PEOPELE'S DEMOCRATIC REPUBLIC OF ALGERIA
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

Faculty of Sciences

Global Optimization

Lecture Notes

Prepared at Department of Mathematics

By

Dr. RAHAL Mohamed

Academic year 2023-2024

Contents

General Introduction

Global optimization is a field with active research. It is the process of finding the global extremum of a function of n variables, with the possibility of being subjected to some constraints. Its importance is due to the increasing needs in many applications in sciences and engineering. Two factors render the global optimization of certain classes of multivariate multi-extremal functions difficult to treat: local irregularity, for example non differentiability and the existence of a large number of local and global extrema of the objective function in the feasible area. One of the key challenges in global optimization is the presence of multiple local optima, which can mislead traditional optimization algorithms into converging to suboptimal solutions. To address this challenge, global optimization algorithms employ various strategies to explore the solution space more thoroughly and efficiently, aiming to find the global optimum or approximate it within a specified tolerance. During the last recent years, numerous works have been realized concerning Lipschitz functions [5, 6, 10, 14]. More recently, some works tackling less regular functions have appeared [4].
Several techniques are used in global optimization, including:
Deterministic methods: These methods systematically explore the solution space to guarantee convergence to the global optimum under certain conditions. Examples include branch and bound, interval analysis, and outer approximation methods.
Stochastic methods: These methods use randomness to explore the solution space, often based on probabilistic principles. Examples include simulated annealing, genetic algorithms, particle swarm optimization, and evolutionary algorithms.
In this work we will consider only Lipschitz and Holder continuous objective functions of which the Lipschitz and the Holder constants are known.
This document is primarily aimed at Master's students in mathematics specializing in Modeling and Decision Support, as well as any student working on global optimization. Chapter 1 introduces some important terminology and definitions and describes the field of local optimization. Also we give some reminders of optimal conditions for an optimization problem without constraints, thus some classical approached algorithms to determine the solution of this problem.
In chapter 2, we are interested in deterministic global optimization methods. Our attention will be paid to one -dimensional covering methods which have the reputation of being effective in dimension 1. Among these methods, there are those based on the use of linear or non-linear support functions, others on covering the feasible domain. We will finish this chapter with a series of exercises taken from Literature and others proposed by the author. The third chapter is devoted to the extension of certain covering methods to mul-

tidimensional cases without and with constraints. Two essential methods were presented. The first method based on a technique for partitioning and eliminating the regions of the feasible set not containing the global minimum known by Branch-And-Bound.
A second method is of Alienor reducing transformation, is presented.
The main idea in this method consists of approximating the objective function of several variables defined on a compact set C of $\mathbb{R}^n$, by a function of a single variable by densifying the feasible set C using a simple curve . This makes it possible to reduce the multidimensional optimization problem to a one-dimensional optimization problem, to be able to use the one-dimensional methods seen in the second chapter. Finally, we will end this chapter with a series of numerical exercises.

Chapter 1

General Review of Classical Optimization

Contents

1.1 Introduction

Optimization involves the search for the minimum (or maximum) of a certain quantity, called cost or objective function [20]. In this document, we will assume that the cost depends on n ($n \geq 1$) real variables, gathered in a vector $x = (x_1, ..., x_n) \in \mathbb{R}^n$, which yield a value $f(x)$ where f is a function from $\mathbb{R}^n$ to $\mathbb{R}$. Generally, the variables $x_1, ..., x_n$ will not be allowed to take on any values, but must satisfy constraints represented by a non-empty subset $C \subset \mathbb{R}^n$. We will write the optimization problems in the following general form:

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

1.2 Local Minimum and Global Minimum

Local Minimum. A point $x^* \in C$ is said to be a local minimizer (or local minimum point) of the function f in a subset C if there exists a neighborhood $V(x^*, r)$ of x^* such that

$$f(x^*) \leq f(x), \quad \forall x \in C \cap V(x^*, r).$$

Recall that a neighborhood $V(x^*, r)$ of center x^* and radius $r > 0$ is defined by:

$$V(x^*, r) = \{x \in \mathbb{R}^n / \quad \|x - x^*\| \leq r\}.$$

Global Minimum A point $x^* \in C$ is called a global minimizer (or global minimum point) of the function f in C if:

$$f(x^*) \leq f(x), \quad \forall x \in C.$$

∘ Every global minimum is a local minimum.

∘ $f(x^*)$ is a minimum of f on C.
∘ A maximization problem can always be reduced to a minimization problem through $\max f(x) = -\min(-f(x))$.

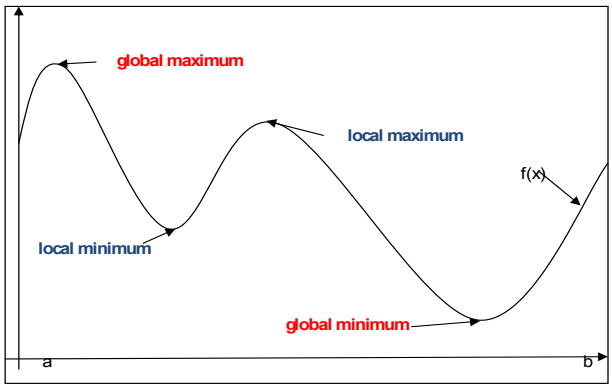

Figure 1.1: Local and Global Minima and Maxima for f(x) on [a,b]

1.3 General Theorems of Existence and Uniqueness

Theorem 1.3.1. [4] *(Weierstrass) If C is a non-empty closed and bounded (compact) set in $\mathbb{R}^n$ and $f : C \to \mathbb{R}$, a continuous function on C, then the function f is bounded and attains its bounds, meaning f has at least one global minimum and at least one global maximum in C.*

Theorem 1.3.2. *If C is a non-empty closed but unbounded set in $\mathbb{R}^n$ and $f : C \to \mathbb{R}$, a continuous function on C.*
If $\lim\limits_{\substack{\|x\| \to +\infty \\ x \in C}} f(x) = +\infty$ (f is coercive on C).
Then f has at least one global minimum x^* in C.

