f_{\min} \neq 0, \quad (1.13)$$

ou

$$f(c_j) \leq d_j \min_{i \in I_2} \frac{f(c_i) - f(c_j)}{d_i - d_j}, \text{ si } f_{\min} = 0. \quad (1.14)$$

Avant de prouver ce lemme, nous montrons comment l'utiliser. Nous avons d'abord identifié tous les intervalles satisfaisant (1.11). Dans les prochains calculs, nous avons besoin seulement de regarder ces intervalles. Pour chacun de ces intervalles, nous calculons le maximum et le

minimum comme indiqué dans l'inégalité (1.12). Si le maximum sur tous les intervalles dans I_1 est supérieur au minimum sur tous les intervalles dans I_2 , l'intervalle courant ne peut pas être potentiellement optimal. Si le maximum est inférieur au minimum, nous utilisons les inégalités (1.13) et (1.14) pour finalement décider si un intervalle est potentiellement optimal. Nous allons maintenant montrer le lemme considéré.

Preuve Premièrement nous regardons l'inégalité (1.9). Utilisant la définition de I_1, I_2 et I_3 , l'inégalité (1.9) devient

$$\tilde{K} \geq \frac{f(c_j) - f(c_i)}{d_j - d_i}, \forall i \in I_1, \quad (1.15)$$

$$\tilde{K} \leq \frac{f(c_i) - f(c_j)}{d_i - d_j}, \forall i \in I_2, \quad (1.16)$$

et

$$f(c_j) \leq f(c_i), \forall i \in I_3. \quad (1.17)$$

De la dernière équation il suit directement que l'intervalle j peut être potentiellement optimal si $f(c_j) = \min_{i \in I_3} f(c_i)$.

Si j vérifie cette condition, alors les inégalités (1.15) et (1.16) donnent l'inégalité (1.10).

Regardons maintenant l'inégalité (1.10). L'intervalle j peut être potentiellement optimal si une constante $\tilde{K} > 0$ existe et vérifie l'inégalité (1.12). Nous voulons choisir \tilde{K} aussi grand que possible. Donc l'intervalle de l'indice j est potentiellement optimal si

$$\epsilon \leq \frac{f_{\min} - f(c_j)}{|f_{\min}|} + \frac{d_j}{|f_{\min}|} \min_{i \in I_2} \frac{f(c_i) - f(c_j)}{d_i - d_j}, \text{ si } f_{\min} \neq 0,$$

ou

$$f(c_j) \leq d_j \min_{i \in I_2} \frac{f(c_i) - f(c_j)}{d_i - d_j}, \text{ si } f_{\min} = 0.$$

Ceci conclut la preuve.

1.7.6 Structure générale de la méthode Branch-and-Bound

D'après ce qui précède, tous les algorithmes de recouvrement en optimisation globale multi-dimensionnelle sont compliqués du point de vue programmation. Il est naturel que certains chercheurs aient pensé à d'autres approches dans le but de ramener le problème considéré à

une forme plus simple. Parmi les méthodes utilisées, il y a la méthode **Branch-and-Bound** (ou séparation et évaluation) qui est fondée sur les techniques de partition et d'élimination de sous-régions ne contenant pas le minimum global. A chaque itération, on divise (séparation) le domaine C en sous-domaines qui peuvent être de tailles différentes, la recherche est effectuée dans le sous-domaine où la fonction objectif a la valeur la plus petite, en partant du principe que les chances de trouver le minimum global sont meilleures dans ce sous-domaine.

A la k -ième itération, on construit une partition de la région d'optimisation (à la première itération cette région est C) en un nombre fini n_k de sous-ensembles E_i pour lesquels les bornes inférieures l_i définies par :

$$l_i = \inf_{x \in E_i} f(x)$$

peuvent être fournies après des évaluations de f en certains points de E_i , $i \in \{1, \dots, n_k\}$. A chaque itération, le record $m_k^* = f(x_k^*)$ est aussi exploité, il correspond à la plus petite valeur de la fonction objectif au fil des itérations.

Les sous-ensembles E_i pour lesquels l_i sont supérieurs ou égaux à m_k^* ne peuvent pas contenir un minimiseur global, donc ils peuvent être exclus du domaine de recherche. On garde le E_j pour lesquels $l_j < m_k^*$. La partition est alors affinée en procédant à une répartition de l'ensemble E_j ayant

$$l_j = \min_{1 \leq i \leq n_k} l_i.$$

Les itérations se succèdent jusqu'à l'obtention du minimum global.

Les aspects de l'implémentation ont été étudiés sous différentes conditions [19], [21] et [47]. La convergence vers le minimum global est assurée, mais la réalisation pratique de la méthode est très compliquée. Ceci est dû à la nécessité du stockage d'un nombre énorme de sous-ensembles de C ainsi qu'à la difficulté liée à leur classement et au choix de l'un d'entre eux pour procéder à sa partition.

1.8 Conclusion

Les méthodes de recouvrement proposées précédemment sont simples, plus développées et ont été profondément étudiées que les méthodes d'optimisation globale basées sur une stratégie

locale. Leur efficacité a été prouvée dans le cas unidimensionnel mais dans le cas multidimensionnel, elles sont encombrantes et trop compliquées pour la réalisation pratique. En plus, le temps de calcul dépend considérablement du choix de la constante de Lipschitz et des calculs auxiliaires. En effet, dans le cas où la constante de Lipschitz est trop grande, les sous-estimateurs de f ne donnent que très peu d'informations sur f .

Chapitre 2

Optimisation globale sans contraintes des fonctions lipschitziennes à plusieurs variables

2.1 Introduction

Dans le premier chapitre on a présenté quelques méthodes de recouvrement s'appliquant aux classes des fonctions lipschitziennes à une seule variable. Plusieurs techniques de généralisation de ces méthodes dans le cas multidimensionnel ont été proposées par les chercheurs. Nous citons les techniques les plus importantes.

1- Algorithmes utilisant les sous-estimateurs.

C'est la généralisation directe de la méthode de Piyavskii au cas multivariables. L'un de ces algorithmes a été élaboré par Mladinéo [32], il est basé sur la détermination du minimum local des fonctions sous-estimateurs. A chaque itération on a besoin de résoudre un système de n équations quadratiques. Plusieurs techniques ont permis d'améliorer cette approche, Mayne et Polak [29], Jaumard et coll.[23], Gergel [13].

Si l'on veut trouver un optimum global des difficultés numériques (temps de calcul) apparaissent dès que la dimension est supérieure ou égale à 3.

2- Branch-and-Bound.

La structure générale de la méthode d'optimisation globale basée sur la séparation et l'évaluation a été développée par Horst et Tuy [19]. Une série de travaux basés sur cette voie dédiée à l'optimisation lipschitzienne sur un pavé P de \mathbb{R}^n a été donnée par Galperin [12], Pinter [37], Meewella et Mayne [30], Baritomba [2], [3], Gourdin et coll.[14]. Sans oublier l'incorporation de l'arithmétique des intervalles dans l'algorithme **Branch-and-Bound** développée par E.R. Hansen [16].

Dans ce chapitre, on va étudier la généralisation de certaines méthodes de recouvrement au cas où la fonction objectif est lipschitzienne à plusieurs variables. On va présenter l'algorithme de Pijavskii qui est un algorithme pour optimiser une fonction d'une seule variable sur un intervalle $[a, b]$, puis on va définir l'extension de cet algorithme au cas multidimensionnel. C'est dans ce contexte qu'Yves Cherruault et Arthur Guillez ont inventé une méthode, appelée méthode Aliénor, qui permet de ramener le problème d'optimisation dans \mathbb{R}^n à un problème d'optimisation univariante à l'aide d'une transformation réductrice utilisant des courbes α -denses. Les courbes α -denses généralisent la classe des courbes remplissant l'espace (space filling curves)[43]. Partant du théorème de Hahn-Mazurkiewicz qui dit que tout espace métrique compact, connexe et localement connexe peut être rempli par une courbe continue, les premières courbes remplissant l'espace ont été construites par Giuseppe Peano (1858-1932) et sont connues sous le nom de courbes de Peano, puis la courbe de Hilbert (1862-1943) qui se base sur une suite itérative, puis la courbe de Lebesgue, etc... Toutes ces courbes sont des limites suivant un certain sens de courbes α -denses. Le principe de notre approche dans ce chapitre consiste à passer d'un problème d'optimisation multidimensionnelle à un problème d'optimisation unidimensionnelle en couplant la méthode de Piyavskii-Shubert avec la nouvelle variante d'Aliénor et avec la méthode **Branch-and-Bound** (qui procède à une minimisation unidimensionnelle sur la diagonale principale). Des résultats numériques sont aussi donnés pour comparer l'efficacité de ces algorithmes.

2.2 La méthode DIRECT multidimensionnelle

Dans ce qui suit, nous décrivons comment étendre l'algorithme DIRECT dans le cas multidimensionnel. Nous faisons d'abord l'hypothèse que P est un hyperrectangle de \mathbb{R}^n . La princi-

pale différence dans cette technique, entre les dimensions supérieures et la dimension un est la procédure de la division de l'espace de recherche. Pour une meilleure compréhension, nous commençons par un hypercube ensuite décrire le cas plus général d'un hyperrectangle.

2.2.1 Division dans les dimensions supérieures

Division d'un hypercube. Soient P un hypercube et c le centre de P . Evaluons la fonction objectif f aux points : $c \pm \delta e_i$, où δ est égal à $1/3$ de la longueur du côté du cube et e_i le $i^{\text{ème}}$ vecteur de la base euclidienne. On définit w_i comme :

$$w_i = \min \{f(c + \delta e_i), f(c - \delta e_i)\}.$$

Divisons d'abord l'hypercube perpendiculairement à la direction de la plus petite valeur w_i . Puis on divise le volume restant perpendiculairement à la direction de la deuxième plus petite valeur w_i , et ainsi de suite jusqu'à ce que l'hypercube est divisé dans toutes les directions. La division qu'on a créée avec cette stratégie rend le point c le centre d'un hypercube de côté δ . Soit

$$c' = \arg \min_{1 \leq i \leq n} \{f(c + \delta e_i), f(c - \delta e_i)\}$$

c' sera le centre d'un hypercube avec un côté de longueur δ , et $n - 1$ côtés de longueur 3δ . Cela signifie que parmi les nouveaux hypercubes, celui qui a la plus petite valeur de la fonction en son centre, a un volume parmi les plus grands.

Division d'un hyperrectangle. Si nous avons besoin de subdiviser un des hyperrectangles, nous le divisons suivant la direction perpendiculaire à son côté le plus long. Ceci assure que nous obtenons une diminution dans la longueur maximale de l'hyperrectangle. Notons que les côtés des hyperrectangles dans DIRECT, ont au plus deux longueurs différentes. L'algorithme Divide est comme suit :

Algorithme Divide

- 1- Identifier l'ensemble I des dimensions, la longueur maximale d des côtés, et poser $\delta = d/3$
- 2- Evaluer f aux points $c \pm \delta e_i, i \in I$
- 3- Calculer w_i et diviser selon les valeurs de w_i .

2.2.2 Hyperrectangles potentiellement optimaux

La seule chose nécessaire qui reste pour l'algorithme DIRECT à n dimensions est la définition des hyperrectangles potentiellement optimaux.

Définition 2.1. Soit $\epsilon > 0$ une constante positive et soit f_{\min} une meilleure valeur courante de la fonction. Un hyperrectangle j est dit être potentiellement optimal s'il existe une constante \tilde{K} tel que :

$$f(c_j) - \tilde{K}\sigma_j \leq f(c_i) - \tilde{K}\sigma_i, \forall i$$

$$f(c_j) - \tilde{K}\sigma_j \leq f_{\min} - \epsilon |f_{\min}|,$$

où σ_i est la caractérisation d'un hyperrectangle i .

La seule différence entre cette définition et la définition 1.4, c'est qu'au lieu de la longueur de l'intervalle dans le cas unidimensionnel, nous avons besoin d'une autre mesure de la taille σ_i de l'hyperrectangle. Jones et al. [24] choisissent σ_i d'être la distance du centre de l'hyperrectangle i à ses sommets. Beaucoup d'extensions et modifications de l'algorithme DIRECT sont données dans la littérature.

Algorithme DIRECT multidimensionnel ($a, b, f, \epsilon, n_i, n_{fe}$)

Normaliser l'espace de recherche d'être l'hypercube unité de point centre c_1 ;

Evaluer $f(c_1)$, $f_{\min} = f(c_1)$, $k = 0$, $m = 1$;

Tant que $k < n_i$ et $m < n_{fe}$ **faire**

 Identifier l'ensemble \mathcal{P} des hyperrectangles potentiellement optimaux;

Tant que $\mathcal{P} \neq \emptyset$ **faire**

 Prendre $j \in \mathcal{P}$;

 Echantillonner des nouveaux points, évaluer f (aux nouveaux points) et diviser

l'hyperrectangle avec la procédure **Divide**

 Mettre à jour f_{\min} , $m = m + \Delta m$;

 Evaluer $f(c_{m+1})$, $f(c_{m+2})$, mettre à jour f_{\min} ;

 Poser $\mathcal{P} = \mathcal{P} \setminus \{j\}$

Fin tant que

$$k = k + 1;$$

Fin tant que

m : compteur pour le nombre d'évaluations de la fonction réalisées par l'algorithme.

k : compteur pour le nombre d'itérations.

Lemme 2.1 Soit $\epsilon > 0$ une constante positive et soit f_{\min} une meilleur valeur courante de la fonction. Notons I l'ensemble des indices de tous les hyperrectangles et

$$I_1 = \{i \in I : \sigma_i < \sigma_j\},$$

$$I_2 = \{i \in I : \sigma_i > \sigma_j\},$$

$$I_3 = \{i \in I : \sigma_i = \sigma_j\}.$$

Un hyperrectangles j est potentiellement optimal si

1)

$$f(c_j) \leq f(c_i), \forall i \in I_3,$$

2) il existe $\tilde{K} > 0$ tel que

$$\max_{i \in I_1} \frac{f(c_j) - f(c_i)}{\sigma_j - \sigma_i} \leq \tilde{K} \leq \min_{i \in I_2} \frac{f(c_i) - f(c_j)}{\sigma_i - \sigma_j},$$

3)

$$\epsilon \leq \frac{f_{\min} - f(c_j)}{|f_{\min}|} + \frac{\sigma_j}{|f_{\min}|} \min_{i \in I_2} \frac{f(c_i) - f(c_j)}{\sigma_i - \sigma_j}, \text{ si } f_{\min} \neq 0,$$

ou

$$f(c_j) \leq \sigma_j \min_{i \in I_2} \frac{f(c_i) - f(c_j)}{\sigma_i - \sigma_j}, \text{ si } f_{\min} = 0.$$

La preuve de ce lemme est identique à la preuve du Lemme 1.2

2.3 La méthode Branch-and-Bound multidimensionnelle et l'arithmétique par intervalles

Dans cette section, nous présentons un algorithme prototype d'optimisation globale par intervalles. Pour cela nous donnons brièvement le concept fondamental de l'analyse d'intervalle.

L'arithmétique par intervalles a été introduite par Ramon Moore [33] qui l'a mentionnée pour la première fois en 1962 et qui l'a définie de façon très complète.

Tout calcul par intervalle est un calcul *garanti*, c'est à dire que le résultat calculé est un intervalle qui garantit contenir la valeur ou l'ensemble des valeurs cherchées. Son atout majeur, est de permettre de faire des calculs dans la classe des ensembles. Cette idée est mis à profit pour déterminer l'optimum global d'une fonction continue, déterminer tous les zéros d'une fonction et de prouver en même temps leur existence et leur éventuelle unicité ou encore déterminer l'image directe ou inverse d'un ensemble par une fonction.

2.3.1 Calcul par intervalles

En arithmétique par intervalles, on ne manipule plus des nombres, qui approchent plus ou moins fidèlement une valeur, mais des intervalles contenant cette valeur. Par exemple, on peut tenir compte d'une erreur de mesure en remplaçant une valeur mesurée x avec une incertitude ϵ par l'intervalle $[x - \epsilon, x + \epsilon]$. On peut également remplacer une valeur non exactement représentable, telle que π , par un intervalle la contenant ; si l'on dispose d'un ordinateur représentant les nombres en base 10 avec 3 chiffres, π sera remplacé par $[3.14, 3.15]$.

Enfin, si l'on désire obtenir un résultat valide pour tout un ensemble de valeurs, on utilise un intervalle contenant ces valeurs. En effet, l'objectif de l'arithmétique par intervalles est de fournir des résultats qui contiennent à coup sûr la valeur ou l'ensemble cherché ; on parle alors de résultats *garantis* ou *validés*.

Notations. Les objets intervalles seront désignés par des caractères gras : \mathbf{x} . On notera \underline{x} le minimum de \mathbf{x} et \bar{x} son maximum, avec l'ordre partiel sur \mathbb{R}^n : $x \leq y$ ssi $x_i \leq y_i$ pour $1 \leq i \leq n$. On a alors

$$\mathbf{x} = [\underline{x}, \bar{x}].$$

Enfin, $w(\mathbf{x})$ est la largeur (ou diamètre) de \mathbf{x} : $\bar{x} - \underline{x}$. Le centre $mid(\mathbf{x})$ et son rayon $rad(\mathbf{x})$ sont définis par $mid(\mathbf{x}) = (\underline{x} + \bar{x})/2$ et $rad(\mathbf{x}) = \frac{w(\mathbf{x})}{2} = (\bar{x} - \underline{x})/2$. On notera \mathbb{I} l'ensemble des intervalles \mathbf{x} de \mathbb{R} , on peut les généraliser aux plusieurs dimensions : un vecteur intervalle $\mathbf{x} \in \mathbb{I}^n$ est un vecteur dont les n composantes sont des intervalles, et une matrice intervalle $\mathbf{A} \in \mathbb{I}^{m \times n}$ est une matrice dont les composantes sont des intervalles.

Définition 2.1 Le résultat d'une opération entre deux intervalles : \mathbf{x} et \mathbf{y} , resp. d'une fonction $F(\mathbf{z})$, est le plus petit intervalle au sens de l'inclusion (ou vecteur intervalle) contenant

$$\{x * y / x \in \mathbf{x}, y \in \mathbf{y}\}, \text{ resp. } \{F(z), z \in \mathbf{z}\},$$

c'est à dire le plus petit intervalle contenant tous les résultats possibles de l'opération appliquée à tous les éléments x de \mathbf{x} et tous les éléments y de \mathbf{y} , resp. tous les résultats possibles de f appliqués à tous les éléments z de \mathbf{z} .

Opérations arithmétiques. Quand on applique la définition précédente aux opérations arithmétiques $=, -, \times, (\cdot)^2, \div$, ou $\sqrt{\cdot}$, on obtient les formules suivantes, plus utilisées en pratique que la définition abstraite :

$$\begin{aligned} [\underline{x}, \bar{x}] + [\underline{y}, \bar{y}] &= [\underline{x} + \underline{y}, \bar{x} + \bar{y}] \\ [\underline{x}, \bar{x}] - [\underline{y}, \bar{y}] &= [\underline{x} - \bar{y}, \bar{x} - \underline{y}] \\ [\underline{x}, \bar{x}] \times [\underline{y}, \bar{y}] &= [\min(\underline{x} \times \underline{y}, \underline{x} \times \bar{y}, \bar{x} \times \underline{y}, \bar{x} \times \bar{y}), \max(\underline{x} \times \underline{y}, \underline{x} \times \bar{y}, \bar{x} \times \underline{y}, \bar{x} \times \bar{y})]. \\ [\underline{x}, \bar{x}]^2 &= [\min(\underline{x}^2, \bar{x}^2), \max(\underline{x}^2, \bar{x}^2)] \text{ si } 0 \notin [\underline{x}, \bar{x}] \text{ et } [0, \max(\underline{x}^2, \bar{x}^2)], \text{ sinon} \\ 1 \div [\underline{y}, \bar{y}] &= [\min(1/\underline{y}, 1/\bar{y}), \max(1/\underline{y}, 1/\bar{y})] \text{ si } 0 \notin [\underline{y}, \bar{y}]. \\ [\underline{x}, \bar{x}] \div [\underline{y}, \bar{y}] &= [\underline{x}, \bar{x}] \times (1 \div [\underline{y}, \bar{y}]) \text{ si } 0 \notin [\underline{y}, \bar{y}]. \\ \sqrt{[\underline{x}, \bar{x}]} &= [\sqrt{\underline{x}}, \sqrt{\bar{x}}] \text{ si } 0 \leq \underline{x}. \end{aligned}$$

On obtient ces formules en utilisant la monotonie (au moins partielle) de ces opérations.