∘ *The uniqueness of the solution generally results from properties of convexity of f and C.*

1.4 Optimality Conditions Case $C = \mathbb{R}^n$

If the exact solution of problem (P) is obtained using the optimality conditions given below, then the resolution method is analytical.

1.4.1 Necessary Conditions

Result 1. $(C.N)$ [20] Suppose a function $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable in the neighborhood of a point x^*. Then:

$$x^* \text{ is a minimum or maximum of } f \text{ on } \mathbb{R}^n \Rightarrow \nabla f(x^*) = 0_{\mathbb{R}^n}.$$

Result 2. $(C.N)$ Suppose a function $f : \mathbb{R}^n \to \mathbb{R}$ is twice differentiable in the neighborhood of a point x^*. Then:

$$x^* \text{ is a minimum of } f \text{ on } \mathbb{R}^n \Rightarrow \begin{cases} \nabla f(x^*) = 0_{\mathbb{R}^n} \\ \nabla^2 f(x^*) \text{ is a positive semidefinite matrix} \end{cases}$$

1.4.2 Sufficient Conditions

Result 3. $(C.S)$ Suppose a function $f : \mathbb{R}^n \to \mathbb{R}$ is twice differentiable in the neighborhood of a point x^*. Then:

$$\text{if } \begin{cases} \nabla f(x^*) = 0_{\mathbb{R}^n} \\ \nabla^2 f(x^*) \text{ is a positive definite matrix} \end{cases} \Rightarrow x^* \text{ is a local minimum of } f \text{ on } \mathbb{R}^n.$$

Result 4. $(C.N.S)$ If $f : \mathbb{R}^n \to \mathbb{R}$ is convex and differentiable on $\mathbb{R}^n$. Then,

$$x^* \text{ is a global minimum of } f \text{ on } \mathbb{R}^n \Leftrightarrow \nabla f(x^*) = 0_{\mathbb{R}^n}.$$

1.5 Approximate Methods for Solving the Problem

1.5.1 Descent Methods

The algorithms presented in this subsection allow for an approximate calculation of the solution to problem (P).
General principle. Given a point x and its value $f(x)$ (f assumed to be of class $C^1(\mathbb{R}^n)$), by locally moving from x (determining whether f increases or decreases), the simplest movement being in a straight line (in Euclidean space), small displacements will be made in a certain number of directions d, aiming to improve the value of f. [20].

Definition 1.5.1. Let d be any direction and $\alpha > 0$, if we have:

$$f(x + \alpha d) \leq f(x), \forall x \in \mathbb{R}^n,$$

then d is called a descent direction.
α : the step size
d : displacement direction or descent direction.

The principle of descent methods is to find a stationary (critical) point $\overline{x} \in \mathbb{R}^n (\nabla f(\overline{x}) = 0)$ by constructing a sequence of vectors $\left\{x^k\right\}, k \in \mathbb{N}$ from a starting point x^0, defined by the following general form:

$$\begin{cases} x^0 \text{ given} \\ x^{k+1} = x^k + \alpha_k d_k, \qquad k \in \mathbb{N} \end{cases}$$

with d_k the descent direction for f satisfying the following condition:

$$f(x^{k+1}) \leq f(x^k), \ \forall \ k \in \mathbb{N}$$

α_k : the displacement step from x^k to x^{k+1}.

Choice of d_k.

If the vector d_k is given in terms of $\nabla f(x^k)$, the method is said to be of first order.
If d_k is given in terms of $\nabla f(x^k)$ and $\nabla^2 f(x^k)$, the method is said to be of second order.

1.5.2 Gradient Method

From the first-order Taylor expansion, we have:

$$f(x^k + \alpha_k d_k) = f(x^k) + \alpha_k \langle \nabla f(x^k), d_k \rangle + O(\alpha_k d_k).$$

Assuming that $\nabla f(x^k) \neq \overrightarrow{o}$ then:

$$f(x^k) - f(x^k + \alpha_k d_k) = -\alpha_k \langle \nabla f(x^k), d_k \rangle + O(\alpha_k d_k).$$

We deduce that each vector d_k such that $\langle \nabla f(x^k), d_k \rangle < 0$ is a descent direction of f in the neighborhood of x^k, since $f(x^k + \alpha_k d_k) < f(x^k)$.
If we take $d_k = -\nabla f(x^k) \Rightarrow \langle \nabla f(x^k), -\nabla f(x^k) \rangle = -\left\|\nabla f(x^k)\right\| < 0$, then $-\nabla f(x^k)$ is a descent direction, and thus we have the general gradient algorithm [4]:

Gradient Algorithm
1) Set $k = 0$, $x^0 \in \mathbb{R}^n$, an arbitrarily chosen initial point, ε : a given precision.
2) While convergence test not satisfied
a) Find a descent direction by calculating $d_k = -\nabla f(x^k)$.
b) Choose a step $\alpha_k > 0$ so that $f(x^k + \alpha_k d_k) < f(x^k)$.
c) Determine $x^{k+1} = x^k + \alpha_k d_k$.
d) Set $k = k + 1$
End while
3) Stop if $\left\|\nabla f(x^k)\right\| \leq \varepsilon$ or $\left|f(x^{k+1}) - f(x^k)\right| \leq \varepsilon$ or $\left\|x^{k+1} - x^k\right\| \leq \varepsilon$
ε is a given precision.

Choice of α_k (the step)

a) **Fixed-Step Gradient Algorithm**
It consists of taking a constant step $\alpha_k = \alpha > 0$ for all iterations, i.e.:

$$x^{k+1} = x^k - \alpha \nabla f(x^k), \quad \alpha > 0.$$

The determination of the parameter α is done manually. If α is chosen too large, the algorithm diverges, if α is too small, convergence is very slow.

1.5.3 Examples

Example 1. Consider the minimization problem:

$$(P)\ \begin{cases} \min f(x,y) = 2(x^2+y^2) - 3xy \\ (x,y)^t \in \mathbb{R}^2 \end{cases}$$

1- Apply 4 iterations using the fixed-step gradient method with $\alpha = \frac{1}{25}$ starting from $(x^0,y^0)^t = (1,1)^t$.
2- Calculate the exact solution $\overline{X}$ of problem (P).
3- Deduce the error $\varepsilon_4 = \left\|\overline{X} - X^{(4)}\right\|$.

Solution
1- f is continuous and differentiable on $\mathbb{R}^2$. Since $\mathbb{R}^2$ is non-compact, it suffices to justify that the function f is coercive on $\mathbb{R}^2$ (hence problem (P) admits at least one global solution, for uniqueness it suffices to see if f is strictly convex on $\mathbb{R}^2$.
The sequence of the fixed-step gradient method is given by:

$$\begin{cases} (x^0,y^0)^t = (1,1)^t \\ X^{k+1} = X^k - \alpha\nabla f(X^k), \quad \alpha = \frac{1}{25}, X = (x,y)^t \end{cases}$$

$\nabla f(x,y) = (4x-3y, -3y+4y)^t$ hence:
$\begin{cases} (x^0,y^0)^t = (1,1)^t \\ X^{k+1} = \begin{pmatrix} x^{k+1} \\ y^{k+1} \end{pmatrix} = \begin{pmatrix} x^k \\ y^k \end{pmatrix} - \frac{1}{25}\begin{pmatrix} 4x^k - 3y^k \\ -3y^k + 4y^k \end{pmatrix} \end{cases}$
i.e. $\begin{cases} x^{k+1} = \frac{21}{25}x^k + \frac{3}{25}y^k \\ y^{k+1} = \frac{3}{25}x^k + \frac{21}{25}y^k \end{cases}$
For
$k=0 \Rightarrow (x^0,y^0)^t = (1,1)^t$
$k=1 \Rightarrow (x^1,y^1)^t = \left(\frac{24}{25},\frac{24}{25}\right)^t = \frac{24}{25}(1,1)^t$
$k=2 \Rightarrow (x^2,y^2)^t = \left(\frac{24}{25}\right)^2(1,1)^t$
$k=3 \Rightarrow (x^3,y^3)^t = \left(\frac{24}{25}\right)^3(1,1)^t$
$k=4 \Rightarrow (x^4,y^4)^t = \left(\frac{24}{25}\right)^4(1,1)^t$

2- Analytical method. We have $\nabla^2 f(x,y) = \begin{pmatrix} 4 & -3 \\ -3 & 4 \end{pmatrix}$ and the principal minors are $D_1 = 4 > 0$, $D_2 = 7 > 0 \Rightarrow \nabla^2 f(x,y)$ is positive definite $\forall (x,y) \in \mathbb{R}^2 \Rightarrow f$ is strictly convex on $\mathbb{R}^2$, thus:

$$\overline{X} \text{ is a solution of } (P) \overset{C.N.S}{\Leftrightarrow} \nabla f(\overline{X}) = 0_{\mathbb{R}^2} \Leftrightarrow \overline{X} = (0,0)^t$$

So $\overline{X} = (0,0)^t$ is the unique global solution (global minimum) of (P).

3- The error $\varepsilon_4 = \left\|\overline{X} - X^{(4)}\right\| = \left\|\left(\frac{24}{25}\right)^4(1,1)^t - (0,0)^t\right\| = \sqrt{2}\left(\frac{24}{25}\right)^4 \approx 1.2.$

Remark 1.5.1. 1- As $(x^k,y^k)^t = \left(\frac{24}{25}\right)^k(1,1)^t \Rightarrow \varepsilon_k = \left\|\overline{X} - X^{(k)}\right\| = \sqrt{2}\left(\frac{24}{25}\right)^k \to 0$ as $k \to \infty$, thus the sequence of the method converges to the unique solution $\overline{X} = (0,0)^t$, but the convergence is very slow.