Propriétés algébriques. On peut 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 $\mathbf{x} = [2, 3]$, $\mathbf{x} - \mathbf{x} = [2, 3] - [2, 3] = [-1, 1] \neq 0$ même s'il le contient. En effet,

$$\mathbf{x} - \mathbf{x} = \{x - y / x \in \mathbf{x}, y \in \mathbf{y}\} \supset \{x - x / x \in \mathbf{x}\} = \{0\}$$

et l'inclusion est stricte.

De la même façon, la division n'est pas la réciproque de la multiplication : si $\mathbf{x} = [2, 3]$, l'inter-

valle $\mathbf{x} \div \mathbf{x} = [2, 3] \div [2, 3] = [2/3, 3/2]$ 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'élevation au carré : si $\mathbf{x} = [-3, 2]$,

$$\mathbf{x} \times \mathbf{x} = [-3, 2] \times [-3, 2] = [-6, 9]$$

alors que

$$\mathbf{x}^2 = \{x^2 / x \in \mathbf{x}\} = [0, 9].$$

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

$$\begin{aligned}\mathbf{x} \times (\mathbf{y} + \mathbf{z}) &= [-2, 3] \times ([1, 4] + [-2, 1]) = [-2, 3] \times [-1, 5] = [-10, 15] \\ \mathbf{x} \times \mathbf{y} + \mathbf{x} \times \mathbf{z} &= [-2, 3] \times [1, 4] + [-2, 3] \times [-2, 1] = [-8, 12] + [-6, 4] = [-14, 16]\end{aligned}$$

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

$$\mathbf{x} \times (\mathbf{y} + \mathbf{z}) \subset \mathbf{x} \times \mathbf{y} + \mathbf{x} \times \mathbf{z}.$$

On peut également définir des fonctions élémentaires (sin, cos, exp, ...) prenant des intervalles pour argument, à l'aide de la définition abstraite ci-dessus. Exemple,

$$\exp([\underline{x}, \bar{x}]) = [\exp \underline{x}, \exp \bar{x}] \text{ puisque } \exp \text{ est croissante.}$$

On prend en considération la monotonie des fonctions. En revanche, il faut être soigneux pour les fonctions périodiques par exemple, mais il est possible d'établir des algorithmes de calcul pour ces fonctions, dès que l'on dispose de ces fonctions sur les réels. Par exemple, $\sin[\pi/3, \pi] = [0, 1]$. Enfin, on ne sait définir les fonctions élémentaires que sur des intervalles inclus dans leur domaine de définition : le logarithme ne sera défini que pour des intervalles strictement positifs.

Evaluation d'une expression. On sait également calculer le résultat d'une expression mêlant opérations arithmétiques ou algébriques et fonctions élémentaires sur des intervalles.

Par exemple, l'expression polynomiale $\mathbf{x}^3 - 2\mathbf{x}^2 + \mathbf{x} - 3$, avec $\mathbf{x} = [-5, 2]$, a pour résultat

$$[-5, 2]^3 - 2[-5, 2]^2 + [-5, 2] - 3 = [-125, 8] - 2[0, 25] + [-5, 2] - 3 = [-183, 7].$$

L'expression en plusieurs variables $\sin \mathbf{x} + 2\mathbf{x} \exp \mathbf{y} - \mathbf{y}^2 \sqrt{\mathbf{z}}$ avec $\mathbf{x} = [-\pi, \pi/4]$, $\mathbf{y} = [-1, 1]$ et $\mathbf{z} = [1, 4]$ a pour résultat

$$\sin[-\pi, \pi/4] + 2[-\pi, \pi/4] \exp[-1, 1] - [-1, 1]^2 \sqrt{[1, 4]} = [-3 - 2\pi e, \sqrt{2}/2 + \pi e/2].$$

Définition 2.2 Soit f une fonction réelle à plusieurs variables réelles. On appelle fonction d'inclusion de f toute fonction intervalle F définie de \mathbb{III} dans \mathbb{III} qui encadre f .

C'est-à-dire on a

$$[\min_{x \in \mathbf{x}} f(x), \max_{x \in \mathbf{x}} f(x)] \subseteq F(\mathbf{x}), \forall \mathbf{x} \in \mathbb{III}.$$

Théorème 2.1 Soit f une fonction continue de \mathbb{R}^n dans \mathbb{R} , définie par une expression dans laquelle chaque variable x_i apparaît au plus une fois, alors pour tout pavé $\mathbf{x} \subset \mathbb{R}^n$, l'évaluation par intervalles obtenue en remplaçant x_i par la composante correspondante de \mathbf{x} , est une fonction d'inclusion.

Ce Théorème est un résultat fondamental de l'analyse des intervalles [33]. Il nous permet de construire des fonctions d'inclusion très efficaces [31] ce qui contribue à améliorer la performance de l'algorithme d'optimisation globale.

2.3.2 Forces et faiblesses de l'arithmétique par intervalles

1. Force : Calcul sur des ensembles.

L'avantage majeur de l'arithmétique par intervalles est qu'elle permet de calculer avec des ensembles. C'est cette force qui a permis de surmonter la mauvaise impression laissée par des débuts peu calculant (estimation grossière des erreurs d'arrondi), de repartir vers des utilisations adaptées et de faire l'effort de développer des algorithmes qui tiennent compte des avantages et difficultés de l'arithmétique par intervalles. Aussi cette arithmétique, en permettant le calcul sur des ensembles, rend effectifs des théorèmes. Cela signifie qu'en calculant par intervalles, on

peut obtenir des preuves, au sens mathématique du terme, de certaines propriétés. On peut également obtenir des preuves d'existence seule ou d'existence et unicité de solutions.

2. Force : informations globales, optimisation globale sous contraintes.

Fréquemment, les problèmes d'optimisation globale sont des problèmes sous contraintes : on cherche le minimum d'une fonction f qui vérifie en outre un certain nombre d'équations et inéquations. Mathématiquement, le problème s'écrit :

$$\begin{array}{l} \min_x f(x) \\ \text{contraintes} \left\{ \begin{array}{ll} p_i(x) \leq 0 & 1 \leq i \leq m \\ q_j(x) = 0 & 1 \leq j \leq r \end{array} \right. \end{array}$$

Maintenant, la difficulté réside en deux points : déterminer des points qui satisfont les contraintes et trouver leur minimiseur x^* , qui lui ne vérifie plus $\nabla f(x^*) = 0$ ni f localement convexe en x^* .

Une approche utilise les contraintes pour rejeter ou réduire le pavé de recherche à chaque étape. Un pavé \mathbf{x} sera rejeté s'il ne satisfait pas les contraintes, par exemple s'il existe un indice i tel que $p_i(\mathbf{x}) > 0$ ou $0 \notin q_j(\mathbf{x})$. Si ce pavé \mathbf{x} n'est pas rejeté, on tentera d'éliminer de \mathbf{x} les parties ne satisfaisant pas les contraintes.

Faiblesse : le résultat dépend de l'expression utilisée.

La propriété fondamentale de garantie des résultats permet d'assurer que pour tout pavé $\mathbf{x} \subset C$, le pavé obtenu en utilisant les variables scalaires par les composants intervalles correspondants de \mathbf{x} est un surencadrement de l'image de \mathbf{x} par f . L'idéal serait de déterminer le plus petit pavé contenant $f(x)$.

Soit la fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ définie par : $f(x) = x^2 - 2x + 1$. Si $\mathbf{x} = [-1, 3]$, en remplaçant x par \mathbf{x} dans l'expression de f , on obtient $\mathbf{x}^2 - 2\mathbf{x} + 1 = [-5, 12]$. Si on utilise plutôt l'expression équivalente : $f(x) = x(x - 2) + 1$ on obtient $\mathbf{x}(\mathbf{x} - 2) + 1 = [-8, 4]$ et enfin en utilisant l'écriture factorisée $f(x) = (x - 1)^2$ on obtient $(\mathbf{x} - 1)^2 = [1, 4] = f(\mathbf{x})$. Cet exemple illustre clairement le fait que des expressions équivalentes en arithmétique réelle ne le sont plus en arithmétique par intervalles, même si chacune donne lieu à un surencadrement de l'image de \mathbf{x} par f .

Nous venons de mettre le doigt sur l'inconvénient majeur de l'arithmétique par intervalles, à savoir la surestimation (ou encadrement trop large) des résultats. Or on souhaite obtenir un résultat précis, et dans l'idéal le plus petit résultat possible, c'est-à-dire le plus petit intervalle (au sens de l'inclusion) qui contient le résultat exact.

Evaluation d'une expression : développement de Taylor.

Déterminer l'expression qui conduira au meilleur résultat et chercher des techniques qui limitera le phénomène de surestimation est un problème difficile. Pour cela des recherches ont été consacrées dans ce sens. A partir d'une expression donnée, on peut montrer que pour des intervalles pas trop larges, l'utilisation de développement de Taylor d'ordre 1 et 2 réduit le phénomène de surestimation. Plusieurs propositions concernant l'évaluation de l'image d'une fonction, à variables scalaires sur un vecteur d'intervalles, ont été données [42], comme fonction d'inclusion naturelle, formes centrées (en utilisant le développement de Taylor), etc...

2.3.3 Un algorithme prototype de Branch-and-Bound par intervalles

Parmi les forces de l'arithmétique par intervalles est de savoir calculer sur des ensembles, même si on est restreint à des ensembles de forme particulière, ceci permet de fournir des informations globales précieuses pour l'*optimisation globale* d'une fonction.

Supposons que l'on cherche à déterminer le minimum global x^* d'une fonction f de \mathbb{R}^n dans \mathbb{R} assez régulière (au moins continue et si possible de classe C^2) sur un pavé \mathbf{x}_0 . On suppose que l'on dispose d'une expression analytique de f , afin de pouvoir l'évaluer sur des intervalles. Voici un algorithme élémentaire, pour déterminer un petit intervalle contenant x^* . C'est un algorithme prototype de **Branch-and-Bound** par intervalles.

Algorithme.

Données : f la fonction à minimiser et \mathbf{x}_0 l'intervalle dans lequel on cherche le minimum,
 ϵ le seuil de la longueur des intervalles résultats.

Résultats : $[\underline{f}, \overline{f}] \ni f^* = \min_{x \in \mathbf{x}_0} f(x)$, un encadrement de la valeur minimale de f sur \mathbf{x}_0 .

Res : une liste de pavés \mathbf{x}^* tels que $f(\mathbf{x}^*) \cap [\underline{f}, \overline{f}] \neq \emptyset$.

Initialisation.

$\mathcal{L} := \{\mathbf{x}_0\}$ la liste des pavés en attente de traitement.

$Res := \emptyset$

$$[\underline{f}, \bar{f}] := f(\mathbf{x}_0)$$

évaluer f en un ou plusieurs points de \mathbf{x}_0 , par exemple $\bar{f} = f(\text{mid}(\mathbf{x}_0))$

on a maintenant un encadrement plus précis de $f(x^*)$ par $[\underline{f}, \bar{f}]$.

Tant que $\mathcal{L} \neq \emptyset$ **faire**

sortir un pavé \mathbf{x} de \mathcal{L}

diviser \mathbf{x} en deux : \mathbf{x}_1 et \mathbf{x}_2

évaluer $f(\mathbf{x}_1) : [\underline{f}_1, \bar{f}_1]$

mettre à jour \bar{f} : si $\bar{f}_1 \leq \bar{f}$ alors $\bar{f} := \bar{f}_1$

si $\underline{f}_1 > \bar{f}$ alors

jeter \mathbf{x}_1 / \mathbf{x}_1 ne peut pas contenir le minimum x^*

sinon si $w(\mathbf{x}_1) \leq \epsilon$ alors ranger \mathbf{x}_1 dans Res

sinon ranger \mathbf{x}_1 dans \mathcal{L}

faire la même chose pour \mathbf{x}_2

Fin.

mettre à jour l'encadrement de $f(x^*) = [\underline{f}, \bar{f}] : \underline{f} = \min_{x \in Res} f(x), \bar{f} = \max_{x \in Res} f(x)$

éliminer les \mathbf{x} de Res tels que $\underline{y} > \bar{f}$ avec $[\underline{y}, \bar{y}] = f(\mathbf{x})$.

Cet algorithme génère une liste d'intervalles, Res , qui contient tous les intervalles susceptibles de contenir un minimiseur global sur chacun de ces intervalles, f prend des valeurs proches de l'optimum. On peut noter une caractéristique de l'arithmétique par intervalles : cet algorithme garantit de n'oublier aucun minimiseur global. Cet algorithme est aussi le seul envisageable si f est de classe C^0 uniquement. Si f est plus régulière, on peut appliquer différentes stratégies pour tenter d'éliminer rapidement ou de réduire un pavé. Voir l'algorithme d'optimisation globale, dû à E. Hansen [16].

En résumé, l'arithmétique par intervalles constitue une approche intéressante pour répondre à l'exigence de fiabilité des calculs. Elle repose sur le principe que tout calcul fournit un encadrement garanti de son résultat. De plus, elle peut être implantée efficacement sur ordinateur et une panoplie d'algorithmes numériques ont été développés spécifiquement pour tirer partie de cette arithmétique. L'approche des intervalles s'applique désormais à l'optimisation avec ou sans contraintes sur des domaines bornés ou non.

2.4 Algorithme de Piyavskii généralisé

2.4.1 Description de l'algorithme

Considérons le problème de minimisation globale

$$\min_{x \in P} f(x) \quad (\mathbf{P})$$

que l'on veut résoudre avec une précision exigée $\varepsilon > 0$, avec f une fonction lipschitzienne de constante $L > 0$ définie sur un pavé $P = \prod_{i=1}^n [a_i, b_i]$ de \mathbb{R}^n et à valeurs dans \mathbb{R} . C'est-à-dire, une fonction telle que :

$$\forall x, y \in P, \quad |f(x) - f(y)| \leq L \|x - y\|$$

où $\|\cdot\|$ est la norme euclidienne.

Le principe de cette méthode est de diviser le pavé P en plusieurs sous pavés et de construire sur chacun d'eux une fonction minorante constante de la fonction objectif.

Considérons un point x_i du pavé P dont la fonction objectif f est évaluée. La meilleure fonction minorante de f est $f_i(x) = f(x_i) - L \|x - x_i\|$, et donc la meilleure fonction constante minorante de f est le minimum de $f_i(x)$ sur le pavé P .

Si $c = \frac{a+b}{2}$ est le centre du pavé P , la constante minorante de f est :

$$F = \min_{x \in P} f_i(x) = f(c) - L \frac{\|b - a\|}{2}.$$

En effet, le minimum de $f_i(x)$ est évalué en un point \bar{x} , où la distance entre \bar{x} et le centre c est maximale. Comme c est le centre du pavé P , alors

$$\|\bar{x} - c\| = \frac{\|b - a\|}{2},$$

où $\|b - a\|$ est la longueur de la diagonale du pavé P .

L'algorithme divise le pavé P en p sous pavés $(P_j)_{1 \leq j \leq p}$. Soit $P_j = \prod_{i=1}^n [a_i^j, b_i^j] = [a^j, b^j]$ un de ces sous pavés et c_j le centre de P_j , la valeur de la fonction constante minorante sur P_j est définie par :

$$F_j = f(c_j) - L \frac{\|b^j - a^j\|}{2}.$$

On note par $l_i^j = b_i^j - a_i^j$, $i = 1, \dots, n$. Soit $D = \|b - a\| = \left(\sum_{i=1}^n l_i^2 \right)^{1/2}$ où $l_i = \|b_i - a_i\|$ la longueur de la diagonale du pavé P , \mathbb{L} la liste des sous problèmes \mathbf{P}_i pour $i = 1, \dots, n$.

Algorithme

$k \leftarrow 1$;

$x^1 \leftarrow (a_1 - \frac{l_1}{2}, a_2 - \frac{l_2}{2}, \dots, a_n - \frac{l_n}{2})$;

$x_{opt} \leftarrow x^1$.

$f_{opt} \leftarrow f(x_{opt})$;

$F_{opt} \leftarrow f_{opt} - \frac{LD}{2}$.

$F_1 \leftarrow F_{opt}$;

$\mathbf{P}_1 \leftarrow (F_1, (a_1, \dots, a_n); (l_1, \dots, l_n))$;

$\mathbb{L} \leftarrow \{\mathbf{P}_1\}$;

Tant que $f_{opt} - F_{opt} > \varepsilon$ faire

$\mathbf{P}_i \leftarrow (F_i, (a_1^j, \dots, a_n^j); (l_1^j, \dots, l_n^j))$;

$l_{io}^j \leftarrow \max_{1 \leq i \leq n} l_i^j$;

$D \leftarrow \left(\sum_{i=1, i \neq io}^n (l_i^j)^2 + \left(\frac{l_{io}^j}{p} \right)^2 \right)^{1/2}$;

Supprimer \mathbf{P}_i de \mathbb{L} ;

Pour $k : 1$ à p faire

$S_k \leftarrow (a_1^j, \dots, a_{io-1}^j, a_{io}^j + \frac{k-1}{p} l_{io}^j, \dots, a_n^j)$;

$L_k \leftarrow (l_1^j, \dots, l_{io-1}^j, \frac{l_{io}^j}{p}, \dots, l_n^j)$;

$F_k \leftarrow f(S_k - \frac{L_k}{2}) - \frac{LD}{2}$;

Si $f(S_k - \frac{L_k}{2}) < f_{opt}$;

$f_{opt} \leftarrow f(S_k - \frac{L_k}{2})$;

$x_{opt} \leftarrow S_k - \frac{L_k}{2}$;

Fin Si;

Ajouter le sous problème à \mathbb{L}

Fin Pour;

Eliminer de \mathbb{L} tous les sous problèmes ayant un minorant supérieur à f_{opt}

Fin Tant que

Remarque. Le test d'élimination n'est pas nécessaire pour la convergence de l'algorithme.

2.4.2 Convergence de l'algorithme

Théorème 2.2. L'algorithme de Piyavskii généralisé est convergent, c'est-à-dire se termine après un nombre fini d'itérations, ou bien on a

$$\lim_{k \rightarrow \infty} F_{opt}^k = \lim_{k \rightarrow \infty} f_{opt}^k = f^* = \min_{x \in P} f(x).$$

Démonstration

Soit $\varepsilon > 0$, pour montrer la convergence de l'algorithme, on doit trouver une suite $(r_1, \dots, r_n) \in \mathbb{N}^n$ telle que :

$$D = \frac{L}{2} \left(\sum_{i=1}^n \left(\frac{l_i}{p^{r_i}} \right)^2 \right)^{1/2} \leq \varepsilon.$$

$$\frac{LD}{2} \leq \varepsilon \Leftrightarrow D \leq \frac{2\varepsilon}{L} \Leftrightarrow \sum_{i=1}^n \left(\frac{l_i}{p^{r_i}} \right)^2 \leq \frac{4\varepsilon^2}{L^2}.$$

Il suffit de choisir :

$$\forall i \quad r_i \geq \log \left(\frac{Ll_i \sqrt{n}}{2\varepsilon} \right).$$

Donc $\exists m \in P, \exists k, \exists (r_1, \dots, r_n)$ tels que

$$\left| f^k(m) - F_{opt}^k \right| \leq \varepsilon$$

où $f^k(m) = F_{opt}^k + \frac{LD}{2}$.

Par définition on a $F_{opt}^k \leq f^* \leq f^k(m)$. D'où $\exists k \in \mathbb{N}$ tel que $|f^* - F_{opt}^k| \leq \varepsilon$.

Puisque ε est arbitraire, on a :

$$\lim_{k \rightarrow \infty} F_{opt}^k = f^*.$$

De plus $f^* \leq f_{opt}^k \leq f^k(m)$, ce qui implique que

$$\left| f^* - f_{opt}^k \right| \leq \varepsilon.$$

2.5 La méthode de la transformation réductrice Aliénor

2.5.1 Présentation de la méthode

Dans cette section, nous allons présenter une méthode déterministe. Il s'agit de la méthode *Aliénor*. Elle a été développée par Y. Cherruault et coll.[1], [6] à la fin des années 1970. L'idée consiste à approcher une fonction de plusieurs variables par une fonction d'une seule variable, il devenait alors assez simple de déterminer les minima globaux car il suffisait de les chercher en suivant l'évolution d'une certaine courbe.

Soit la fonction $f(x_1, x_2, \dots, x_n)$ de n variables x_1, x_2, \dots, x_n , la méthode donc consiste à exprimer ces variables à l'aide d'une seule en densifiant l'espace \mathbb{R}^n à l'aide d'une simple courbe de classe C^∞ . Nous allons mettre en évidence ce que nous appellerons la transformation réductrice :

$$x_i = h_i(\theta), \theta > 0, i = 1, \dots, n$$

où θ est une variable réelle et les h_i des fonctions très régulières.