b) **Optimal Step Gradient Method**

The idea in this method is to calculate the step α_k such that:

$$\begin{aligned} \alpha_k &= \arg\min_{\alpha>0}\varphi_k(\alpha) \\ \text{where } \varphi_k(\alpha) &= f(x^k - \alpha\nabla f(x^k)) \text{ (a function of a single variable).} \end{aligned}$$

i.e., starting from an initial point x^0, we have

$$\begin{aligned} x^1 &= x^0 - \alpha_0\nabla f(x^0) \text{ with } \alpha_0 = \arg\min_{\alpha>0}\varphi_0(\alpha) \text{ and } \varphi_0(\alpha) = f(x^0 - \alpha\nabla f(x^0)) \\ &\qquad\qquad \text{(optimal value of } \alpha_0) \\ x^2 &= x^1 - \alpha_1\nabla f(x^1) \text{ with } \alpha_1 = \arg\min_{\alpha>0}\varphi_1(\alpha) \text{ and } \varphi_1(\alpha) = f(x^1 - \alpha\nabla f(x^1)) \end{aligned}$$

that is, at each iteration k, we obtain an optimal value of the step α_k.

Example 2. Consider the following problem

$$(P)\ \begin{cases} \min f(x,y) = x^2y^2 + y^2 + 2xy + 2y + 2 \\ (x,y)^t \in \mathbb{R}^2 \end{cases}$$

Use the optimal step gradient method starting from $X^0 = (x^0, y^0)^t = (0,0)^t$.
We have $\nabla f(x,y) = (2xy^2 + 2y, 2yx^2 + 2y + 2x + 2)^t$.
$\nabla f(x^0, y^0) = \nabla f(0,0) = (0,2)^t$, we seek $\alpha_0 = \arg\min_{\alpha>0}\varphi_0(\alpha)$ with $\varphi_0(\alpha) = f(x^0 - \alpha\nabla f(x^0)) = f((0,0) - \alpha(0,2)) = f(0,-2\alpha) = 4\alpha^2 - 4\alpha + 2$.
$\varphi_0'(\alpha) = 8\alpha - 4 = 0 \Leftrightarrow \alpha = \frac{1}{2}$ and since $\varphi_0''(\alpha) = 8 > 0 \Rightarrow \alpha_0 = \frac{1}{2}$ is an optimal point (minimizer) of φ_0. Therefore,

$$X^1 = X^0 - \alpha_0\nabla f(X^0) = (0,0) - \frac{1}{2}(0,2) = (0,-1)^t.$$

Is X^1 a solution? We have $\nabla f(X^1) = \nabla f(0,-1)^t = (-2,0)$ and $\|\nabla f(X^1)\| = 2 > \varepsilon$ (where ε is a given precision).
Thus, X^1 is not a solution, so we need to calculate X^2.
First, we find $\alpha_1 = \arg\min_{\alpha>0}\varphi_1(\alpha)$, the second optimal step, with $\varphi_1(\alpha) = f(X^1 - \alpha\nabla f(X^1)) = f((0,-1) - \alpha(-2,0)) = f(2\alpha,-1) = 4\alpha^2 - 4\alpha + 1$.
$\varphi_1'(\alpha) = 8\alpha - 4 = 0 \Leftrightarrow \alpha = \frac{1}{2}$ and since $\varphi_1''(\alpha) = 8 > 0 \Rightarrow \alpha_1 = \frac{1}{2}$ is an optimal point (minimum) of φ_1.
Therefore,

$$X^2 = X^1 - \alpha_1\nabla f(X^1) = (0,-1) - \frac{1}{2}(-2,0) = (1,-1)^t.$$

We stop at this iteration because $\nabla f(X^2) = (0,0)^t$ (X^2 is a critical point), so the solution is $\overline{X} = (1,-1)^t$.

See the Newton and relaxation methods and others in [20].

1.6 Limitations of classic optimization methods

The local optimization methods presented aim to find the minimum of a function based on the knowledge of a search direction. These methods are applicable (and even recommended) when the desired solution is close to the known solution (starting point) or if the objective function is convex, differentiable, etc. However, they are not suitable for multi-modal, non-convex, non-differentiable problems where there is a risk of getting stuck around a local optimum.

Chapter 2

Univariate Lipschitz global optimization

Contents

2.1 Introduction

Global optimization is a rapidly developing research field over the past few years. Its importance is primarily related to the growing needs of applications in engineering, finance, chemistry, medicine, and other fields [6]. This topic studies the question of finding extrema (local and global) of a function of one or several variables. Naturally, if we want to optimize in the best possible way, we are interested in the absolute optima, hence the importance of global optimization, which aims to obtain the global optimum. The main quality of deterministic global optimization methods (which do not use any stochastic concept) is that they do not always require the starting point in the search for the optimum. These methods allow, at convergence, to obtain the exact solution of the considered optimization problem with an absolute guarantee.

2.2 Global optimization methods based on local search techniques.

2.2.1 Methods of multiple initializations (Multi-start)

This method consists of performing multiple searches for local minima starting from different initial points (starting points) with different descent directions. These points are generally chosen from the elements of a uniform grid. The difficulty with this method is that to be sure of finding the global minimum with the desired precision, one must take a number of initial points much larger than the number of local minimizers, which is almost always unknown. The "multi-start" algorithm [22] would be efficient if multiple determinations of the same local minimum could be avoided.