Cette transformation permet d'approcher l'espace \mathbb{R}^n par une courbe et ceci avec la précision que l'on souhaite c'est à dire elle consiste à ramener une fonction multivariées à une fonction d'une seule variable $f^*(\theta) = f(h_1(\theta), h_2(\theta), \dots, h_n(\theta))$. La méthode de la transformation réductrice Aliénor a été essentiellement développée pour l'optimisation globale mais elle s'est révélée être une approche tout à fait universelle car elle peut être appliquée à l'approximation des fonctions de plusieurs variables, à la résolution numérique d'équations fonctionnelles et à beaucoup d'autres problèmes multidimensionnels.

On désigne par μ la mesure de Lebesgue, A un intervalle fermé et borné de \mathbb{R} (en général $A = [0, T]$, avec $T > 0$) et α un nombre réel, strictement positif et supposé être très petit par rapport aux dimensions du pavé $P = \prod_{i=1}^n [a_i, b_i]$, ($\alpha \ll \min_{1 \leq i \leq n} (b_i - a_i)$). Où $n \geq 2$, est un entier.

Définition 2.. On dit qu'une courbe paramétrée de \mathbb{R}^n définie par :

$$h : A \rightarrow P = \prod_{i=1}^n [a_i, b_i]$$

est α -dense dans P ou bien qu'elle α -densifie P si pour chaque $x \in P$, il existe $\theta \in A$ tel que

$$d(x, h(\theta)) \leq \alpha,$$

où d est la distance euclidienne dans \mathbb{R}^n .

La première transformation réductrice. Dans les travaux de Y. Cherruault et A. Guillez [6] où la méthode a été décrite originellement en utilisant la spirale d'Archimède, la fonction objectif $f(x_1, x_2, \dots, x_n)$ est supposée continue sur \mathbb{R}^n et ayant un minimum global atteint à une distance finie. Dans le plan \mathbb{R}^2 , on considère le vecteur à deux composantes x, y . Nous allons d'abord passer en coordonnées polaires avec

$$x = r \cos \theta, \quad y = r \sin \theta, \quad \theta \geq 0.$$

Ce faisant nous obtenons deux nouvelles variables r et θ que nous allons relier grâce à la spirale d'Archimède d'équation $r = a\theta$, où a est un paramètre positif fixé, destiné à tendre vers 0 pour permettre d'approcher tout point du plan euclidien par un point de la courbe. La représentation paramétrique de la spirale d'Archimède est donc :

$$h(\theta) = (h_1(\theta), h_2(\theta)) = (a\theta \cos \theta, a\theta \sin \theta).$$

Nous avons exprimé les variables x et y avec une variable unique θ . On peut montrer facilement que la courbe obtenue est α -dense dans \mathbb{R}^2 avec $\alpha = \pi a$.

D'une façon générale, si nous avons n variables x_1, x_2, \dots, x_n , on relie deux à deux les variables par des spirales d'Archimède de paramètres θ_i , jusqu'à l'utilisation d'une dernière spirale de paramètre θ .

On a ainsi une structure d'arbre, à chaque étape on divise le nombre de variables par 2. Pour la transformation réductrice :

$$x_i = h_i(\theta), \quad i = 1, \dots, n.$$

On voit que les fonctions h_i font intervenir les fonctions trigonométriques en sinus et cosinus. Ces $h_i(\theta)$ sont de classe C^∞ . Le résultat d' α -densification suivant est donné dans [51].

Remarques.

- La fonction initiale $f(x_1, x_2, \dots, x_n)$ est alors approchée par la fonction d'une seule variable $f(h_1(\theta), h_2(\theta), \dots, h_n(\theta))$ où $x_i = h_i(\theta)$, $i = 1, \dots, n$, sont obtenus comme précédemment en reliant deux à deux les variables à l'aide de spirales d'Archimède.

La difficulté avec cette méthode de base est que les fonctions composantes $h_i(\theta)$ ne s'obtiennent pas en une seule étape. Aussi, on peut voir que lorsque la dimension de l'espace \mathbb{R}^n augmente, les fonctions $h_i(\theta)$ prennent des expressions de plus en plus sophistiquées; ce qui accroît le temps d'évaluation de la fonction $f^*(\theta)$ lors de la résolution du problème de minimisation unidimensionnel.

Quelques variantes de la méthode Aliénor. Plutôt que de "réduire" deux à deux les variables (structure d'arbre) par la méthode **Aliénor** de base, il est possible d'exprimer toutes les variables en fonction d'une seule variable, en une seule étape. Les premières courbes ont été proposées dans [1]. Pour un problème d'optimisation dans le pavé $\prod_{i=1}^n [a_i, b_i]$, la courbe considérée possède la représentation paramétrique suivante :

$$\begin{aligned}
 h_1(\theta) &= \frac{\alpha\alpha_2(\theta)}{2\pi} \cos(\alpha_2(\theta)) \\
 h_2(\theta) &= \frac{\alpha\alpha_2(\theta)}{2\pi} \sin(\alpha_2(\theta)) \\
 h_i(\theta) &= a_i + \frac{\alpha_i(\theta)}{\theta_i} (b_i - a_i), \text{ pour } i = 3, \dots
 \end{aligned}$$

$$\text{avec } \begin{cases} \theta_2 = \frac{\pi}{\alpha} \left((b_1 - a_1)^2 + (b_2 - a_2)^2 \right)^{\frac{1}{2}} \\ \theta_i = \left(\frac{b_i - a_i}{\alpha} \right) \theta_{i-1} \text{ pour } i = 3, \dots, n \end{cases}$$

$$\text{et pour } i = 2, \dots, n \text{ on a } \begin{cases} \beta_i(\theta) = Ent\left(\frac{\theta}{\theta_i}\right) \\ \alpha_i(\theta) = (-1)^{\beta_i(\theta)} \left[\theta - \left(\beta_i(\theta) - \frac{1}{2} (-1)^{\beta_i(\theta)} + \frac{1}{2} \right) \theta_i \right] \end{cases}$$

On peut montrer que cette courbes est $\sqrt{n-1}\alpha$ -dense dans $\prod_{i=1}^n [a_i, b_i]$, [1].

Nous avons vu que la méthode d'Aliénor ou ses variantes permettent de construire des courbes "remplissante" au sens de l' α -densité. Ces courbes ont l'avantage d'être très régulières (C^∞) contrairement à celles qui sont proposées dans la littérature (courbe de Peano et de Hilbert) Sagan [43], où les courbes remplissant l'espace sont seulement continues. De plus, les courbes de

types "Alienor" sont facile à générer. Mais deux questions essentielles se posent, existe-t-il une méthode générale permettant de générer ces courbes α -denses? et peut-on trouver des classes de transformation minimisant le temps de calcul du minimum global?.

Des réponses positives ont été données par A. Ziadi et Y. Cherruault [53] et par G. Mora et Y. Cherruault [34]. Mais par la suite, A. Ziadi et Y. Cherruault [54] l'ont améliorée par l'obtention de nouveaux résultats qui permettent de construire de grandes classes de courbes α -denses. Cette dernière voie s'est révélée être la plus intéressante car les courbes obtenues possèdent des représentations paramétriques plus simples. Une étude complète et détaillée sur quelques aspects de la théorie des courbes α -denses est donnée dans [51], [52], [53].

Une nouvelle transformation réductrice

La caractérisation et la génération des courbes α -denses dans un pavé de \mathbb{R}^n ($n \geq 2$) est un sujet fondamental. D'autre part, l'un des objectif les plus importants est l'application des courbes α -denses à l'optimisation globale. Mais, tout cela dépend essentiellement de la longueur de la courbe qui densifie le domaine de la fonction objectif. Des résultats intéressants, sur l'existence des courbes avec une densité fixée et une longueur minimale sur des compacts densifiables dans un espace métrique général ont été donnés par A. Ziadi et Y. Cherruault [52].

En se basant sur un résultat donné en [53]. On définit d'une manière constructive une courbe α -dense dans un pavé quelconque de \mathbb{R}^n .

Théorème 2.3 Soient h_1, h_2, \dots, h_n des fonctions continues surjectives, respectivement définies de A dans $[a_i, b_i]$ pour $i = 1, 2, \dots, n$, et soient aussi $\theta_1, \theta_2, \dots, \theta_{n-1}$ et α des nombres strictement positifs tels que pour tout $i = 1, 2, \dots, n - 1$, il existe une partition finie de A composée d'intervalles $(I_{i,j})_{1 \leq j \leq m_i}$ et vérifiant :

- (a) $\mu(I_{i,j}) = \theta_i, \quad \forall j = 1, 2, \dots, m_i.$
- (b) $h_i(\bar{I}_{i,j}) = [a_i, b_i], \quad \forall j = 1, 2, \dots, m_i$
- (c) $\mu(h_{i+1}(I_{i,j})) \leq \alpha, \quad \forall j = 1, 2, \dots, m_i.$

Alors la courbe paramétrée définie par la fonction $h = (h_1, h_2, \dots, h_n)$, est $\sqrt{n-1}\alpha$ -dense dans $\prod_{i=1}^n [a_i, b_i]$. On peut voir la preuve dans [53].

Théorème 2.4 On considère dans \mathbb{R}^n le pavé $\prod_{i=1}^n [a_i, b_i]$ et α un nombre strictement positif tel

que le nombre $\frac{b_i - a_i}{\alpha}$ soit un entier pour tout $i = 1, 2, \dots, n$. Posons

$$T = \frac{\prod_{i=1}^n (b_i - a_i)}{\alpha^n}$$

et soient h_1, h_2, \dots, h_n des fonctions, respectivement définies de $[0, T]$ dans $[a_i, b_i]$ par :

$$\begin{aligned} h_i(\theta) &= a_i + \frac{\sigma_i(\theta)}{\theta_i} (b_i - a_i) \quad \text{pour } i = 1, 2, \dots, n ; \\ \text{avec} \quad \theta_0 &= 1, \quad \theta_i = \frac{b_i - a_i}{\alpha} \theta_{i-1} \\ \sigma_i(\theta) &= (-1)^{\beta_i(\theta)} \left[\theta - (\beta_i(\theta) + \frac{1}{2}) \theta_i \right] \\ \beta_i(\theta) &= \text{Ent} \left(\frac{\theta}{\theta_i} \right) \quad \text{où Ent est l'application " partie entière" } \end{aligned}$$

Alors la fonction définie par $h(\theta) = (h_1(\theta), h_2(\theta), \dots, h_n(\theta))$ pour $\theta \in [0, T]$ représente une courbe paramétrée $\sqrt{n-1}\alpha$ -dense dans $\prod_{i=1}^n [a_i, b_i]$.

Preuve. (a) Posons

$$m_i = \frac{\prod_{k=i+1}^n (b_k - a_k)}{\alpha^{n-i}} \quad \text{pour } i = 1, 2, \dots, n-1.$$

Les hypothèses impliquent que pour tout $i = 1, 2, \dots, n-1$, le nombre m_i est un entier et $\theta_n = m_i \theta_i = T$. Considérons les intervalles

$$\begin{aligned} I_{i,j} &= [(j-1)\theta_i, j\theta_i[\quad \text{pour } j = 1, 2, \dots, m_i - 1 \\ \text{et } I_{i,m_i} &= [(m_i-1)\theta_i, m_i\theta_i]. \end{aligned}$$

Il est facile de voir que pour tout $i = 1, 2, \dots, n-1$, la famille $(I_{i,j})_{1 \leq j \leq m_i}$ forme une partition de l'intervalle $[0, T]$ et $\mu(I_{i,j}) = \theta_i$, $\forall j = 1, 2, \dots, m_i$.

En outre, on peut montrer que :

(b) pour tout $i = 1, 2, \dots, n-1$,

$$h_i(\bar{I}_{i,j}) = [a_i, b_i], \quad \forall j = 1, 2, \dots, m_i,$$

(c) pour tout $i = 1, 2, \dots, n - 1$,

$$\mu(h_{i+1}(I_{i,j})) \leq \alpha, \quad \forall j = 1, 2, \dots, m_i.$$

Il en résulte que toutes les hypothèses du théorème 3.3 sont satisfaites. Donc la courbe paramétrée définie par $h(\theta) = (h_1(\theta), h_2(\theta), \dots, h_n(\theta))$, pour $\theta \in [0, T]$, est $\sqrt{n-1}\alpha$ -dense dans $\prod_{i=1}^n [a_i, b_i]$.

Remarque.

Dans cette construction, la courbe obtenue densifie le pavé sans en sortir. Pour le même pavé et le même paramètre de densification α , on peut montrer qu'elle est plus courte que plusieurs d'autres courbes. On a pris $\theta_0 = 1$, mais θ_0 peut prendre n'importe quelle autre valeur strictement positive.

2.5.2 Application de la méthode Aliénor à l'optimisation globale

La transformation Aliénor permet de résoudre des problèmes d'optimisation globale multidimensionnelle en approximant les fonctions de n variables par des fonctions d'une seule variable.

Soit à résoudre le problème d'optimisation global suivant :

$$\min_{(x_1, \dots, x_n) \in P} f(x_1, \dots, x_n), \quad (\mathbf{P})$$

où f est une fonction lipschitzienne de constante l_1 , définie sur $P = \prod_{i=1}^n [a_i, b_i]$ et à valeurs dans \mathbb{R} . La transformation Aliénor

$$x_i = h_i(\theta), \quad i = \overline{1, n}$$

sert à approcher la fonction f par la fonction à une seule variable f^* définie par :

$$f^*(\theta) = f(h_1(\theta), \dots, h_n(\theta)).$$

Le problème de minimisation (P) est alors approximé par le problème de minimisation

$$\min_{\theta \in [0, \theta_{\max}]} f^*(\theta) \quad (\mathbf{P}^*)$$

que nous pouvons résoudre simplement en discrétisant l'intervalle $[0, \theta_{\max}]$ via un pas $\Delta\theta$ puis en cherchant le minimum de l'ensemble $\{f^*(\theta_i) : i = 0, 1, \dots, N\}$, où $\theta_i = i\Delta\theta$, $i = \overline{1, N}$ sont les points de la discrétisation et $\theta_N = \theta_{\max}$.

La difficulté dans cette méthode de base réside dans le fait que le pas $\Delta\theta$ doit être choisi inversement proportionnel à θ (car la distance entre deux points successifs M_θ et $M_{\theta+\Delta\theta}$ de la spirale d'Archimède dépend de la valeurs θ). Mais si on veut que la distance entre deux points successifs sur la spirale reste une constante C , il faut choisir $\Delta\theta = \frac{C}{a\theta}$, ce qui va augmenter le nombre de points à considérer.

Théorème 2.5 *Le minimum global de $f(x_1, \dots, x_n)$ peut être approché par le minimum global de $f^*(\theta)$.*

Preuve. Soient $(x_1^*, x_2^*, \dots, x_n^*)$ un point de $P = \prod_{i=1}^n [a_i, b_i]$ réalisant le minimum global de f et θ^* un point de $[0, \theta_{\max}]$ réalisant le minimum global de f^* .

Puisque la courbe $h(\theta)$ est α -dense dans P , il existe $\theta^+ \in [0, \theta_{\max}]$ tel que :

$$\|(x_1^*, x_2^*, \dots, x_n^*) - h(\theta^+)\| \leq \alpha$$

Il en résulte que

$$\|f(x_1^*, x_2^*, \dots, x_n^*) - f(h(\theta^+))\| \leq l_1 \alpha$$

donc

$$f(h(\theta^+)) - f(x_1^*, x_2^*, \dots, x_n^*) \leq l_1 \alpha$$

Et comme on a

$$f(x_1^*, x_2^*, \dots, x_n^*) \leq f^*(\theta^*) \leq f(h(\theta^+)),$$

on déduit que

$$0 \leq f^*(\theta^*) - f(x_1^*, x_2^*, \dots, x_n^*) \leq l_1 \alpha$$

Donc, pour obtenir la précision désirée ε , il suffit de choisir $\alpha = \frac{\varepsilon}{l_1}$. ■

2.6 La méthode hybride Branch-and-Bound-Piyavskii

2.6.1 Description de la méthode

Considérons le problème d'optimisation globale suivant :

$$\min_{x \in P} f(x), \quad (\mathbf{P})$$

où f est une fonction lipschitzienne de constante l_1 définie sur le pavé $P = \{x \in \mathbb{R}^n : a \leq x \leq b\}$ où $a, b \in \mathbb{R}^n$, et à valeurs dans \mathbb{R} .

On a déjà noté que la méthode de Piyavskii est numériquement inefficace pour $n \geq 3$ (dans le premier chapitre) puisque l'effort de calcul des sous-problèmes correspondants est énorme. Pour cela, nous préférons présenter l'extension de Branch-and-Bound en appliquant essentiellement la version unidimensionnelle de la méthode de Piyavskii-Shubert à la diagonale principale des ensembles rectangulaires de la partition de P .

Notons par $a_M, b_M \in \mathbb{R}^n$, les bornes inférieures et supérieures du pavé M , i.e.,

$$M = \{x \in \mathbb{R}^n : a_M \leq x \leq b_M\}.$$

L'algorithme hybride B&B-Piyavskii-Shubert

1) Initialisation :

$$M_0 = X, \quad \mathcal{M}_0 = \{M_0\}.$$

$$\alpha_0 = \min \{f(a), f(b)\}, \quad x^0 \in \{a, b\} \text{ tel que } f(x^0) = \alpha_0.$$

$$\beta_0 = \max \{f(a), f(b)\} - L \|b - a\|.$$

Aller à l'étape k .

2) Etape $k = 1, 2, \dots$

Au début de l'étape k , nous avons la partition actuelle \mathcal{M}_{k-1} d'un sous ensemble de $M_0 = X$ qui est encore digne d'intérêt et pour tout $M \in \mathcal{M}_{k-1}$, il existe des bornes $\beta(M), \alpha(M)$ telles que :

$$\beta(M) \leq \min f(M) \leq \alpha(M).$$

de plus, nous avons les bornes $\beta_{k-1}, \alpha_{k-1}$ telles que :

$$\beta_{k-1} \leq \min f(X) \leq \alpha_{k-1},$$

avec $\alpha_{k-1} = f(x^{k-1})$, où x^{k-1} est le record obtenu dans l'étape $(k-1)$.

k.1. On exclut tout sous ensemble $M \in \mathcal{M}_{k-1}$ vérifiant :

$$\beta(M) \geq \alpha_{k-1}.$$

Considérons la classe \mathfrak{R}_k des sous ensembles restants de \mathcal{M}_{k-1} .

k.2. Prenons une classe $\mathfrak{S}_k \subset \mathfrak{R}_k$, telle que :

$$\mathfrak{S}_k \cap \arg \min \{\beta(M) : M \in \mathfrak{R}_k\} \neq \emptyset$$

Puis, considérons dans chaque $\widetilde{M} \in \mathfrak{S}_k$, le point $x_{\widetilde{M}}$ donné par :

$$x_{\widetilde{M}} = \frac{1}{2}(a_{\widetilde{M}} + b_{\widetilde{M}}) + \frac{f(a_{\widetilde{M}}) - f(b_{\widetilde{M}})}{2L \|b_{\widetilde{M}} - a_{\widetilde{M}}\|} (b_{\widetilde{M}} - a_{\widetilde{M}})$$

et subdivisons \widetilde{M} en deux par l'hyperplan qui contient $x_{\widetilde{M}}$ et qui est orthogonal à la face la plus longue de \widetilde{M} .

Soit \mathcal{M}'_k l'ensemble de tous les nouveaux éléments de la partition.

Pour tout $M' \in \mathcal{M}'_k$, notons par $\widetilde{M}' \in \mathfrak{S}_k$, le pavé dont la subdivision engendre M' .

k.3. Pour chaque $M' \in \mathcal{M}'_k$, poser :

$$\begin{aligned} \alpha_k(M') &= \min \{f(a_{M'}), f(b_{M'}), f(x_{\widetilde{M}'})\}. \\ \beta_k(M') &= \max \left\{ \beta(\widetilde{M}'), \max \{f(a_{M'}), f(b_{M'}), f(x_{\widetilde{M}'})\} - L \|b_{M'} - a_{M'}\| \right\} \end{aligned}$$

k.4. Poser :

$$\mathcal{M}_k = (\mathfrak{R}_k \setminus \mathfrak{S}_k) \cup \mathcal{M}'_k$$

déterminer

$$\begin{aligned}\alpha_k &= \min \{ \alpha(M) : M \in \mathcal{M}_k \}, \\ \beta_k &= \min \{ \beta(M) : M \in \mathcal{M}_k \}, \\ x^k &\in X \text{ tel que : } f(x^k) = \alpha_k\end{aligned}$$

k.5. Si $\alpha_k - \beta_k \leq \varepsilon$, alors arrêter.

Sinon, aller à l'étape $k + 1$.