2.2.2 Tunnel excavation technique (Tunneling)

The tunneling method, presented by Levy et al.[22], consists of a series of cycles, each consisting of 2 phases:

• minimization phase;
• "tunneling" phase.

◦ The minimization phase is executed to obtain a point x^* such that the following conditions are satisfied:

$$\nabla f(x^*) = 0_{\mathbb{R}^n} \quad \text{and} \quad \nabla^2 f(x^*) \text{ is a positive definite matrix.}$$

For a given starting point x_0, any local minimization algorithm (local descent, gradient, Newton, etc.) can be used.

◦ The tunneling phase is known to reach a "good" starting point for the next minimization phase: starting from the point x^* just obtained, a solution of the "tunneling" equation given by:

$$f(x) - f(x^*) = 0 \tag{1.1}$$

is sought.

As soon as a zero (root) of equation (1.1) is reached for an $x_0 \neq x^*$, this point is taken as the new starting point for the next minimization phase. If, after a sufficient computation time (arbitrarily chosen), an $x_0 \neq x^*$ solution of (1.1) is not found, it is considered that the equation:

$$f(x) - f(x^*) > 0, \quad \forall x \in S,\ x \neq x^*,$$

is satisfied and the algorithm is terminated.

The figure below provides a geometric illustration of this algorithm for a one-dimensional function (single variable).
Starting from the point x_1^0, the local minimization phase leads to the local minimum x_1^*, Then, the tunneling phase provides the point x_2^0, which serves as the starting point for a second iteration in two phases: minimization until x_2^* followed by tunneling until x_3^* . The final iteration brings, in the minimization phase, to x_3^*, Since the tunneling phase cannot provide a zero (root) of equation (1.1), the program stops. We consider x_3^* as the global minimum of the function f.

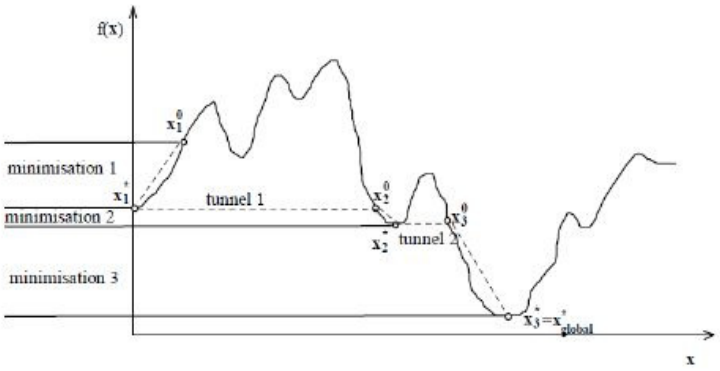

Figure 2.1: The geometric interpretation of the tunneling method

Remark 2.2.1. This procedure avoids the attraction of local minima that have already been encountered. However, it has a drawback: the auxiliary piercing function, progressively constructed from f and other coefficients, becomes increasingly flat, which complicates its minimization.

2.3 Covering methods

First, consider the global minimization problem:

$$f^* = \operatorname{glob}\min_{x \in C} f(x), \tag{P}$$

where $C \subset \mathbb{R}^n$ is a compact set, and $f : C \to \mathbb{R}$ is a continuous function on D. (Weierstrass).
In numerical computations, this problem is usually simplified by introducing what is called the set of ε-approximate (or approximated) solutions:

$$C_\epsilon^* = \{x \in C : \ f(x) - \epsilon \leq f^*\},$$

where $\epsilon > 0$ is a given precision. We need to find some point $x^* \in C_\epsilon^*$. In other words, it is necessary to find, with precision ϵ, the value of the global minimum of the function f and at least one point $x^* = \arg\min_C f(x)$ at which the approximate value is attained, that is: $\min f(x) = f^* = f(x^*)$.

General principle of covering methods.

First, let's give the key idea of these methods (we will specify the description later):[5]

Suppose that the objective function f is evaluated at the points $x_1, x_2, ..., x_k$ of C.
We call **record of order** k the following value:

$$R_k = \min\{f(x_1), f(x_2), ..., f(x_k)\}; \tag{2.1}$$

and we call **record point of order** k any point $x_i, i = 1, ..., k$ verifying:

$$f(x_i) = R_k. \quad \text{that is, } x_i = \arg\min R_k.$$

Define the following set:

$$E_k = \{x \in C : \ R_k - \epsilon \leq f(x)\}, \tag{2.2}$$

obviously,

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

then:

$$f^* = \operatorname{globmin}_{x \in C} f(x) = \min_{x \in C - E_k} f(x), \text{ if } C \neq E_k.$$

Thus, the points of E_k are not of interest in the search for the global minimum and consequently they can be omitted (or eliminated) from the feasible set C. The search for the global minimum continues only on the set $C - E_k$. In particular, if:

$$C \subseteq E_k \ , \qquad (\text{i.e., } E_k \text{ covers } C) \tag{2.3}$$

then the initial problem is solved and the record point x_k is considered as an approximate solution (R_k is the approximated global minimum of f), with the guarantee that $x_k \in C$.

Now, if f is evaluated at a new point x_{k+1}, then another record R_{k+1} is obtained, and consequently a larger set E_{k+1} can be omitted. The real sequence (R_k) is decreasing, while the sequence of sets $\{E_k\}$ is increasing in the sense of inclusion, i.e.,

$$R_{k+1} \leq R_k \qquad \text{and} \qquad E_k \subset E_{k+1}. \tag{2.4}$$

Thus, the minimization problem (P) is reduced to the construction of a sequence of points $x_1, x_2, ..., x_k$ verifying (2.4).