Dans le cas unidimensionnel ($X = [a, b]$), à la k -ième itération de la méthode de Piyavskii-Shubert, on obtient les points d'évaluation de $f : x_1, x_2, \dots, x_k$. En ordonnant ces points, on obtient :

$$a = y^1 \leq y^2 \leq \dots \leq y^k = b$$

avec $\{y^1, y^2, \dots, y^k\} = \{x_1, x_2, \dots, x_k\}$.

Alors, l'ensemble $\{M_{k,i} = [y^i, y^{i+1}] : i = 1, \dots, k - 1\}$ est une partition de $X = [a, b]$.

Pour tout intervalle $[y^i, y^{i+1}]$ de X , on prend :

$$\alpha(M_{k,i}) = \min \{ f(y^i), f(y^{i+1}) \} = f(x^{k,i})$$

et

$$\beta(M_{k,i}) = \frac{f(y^{i+1}) + f(y^i)}{2} - L \frac{y^{i+1} - y^i}{2}$$

avec $i = 1, \dots, k - 1$.

Il en résulte que $\beta(M_{k,i})$ est une borne inférieure de f dans $M_{k,i}$. En effet, d'après la méthode de Piyavskii-Shubert, on a :

$$\min F_k(X) \leq \min f(X),$$

mais $\forall x \in M_{k,i}$, on a :

$$F_k(x) = \max_{j=i, i+1} \{ f(y^j) - L|x - y^j| \}.$$

Posons

$$z^{k,i} = \arg \max_{x \in [y^i, y^{i+1}]} F_k(x).$$

Donc $z^{k,i}$ vérifie l'égalité suivante :

$$f(y^i) - L(z^{k,i} - y^i) = f(y^{i+1}) - L(y^{i+1} - z^{k,i})$$

d'où

$$z^{k,i} = \frac{f(y^i) - f(y^{i+1})}{2L} + \frac{y^i + y^{i+1}}{2}$$

D'autre part, on a :

$$\max_{x \in [y^i, y^{i+1}]} F_k(x) = F_k(z^{k,i})$$

d'où

$$\max_{x \in [y^i, y^{i+1}]} F_k(x) = \frac{f(y^{i+1}) + f(y^i)}{2} - L \frac{y^{i+1} - y^i}{2} = \beta(M_{k,i}).$$

Alors, il est clair que pour tout $M_{k,i}$, on a :

$$\beta(M_{k,i}) \leq \min f(M_{k,i}) \leq \alpha(M_{k,i})$$

Considérons maintenant les bornes suivantes :

$$\begin{aligned} \alpha_k &= f(x^k) = \min \{ \alpha(M_{k,i}) : i = 1, \dots, k-1 \}, \\ \beta_k &= \beta(M_{k,j}) = \min \{ \beta(M_{k,i}) : i = 1, \dots, k-1 \}, \end{aligned}$$

Si $\alpha_k - \beta_k \leq \varepsilon$, alors arrêter. Le point x^k est une approximation de la solution optimale de f .

Sinon, on subdivise $M_{k,j}$ en deux intervalles $[y^j, z^k]$ et $[z^k, y^{j+1}]$ tels que :

$$z^k = \frac{f(y^j) - f(y^{j+1})}{2L} + \frac{y^j + y^{j+1}}{2}.$$

On obtient les points d'évaluation de $f : y^1, \dots, y^k, z^k$ et on considère $a = y^1 \leq y^2 \leq \dots \leq y^{k+1} = b$ avec $\{y^1, \dots, y^{k+1}\} = \{y^1, \dots, y^k, z^k\}$.

On répète la même itération avec $k = k + 1$. On continue la procédure jusqu'à l'obtention du minimum global.

Remarque :

En pratique, un ensemble M ne sera pas subdivisé plus loin quand la longueur de sa diagonale $\|b_M - a_M\|$ est inférieure à un paramètre fixé $\delta > 0$.

2.6.2 Convergence de la méthode

Nous donnons le théorème suivant :

Théorème 2.6 : *Considérons l'algorithme BB-Piyavskii décrit précédemment. Alors nous avons :*

$$\lim_{k \rightarrow \infty} \alpha_k = \lim_{k \rightarrow \infty} \beta_k = \min f(X)$$

et tout point d'accumulation de $\{x^k\}$ est un minimiseur global de f .

Preuve :

Soit $\{M_{k_q}\}$ une suite emboîtée d'intervalles issus des subdivisions successives générées par l'algorithme. La suite $\{\alpha(M_{k_q})\}$ des bornes supérieures associées est décroissante et minorée par $\min f(X)$ alors que la suite $\{\beta(M_{k_q})\}$ des bornes inférieures associées est croissante et majorée par $\min f(X)$.

Donc, il existe α, β , tels que :

$$\beta = \lim_{q \rightarrow \infty} \beta_{k_q} \leq \min f(X) \leq \lim_{q \rightarrow \infty} \alpha_{k_q} = \alpha. \tag{2.1}$$

Considérons maintenant la suite de points $\{x^k\}$ dont le terme général x^k est généré à l'étape k par l'algorithme. Puisque M_{k_q} est subdivisé à l'itération k_{q+1} , on a :

$$x^{k_{q+1}} \in M_{k_q} \text{ et } \alpha(M_{k_{q+1}}) \leq f(x^{k_{q+1}}). \tag{2.2}$$

La suite $\{x^k\}$ a des points d'accumulation puisque X est un compact. Soit \bar{x} un point d'accumulation de la sous suite $\{x^{k_{q+1}}\}$ de $\{x^k\}$. A partir de la continuité de f et de la relation (2.1), on a l'inégalité suivante :

$$\alpha \leq f(\bar{x}). \tag{2.3}$$

Supposons qu'il existe $\varepsilon > 0$, tel que :

$$f(\bar{x}) > \beta + \varepsilon$$

par définition de \bar{x} , il existe une sous suite $\{x^{k_{q'+1}}\}$ de $\{x^{k_{q+1}}\}$ et un $q_\varepsilon \in \mathbb{N}$ tels que : pour $q' > q_\varepsilon$

$$\|x^{k_{q'+1}} - \bar{x}\| < \frac{\varepsilon}{2L}.$$

d'autre part, puisque f est lipschitzienne, on a :

$$f(x^{k_{q'+1}}) \geq f(\bar{x}) - L\|x^{k_{q'+1}} - \bar{x}\| > \beta + \frac{\varepsilon}{2} \quad \forall q' > q_\varepsilon .$$

donc pour tout $x \in X$ vérifiant $\|x - x^{k_{q'+1}}\| < \frac{\varepsilon}{2L} \quad \forall q' > q_\varepsilon$, on a :

$$f(x^{k_{q'+1}}) - L\|x - x^{k_{q'+1}}\| > \left(\beta + \frac{\varepsilon}{2}\right) - \frac{\varepsilon}{2} = \beta. \quad (2.4)$$

Mais, par la construction de la procédure de Branch-and-Bound, on a pour tout $k : \beta_k \leq \beta$.

Donc, à partir de (2.4), on déduit que :

$$x^k \notin \left\{ x : \|x - x^{k_{q'+1}}\| < \frac{\varepsilon}{2L}, q' > q_\varepsilon \right\}, \quad \forall k,$$

ce qui contredit la supposition que \bar{x} est un point d'accumulation de $\{x^k\}$. Alors il n'existe pas un $\varepsilon > 0$ tel que :

$$f(\bar{x}) > \beta + \varepsilon.$$

Par conséquent :

$$\alpha \leq f(\bar{x}) \leq \beta \quad (2.5)$$

Ainsi, à partir de (2.1), (2.3) et (2.5), on a :

$$\alpha = \min f(X) = f(\bar{x}) = \beta. \quad \blacksquare$$

2.7 La méthode hybride Aliénor-Piyavskii

Dans l'optimisation globale, pratiquement tous les algorithmes de recouvrement sont très compliqués à réaliser pour $n \geq 3$. Le couplage de la méthode Aliénor avec certaines méthodes unidimensionnelles de recouvrement (méthode d'Evtushenko, méthode de Piyavskii-Shubert et méthode de Brent) s'est avéré très efficace pour trouver les minima globaux des fonctions de plusieurs variables ayant un minimum global difficile à trouver par les méthodes classiques. Ce couplage consiste à approcher la fonction objectif f à plusieurs variables par une fonction à une seule variable f^* , puis à appliquer un algorithme de recouvrement unidimensionnel à f^* . Dans ce qui suit, nous étudions le couplage de la méthode Aliénor avec l'algorithme unidimensionnel de Piyavskii-Shubert.

2.7.1 Présentation de la méthode

Soit à résoudre le problème :

$$\min_{x \in P} f(x), \quad (\mathbf{P})$$

où $P = \prod_{i=1}^n [a_i, b_i]$ et f une fonction lipschitzienne de constante l_1 sur P .

La transformation réductrice Aliénor permet de remplacer les variables x_1, \dots, x_n par la variable $\theta \geq 0$. La fonction $f(x_1, \dots, x_n)$ devient une fonction $f^*(\theta)$ qui s'exprime facilement à partir de f . Le problème (P) se ramène à :

$$\min_{\theta} f^*(\theta). \quad (\mathbf{P}^*)$$

Autrement dit, on obtient un problème de minimisation par rapport à une variable où l'algorithme de Piyavskii-Shubert peut être facilement appliqué.

La méthode mixte Aliénor avec Piyavskii-Shubert est simple et ne demande pas des calculs auxiliaires. Elle est composée de deux étapes : l'étape de la transformation réductrice et l'étape de l'application de l'algorithme de Piyavskii-Shubert.

Première étape :

On définit l'application $h = (h_1, h_2, \dots, h_n) : \left[0, \frac{\pi}{\alpha_n}\right] \rightarrow \prod_{i=1}^n [a_i, b_i]$ par :

$$h_1(\theta) = \frac{a_1 - b_1}{2} \cos(\alpha_1 \theta) + \frac{a_1 + b_1}{2}$$

$$h_2(\theta) = \frac{a_2 - b_2}{2} \cos(\alpha_2 \theta) + \frac{a_2 + b_2}{2}$$

...

$$h_n(\theta) = \frac{a_n - b_n}{2} \cos(\alpha_n \theta) + \frac{a_n + b_n}{2},$$

où $\alpha_1, \alpha_2, \dots, \alpha_n$ sont des paramètres donnés par :

$$\alpha_1 = 2^n$$

$$\alpha_2 = \frac{2^n \alpha}{\pi(b_2 - a_2)}$$

$$\alpha_3 = \frac{2^n \alpha^2}{\pi^2(b_2 - a_2)(b_3 - a_3)}$$

...

$$\alpha_n = \frac{2^n \alpha^{(n-1)}}{\pi^{(n-1)}(b_2 - a_2)(b_3 - a_3) \dots (b_n - a_n)}$$

avec

$$\alpha = \frac{1}{\sqrt{n-1}} \frac{\varepsilon}{2l_1}.$$

La courbe paramétrée $h(\theta) = (h_1(\theta), h_2(\theta), \dots, h_n(\theta))$ est $\frac{\varepsilon}{2l_1}$ -dense dans le pavé $\prod_{i=1}^n [a_i, b_i]$. De plus, il est facile de montrer que la fonction h est lipschitzienne de constante

$$l_2 = \frac{1}{2} \left(\sum_{i=1}^n (b_i - a_i)^2 \alpha_i^2 \right)^{\frac{1}{2}}.$$

Deuxième étape :

On applique maintenant l'algorithme de Piyavskii-Shubert à la fonction $f^*(\theta) = f(h(\theta))$ qui est une fonction lipschitzienne de constante $L = l_1 l_2$ pour obtenir le minimum global de $f^*(\theta)$.

1. Initialisation

$$\theta_1 = \frac{\pi}{2\alpha_n}$$

$$\theta_\varepsilon = \theta_1$$

$$f_\varepsilon^* = f^*(\theta_\varepsilon)$$

$$F_\varepsilon = f_\varepsilon^* - \frac{L\pi}{2\alpha_n}$$

$$F_1(\theta) = f^*(\theta_1) - L|\theta - \theta_1|$$

2. Pour $k = 1, 2, \dots$

Si $f_\varepsilon^* - F_\varepsilon \leq \frac{\varepsilon}{2}$, alors arrêter

Sinon, poser $\theta_{k+1} = \arg \min_{\theta \in [0, \frac{\pi}{\alpha_n}]} F_k(\theta)$.

Si $f^*(\theta_{k+1}) < f_\varepsilon^*$, alors poser $f_\varepsilon = f(\theta_{k+1})$, $\theta_\varepsilon = \theta_{k+1}$.

Poser $F_{k+1}(\theta) = \max_{j=1, \dots, k+1} \{f^*(\theta_j) - L|\theta - \theta_j|\}$,

et $F_\varepsilon = \min F_{k+1} \left(\left[0, \frac{\pi}{\alpha_n} \right] \right)$.

Poser $k = k + 1$.

Aller à 2.

2.7.2 Convergence de la méthode

Nous donnons le théorème suivant :

Théorème 2.7 *La méthode mixte Aliénor couplée avec Piyavskii-Shubert appliquée au problème (P) converge vers le minimum global avec une précision inférieure ou égale à ε .*

Preuve. Notons par M et M' , les minima globaux de f et f^* respectivement. D'autre part, désignons par f_ε^* le minimum global du problème (\mathbf{P}^*) obtenu par la méthode mixte Aliénor couplée avec Piyavskii-Shubert. Montrons que

$$f_\varepsilon^* - M \leq \varepsilon.$$

a) Puisque f est continue sur P , il existe un point $y \in P$ tel que : $M = f(y)$. Par ailleurs, il existe $\theta_0 \in \left[0, \frac{\pi}{\alpha_n} \right]$ tel que $\|y - h(\theta_0)\| \leq \frac{\varepsilon}{2l_1}$ donc $\|f(y) - f(h(\theta_0))\| \leq \frac{\varepsilon}{2}$.

Il en résulte que :

$$f(h(\theta_0)) - M \leq \frac{\varepsilon}{2}.$$

Puisque $M \leq M' \leq f(h(\theta_0))$, on déduit que :

$$M' - M \leq \frac{\varepsilon}{2}. \quad (2.6)$$

b) Puisque f_ε^* est le minimum global du problème (P) obtenu par la méthode mixte Aliénor couplée avec Piyavskii-Shubert, on obtient :

$$f_\varepsilon^* - F_\varepsilon \leq \frac{\varepsilon}{2},$$

où $F_\varepsilon = \min_{\theta \in [0, \alpha_n]} \left\{ \max_{i=1, \dots, k} [f^*(\theta_i) - L|\theta - \theta_i|] \right\}$ et f_ε^* est le minimum obtenu après k évaluations.

D'autre part, pour tout $\theta \in \left[0, \frac{\pi}{\alpha_n} \right]$, on a : $F_\varepsilon \leq f^*(\theta)$, donc $F_\varepsilon \leq M'$.

Il en résulte que :

$$f_\varepsilon^* - M' \leq \frac{\varepsilon}{2}. \quad (2.7)$$

Finalement, on déduit de (2.6) et (2.7) que :

$$f_\varepsilon^* - M \leq \varepsilon \quad \blacksquare$$

Remarque.

La méthode **Aliénor** est une méthode qui permet de ramener un problème multidimensionnelle à un problème unidimensionnel de même type.

Une fois le problème multidimensionnel converti en un problème unidimensionnel, on utilise la méthode de **Piyavskii-Shubert** basée sur la construction des sous-estimateurs de la fonction objectif f^* . Cette méthode présente l'avantage de ne pas utiliser de dérivées de f^* et la convergence de l'algorithme ne dépend pas du choix du point de départ.

2.8 Applications numériques

La méthode de **Piyavskii-Shubert** est une méthode d'optimisation globale. Elle s'applique aux fonctions lipschitziennes définies sur un intervalle fermé et borné de \mathbb{R} et consiste à construire une suite croissante de sous-estimateurs générant une suite de points qui convergent vers le minimum global.

En couplant la méthode de **Piyavskii-Shubert** unidimensionnelle avec la nouvelle variante d'**Aliénor** d'une part et avec la méthode **Branch-and-Bound** d'autre part, nous avons procédé à une série d'applications sur des fonctions tests tirées de la littérature, ayant la particularité de posséder plusieurs minima locaux.

Nous étudierons les fonctions suivantes :

1- **Rosenbrok** : $100(x_2 - x_1^2)^2 + (1 - x_1)^2$

$$(x_1, x_2) \in [-5; 5] \times [-5; 5]$$

2- **Goldstein** : $\left[1 + (1 + x_2 - x_1)^2 f(x_1, x_2)\right] \left[30 + (2x_1 - 3x_2)^2 g(x_1, x_2)\right]$

$$(x_1, x_2) \in [-2; 2] \times [-2; 2]$$

avec $f(x_1, x_2) = 19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2$

$$g(x_1, x_2) = 18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2$$

3- Branin : $4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 + 4x_2^2$

$$(x_1, x_2) \in [-5; 5] \times [-5; 5]$$

4- Wood : $100(x_2 - x_1^2)^2 + (1 - x_1^2)^2 + 90(x_4 - x_3^2)^2 + (1 - x_3)^2 +$

$$10.1 \left[(x_2 - 1)^2 + (x_4 - 1)^2 \right] + 19.8(x_2 - 1)(x_4 - 1)$$

$$X \in [-2; 2] \times [-2; 2] \times [-2; 2] \times [-2; 2]$$

5- Exp(10) : $\exp\left(-\frac{1}{10} \sum_{i=1}^{10} x_i^2\right)$

$$X \in [-1; 1]^{10}$$

6- Exp(2) : $\exp\left(-\frac{1}{2}(x_1^2 + x_2^2)\right)$

$$(x_1, x_2) \in [-1; 1] \times [-1; 1]$$

7- Cos(6) : $\sum_{i=1}^6 \left(x_i^2 - \frac{1}{10} \cos(5\pi x_i)\right)$

$$X \in [-1; 1]^6$$

8- Cos(2) : $x_1^2 + x_2^2 - \frac{1}{10}(\cos(5\pi x_1) + \cos(5\pi x_2))$

$$(x_1, x_2) \in [-1; 1] \times [-1; 1]$$

Tous les résultats donnés dans les tableaux ci dessous ont été effectués sur un ordinateur Pentium (R) 4 CPU pour l'obtention du minimum global avec une précision de l'ordre de 0.1.

Dans les deux tableaux de comparaison de ces méthodes, on a utilisé les notations suivantes :

F_{\min} : le minimum global

Nbre : le nombre d'itération de l'algorithme

t : le temps de calcul

AS : méthode mixte Aliénor-Shubert

BB-S : méthode hybride Branch-and-Bound-Shubert

Fonctions	Solution exacte	<i>AS</i>	<i>BB-S</i>
Rosenbrok	$(x_1, x_2) = (1, 1)$	$(x_1, x_2) = (1.039, 1.08)$	$(x_1, x_2) = (0.99, 1.00225)$
	$F_{\min} = 0$	$F_{\min} = 0.000628$	$F_{\min} = 0.0006$
		$t = 2 \text{ sec}$	$t = 609 \text{ sec}$
		$Nbre = 23319$	$Nbre = 18$
Goldstein	$(x_1, x_2) = (0, -1)$	$(x_1, x_2) = (0.02067, -0.98417)$	$(x_1, x_2) = (-0.015, -1.020)$
	$F_{\min} = 3$	$F_{\min} = 3.0038$	$F_{\min} = 3.0007$
		$t = 2 \text{ sec}$	$t = 240 \text{ sec}$
		$Nbre = 2006$	$Nbre = 17$
Branin	$(x_1, x_2) = (0.08983, -0.7126)$	$(x_1, x_2) = (0.08959, -0.71316)$	$(x_1, x_2) = (-0.1134, 0.7023)$
	$F_{\min} = -1.031628$	$F_{\min} = -1.031625$	$F_{\min} = -1.031$
		$t = 145 \text{ sec}$	$t = 459 \text{ sec}$
		$Nbre = 3113672$	$Nbre = 18$
Cos (2)	$(x_1, x_2) = (0, 0)$	$(x_1, x_2) = (-0.00001, -0.0003)$	$(x_1, x_2) = (-0.01, -0.005)$
	$F_{\min} = -0.2$	$F_{\min} = -0.199$	$F_{\min} = -0.19981$
		$t = 35 \text{ sec}$	$t = 1 \text{ sec}$
		$Nbre = 2591630$	$Nbre = 12$
Exp (2)	$(x_1, x_2) = (\pm 1, \pm 1)$	$(x_1, x_2) = (-1, -1)$	$(x_1, x_2) = (-1, -1)$
	$F_{\min} = 0.3678$	$F_{\min} = 0.36787$	$F_{\min} = 0.36787$
		$t = 4 \text{ sec}$	$t = 1 \text{ sec}$
		$Nbre = 448326$	$Nbre = 8$
Wood	$x_1 = 1$	$x_1 = 1.009$	$x_1 = 0.99$
	$x_2 = 1$	$x_2 = 1.023$	$x_2 = 0.99$
	$x_3 = 1$	$x_3 = 0.9888$	$x_3 = 0.99$
	$x_4 = 1$	$x_4 = 0.951$	$x_4 = 0.99$
	$F_{\min} = 0$	$F_{\min} = 0.00297$	$F_{\min} = 9.03 \times 10^{-9}$
		$t = 10 \text{ min}$	$t = 36 \text{ sec}$
		$Nbre = 24571189$	$Nbre = 16$

Fonctions	Solution exacte	<i>AS</i>	<i>BB-S</i>
Exp (10)	$x_1 = \pm 1$	$x_1 = -0.96$	$x_1 = -1$
	$x_2 = \pm 1$	$x_2 = -0.98$	$x_2 = -1$
	$x_3 = \pm 1$	$x_3 = -0.99$	$x_3 = -1$
	$x_4 = \pm 1$	$x_4 = -0.99$	$x_4 = -1$
	$x_5 = \pm 1$	$x_5 = 0.99$	$x_5 = -1$
	$x_6 = \pm 1$	$x_6 = 0.99$	$x_6 = -1$
	$x_7 = \pm 1$	$x_7 = 0.99$	$x_7 = -1$
	$x_8 = \pm 1$	$x_8 = 0.97$	$x_8 = -1$
	$x_9 = \pm 1$	$x_9 = 0.99$	$x_9 = -1$
	$x_{10} = \pm 1$	$x_{10} = 0.97$	$x_{10} = -1$
	$F_{\min} = 0.3678$	$F_{\min} = 0.377$	$F_{\min} = 0.3678$
		$t = 10 \text{ min}$	$t = 1 \text{ sec}$
		$Nbre = 1543919$	$Nbre = 3$
Cos (6)	$x_1 = 0$	$x_1 = 0.017$	$x_1 = -1.03 \times 10^{-5}$
	$x_2 = 0$	$x_2 = 0.028$	$x_2 = -1.03 \times 10^{-5}$
	$x_3 = 0$	$x_3 = -0.354$	$x_3 = -1.03 \times 10^{-5}$
	$x_4 = 0$	$x_4 = -0.021$	$x_4 = -1.03 \times 10^{-5}$
	$x_5 = 0$	$x_5 = -0.026$	$x_5 = -1.03 \times 10^{-5}$
	$x_6 = 0$	$x_6 = 0.032$	$x_6 = -1.03 \times 10^{-5}$
	$F_{\min} = -0.6$	$F_{\min} = -0.940$	$F_{\min} = -0.6$
		$t = 10 \text{ min}$	$t = 19 \text{ sec}$
		$Nbre = 17380198$	$Nbre = 15$

2.8.1 Etude comparative

Ces algorithmes ont été testés sur plusieurs exemples. La comparaison entre eux est basée sur : l'occupation de mémoire nécessaire, le temps de calcul, le nombre d'itérations et le nombre d'évaluations.

On constate que :

- lorsque nous augmentons la constante de Lipschitz de la fonction objectif le nombre d'itérations

augmente dans les deux algorithmes et nous arrivons à la même solution approchée après un temps très long. On doit pouvoir estimer la plus petite valeur de la constante de Lipschitz pour assurer la convergence en un nombre raisonnable de points d'évaluation de f , car si on prend une constante L inférieure à la plus petite constante de Lipschitz, les algorithmes ne convergent pas nécessairement vers le minimum global.

- La recherche à chaque itération d'un nouveau sommet d'un pavé de la partition et d'un point de la subdivision dans la méthode mixte **Branch and Bound-Shubert** augmente le temps de calcul par rapport à la méthode mixte **Aliénor-Shubert**.

- Dans la méthode hybride **Branch-and-Bound-Shubert**, le nombre d'évaluations de l'algorithme est beaucoup plus grand que le nombre d'itérations. Par contre, dans la méthode mixte **Aliénor-Shubert** le nombre d'itérations de l'algorithme est égal que le nombre d'évaluations.

- La taille de la mémoire nécessaire dans l'algorithme hybride **Branch-and-Bound-Shubert** est moindre que celle de la méthode mixte **Aliénor-Shubert**.

- Le temps de calcul est très grand pour les deux algorithmes ; ceci peut être expliqué par le choix de la constante de Lipschitz, le calcul de α_k , β_k et les sous-estimateurs de la fonction objectif.

- Malgré certains avantages, nous avons indiqué que plus la constante de Lipschitz est grande plus la convergence est lente. La convergence est quasiment immédiate quand la constante de Lipschitz est petite (Exp(2), Exp(10), Cos(2), Cos(6)).

- L'algorithme de **Piyavskii** converge bien vers l'optimum global, mais il est lent au voisinage de l'optimum.

Chapitre 3

Problèmes d'optimisation globale lipschitziennes avec contraintes

3.1 Introduction

Considérons le problème d'optimisation globale suivant :

$$\min_{x \in X} f(x). \quad (\mathbf{P})$$

Nous allons étudier le cas de la minimisation d'une fonction multi-extrémale lorsque les variables de X sont soumises à des contraintes.

Rappelons quelques classes importantes des problèmes d'optimisation globale avec contraintes :

1) **Les problèmes de minimisation concave** : sont des problèmes qui consistent à minimiser une fonction concave sur un ensemble compact convexe (ou à maximiser une fonction convexe sur un ensemble compact convexe).

Parfois, X est défini par des inégalités de la forme $g_i(x) \leq 0$ ($i = 1, \dots, m$), où g_i sont des fonctions convexes sur un ensemble contenant X .

La convexité de l'ensemble faisable X est bien exploitée dans la conception des algorithmes mais la propriété la plus intéressante est que la fonction concave atteint son minimum global sur la frontière de X .

Ces problèmes sont issus généralement de la recherche opérationnelle et des sciences écono-

miques. Cependant, le champ d'application des algorithmes correspondants est très large [18], [19]. Il existe même plusieurs modèles de problèmes importants dont la fonction objectif n'est pas nécessairement concave mais qui peut être transformée en fonction concave (Programmation Bilinéaire, Programmes Multiplicatifs,...).

2) **Les problèmes de minimisation d.c.** : Dans plusieurs domaines des mathématiques ou de la physique, on est amené à résoudre des programmes de mathématiques où la fonction objectif est d.c et les fonctions définissant les contraintes sont aussi d.c (i.e : différence de deux fonctions convexes)

Un problème d'optimisation globale est appelé problème de minimisation d.c., s'il est de la forme suivante :

$$\begin{cases} \min_{x \in C \cap X} f(x) \\ \text{tel que } X = \{x \in C : g_j(x) \leq 0, j = 1, \dots, m\} \end{cases}$$

où C est un convexe et les fonctions f, g_j ($j = 1, \dots, m$) sont d.c. sur C .

Pour résoudre des problèmes de programmation d.c, plusieurs travaux ont été réalisés par Pham Dinh Tao, Toland, Tuy...([19], [22], [36]) dans le but de trouver des conditions d'optimalité plus adaptées et de construire des algorithmes de résolution plus efficaces.

3) **Les problèmes de minimisation lipschitzienne** : sont des problèmes qui consistent à minimiser une fonction f lipschitzienne sur un sous-ensemble compact X de \mathbb{R}^n défini par un nombre fini de contraintes $g_j(x) \leq 0$, où g_j sont des fonctions lipschitziennes sur un ensemble compact P contenant X , i. e.,

$$\forall x, y \in P, \quad |g_i(x) - g_i(y)| \leq L_i \|x - y\|, \quad i = 1, \dots, k.$$

Pour ce problème avec contraintes, on peut utiliser l'algorithme mixte Branch-and-Bound-Piyavskii qu'on a vu dans le deuxième chapitre, on ajoutant deux tests :

i) Test d'élimination :

L'algorithme calcule une borne inférieure pour chaque contrainte dans le sous pavé P_j :

$$LB_{ij} = g_i(c_j) - \frac{L_i D_j}{2} \leq g_i(x), \forall x \in P_j$$

où c_j est le point milieu du sous pavé P_j et D_j la longueur de la diagonale de P_j .

Puis, il élimine les sous-problèmes qui correspondent à une région surface non admissible, c'est-à-dire les sous problèmes correspondant à $LB_{ij} > 0 \quad \forall i$.

ii) Test de mise à jour :

L'algorithme vérifie la nouvelle valeur de x_{opt} et la met à jour si elle satisfait les contraintes.

3.2 La méthode de la transformation réductrice et l'optimisation globale avec contraintes

Considérons le problème d'optimisation global suivant :

$$\min_{x \in X} f(x), \tag{P}$$

où f est une fonction lipschitzienne et X l'ensemble défini par les contraintes suivantes :

$$\left\{ \begin{array}{l} g_1(x_1, x_2, \dots, x_n) \leq 0 \\ g_2(x_1, x_2, \dots, x_n) \leq 0 \\ \vdots \\ g_m(x_1, x_2, \dots, x_n) \leq 0. \end{array} \right.$$

Dans un grand nombre de cas, l'ensemble faisable X peut être exprimé (après certaines transformations) sous la forme suivante :

$$X = \left\{ x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n / \begin{array}{l} a \leq x_1 \leq b \\ \varphi_2(x_1) \leq x_2 \leq \psi_2(x_1) \\ \vdots \\ \varphi_n(x_1, x_2, \dots, x_{n-1}) \leq x_n \leq \psi_n(x_1, x_2, \dots, x_{n-1}) \end{array} \right\}$$

avec $\varphi_2, \dots, \varphi_n, \psi_2, \dots, \psi_n$ sont des fonctions lipschitziennes de constantes de Lipschitz respectivement $l_2, \dots, l_n, L_2, \dots, L_n$ par rapport à la norme du maximum.

Pour résoudre le problème (P), nous allons utiliser la méthode de la transformation réductrice qui nous permet de ramener ce problème multidimensionnel à un problème unidimensionnel.

3.3 Construction des courbes α -denses dans un compact de \mathbb{R}^n

Commençons par construire des courbes α -denses sur des ensembles de type X . On donne des conditions suffisantes pour qu'une fonction h définie sur un intervalle fermé et borné A de \mathbb{R} , génère une courbe α -dense dans le compact X .

3.3.1 Construction des courbes α -denses à partir de certaines classes de fonctions

Notons par μ la mesure de Lebesgue. Le nombre α est supposé très petit par rapport à la longueur de l'intervalle $[a, b]$.

Pour $n = 2$, on a le théorème suivant :

Théorème 3.1. *Soient $h = (h_1, h_2) : A \rightarrow X$ une fonction continue et θ_2, α des nombres strictement positifs tels que :*

- 1) h_1 est surjective.
- 2) Pour tout intervalle I de longueur θ_2 , il existe $\theta'_2, \theta''_2 \in I$ tels que :

$$\begin{cases} h_2(\theta'_2) = \varphi_2(h_1(\theta'_2)) \\ h_2(\theta''_2) = \psi_2(h_1(\theta''_2)) \end{cases}$$

- 3) Pour tout intervalle J de A , on a :

$$\mu(J) < \theta_2 \Rightarrow \mu(h_1(J)) < \frac{\alpha}{\max(1, l_2, L_2)}.$$

Alors, la courbe paramétrée définie par $h(\theta) = (h_1(\theta), h_2(\theta))$ pour $\theta \in A$, est α -dense dans X .

Preuve. Soit $(x_1, x_2) \in X$. Considérons l'intervalle $\left[x_1 - \frac{\alpha}{\max(1, l_2, L_2)}, x_1 + \frac{\alpha}{\max(1, l_2, L_2)}\right]$. Du fait que la fonction h_1 est surjective, il existe un intervalle fermé $I_1 \subset A$ tel que :

$$h_1(I_1) = \left[x_1 - \frac{\alpha}{\max(1, l_2, L_2)}, x_1 + \frac{\alpha}{\max(1, l_2, L_2)}\right] \cap [a, b].$$

Mais $\mu(h_1(I_1)) \geq \frac{\alpha}{\max(1, l_2, L_2)}$, donc $\mu(I_1) \geq \theta_2$.

Ainsi, on a pour tout θ de I_1 : $|x_1 - h_1(\theta)| \leq \alpha$ et il existe $\theta'_2, \theta''_2 \in I_1$ tels que :

$$\begin{cases} h_2(\theta'_2) = \varphi_2(h_1(\theta'_2)) \\ h_2(\theta''_2) = \psi_2(h_1(\theta''_2)). \end{cases}$$

D'autre part, on a :

$$\varphi_2(x_1) \leq x_2 \leq \psi_2(x_1);$$

et puisque les fonctions φ_2, ψ_2 sont lipschitziennes de constantes l_2, L_2 respectivement, on montre que :

$$h_2(\theta'_2) - \alpha \leq x_2 \leq h_2(\theta''_2) + \alpha.$$

On a ainsi trois cas :

$$\begin{cases} (1) & h_2(\theta'_2) - \alpha \leq x_2 < h_2(\theta'_2) \\ (2) & h_2(\theta'_2) \leq x_2 \leq h_2(\theta''_2) \\ (3) & h_2(\theta'_2) < x_2 \leq h_2(\theta''_2) + \alpha. \end{cases}$$

i) Si c'est (1) qui est vérifiée, on a :

$$\left| x_2 - h_2(\theta'_2) \right| \leq \alpha,$$

d'où

$$\left\| x - h(\theta'_2) \right\|_{\max} = \max \left\{ \left| x_1 - h_1(\theta'_2) \right|, \left| x_2 - h_2(\theta'_2) \right| \right\} \leq \alpha.$$

ii) Si c'est (2) qui est vérifiée, il existe $\theta_0 \in [\theta'_2, \theta''_2]$ tel que :

$$x_2 = h_2(\theta_0),$$

d'où

$$\left\| x - h(\theta_0) \right\|_{\max} \leq \alpha.$$

iii) Si c'est (3) qui est vérifiée, on a :

$$\left| x_2 - h_2(\theta''_2) \right| \leq \alpha,$$

d'où

$$\left\| x - h(\theta_2'') \right\|_{\max} \leq \alpha.$$

Donc la courbe paramétrée définie par $h(\theta) = (h_1(\theta), h_2(\theta))$ pour $\theta \in A$, est α -dense dans X

. ■

Si les constantes de Lipschitz de $\varphi_2, \dots, \varphi_n, \psi_2, \dots, \psi_n$ sont choisies tels que $l_2 \leq l_3 \leq \dots \leq l_n$, $L_2 \leq L_3 \leq \dots \leq L_n$ et le nombre α est plus petit que la longueur de l'intervalle $\left[h_i(\theta_i'), h_i(\theta_i'') \right]$ pour tout $i = 2, \dots, n$. On a le théorème suivant pour $n \geq 2$.

Théorème 3.2. Soient $h = (h_1, \dots, h_n) : A \rightarrow X$ une fonction continue et $\theta_2, \dots, \theta_n, \alpha$ des nombres strictement positifs tels que :

1) h_1 est surjective.

2) Pour tout $i \in \{2, \dots, n\}$ et pour tout intervalle I de longueur θ_i , il existe θ_i' et $\theta_i'' \in I$ tels que :

$$\begin{cases} h_i(\theta_i') = \varphi_i(h_1(\theta_i'), \dots, h_{i-1}(\theta_i')) \\ h_i(\theta_i'') = \psi_i(h_1(\theta_i''), \dots, h_{i-1}(\theta_i'')) \end{cases}$$

3) Pour tout $i \in \{2, \dots, n\}$, et pour tout intervalle fermé J , on a :

$$\mu(J) < \theta_i \Rightarrow \mu(h_{i-1}(J)) < \alpha.$$

Alors, la courbe paramétrée définie par $h(\theta) = (h_1(\theta), \dots, h_n(\theta))$ pour $\theta \in A$, est $\eta\alpha$ -dense dans X , avec $\eta = \max\{1, l_n, L_n\} \left[1 + \sum_{i=2}^{n-1} \prod_{j=i}^{n-1} \max(l_j, L_j) \right]$.

Preuve. Soit $x = (x_1, \dots, x_n) \in X$. Considérons l'intervalle $[x_1 - \alpha, x_1 + \alpha]$, puisque h_1 est surjective, il existe un intervalle fermé $I_1 \subset A$ tel que :

$$h_1(I_1) = [x_1 - \alpha, x_1 + \alpha] \cap [a, b],$$

mais $\mu(h_1(I_1)) \geq \alpha$, donc $\mu(I_1) \geq \theta_2$.

Il en résulte que $|x_1 - h_1(\theta)| \leq \alpha$ pour tout $\theta \in I_1$. En outre, il existe $\theta_2', \theta_2'' \in I_1$ tels que :

$$\begin{cases} h_2(\theta_2') = \varphi_2(h_1(\theta_2')) \\ h_2(\theta_2'') = \psi_2(h_1(\theta_2'')) \end{cases}$$

Puisque $\left| h_2(\theta_2'') - h_2(\theta_2') \right| \geq \alpha$, on a

$$\mu\left(h_2\left([\theta_2', \theta_2'']\right)\right) \geq \alpha,$$

d'où

$$\mu\left([\theta_2', \theta_2'']\right) \geq \theta_3.$$

D'autre part, on a $\varphi_2(x_1) \leq x_2 \leq \psi_2(x_1)$,

et comme

$$\begin{cases} \left| \varphi_2(x_1) - \varphi_2\left(h_1\left(\theta_2'\right)\right) \right| \leq l_2 \left| x_1 - h_1\left(\theta_2'\right) \right| \leq l_2 \alpha \\ \left| \psi_2(x_1) - \psi_2\left(h_1\left(\theta_2''\right)\right) \right| \leq L_2 \left| x_1 - h_1\left(\theta_2''\right) \right| \leq L_2 \alpha, \end{cases}$$

alors

$$h_2\left(\theta_2'\right) - l_2 \alpha \leq x_2 \leq h_2\left(\theta_2''\right) + L_2 \alpha.$$

On a ainsi trois cas

$$\begin{cases} (1) & h_2\left(\theta_2'\right) - l_2 \alpha \leq x_2 \leq h_2\left(\theta_2'\right) \\ (2) & h_2\left(\theta_2'\right) \leq x_2 \leq h_2\left(\theta_2''\right) \\ (3) & h_2\left(\theta_2''\right) \leq x_2 \leq h_2\left(\theta_2''\right) + L_2 \alpha. \end{cases}$$

Si c'est (1) qui est vérifiée, on prend $I_2 = [\theta_2', \theta_2' + \theta_3]$.

Si c'est (2) qui est vérifiée, il existe $I_2 \subset [\theta_2', \theta_2'']$ tel que :

$$h_2(I_2) = [x_2 - \alpha, x_2 + \alpha] \cap [h_2(\theta_2'), h_2(\theta_2'')].$$

Si c'est (3) qui est vérifiée, on prend $I_2 = [\theta_2'' - \theta_3, \theta_2'']$.

Il est clair que $I_2 \subset [\theta_2', \theta_2'']$ et $\mu(I_2) \geq \theta_3$ dans les trois cas. En effet, pour (2) on a

$$\mu(h_2(I_2)) \geq \alpha,$$

et donc $\mu(I_2) \geq \theta_3$ d'après l'hypothèse 3) du théorème.