These are called covering methods because they generate a finite family of balls $(B(x_i, r_i))_{1 \leq i \leq l}$ whose union covers C, that is

$$C \subseteq \bigcup_{i=1}^{l} (B(x_i, r_i)).$$

Covering methods are divided into two types:

a) Passive covering methods.
b) Iterative covering methods.

2.3.1 Passive covering method

Generally, the functions encountered in most practical problems have a bounded rate of variation (a **less restrictive** assumption than **differentiability**), which is why the assumption that the objective function f is **Lipschitz** is very common in numerical analysis. Covering methods have been theoretically constructed to exploit this property.

Definition 2.3.1. Let (C, d) be a metric space, a function $f : C \subset \mathbb{R}^n \to \mathbb{R}$ is said to be Lipschitz on C if there exists a constant $L = L(f, C) > 0$ such that

$$|f(x) - f(y)| \leq L\mathbf{d}(x, y), \quad \forall x, y \in C. \tag{5}$$

Obviously, if f is Lipschitz with constant L on C, then it is also Lipschitz for any constant $L' > L$. Furthermore, from inequality (5), we deduce that any Lipschitz function on C is continuous on C (the converse is not always true).

In most cases, the Euclidean distance is used. However, in several applications, inequality (5) is considered with respect to other distances:

$\mathbf{d}_p(x,y) = \left(\sum_{j=1}^{n} |x_j - y_j|^p\right)^{1/p}$ for $1 \leq p < \infty$.

If $p = 2$: $\mathbf{d}_2(x,y) = \left(\sum_{j=1}^{n} |x_j - y_j|^2\right)^{1/2} = \|x-y\|_2$ (the Euclidean distance and norm in $\mathbb{R}^n$).

And $\mathbf{d}_{\max}(x,y) = \max_{1\leq j\leq n}(|x_j - y_j|)$ with $x = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$ and $y = (y_1, y_2, ..., y_n) \in \mathbb{R}^n$. (Chebyshev distance).

∘ While considering the equivalence between the norms $\|.\|_1$, $\|.\|_2$ and $\|.\|_\infty$.

∘ Any Lipschitz function with respect to one norm is also Lipschitz with respect to the other norms.

Remark 2.3.1. Evaluating the objective function f only at the points of a sufficiently (or α-**dense**) α-**mesh** (or grid) of the domain C guarantees the detection of the global minimum with a given precision.

Definition 2.3.2. A sequence of points $(x_k)_{k\in\mathbb{N}} \subset C$ is said to be an α-mesh (or α-grid), $(\alpha > 0)$ in C if:

$$\forall x \in C,\ \exists k \in \mathbb{N} \text{ such that } d(x, x_k) \leq \alpha.$$

In passive covering methods, there are several ways to choose the sequence of points $x_1, x_2, ..., x_k$ that satisfies condition (4). The selection of points $x_1, x_2, ..., x_k$ is done in advance among the elements of the feasible set and independently of the information that arises during the search of the algorithm. The meshes are generally uniform. These methods are simple but of great theoretical and practical interest.

Remark 2.3.2. The most common and easiest to construct uniform meshes in a cube of $\mathbb{R}^n$ are those where the space is equipped with the maximum distance (the Chebyshev distance). Naturally, the Lipschitz constant must be estimated with respect to this distance. However, the construction of a passive mesh is easy only for certain sets such as parallelepipeds and balls in $\mathbb{R}^n$. If the set C does not have a simple structure, one may encounter enormous difficulties in constructing a mesh with good uniformity characteristics. Moreover, this is the most important step in the algorithm, given the simplicity of investigating its realization. Cubic meshes are obtained by equipping C with the distances $\mathbf{d}_p(x,y)$ and $\mathbf{d}_{\max}(x,y)$.

Rectangular meshes are obtained from the following distances:

$\partial_p(x,y) = \left(\sum_{j=1}^{n} \lambda_j |x_j - y_j|^p\right)^{\frac{1}{p}}$ for $p \geq 1$

$\partial_{\max}(x,y) = \max_{1\leq j\leq n}(\lambda_j |x_j - y_j|)$ with $\lambda_1, \lambda_2, ..., \lambda_n$ being strictly positive numbers not all

equal.

Examples of uniform meshes.

Example 2.3.1. If $n = 1$, $n \in \mathbb{N}$ and $p \geq 2$ is an integer. Let the set $\mathbf{C} = [a, b] \subset \mathbb{R}$ be an interval, then we define the uniform grid of points in $\mathbf{C}$ as follows:

$$G_p = \left\{ x_i = a + (i + \frac{1}{2})\frac{b-a}{p} \;\; / \;\; i = 0, 1, ..., p-1 \right\}$$

where p is the number of sub-intervals $[a_i, b_i]$ with centers x_i and the same radii $\frac{1}{2}\frac{b-a}{p}$ such that:

$$[a, b] = \bigcup_{i=0}^{p-1} [a_i, b_i] .$$

Example 2.3.2. If $n \geq 2$, let the hypercube in $\mathbb{R}^n$ be defined by $\mathbf{C} = [0, a]^n = \underbrace{[0, a] \times ... \times [0, a]}_{n \text{ times}}$; then the uniform grid of points in $\mathbf{C}$ is given as follows:

$$\mathbf{g}_p = \left\{ \left((i_1 + \frac{1}{2})\frac{a}{p}, (i_2 + \frac{1}{2})\frac{a}{p}, ..., (i_n + \frac{1}{2})\frac{a}{p} \right) \;\; / \; i_1, i_2, ..., i_n \in \{0, 1, ..., p-1\} \right\}$$

is a uniform mesh of $\mathbf{C}$ consisting of p^n points. We can easily show that for any $x \in \mathbf{C}$, there exists $\xi \in \mathbf{g}_p$ such that:

$$\mathbf{d}_{\max}(x, \xi) \leq \frac{a}{p}.$$

Thus, the balls centered at the mesh points with radii $\frac{a}{p}$ form a covering of $\mathbf{C}$.
See an example for $n = 2$ and $p = 3$.

Example 2.3.3. If $\mathbf{C}$ is a hyper-rectangle, i.e., $\mathbf{C} = \prod_{i=1}^{n} [a_i, b_i] = [a_1, b_1] \times [a_2, b_2] \times ... \times [a_n, b_n]$. Then we divide each side $[a_i, b_i]$ of $\mathbf{C}$ into p_i segments of length $l_i = \frac{b_i - a_i}{p_i}$ and we choose the mesh points as follows:

$$\mathbf{g} = \left\{ \begin{array}{c} \left(a_1 + (i_1 + \frac{1}{2})l_1, \; a_2 + (i_2 + \frac{1}{2})l_2, ..., \; a_n + (i_n + \frac{1}{2})l_n \right) \\ / \; i_1, i_2, ..., i_n \in \{0, 1, ..., p_k - 1\}, \; k = 1, 2, ..., n \end{array} \right\}$$

Passive covering algorithms

The idea of the passive covering algorithm is simple. Indeed, for a Lipschitz function f with constant L defined on a compact set C of the metric space $(\mathbb{R}^n, \mathbf{d})$, we take a mesh $\mathbf{g}_k = \{x_1, x_2, ..., x_k\} \subset C$. In the construction of the covering of C, one of the most commonly used methods is covering by balls $B(x_i, r)$ where r is the radius of the balls satisfying $C \subseteq \bigcup_{i=1}^{k} B(x_i, r)$. Then we have:

$$\forall x \in C, \; \exists i \in \{0, 1, ..., k\} \text{ such that } x \in B(x_i \, , r).$$

Suppose that $x_k^* = \arg\min_{1 \leq i \leq k} \{f(x_i)\} = \arg\min_{\mathbf{g}_k} f(x)$ (**record point**), i.e., x_k^* is the point that achieves the minimum of f on the grid $\mathbf{g}_k$, so x_k^* is an approximation of a global minimizer x^* of f on the grid $\mathbf{g}_k$, thus $f_k^* = f(x_k^*)$ is an approximation of $m^* = f(x^*) = \min_{x \in C} f(x)$ (exists according to Weierstrass), that is $f_k^* = \min_{1 \leq i \leq k} \{f(x_i)\}$ (**record**).

Furthermore, there exists $j \in \{1, ..., k\}$ such that $x^* \in B(x_j, r)$, and since f is Lipschitz, we have

$$|f_k^* - m^*| = |f(x_k^*) - f(x^*)| \leq |f(x_j) - f(x^*)| \leq L\mathbf{d}(x_j, x^*) \leq Lr \leq \epsilon.$$

Thus, by taking the radius of the balls $r \leq \frac{\epsilon}{L}$, the global minimum of f can be obtained with the precision ϵ and the minimizer x_k^* with the precision $\frac{\epsilon}{L}$.
Passive covering algorithms have interesting theoretical properties but they have the drawback of completely neglecting the information about the objective function obtained during the search. This drawback makes grid algorithms inefficient, especially in the case where C is of high dimension (since the number of evaluation points of f increases enormously with the dimension of C). Below, we provide a passive covering algorithm in the one-dimensional case, i.e., $C = [a, b] \subset \mathbb{R}$.

Algorithm.

1- Initialization.
Set $k = 1,\ x_1 = a + \frac{\epsilon}{L},\ x_\epsilon \to x_1$ and $f_\epsilon \to f(x_\epsilon)$ **(record)**.
2- Steps $k = 2, 3, ...$
Set $x_{k+1} = x_k + 2\frac{\epsilon}{L}$.
If $x_{k+1} > b - \frac{\epsilon}{L}$ (or if $x_{k+1} > b$) then stop.
Otherwise, determine $f(x_{k+1})$.
If $f(x_{k+1}) < f_\epsilon$ then set $x_\epsilon \to x_{k+1},\ f_\epsilon \to f(x_{k+1})$.
Set $k \to k + 1$ and go to step 2.
End of the algorithm.

Remark 2.3.3. 1) In the case where $C = [a, b]$, we can cover $[a, b]$ with sub-intervals $\{I_i\}_{1 \leq i \leq k}$ centered at $\{x_i\}_{1 \leq i \leq k}$ and having the same radius $r_i = \frac{\epsilon}{L}, \forall i$ such that $[a, b] = \bigcup_{i=1}^{k} I_i$. (see the first example of uniform grids).