De plus, pour tout $\theta \in I_2$, on a :

i) Si (1) est vérifiée, alors

$$|x_2 - h_2(\theta)| \leq \left| x_2 - h_2(\theta'_2) \right| + \left| h_2(\theta'_2) - h_2(\theta) \right| \leq (1 + l_2) \alpha.$$

ii) Si (2) est vérifiée, alors

$$|x_2 - h_2(\theta)| \leq \alpha.$$

iii) Si (3) est vérifiée, alors

$$|x_2 - h_2(\theta)| \leq \left| x_2 - h_2(\theta''_2) \right| + \left| h_2(\theta''_2) - h_2(\theta) \right| \leq (1 + L_2) \alpha.$$

Donc pour tout $\theta \in I_2$, on a :

$$|x_2 - h_2(\theta)| \leq (1 + \max(l_2, L_2)) \alpha.$$

Ensuite, puisque $\mu(I_2) \geq \theta_3$ il existe $\theta'_3, \theta''_3 \in I_2$ tels que :

$$\begin{cases} h_3(\theta'_3) = \varphi_3(h_1(\theta'_3), h_2(\theta'_3)) \\ h_3(\theta''_3) = \psi_3(h_1(\theta''_3), h_2(\theta''_3)). \end{cases}$$

Et puisque

$$\left| h_3(\theta''_3) - h_3(\theta'_3) \right| \geq \alpha,$$

alors

$$\mu(h_3([\theta'_3, \theta''_3])) \geq \alpha.$$

Donc

$$\mu([\theta'_3, \theta''_3]) \geq \theta_4.$$

D'autre part, on a : $\varphi_3(x_1, x_2) \leq x_3 \leq \psi_3(x_1, x_2)$ et φ_3, ψ_3 sont des fonctions lipschitziennes,

ie :

$$\begin{aligned} \left| \varphi_3(x_1, x_2) - \varphi_3(h_1(\theta'_3), h_2(\theta'_3)) \right| &\leq l_3 \max \left\{ \left| x_1 - h_1(\theta'_3) \right|, \left| x_2 - h_2(\theta'_3) \right| \right\} \\ &\leq l_3 \{1 + \max(l_2, L_2)\} \alpha \end{aligned}$$

$$\begin{aligned} \left| \psi_3(x_1, x_2) - \psi_3\left(h_1(\theta_3''), h_2(\theta_3'')\right) \right| &\leq L_3 \max \left\{ \left| x_1 - h_1(\theta_3'') \right|, \left| x_2 - h_2(\theta_3'') \right| \right\} \\ &\leq L_3 \{1 + \max(l_2, L_2)\} \alpha. \end{aligned}$$

D'où

$$h_3(\theta_3') - l_3 \{1 + \max(l_2, L_2)\} \alpha \leq x_3 \leq h_3(\theta_3'') + L_3 \{1 + \max(l_2, L_2)\} \alpha.$$

On a ainsi trois cas :

$$\begin{cases} (1') & h_3(\theta_3') - l_3 \{1 + \max(l_2, L_2)\} \alpha \leq x_3 \leq h_3(\theta_3') \\ (2') & h_3(\theta_3') \leq x_3 \leq h_3(\theta_3'') \\ (3') & h_3(\theta_3'') \leq x_3 \leq h_3(\theta_3'') + L_3 \{1 + \max(l_2, L_2)\} \alpha. \end{cases}$$

Si c'est (1') qui est vérifiée, on prend $I_3 = [\theta_3', \theta_3' + \theta_4]$.

Si c'est (2') qui est vérifiée, il existe $I_3 \subset [\theta_3', \theta_3'']$ tel que :

$$h_3(I_3) = [x_3 - \alpha, x_3 + \alpha] \cap [h_3(\theta_3'), h_3(\theta_3'')].$$

Si c'est (3') qui est vérifiée, on prend $I_3 = [\theta_3'' - \theta_4, \theta_3'']$.

Donc, on a $I_3 \subset [\theta_3', \theta_3'']$ et $\mu(I_3) \geq \theta_4$ dans les trois cas. De plus, pour tout $\theta \in I_3$ on a :

1) Si (1') est vérifiée :

$$|x_3 - h_3(\theta)| \leq \left| x_3 - h_3(\theta_3') \right| + \left| h_3(\theta_3') - h_3(\theta) \right| \leq [1 + l_3(1 + \max(l_2, L_2))] \alpha.$$

2) Si (2') est vérifiée :

$$|x_3 - h_3(\theta)| \leq \alpha.$$

3) Si (3') est vérifiée :

$$|x_3 - h_3(\theta)| \leq \left| x_3 - h_3(\theta_3'') \right| + \left| h_3(\theta_3'') - h_3(\theta) \right| \leq [1 + L_3(1 + \max(l_2, L_2))] \alpha.$$

Donc pour tout $\theta \in I_3$, on a :

$$|x_3 - h_3(\theta)| \leq \{1 + \max(l_3, L_3) [1 + \max(l_2, L_2)]\} \alpha,$$

c'est-à-dire

$$|x_3 - h_3(\theta)| \leq \{1 + \max(l_3, L_3) + \max(l_3, L_3) \max(l_2, L_2)\} \alpha.$$

En procédant ainsi jusqu'à l'étape $(n-1)$, on obtient pour tout $\theta \in I_{n-1}$,

$$|x_{n-1} - h_{n-1}(\theta)| \leq \left[1 + \sum_{i=2}^{n-1} \prod_{j=i}^{n-1} \max(l_j, L_j) \right] \alpha,$$

avec $\mu(I_{n-1}) \geq \theta_n$ et $I_{n-1} \subset I_{n-2} \subset \dots \subset I_1 \subset A$.

Alors, il existe $\theta'_n, \theta''_n \in I_{n-1}$ tels que :

$$\begin{cases} h_n(\theta'_n) = \varphi_n(h_1(\theta'_n), \dots, h_{n-1}(\theta'_n)) \\ h_n(\theta''_n) = \psi_n(h_1(\theta''_n), \dots, h_{n-1}(\theta''_n)). \end{cases}$$

D'autre part, on a :

$$\varphi_n(x_1, \dots, x_{n-1}) \leq x_n \leq \psi_n(x_1, \dots, x_{n-1}),$$

et

$$\begin{aligned} & \left| \varphi_n(x_1, \dots, x_{n-1}) - \varphi_n(h_1(\theta'_n), \dots, h_{n-1}(\theta'_n)) \right| \\ & \leq l_n \max \left\{ |x_1 - h_1(\theta'_n)|, \dots, |x_{n-1} - h_{n-1}(\theta'_n)| \right\} \\ & \leq l_n \left[1 + \sum_{i=2}^{n-1} \prod_{j=i}^{n-1} \max(l_j, L_j) \right] \alpha \\ & \left| \psi_n(x_1, \dots, x_{n-1}) - \psi_n(h_1(\theta''_n), \dots, h_{n-1}(\theta''_n)) \right| \\ & \leq L_n \max \left\{ |x_1 - h_1(\theta''_n)|, \dots, |x_{n-1} - h_{n-1}(\theta''_n)| \right\} \\ & \leq L_n \left[1 + \sum_{i=2}^{n-1} \prod_{j=i}^{n-1} \max(l_j, L_j) \right] \alpha. \end{aligned}$$

Donc

$$h_n(\theta'_n) - l_n \left[1 + \sum_{i=2}^{n-1} \prod_{j=i}^{n-1} \max(l_j, L_j) \right] \alpha \leq x_n \leq h_n(\theta''_n) + L_n \left[1 + \sum_{i=2}^{n-1} \prod_{j=i}^{n-1} \max(l_j, L_j) \right] \alpha.$$

On a trois cas :

$$\begin{cases} (1'') & h_n(\theta'_n) - l_n \left[1 + \sum_{i=2}^{n-1} \prod_{j=i}^{n-1} \max(l_j, L_j) \right] \alpha \leq x_n \leq h_n(\theta'_n) \\ (2'') & h_n(\theta'_n) \leq x_n \leq h_n(\theta''_n) \\ (3'') & h_n(\theta''_n) \leq x_n \leq h_n(\theta''_n) + L_n \left[1 + \sum_{i=2}^{n-1} \prod_{j=i}^{n-1} \max(l_j, L_j) \right] \alpha. \end{cases}$$

Si c'est (1'') qui est vérifiée, on a :

$$\left| x_n - h_n(\theta'_n) \right| \leq l_n \left[1 + \sum_{i=2}^{n-1} \prod_{j=i}^{n-1} \max(l_j, L_j) \right] \alpha,$$

et

$$\begin{aligned} \|x - h(\theta'_n)\|_{\max} &= \max \left\{ |x_1 - h_1(\theta'_n)|, \dots, |x_{n-1} - h_{n-1}(\theta'_n)|, |x_n - h_n(\theta'_n)| \right\} \\ &\leq \max \left\{ \alpha, \dots, \left[1 + \sum_{i=2}^{n-1} \prod_{j=i}^{n-1} \max(l_j, L_j) \right] \alpha, l_n \left[1 + \sum_{i=2}^{n-1} \prod_{j=i}^{n-1} \max(l_j, L_j) \right] \alpha \right\} \\ &\leq \max \left\{ \left[1 + \sum_{i=2}^{n-1} \prod_{j=i}^{n-1} \max(l_j, L_j) \right] \alpha, l_n \left[1 + \sum_{i=2}^{n-1} \prod_{j=i}^{n-1} \max(l_j, L_j) \right] \alpha \right\}. \end{aligned}$$

Si c'est (2'') qui est vérifiée, il existe alors $\theta_0 \in I_{n-1}$ tel que $x_n = h_n(\theta_0)$ et on a :

$$\begin{aligned} \|x - h(\theta_0)\|_{\max} &= \max \{ |x_1 - h_1(\theta_0)|, \dots, |x_{n-1} - h_{n-1}(\theta_0)|, |x_n - h_n(\theta_0)| \} \\ &\leq \max \left\{ \alpha, [1 + \max(l_2, L_2)] \alpha, \dots, \left[1 + \sum_{i=2}^{n-1} \prod_{j=i}^{n-1} \max(l_j, L_j) \right] \alpha \right\} \\ &\leq \left[1 + \sum_{i=2}^{n-1} \prod_{j=i}^{n-1} \max(l_j, L_j) \right] \alpha \end{aligned}$$

Si c'est (3'') qui est vérifiée, on a :

$$\left| x_n - h_n(\theta''_n) \right| \leq L_n \left[1 + \sum_{i=2}^{n-1} \prod_{j=i}^{n-1} \max(l_j, L_j) \right] \alpha.$$

Ainsi

$$\begin{aligned}
\|x - h(\theta''_n)\|_{\max} &= \max \left\{ |x_1 - h_1(\theta''_n)|, \dots, |x_{n-1} - h_{n-1}(\theta''_n)|, |x_n - h_n(\theta''_n)| \right\} \\
&\leq \max \left\{ \alpha, \dots, \left[1 + \sum_{i=2}^{n-1} \prod_{j=i}^{n-1} \max(l_j, L_j) \right] \alpha, L_n \left[1 + \sum_{i=2}^{n-1} \prod_{j=i}^{n-1} \max(l_j, L_j) \right] \alpha \right\} \\
&\leq \max \left\{ \left[1 + \sum_{i=2}^{n-1} \prod_{j=i}^{n-1} \max(l_j, L_j) \right] \alpha, L_n \left[1 + \sum_{i=2}^{n-1} \prod_{j=i}^{n-1} \max(l_j, L_j) \right] \alpha \right\}.
\end{aligned}$$

Donc la courbe paramétrée définie par $h(\theta) = (h_1(\theta), \dots, h_n(\theta))$ pour $\theta \in A$, est $\eta\alpha$ -dense dans X , où $\eta = \max\{1, l_n, L_n\} \left[1 + \sum_{i=2}^{n-1} \prod_{j=i}^{n-1} \max(l_j, L_j) \right]$. ■

Remarque.

Si dans le théorème 3.2, on prend $\varphi_2, \dots, \varphi_n, \psi_2, \dots, \psi_n$ des fonctions lipschitziennes ayant la même constante de Lipschitz ℓ , la courbe paramétrée $h(\theta) = (h_1(\theta), \dots, h_n(\theta))$ est γ -dense dans X avec $\gamma = \max\{1, \ell\} \left[1 + \frac{\ell^{n-2} - 1}{\ell - 1} \right] \alpha$.

Si $\varphi_2, \dots, \varphi_n, \psi_2, \dots, \psi_n$ sont des fonctions contractantes par rapport à la norme du maximum.

On obtient le résultat plus simple suivant :

Corollaire 3.1. Soient $h = (h_1, \dots, h_n) : A \rightarrow X$ une fonction continue et $\theta_2, \dots, \theta_n, \alpha$ des nombres strictement positifs tels que :

1) h_1 est surjective.

2) Pour tout $i \in \{2, \dots, n\}$ et pour tout intervalle I de longueur θ_i , il existe θ'_i et $\theta''_i \in I$ tels que :

$$\begin{cases} h_i(\theta'_i) = \varphi_i(h_1(\theta'_i), \dots, h_{i-1}(\theta'_i)) \\ h_i(\theta''_i) = \psi_i(h_1(\theta''_i), \dots, h_{i-1}(\theta''_i)) \end{cases}$$

3) Pour tout $i \in \{2, \dots, n\}$ et pour tout intervalle fermé J , on a :

$$\mu(J) < \theta_i \Rightarrow \mu(h_{i-1}(J)) < \alpha.$$

Alors, la courbe définie par la représentation paramétrique $h(\theta) = (h_1(\theta), \dots, h_n(\theta))$ pour $\theta \in A$ est $(n-1)\alpha$ -dense dans X .

Preuve. C'est une conséquence du théorème 3.2, il suffit de prendre $l_i = L_i = 1$ pour tout $i = 2, \dots, n$. ■

Corollaire 3.2. Soit $h = (h_1, h_2, \dots, h_n)$, une fonction de A dans X , avec h_1, h_2, \dots, h_n des fonctions lipschitziennes de constantes c_1, c_2, \dots, c_n respectivement. Et soient $\theta_2, \theta_3, \dots, \theta_n, \alpha$ des

nombre strictement positifs tels que :

a) h_1 est surjective,

b) Pour tout $i \in \{2, \dots, n\}$ et pour tout intervalle I de longueur θ_i , il existe θ'_i et $\theta''_i \in I$ tels que :

$$\begin{cases} h_i(\theta'_i) = \varphi_i(h_1(\theta'_i), \dots, h_{i-1}(\theta'_i)). \\ h_i(\theta''_i) = \psi_i(h_1(\theta''_i), \dots, h_{i-1}(\theta''_i)). \end{cases}$$

c) Pour tout $i \in \{2, \dots, n\}$

$$c_{i-1} < \frac{\alpha}{\theta_i}.$$

Alors, la courbe paramétrée définie par $h(\theta) = (h_1(\theta), h_2(\theta), \dots, h_n(\theta))$ pour $\theta \in A$ est $(n-1)\alpha$ -dense dans X .

Preuve. C'est une conséquence du corollaire 3.1. Les conditions 1) et 2) du corollaire sont satisfaites. Il reste à prouver la condition 3). En effet, soient $i \in \{2, 3, \dots, n\}$ et I un intervalle de A vérifiant $\mu(I) < \theta_i$. Pour $\theta', \theta'' \in I$, on a :

$$\left| h_{i-1}(\theta'') - h_{i-1}(\theta') \right| \leq c_{i-1} \left| \theta'' - \theta' \right| \leq c_{i-1} \mu(I);$$

et donc

$$\mu(h_{i-1}(I)) \leq c_{i-1} \mu(I) < c_{i-1} \theta_i < \alpha. \quad \blacksquare$$

3.3.2 Construction des courbes α -denses à partir des fonctions périodiques

Théorème 3.3. *Considérons la fonction $h = (h_1, h_2, \dots, h_n) : A \rightarrow X$ et α un nombre strictement positif, tel que :*

1) h_1, \dots, h_n sont continues et surjectives ;

2) h_2, \dots, h_n sont périodiques de périodes $\theta_2, \theta_3, \dots, \theta_n$;

3) pour tout intervalle I de A , et pour tout $i \in \{2, 3, \dots, n\}$, on a :

$$\mu(I) < \theta_i \Rightarrow \mu(h_{i-1}(I)) < \alpha.$$

Alors la courbe définie par $h(\theta) = (h_1(\theta), h_2(\theta), \dots, h_n(\theta))$, pour $\theta \in A$, est $(n-1)\alpha$ -dense dans X .

Preuve. C'est une conséquence du corollaire 3.1. Pour tout $i \in \{1, \dots, n\}$, h_i est continue. Les conditions 1) et 3) du corollaire 3.2.26 sont satisfaites. Il reste à montrer la condition 2). En effet, pour tout $i \in \{2, \dots, n\}$, h_i étant surjective et périodique de période θ_i et pour tout $(x_1, x_2, \dots, x_n) \in X$:

$$\varphi_i(x_1, \dots, x_{i-1}) \leq x_i \leq \psi_i(x_1, \dots, x_{i-1}).$$

Alors, pour tout intervalle I de longueur θ_i il existe θ'_i et $\theta''_i \in I$ tels que :

$$\begin{cases} h_i(\theta'_i) = \varphi_i(h_1(\theta'_i), \dots, h_{i-1}(\theta'_i)) \\ h_i(\theta''_i) = \psi_i(h_1(\theta''_i), \dots, h_{i-1}(\theta''_i)). \end{cases}$$

Donc la courbe h est $(n-1)\alpha$ -dense dans X . ■

Corollaire 3.3. Soit $h = (h_1, h_2, \dots, h_n)$ une fonction continue définie de A dans X , avec h_1, h_2, \dots, h_n des fonctions dérivables et surjectives telles que :

- a) h_2, h_3, \dots, h_n sont périodiques, respectivement de périodes $\theta_2, \theta_3, \dots, \theta_n$;
- b) il existe un nombre $\alpha > 0$ tel que

$$\frac{d}{d\theta} h_{i-1}(\theta) < \frac{\alpha}{\theta_i} \quad \forall i = 2, 3, \dots, n \quad \text{et } \forall \theta \in A.$$

Alors la courbe définie par $h(\theta) = (h_1(\theta), h_2(\theta), \dots, h_n(\theta))$ pour $\theta \in A$, est $(n-1)\alpha$ -dense dans X .

Preuve. On applique le théorème 3.3. Il suffit de vérifier que pour tout intervalle I de A , et pour tout $i \in \{2, 3, \dots, n\}$:

$$\mu(I) < \theta_i \Rightarrow \mu(h_{i-1}(I)) < \alpha.$$

Soit $i \in \{2, 3, \dots, n\}$ et I intervalle de A , tel que $\mu(I) < \theta_i$. En utilisant le théorème des valeurs intermédiaires, on montre facilement que $\forall \theta', \theta'' \in I$, on a :

$$\left| h_{i-1}(\theta'') - h_{i-1}(\theta') \right| \leq \frac{\alpha}{\theta_i} \mu(I).$$

Donc

$$\mu(h_{i-1}(I)) \leq \frac{\alpha}{\theta_i} \mu(I) < \alpha. \quad \blacksquare$$

3.3.3 Génération concrète des courbes α -denses dans des compacts de \mathbb{R}^n

Corollaire 3.4. On considère $X = \{(x_1, x_2) \in \mathbb{R}^2 \mid a \leq x_1 \leq b$

$$\varphi_2(x_1) \leq x_2 \leq \psi_2(x_1)\}$$

où φ_2, ψ_2 sont des fonctions contractantes, si on prend la fonction $h = (h_1, h_2) : \left[0, \frac{\pi}{\alpha_1}\right] \rightarrow X$ définie par :

$$h_1(\theta) = \frac{a-b}{2} \cos(\alpha_1 \theta) + \frac{a+b}{2}$$

$$h_2(\theta) = \frac{\varphi_2(h_1(\theta)) - \psi_2(h_1(\theta))}{2} \cos(\alpha_2 \theta) + \frac{\varphi_2(h_1(\theta)) + \psi_2(h_1(\theta))}{2},$$

où $\alpha_1, \alpha_2, \alpha$ des nombres strictement positifs tels que :

$$\alpha_2 = \frac{\pi(b-a)}{2\alpha} \alpha_1.$$

La courbe paramétrée $h(\theta) = (h_1(\theta), h_2(\theta))$ est α -dense dans le compact X .

Preuve. Ce résultat est une conséquence du théorème 3.1. En effet :

1) h_1 est surjective par définition.

2) Si on prend $\theta_2 = \frac{\pi}{\alpha_2}$, alors pour tout intervalle I de longueur θ_2 , il existe $\theta', \theta'' \in I$ tels que :

$$\begin{cases} h_2(\theta') = \varphi_2(h_1(\theta')) \\ h_2(\theta'') = \psi_2(h_1(\theta'')) \end{cases}$$

$$\begin{cases} h_2(\theta') = \varphi_2(h_1(\theta')) \\ h_2(\theta'') = \psi_2(h_1(\theta'')) \end{cases} \Rightarrow \begin{cases} \cos(\alpha_2 \theta') = 1 \\ \cos(\alpha_2 \theta'') = -1 \end{cases}$$

$$\Rightarrow \begin{cases} \theta' = \frac{2\pi k}{\alpha_2} \\ \theta'' = \frac{\pi}{\alpha_2} + \frac{2\pi k}{\alpha_2} \end{cases}, \quad k = 0, 1, 2, \dots$$

3) Pour tout intervalle I vérifiant $\mu(I) < \frac{\pi}{\alpha_2}$, on a $\mu(h_1(I)) < \alpha$:

$\forall \theta', \theta'' \in I$:

$$\begin{aligned}
|h_1(\theta'') - h_1(\theta')| &= \left| \frac{a-b}{2} \cos(\alpha_1 \theta'') - \frac{a-b}{2} \cos(\alpha_1 \theta') \right| \\
&= \left| \frac{a-b}{2} \right| |\cos(\alpha_1 \theta'') - \cos(\alpha_1 \theta')| \\
&= \frac{b-a}{2} \left| -2 \sin \frac{\alpha_1 \theta'' + \alpha_1 \theta'}{2} \sin \frac{\alpha_1 \theta'' - \alpha_1 \theta'}{2} \right| \\
&\leq (b-a) \left| \sin \frac{\alpha_1 \theta'' - \alpha_1 \theta'}{2} \right| \\
&\leq \frac{(b-a)}{2} \alpha_1 |\theta'' - \theta'| < \frac{(b-a)\pi}{2\alpha_2} \alpha_1 = \alpha.
\end{aligned}$$

D'après le théorème 3.2., la courbe paramétrée définie par la fonction h est α -dense dans X . ■

Corollaire.3.5. *On considère le compact*

$$X = \left\{ (x_1, x_2, \dots, x_n) \in \mathbb{R}^n / \begin{array}{l} a \leq x_1 \leq b \\ \varphi_2(x_1) \leq x_2 \leq \psi_2(x_1) \\ \dots \\ \varphi_n(x_1, x_2, \dots, x_{n-1}) \leq x_n \leq \psi_n(x_1, x_2, \dots, x_{n-1}) \end{array} \right\}$$

avec $\varphi_2, \dots, \varphi_n, \psi_2, \dots, \psi_n$ des fonctions contractantes par rapport à la norme du maximum; et

soit la fonction $h = (h_1, h_2, \dots, h_n) : [0, \frac{\pi}{\alpha_1}] \rightarrow X$, définie par :

$$\begin{aligned}
h_1(\theta) &= \frac{a-b}{2} \cos(\alpha_1 \theta) + \frac{a+b}{2} \\
h_2(\theta) &= \frac{\varphi_2(h_1(\theta)) - \psi_2(h_1(\theta))}{2} \cos(\alpha_2 \theta) + \frac{\varphi_2(h_1(\theta)) + \psi_2(h_1(\theta))}{2} \\
&\dots \\
h_n(\theta) &= \frac{\varphi_n(h_1(\theta), \dots, h_{n-1}(\theta)) - \psi_n(h_1(\theta), \dots, h_{n-1}(\theta))}{2} \cos(\alpha_n \theta) + \\
&\quad \frac{\varphi_n(h_1(\theta), \dots, h_{n-1}(\theta)) + \psi_n(h_1(\theta), \dots, h_{n-1}(\theta))}{2}
\end{aligned}$$

où $\alpha_1, \dots, \alpha_n, \alpha$ sont des nombres strictement positifs vérifiant la relation

$$\alpha_i \geq \frac{\pi}{2\alpha} \left[\sum_{k=1}^{i-1} 2^{(i-k-1)} (M_k - m_k) \alpha_k \right],$$

avec $m_k = \min \varphi_k$, $M_k = \max \psi_k$ et $m_1 = a$, $M_1 = b$.

Alors, la courbe paramétrée $h(\theta) = (h_1(\theta), \dots, h_n(\theta))$ est $(n-1)\alpha$ -dense dans le compact X .

Preuve. Ce résultat est une conséquence du théorème 3.2. En effet :

1) h_1 est surjective par définition.

2) Pour tout $i = 2, \dots, n$ et pour tout intervalle I de longueur $\theta_i = \frac{\pi}{\alpha_i}$, il existe $\theta'_i = \frac{2\pi k}{\alpha_i}$ et

$\theta''_i = \frac{\pi}{\alpha_i} + \frac{2\pi k}{\alpha_i} \in I$; $k \in \mathbb{N}$, tels que :

$$\begin{cases} h_i(\theta') = \varphi_i(h_1(\theta'), \dots, h_{i-1}(\theta')) \\ h_i(\theta'') = \psi_i(h_1(\theta''), \dots, h_{i-1}(\theta'')) \end{cases}$$

3) Pour tout $i = 2, \dots, n-1$, et pour tout intervalle fermé J ; on a :

$$\mu(J) < \theta_i \Rightarrow \mu(h_{i-1}(J)) < \alpha.$$

On procède par récurrence sur i . Pour $i = 2$, l'inégalité est vérifiée.

Supposons que l'inégalité est vérifiée pour i , c'est-à-dire; pour tout intervalle fermé J de longueur inférieure à $\frac{\pi}{\alpha_i}$, on a :

$$\forall \theta', \theta'' \in J : \left| h_{i-1}(\theta'') - h_{i-1}(\theta') \right| < \left[\sum_{k=1}^{i-1} 2^{(i-k-1)} \left(\frac{M_k - m_k}{2} \right) \alpha_k \right] \left| \theta'' - \theta' \right| \leq \alpha.$$

Puis montrons-la pour $i+1$. Soit l'intervalle fermé J , tel que :

$$\mu(J) < \frac{\pi}{\alpha_{i+1}}, \text{ alors}$$

$$\forall \theta', \theta'' \in J : \left| h_i(\theta'') - h_i(\theta') \right| < \left[\sum_{k=1}^i 2^{(i-k)} \left(\frac{M_k - m_k}{2} \right) \alpha_k \right] \left| \theta'' - \theta' \right|.$$

En effet, on pose :

$$\begin{cases} \xi_i(\theta) = \varphi_i(h_1(\theta), \dots, h_{i-1}(\theta)) \\ \gamma_i(\theta) = \psi_i(h_1(\theta), \dots, h_{i-1}(\theta)) \end{cases},$$

et

$$h_i(\theta) = \frac{\xi_i(\theta) - \gamma_i(\theta)}{2} \cos(\alpha_i \theta) + \frac{\xi_i(\theta) + \gamma_i(\theta)}{2}.$$

Pour $\theta', \theta'' \in J$, on a :

$$\begin{aligned} & \left| h_i(\theta'') - h_i(\theta') \right| = \\ & \left| \frac{\xi_i(\theta'') - \gamma_i(\theta'')}{2} \cos(\alpha_i \theta'') + \frac{\xi_i(\theta'') + \gamma_i(\theta'')}{2} - \frac{\xi_i(\theta') - \gamma_i(\theta')}{2} \cos(\alpha_i \theta') - \frac{\xi_i(\theta') + \gamma_i(\theta')}{2} \right| \\ & = \left| \frac{\xi_i(\theta'') - \gamma_i(\theta'')}{2} \cos(\alpha_i \theta'') + \frac{\xi_i(\theta'') + \gamma_i(\theta'')}{2} - \frac{\xi_i(\theta') - \gamma_i(\theta')}{2} \cos(\alpha_i \theta') - \frac{\xi_i(\theta') + \gamma_i(\theta')}{2} \right| \end{aligned}$$

$$\begin{aligned}
& + \frac{\xi_i(\theta'') - \gamma_i(\theta'')}{2} \cos(\alpha_i \theta') - \frac{\xi_i(\theta'') - \gamma_i(\theta'')}{2} \cos(\alpha_i \theta') \Big| \\
& \leq \left| \frac{\xi_i(\theta'') - \gamma_i(\theta'')}{2} \right| \left| \cos(\alpha_i \theta'') - \cos(\alpha_i \theta') \right| + \left| \xi_i(\theta'') - \xi_i(\theta') \right| + \left| \gamma_i(\theta'') - \gamma_i(\theta') \right| \\
& \leq \left(\frac{M_i - m_i}{2} \right) \alpha_i \left| \theta'' - \theta' \right| + 2 \max \left\{ \left| h_1(\theta'') - h_1(\theta') \right|, \dots, \left| h_{i-1}(\theta'') - h_{i-1}(\theta') \right| \right\}. \text{ (puisque } \\
& \varphi_i \text{ et } \psi_i \text{ sont contractantes par rapport à la norme du maximum).} \\
& \leq \left(\frac{M_i - m_i}{2} \right) \alpha_i \left| \theta'' - \theta' \right| + 2 \left[\sum_{k=1}^{i-1} 2^{(i-k-1)} \left(\frac{M_k - m_k}{2} \right) \alpha_k \right] \left| \theta'' - \theta' \right| = \left[\sum_{k=1}^i 2^{(i-k)} \left(\frac{M_k - m_k}{2} \right) \alpha_k \right] \left| \theta'' - \theta' \right| \\
& < \left[\sum_{k=1}^i 2^{(i-k)} \left(\frac{M_k - m_k}{2} \right) \alpha_k \right] \frac{\pi}{\alpha_{i+1}} < \alpha
\end{aligned}$$

Alors, la courbe paramétrée définie par la fonction h est $(n-1)\alpha$ -dense dans X .

3.3.4 Exemple de courbes α -dense

Exemple 3.1. Déterminons une courbe α -dense dans la boule unité de \mathbb{R}^3 ,

$$X = \{(x_1, x_2, x_3) \in \mathbb{R}^3 / x_1^2 + x_2^2 + x_3^2 \leq 1\}.$$

X est la boule de centre $(0, 0, 0)$ et de rayon 1 de \mathbb{R}^3 . X peut s'écrire sous la forme suivante :

$$X = \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 / \begin{array}{l} -1 \leq x_1 \leq 1 \\ -\sqrt{1-x_1^2} \leq x_2 \leq \sqrt{1-x_1^2} \\ -\sqrt{1-x_1^2-x_2^2} \leq x_3 \leq \sqrt{1-x_1^2-x_2^2} \end{array} \right\}$$

La courbe paramétrée $h(\theta)$ ($\theta \in [0, \frac{\pi}{\alpha_1}]$) définie par :

$$\begin{cases} h_1(\theta) = -(1-\alpha) \cos(\alpha_1 \theta), \\ h_2(\theta) = -(1-\alpha) \cos(\alpha_2 \theta) \sqrt{1 - (1-\alpha)^2 \cos^2(\alpha_1 \theta)}, \\ h_3(\theta) = -\cos(\alpha_3 \theta) \sqrt{\left(1 - (1-\alpha)^2 \cos^2(\alpha_1 \theta)\right) \left(1 - (1-\alpha)^2 \cos^2(\alpha_2 \theta)\right)}, \end{cases}$$

où $\alpha_1, \alpha_2, \alpha_3, \alpha$ sont des nombres strictement positifs tels que :

$$\begin{aligned}
\alpha_2 &= \frac{\pi}{\alpha} (1-\alpha) \alpha_1 \\
\alpha_3 &= \frac{\pi}{\alpha} (1-\alpha) \left[\frac{(1-\alpha)^2}{\sqrt{\alpha(2-\alpha)}} \alpha_1 + \alpha_2 \right]
\end{aligned}$$

est γ -dense dans la boule X , où $\gamma = \max \left\{ 1, \frac{(1-\alpha)}{\alpha(2-\alpha)} \right\} \left[1 + \frac{(1-\alpha)^2}{\sqrt{\alpha(2-\alpha)}} \right] \alpha$.

Preuve. Ce résultat est une conséquence du théorème 3.2. Il suffit de prendre $\theta_2 = \frac{\pi}{\alpha_2}$, $\theta_3 = \frac{\pi}{\alpha_3}$, $\varphi_2(x_1) = -(1-\alpha) \sqrt{1-x_1^2}$, $\psi_2(x_1) = (1-\alpha) \sqrt{1-x_1^2}$, $\varphi_3(x_1, x_2) = -\sqrt{1-x_1^2-x_2^2}$, $\psi_3(x_1, x_2) = \sqrt{1-x_1^2-x_2^2}$, et $-(1-\alpha) \leq x_1 \leq (1-\alpha)$, on a :

φ_2, ψ_2 sont des fonctions lipschitziennes de constante $\frac{(1-\alpha)^2}{\sqrt{\alpha(2-\alpha)}}$, et φ_3, ψ_3 sont des fonctions lipschitziennes de constante $\frac{(1-\alpha)}{\alpha(2-\alpha)}$.

i) h_1 est surjective par définition.

ii) Pour tout $i \in \{2, 3\}$ et pour tout intervalle I de longueur θ_i , il existe $\theta'_i = \frac{2\pi k}{\alpha_i}, \theta''_i = \frac{\pi}{\alpha_i} + \frac{2\pi k}{\alpha_i} \in I$ tels que :

$$\begin{cases} h_i(\theta'_i) = \varphi_i(h_1(\theta'_i), \dots, h_{i-1}(\theta'_i)). \\ h_i(\theta''_i) = \psi_i(h_1(\theta''_i), \dots, h_{i-1}(\theta''_i)). \end{cases}$$

iii) Pour tout $i \in \{2, 3\}$, et pour tout intervalle fermé J , on a :

$$\mu(J) < \theta_i \Rightarrow \mu(h_{i-1}(J)) < \alpha.$$

En effet, $\forall \theta', \theta'' \in J$:

$$\begin{aligned} |h_1(\theta'') - h_1(\theta')| &= \left| -(1-\alpha) \cos(\alpha_1 \theta'') + (1-\alpha) \cos(\alpha_1 \theta') \right| \\ &\leq \left| -2(1-\alpha) \sin\left(\alpha_1 \frac{\theta'' + \theta'}{2}\right) \sin\left(\alpha_1 \frac{\theta'' - \theta'}{2}\right) \right| \\ &\leq (1-\alpha) \alpha_1 |\theta'' - \theta'| \\ &< (1-\alpha) \alpha_1 \theta_2 = \alpha \end{aligned}$$

$$\begin{aligned} \forall \theta', \theta'' \in J : & \left| h_2(\theta'') - h_2(\theta') \right| \\ &= (1-\alpha) \left| -\cos(\alpha_2 \theta'') \sqrt{1 - (1-\alpha)^2 \cos^2(\alpha_1 \theta'')} + \cos(\alpha_2 \theta') \sqrt{1 - (1-\alpha)^2 \cos^2(\alpha_1 \theta')} \right| \\ &\leq (1-\alpha) \left| -\cos(\alpha_2 \theta'') \sqrt{1 - (1-\alpha)^2 \cos^2(\alpha_1 \theta'')} + \cos(\alpha_2 \theta') \sqrt{1 - (1-\alpha)^2 \cos^2(\alpha_1 \theta')} \right. \\ &\quad \left. - \cos(\alpha_2 \theta') \sqrt{1 - (1-\alpha)^2 \cos^2(\alpha_1 \theta'')} + \cos(\alpha_2 \theta'') \sqrt{1 - (1-\alpha)^2 \cos^2(\alpha_1 \theta'')} \right| \\ &\leq (1-\alpha) \left[\left| \cos(\alpha_2 \theta'') - \cos(\alpha_2 \theta') \right| + \left| \sqrt{1 - (1-\alpha)^2 \cos^2(\alpha_1 \theta'')} - \sqrt{1 - (1-\alpha)^2 \cos^2(\alpha_1 \theta')} \right| \right] \\ &\leq (1-\alpha) \left[|\theta'' - \theta'| + \frac{(1-\alpha)^3}{\sqrt{\alpha(2-\alpha)}} \alpha_1 |\theta'' - \theta'| \right] \\ &< (1-\alpha) \left[\alpha_2 + \frac{(1-\alpha)^2}{\sqrt{(2-\alpha)\alpha}} \alpha_1 \right] \theta_3 = \alpha. \end{aligned}$$

Donc, d'après le théorème 3.2, la courbe paramétrée définie par la fonction $h(\theta)$ est γ -dense dans X , où $\gamma = \max \left\{ 1, \frac{(1-\alpha)}{\alpha(2-\alpha)} \right\} \left[1 + \frac{(1-\alpha)^2}{\sqrt{\alpha(2-\alpha)}} \right] \alpha$.

3.4 Transformation à une dimension de quelques problèmes d'optimisation globale avec contraintes par Alienor

3.4.1 Quelques problèmes tests

Problème 1 : Fonction cos2

Considérons le problème d'optimisation globale suivant :

$$\min_{x \in X} f(x) = x_1^2 + x_2^2 - \frac{1}{10} (\cos(5\pi x_1) + \cos(5\pi x_2)),$$

avec X l'ensemble des points $x = (x_1, x_2) \in \mathbb{R}^2$ vérifiant :

$$\begin{aligned} -1 &\leq x_1 \leq 1 \\ \frac{1}{2}x_1 - \frac{1}{2} &\leq x_2 \leq \frac{1}{2}x_1 + \frac{1}{2} \end{aligned}$$

Cette fonction admet plusieurs minima locaux dans X . Le minimum global est $f^* = -0.2$ est atteint au point $x^* = (0, 0)$.

On construit une courbe $h(\theta)$ α -dense dans le compact X , comme suit :

$$\begin{cases} h_1(\theta) = -\cos(\alpha_1\theta) \\ h_2(\theta) = -\frac{1}{2}(\cos(\alpha_1\theta) + \cos(\alpha_2\theta)) \end{cases}$$

avec $\alpha_2 \geq \frac{2\pi}{\alpha}\alpha_1$ (pour n'importe quelle valeur de α_1).

Alors, notre problème se transforme au problème unidimensionnel suivant :

$$\min_{\theta \in [0, \frac{\pi}{\alpha_1}]} f^*(\theta),$$

avec $f^*(\theta) = f(h(\theta))$.

Problème 2 : Fonction de Branin

Considérons le problème d'optimisation globale suivant :

$$\min_{x \in X} f(x) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4$$

avec X est l'ensemble des points $x = (x_1, x_2) \in \mathbb{R}^2$ vérifiant :

$$\begin{aligned} -1 &\leq x_1 \leq 1 \\ -\cos(x_1 - 6) - \frac{1}{3} &\leq x_2 \leq \ln(x_1 + 2) \end{aligned}$$

Cette fonction est symétrique par rapport à l'origine et admet plusieurs minima locaux (> 3) dans X . Le minimum global est $f^* = -1.031628$, il est atteint aux points $x^* = (0.08983, -0.7126)$ et $x^* = (-0.08983, 0.7126)$

On construit une courbe $h(\theta)$ α -dense dans le compact X comme suit :

$$\begin{aligned} h_1(\theta) &= -\cos(\alpha_1\theta) \\ h_2(\theta) &= -\frac{1}{2} \left(\cos(\cos(\alpha_1\theta) + 6) + \ln(2 - \cos(\alpha_1\theta)) + \frac{1}{3} \right) \cos(\alpha_2\theta) + \\ &\quad \frac{1}{2} \left(\ln(2 - \cos(\alpha_1\theta)) - \cos(\cos(\alpha_1\theta) + 6) - \frac{1}{3} \right) \end{aligned}$$

où $\theta \in \left[0, \frac{\pi}{\alpha_1}\right]$, et α_1, α_2 sont des nombres strictement positifs tels que :

$$\alpha_2 \geq \frac{2\pi}{\alpha} \alpha_1.$$

Problème 3 : Fonction de Rastrigin

Considérons le problème d'optimisation global suivant :

$$\min_{x \in X} f(x) = x_1^2 + x_2^4 - \cos(18x_1) - \cos(18x_2)$$

où X est l'ensemble des points $x = (x_1, x_2) \in \mathbb{R}^2$ vérifiant :

$$\begin{aligned} -1 &\leq x_1 \leq 1 \\ -\ln(x + 2) + 1/3 &\leq x_2 \leq \sqrt{x_1^2 + 4} \end{aligned}$$

Cette fonction admet au moins 50 minima locaux dans X . Le minimum global est $f^* = -2$ il est atteint au point $x^* = (0, 0)$

On construit la courbe $h(\theta)$ α -dense dans le compact X comme suit :

$$\begin{aligned} h_1(\theta) &= -\cos(\alpha_1\theta) \\ h_2(\theta) &= -\frac{1}{2} \left(\ln(2 - \cos(\alpha_1\theta)) + \sqrt{\cos^2(\alpha_1\theta) + 4} - \frac{1}{3} \right) \cos(\alpha_2\theta) + \\ &\quad \frac{1}{2} \left(\sqrt{\cos^2(\alpha_1\theta) + 4} - \ln(2 - \cos(\alpha_1\theta)) + \frac{1}{3} \right) \end{aligned}$$

avec $\theta \in \left[0, \frac{\pi}{\alpha_1}\right]$, et α_1, α_2 sont des nombres strictement positifs tels que :

$$\alpha_2 \geq \frac{2\pi}{\alpha} \alpha_1.$$

Problème 4 : Fonction ExpCos

Considérons le problème d'optimisation global suivant :

$$\min_{x \in X} f(x) = \exp\left(\frac{x_1^2 + x_2^2}{2}\right) \cos(3\pi x_1) \cos(3\pi x_2)$$

où X est l'ensemble défini par les contraintes suivantes :

$$\begin{aligned} -1 &\leq x_1 \leq 1 \\ -\frac{1}{4}x_1^2 + \frac{1}{2}x_1 - 1 &\leq x_2 \leq \frac{1}{4}x_1^2 + \frac{1}{2}x_1 + 1 \end{aligned}$$

Cette fonction admet plusieurs minima locaux dans X , le minimum global est $f^* = -2.4624$, il est atteint au point $x^* = (-1, -1.2475)$.

On construit la courbe $h(\theta)$ α -dense dans le compact X comme suit :

$$\begin{aligned} h_1(\theta) &= -\cos(\alpha_1\theta) \\ h_2(\theta) &= -\frac{1}{4}(\cos^2(\alpha_1\theta) + 4) \cos(\alpha_2\theta) - \frac{1}{2} \cos(\alpha_1\theta) \end{aligned}$$

où $\theta \in \left[0, \frac{\pi}{\alpha_1}\right]$, et α_1, α_2 sont des nombres strictement positifs tels que :

$$\alpha_2 \geq \frac{2\pi}{\alpha} \alpha_1.$$

Problème 5 : Fonction Cos6

Considérons le problème d'optimisation globale :

$$\min_{x \in X} f(x) = \sum_{i=1}^6 \left(x_i^2 - \frac{1}{10} \cos(5\pi x_i) \right),$$

avec X la boule unité (de centre 0 et rayon 1) de \mathbb{R}^6 .

On construit une courbe $h(\theta)$ α -dense dans le compact suivant :

$$S = \left\{ (x_1, \dots, x_6) \in \mathbb{R}^6 / \begin{array}{l} -(\frac{1}{2} - \alpha) \leq x_1 \leq (\frac{1}{2} - \alpha) \\ -(\frac{1}{2} - \alpha) \sqrt{\frac{1}{2} - \sum_{i=1}^{k-1} x_i^2} \leq x_k \leq (\frac{1}{2} - \alpha) \sqrt{\frac{1}{2} - \sum_{i=1}^{k-1} x_i^2} \text{ pour } k = 2, \dots, 5 \\ -\sqrt{\frac{1}{2} - \sum_{i=1}^5 x_i^2} \leq x_6 \leq \sqrt{\frac{1}{2} - \sum_{i=1}^5 x_i^2} \end{array} \right\}$$

Cette fonction admet plusieurs minima locaux dans S . Le minimum global est $f^* = -0.6$, il est atteint au point $x^* = (0, 0, 0, 0, 0, 0)$

Notre problème se transforme au problème unidimensionnel suivant :

$$\min_{\theta \in [0, \frac{\pi}{\alpha_1}]} f^*(\theta),$$

Avec $f^*(\theta) = f(h(\theta))$, où $h(\theta)$ est la courbe paramétrée donnée au corollaire 3.5.

3.4.2 Résultats numériques

Dans cette partie, nous présentons les résultats numériques des problèmes 1-5. Nous avons appliqué la méthode unidimensionnelle de Piyavskii-Shubert, car elle est la plus connue et la plus étudiée dans la littérature, mais d'autres méthodes unidimensionnelles peuvent aussi bien être utilisées. L'important est de résoudre une nouvelle classe de problèmes d'optimisation globale sous contraintes. Nous effectuons l'algorithme de Piyavskii-shubert avec une précision de $\varepsilon = 0.1$.

Tableau des résultats des tests pour les problèmes de 1-5

Problème	un minimiseur global	Le minimum global	Nombres d'évaluations
1	(0.0048, -0.0010)	-0.1996	1271
2	(0.033, -0.724)	-1.0172	4230
3	(0.015, 0.34)	-1.8347	410
4	(-1, -1.246)	-2.4207	257
5	(0, 0, 0, 0, 0)	-0.6	129

3.5 Conclusion

Les résultats présentés précédemment ont élargi la classe des fonctions permettant de générer des courbes α -denses dans un compact quelconque de \mathbb{R}^n [53], [54], [55]. Ils donnent des nouvelles possibilités à la méthode de la transformation réductrice **Aliénor** pour résoudre les problèmes d'optimisation globale avec contraintes [15], [19]. Nous avons donné des conditions suffisantes sur les coordonnées de la paramétrisation pour obtenir de nouvelles familles de courbes α -denses d'expression analytique plus simple qui améliorent le temps de calcul dans la méthode **Aliénor** et qui évitent ainsi les difficultés pratiques concernant les raccordements sur les éléments d'un compact de \mathbb{R}^n .

Conclusion générale

La plupart des problèmes d'optimisation globale sont de type NP-difficile. Pour résoudre ces problèmes il nous faut certaines hypothèses qui doivent être exigées sur la fonction objectif (différentiabilité, condition de Lipschitz,...). La difficulté devient beaucoup plus considérable lorsqu'on a besoin du minimum global d'une fonction à plusieurs variables, non convexe ou non différentiable.

Dans la première partie de cette thèse on a étudié quelques méthodes de recouvrement pour résoudre des problèmes d'optimisation des fonctions lipschitziennes sans contraintes. Nous avons concentré notre étude à la résolution de problèmes explicites, c'est-à-dire des problèmes dans lesquels toutes les expressions de la fonction objectif et des contraintes sont connues de façon explicite.

Dans le cas unidimensionnel, ces méthodes sont simples, ont été profondément étudiées et leur efficacité a été prouvée. Dans cette partie, on a étudié et développé essentiellement deux méthodes de recouvrement, présentées dans le cas unidimensionnel et dans le cas des fonctions lipschitziennes à plusieurs variables, définies sur un pavé $P = \prod_{i=1}^n [a_i, b_i]$ de \mathbb{R}^n et à valeurs dans \mathbb{R} . Parmi ces méthodes il y a la méthode de Piyavskii, très utilisée dans la littérature ; elle est basée sur la construction des fonctions minorantes de la fonction objectif sur le domaine faisable P . Mais, cette méthode est numériquement inefficace pour $n \geq 3$, car elle est encombrante et trop compliquée pour la réalisation pratique. En plus, le temps de calcul dépend considérablement du choix de la constante de Lipschitz et d'autres calculs auxiliaires. En effet, dans le cas où la constante de Lipschitz est trop grande, les fonctions minorantes ne donnent que très peu d'informations sur la fonction objectif f . En plus les choses se compliquent dès que l'on passe aux grandes dimensions, car l'effort de calcul des sous problèmes correspondants est énorme. Pour cela, nous avons considéré la technique de partition et d'élimination (**Branch-and-Bound**), en appliquant essentiellement la version unidimensionnelle de la méthode de Piyavskii à la diagonale principale des sous pavés de P .

Concernant la méthode de la transformation réductrice, nous avons apporté des nouvelles idées pour pouvoir appliquer cette méthode aux problèmes d'optimisation des fonctions lipschitziennes à plusieurs variables. Le principe fondamental de cette méthode consiste à effectuer une transformation qui permet de ramener le problème d'optimisation multidimensionnel à un pro-

blème unidimensionnel afin d'appliquer les méthodes d'optimisation plus efficaces au cas d'une seule variable. Ainsi notre problème est ramené à un problème plus facile à résoudre car il y a une seule direction à explorer.

Nous avons couplé la méthode de **Piyavskii** d'une part avec la nouvelle variante d'**Aliénor** [56], et d'autre part avec l'algorithme de **Branch-and-Bound** [18, 19]. Il semble à première vue qu'il y a une relation entre les deux, en effet, le dernier algorithme est appliqué sur un hyper-rectangle mais il procède à une minimisation unidimensionnelle de la fonction objectif sur la diagonale principale de l'hyper-rectangle, et l'algorithme **Aliénor-Piyavskii** procède à une minimisation unidimensionnelle de la fonction objectif sur une courbe α -dense qui vérifie certaines propriétés dans l'hyper-rectangle. Les algorithmes obtenus à partir de ces couplages sont simples et efficaces, ils n'exigent pas de la fonction objectif d'être dérivable mais seulement lipschitzienne. Par ailleurs, la convergence est quasiment immédiate si la constante de Lipschitz est relativement petite.

La deuxième partie de cette thèse avait pour but de développer et d'étendre la méthode de la transformation réductrice **Aliénor**. Les résultats présentés ont élargi la classe des fonctions permettant de générer des courbes α -denses dans un compact quelconque de \mathbb{R}^n [52], [53], [54]. Il donnent des nouvelles possibilités à la méthode d'**Aliénor** pour résoudre les problèmes d'optimisation globale avec contraintes [15], [19]. Nous avons donné des conditions suffisantes sur les coordonnées de la paramétrisation pour obtenir de nouvelles familles de courbes α -denses d'expression analytique plus simple qui améliorent le temps de calcul dans la méthode **Aliénor** et qui évitent ainsi les difficultés pratique concernant les raccordements sur les éléments d'un compact de \mathbb{R}^n .

En perspective, beaucoup de questions peuvent être envisagées :

- Dans l'optimisation globale, le temps de calcul nécessaire pour atteindre le minimum global représente le critère d'efficacité le plus commun d'une méthode. Ce temps dans la plupart des cas, est lié au nombre d'évaluations de la fonction objectif obtenues après la transformation qui dépend de la longueur de la courbe α -dense. De ce fait, un intérêt particulier est accordé au choix de la courbe α -densifiant le domaine faisable. Il est donc naturel d'étudier d'autres aspects des courbes α -denses qui nous permettraient, peut être, de construire des transformations réductrices plus simples pour améliorer le temps de calcul.

- Une autre voie consiste aussi à chercher des courbes α -denses dans des compacts de \mathbb{R}^n de structures plus compliquées pour pouvoir appliquer cette approche à des problèmes d'optimisation globale avec contraintes.

Bibliographie

- [1] H. Ammar, and Y. Cherruault, *Approximation of a Several Variables function by a one Variable function and Applications to Global Optimization*, Math. Comput. Modelling, Vol. 18, No. 2, pp. 17-21. (1993).
- [2] W. Baritomba, *Customized Method for Global Optimization-A Geometric View point*, J. of Global Optimization 3, pp. 337-358. (1993).
- [3] W. Baritomba and A. Cutler, *Acceleration for global optimization covering methods using second derivatives*, J. of Global Optimization 4, pp. 329-341. (1994).
- [4] O. Bendiab, and Y. Cherruault, *Divanu : A New Method for Global Optimization in Dimension n*, International J. of Bio-Medical Computing, No. 38, pp. 177-80. (1995).
- [5] R. P. Brent, *Algorithm for Minimization without Derivatives*, Prentice-Hall, Englewood Cliffs, New Jersey.195p, (1973).
- [6] Y. Cherruault and A. Guillez, *Une méthode pour la recherche du minimum global d'une fonctionnelle*, C.R.A.S, Paris, Série I, Vol. 296, pp. 175-178. (1983).
- [7] Y. Cherruault, *Modèles et Méthodes mathématiques pour les sciences du vivant*, P.-U.-F. (1998).
- [8] Y. Cherruault, *Optimisation, Méthodes locales et globales*, P.-U.-F. (1999).
- [9] Y. Cherruault et G. Mora, *Optimisation Globale, Théorie des courbes α -denses*, Ed Economica. (2005).
- [10] G. Choquet, *Cours d'Analyse Tome 2, Topologie*, Masson, Paris. (1973).
- [11] J. C. Culioli, *Introduction à l'optimisation*, Ecole Nationale Supérieure des Mines de Paris, Ellipses. (1994).

- [12] E. A. Galperin and Q. Zheng, *Nonlinear observation via global optimization methods : measure theory approach*, J. Optimiz. Theory and Applic., 54, No. 1, 63-92. (1987).
- [13] V. P. Gergel, *A Global Optimization Algorithm for Multivariate functions with Lipschitzian First Derivatives*, J. of Global Optimization, 10, No. 3, pp. 256-281. (1997).
- [14] E. Gourdin, B. Jaumard, and R. Ellaia , *Global Optimization of Hölder function*, J. of Global Optimization, Vol. 8, pp. 323-348. (1996).
- [15] D. Guettal and A. Ziadi, *Reducing transformation and global optimization*, Applied Mathematics and Computation. 218(2012), 5848-5860.
- [16] E. Hansen, *Global Optimization using Interval analysis, the multidimensionnal case*, Numer. Math., 34, pp. 247-270. (1980).
- [17] J. B. Hiriart-Urruty, *L'Optimisation : deux ou trois choses que je sais d'elle*, Math. Appli. et Applications des maths, pp. 31-53. (2003).
- [18] R. HORST, *A General class of branch-and-bound methods in global optimization with some new approach for concave minimization*, J. Optimz. Theory and Applic., 51, No.2, pp. 271-291,1986.
- [19] R. Horst and H. Tuy, *Global Optimization, Deterministic Approach*, Springer-Verlag, Berlin. (1993).
- [20] R. Horst and P. M. Pardalos, *Handbook of Global Optimization*, Kluwer Academic Publishers, Dordrecht. (1995).
- [21] R. Horst, P. M. Pardalos and N. V. Thoai, *Introduction to Global Optimization*, Kluwer Academic Publishers, Dordrecht. (1995).
- [22] L. K. HUA and H. TUY, *Applications of Number Theory to Numerical Analysis*, springer, Berlin, (1993).
- [23] B. Jaumard, T. Hermann and H. Ribault, *An on-line Intersection Algorithm for Global Optimization of Multivariate Lipschitz Functions*. (1996).
- [24] D. R. Jones, C. D. Perttunen, and B. E. Stuckmann. *Lipschitzian optimization without the lipschitz constant*, Journal of Optimization Theory and Applications, 79 :157, October (1993).

- [25] S. Kiatsupaibul and R. L. Smith, *On the solution of infinite horizon optimization problems through global optimization*, Tech. Report 98-19, DIOE, University of Michigan, Ann Arbor.(1998).
- [26] D. Lera and Ya. D. Sergeyev, *Global Minimization Algorithms for Hölder functions*, BIT, Vol 42, No. 1, pp. 119-133. (2002).
- [27] A. V. Levy and A. Montalvo, *The tunneling algorithm for the global minimization of functions*, SIAM J. Sci. Statist. Comput., 6, No. 1, 15-29, (1985).
- [28] E. Lutton and J. Levy Vehel, *Some remarks on the optimization of Hölder functions with Genetic Algorithms*, Rapport de recherche No 2627 , INRIA. (1995).
- [29] D. Q. Mayne and E. Polak, *Outer Approximation Algorithm for Nondifferentiable Optimization Problem*, J. of Optimiz. Theory and applic. 42(1), pp. 19–30. (1985).
- [30] C. C. Meewella and D. Q. Mayne, *Efficient Domain Partition in Algorithm for Global Optimization of Rational and Lipschitz Continuous functions*, J. of Optimiz. Theory and applic. 61(2), pp. 247-270. (1988).
- [31] F. Messine, *Méthodes d'optimisation globale basées sur l'analyse d'intervalle pour la résolution de problèmes avec contraintes*. Thèse de Ph. D., Institut National Polytechnique de Toulouse. (1997).
- [32] R. H. Mladineo, *An algorithm for finding the Global Maximum of a Multimodal, Multivariate function*, Mathematical Programming 34, pp. 188-200. (1986).
- [33] R. E. Moore, *Interval analysis*, Prentice Hall, Inc. Englewood Cliffs, N. J. (1966).
- [34] G. Mora and Y. Cherruault, *Characterization and Generation of α -dense Curves*, Comp. and math. with Applic. 33(9), pp. 83-91. (1997).
- [35] G. Mora and Y. Cherruault, *The theoretic calculation time associated with α -dense Curves*, Kybernetes, 27(8-9), pp. 919–39. (1998).
- [36] D. PHAM and TAO-EL BERNOUSSI, *D.C Optimization. Algorithms for solving the trust region subproblems*. SIAM. Optim. pp. 476-505. (1998).
- [37] J. Pinter, *Extented Univariate Algorithms for n-dimensional Global Optimization*, Computing 36, pp. 91-103. (1986).

- [38] S. A. Piyavskii, *An algorithm for finding the absolute minimum for a function*, Theory of Optimal Solution, No. 2, pp. 13-24.(1967).
- [39] M. Rahal. *Extention de certaines méthodes de recouvrement en optimisation globale*. Thèse de Doctorat en sciences, Université Ferhat Abbas, Sétif. (2009).
- [40] M. Rahal and A. Ziadi, *A new extension of Piyavskii's method to Hölder functions of several variables*, Applied mathematics and Computation. Vol. 197, pp. 478-488. (2008).
- [41] M. Rahal et A. Ziadi, *Optimisation globale des fonctions höldériennes à plusieurs variables*, Actes du Colloque International sur l'Optimisation et les Systèmes d'Information, COSI'08. Proceeding, pp. 299-313. Tizi-Ouzou, (2008).
- [42] N. Revol, *Introduction à l'arithmétique par intervalle*, Rapport de recherche, INRIA, FRANCE, (2004).
- [43] H. Sagan, *Space-Filling Curves*, Springer, New york, (1994).
- [44] B. O. Shubert, *A sequential method seeking the global maximum of a function*, SIAM, Journal of Numerical Analysis, Vol. 3, pp. 379-388, (1972).
- [45] R. G. Strongin, *Algorithm for multiextremal mathematical programming problems employing the set of joint space-filling curves*, J. of Global Optimization, vol. 2, pp. 357-378. (1992).
- [46] R. G. Strongin and Ya. D. Sergeyev, *Global multidimensional optimization on parallel computer*, Parallel Computing, vol. 18, pp. 1259-1273. (1992).
- [47] A. Törn, and A. Zilinska, *Global Optimization*, Springer-Verlag, Lecture Notes in Computer Sciences, 350, (1989).
- [48] R. J. Vanderbei, *Extension of Piyavskii's Algorithm to Continuous Global Optimization*, Journal of Global Optimization, Vol. 14, pp. 205-216. (1999).
- [49] G. R. Wood, *Multidimensional bisection applied to global optimization*, Comp. and math. with Applic. 21, pp. 161-172. (1991).
- [50] A. A. Zhigljavsky, *Theory of Global Random Search*, Kluwer Academic Publishers, Dordrecht, (1991).
- [51] A. Ziadi, *Optimisation Globale : Contribution à l'étude de la méthode de la transformation réductrice Alienor*. Thèse de Doctorat d'Etat, Université Ferhat Abbas, Sétif, (2000).

- [52] A. Ziadi, Y. Cherruault, and G. Mora, *The existence of α -dense Curves with minimal length in a metric space*, Kybernetes, Vol. 29, No. 2, pp. 219-230. (2000).
- [53] A. Ziadi, and Y. Cherruault, *Generation of α -dense Curves in a cube of \mathbb{R}^n* , Kybernetes, Vol. 27, No. 4, pp. 416-425. (1998).
- [54] A. Ziadi, and Y. Cherruault, *Generation of α -dense Curves and Applications to Global Optimization*, Kybernetes, Vol. 29, No.1, pp. 71-82. (2000).
- [55] A. Ziadi, Y. Cherruault and G. Mora, *Global Optimization, a New Variant of the Alienor Method*, Comp. Math. Appl., Vol. 41, pp. 63-71. (2001).
- [56] A. Ziadi, D. Guettal and Y. Cherruault, *Global Optimization : Alienor mixed method with Piyavskii-Shubert technique*, Kybernetes, Vol. 34, No.7/8, pp.1049–1058. (2005).

Résumé :

Dans la première partie de cette thèse, nous étudions quelques méthodes d'optimisation globale multidimensionnelles sans contraintes en utilisant la technique Branch-and-Bound. Cette étude permet d'améliorer l'efficacité et la fiabilité de ces méthodes et de déterminer les classes de problèmes pour lesquelles elles seraient efficaces.

La deuxième partie a pour but de présenter et d'appliquer une nouvelle approche d'optimisation globale basée sur la réduction de la dimension. Il s'agit de la méthode de la transformation réductrice Alienor qui utilise les courbes α -denses. Le domaine faisable dans cette partie est défini à partir des contraintes non linéaires.

Mots clés Optimisation globale, Méthode Branch-and-Bound, méthode de Piyavskii, Méthode de la transformation réductrice Alienor, Courbes α -denses

Abstract.

In the first part of this thesis, we study some unconstrained multidimensional global optimization methods using the Branch-and-Bound technique. This study improves the efficiency and reliability of these methods and determines the classes of problems for which they would be effective.

The goal of the second part is to present and implement a new approach of global optimization, based on the reduction of the dimension. It involves the Alienor reducing transformation method which uses α -dense curves. The feasible domain in this part is defined from the non-linear constrained.

Key-words: Global optimization, Branch-and-Bound method, Piyavskii's method, Alienor Reducing transformation method, α -Dense curves.

ملخص.

في الجزء الأول من هذه الأطروحة ندرس بعض طرق التجويد الشامل ذات عدة متغيرات وبدون قيود باستخدام تقنية "برانش اند باوند"، هذه الدراسة تسمح بتحسين فعالية و جودة هذه الطرق وتحديد أصناف المسائل التي من أجلها تكون فعالة.

يهدف الجزء الثاني إلى تقديم و تطبيق طريقة جديدة للتجويد الشامل و التي تركز على تقليص البعد. نقصد بها طريقة التحويل التقليصي "الينور" التي تستعمل المنحنيات α -كثيفة. ميدان التجويد في هذا الجزء معرف عن طريق قيود غير خطية.

كلمات فاتحة. التجويد الشامل تقنية برانش اند باوند، طريقة بيافسكي، طريقة التحويل التقليصي أليينور، المنحنيات α -كثيفة.