2) At the initialization of the algorithm, we set $x_1 = a + \frac{\epsilon}{L}$ which is the center of the interval $I_1 = \left[a, a + 2\frac{\epsilon}{L}\right]$ with radius $r_1 = \frac{\epsilon}{L}$, because at the initialization $f(x_1) = R_1$ (record of order 1). With this value of x_1, we save time and ensure that we do not miss the global minimum in the vicinity of this point. Indeed, if $x^* = \arg\min_{[a,b]} f(x) \in \left[a, a + 2\frac{\epsilon}{L}\right]$ then:

$$|f(x_1) - f(x^*)| \leq L\mathbf{d}(x_1, x^*) \leq L\frac{\epsilon}{L}$$

$\Rightarrow x_1$ is an approximation of the global minimum x^*.

3) The passive covering strategy is independent of the use of information about the objective function f (i.e., the evaluations of f during the search).

4) Passive covering algorithms have interesting theoretical properties but they have the drawback of completely ignoring the information about the objective function obtained during the search. This drawback makes grid algorithms inefficient, particularly when C is of high dimension (since the number of evaluation points of f increases enormously with the dimension of C).

Theorem 2.3.1. *The number of evaluation points N for a Lipschitz continuous function with constant L over an interval $[a, b] \subset \mathbb{R}$ is given by:*

$$N = \left\lfloor \frac{(b-a)L}{2\varepsilon} \right\rfloor + 1,$$

where $\lfloor . \rfloor$ denotes the floor function (the greatest integer less than or equal to a given real number) and ε is a given precision.

Proof. According to the passive covering algorithm, we have:

$$\begin{aligned}
x_1 &= a + \frac{\varepsilon}{L} \\
x_2 &= x_1 + \frac{2\varepsilon}{L} = a + \frac{\varepsilon}{L} + \frac{2\varepsilon}{L} = a + \frac{(2 \times 2 - 1)\varepsilon}{L} \\
x_3 &= x_2 + \frac{2\varepsilon}{L} = a + \frac{(2 \times 3 - 1)\varepsilon}{L} \\
&\vdots \\
x_k &= a + \frac{(2 \times k - 1)\varepsilon}{L} \text{ (this can be shown by induction).}
\end{aligned}$$

For the algorithm to converge to the global minimum with the desired precision ε, the number k must satisfy the following condition:

$$x_k \leq b + \frac{\varepsilon}{L} < x_{k+1}.$$

Assuming the algorithm stops at iteration $k = N$, then at this iteration we have the following condition:

$$x_N \leq b + \frac{\varepsilon}{L} < x_{N+1}.$$

Since the sequence $(x_k)_{k \geq 1}$ is increasing, it follows that N is the smallest integer $j \in \mathbb{N}^*$ that satisfies:

$$N = \min \left\{ j \in \mathbb{N}^* \ / \ b + \frac{\varepsilon}{L} < x_{j+1} \right\},$$

where $x_{j+1} = a + \frac{(2 \times j - 1)\varepsilon}{L}$.

$$\begin{aligned}
N &= \min \left\{ j \in \mathbb{N}^* \ / \ b + \frac{\varepsilon}{L} < a + \frac{(2 \times j - 1)\varepsilon}{L} \right\} \\
&= \min \left\{ j \in \mathbb{N}^* \ / \ j > \frac{(b-a)L}{2\varepsilon} \right\},
\end{aligned}$$

thus the smallest integer $j \in \mathbb{N}^*$ that satisfies the last inequality is

$$N = \left\lfloor \frac{(b-a)L}{2\varepsilon} \right\rfloor + 1 \in \mathbb{N}^*.$$

□

Passive overlap algorithm in the multi-dimensional case. In the case where the feasible set $\mathbf{C}$ is of dimension $n \geq 2$ and $\mathbf{C}$ is a hyperrectangle in $\mathbb{R}^n$, i.e., $\mathbf{C} = \prod_{i=1}^{n} [a_i, b_i] = [a_1, b_1] \times [a_2, b_2] \times ... \times [a_n, b_n]$, the overlap intervals are replaced by cubes in $\mathbb{R}^n$ whose faces are parallel to the coordinate axes. For this purpose, it suffices to assume that the Lipschitz condition is satisfied with respect to the Chebyshev norm ($\|.\|_\infty$). In this case, we have:

$$|f(x) - f(y)| \leq L \|x - y\|_\infty , \qquad \forall x, y \in \mathbf{C}.$$

If the norm used is not the Chebyshev norm, for example the Euclidean norm, we obtain:

$$|f(x) - f(y)| \leq L \|x - y\|_2 \leq L\sqrt{n} \|x - y\|_\infty , \qquad \forall x, y \in \mathbf{C}.$$

Then, taking the radius of the cubes $P(x_i, r)$, $r \leq \frac{\varepsilon}{L\sqrt{n}}$, verifying $\mathbf{C} \subset \bigcup_{i=1}^{k} P(x_i, r)$.

Passive covering algorithm if $n = 2$.

If $n = 2$ then $\mathbf{C} = [a_1, b_1] \times [a_2, b_2]$.

1- Initialization. Set $x_1 = (x_1^1, x_1^2) = (a_1, a_2) + \frac{\varepsilon}{L\sqrt{2}}(1, 1)$.

$m = f(x_1)$, $R_1 = m$ (record of order 1) (x_1 is an initial approximation of $\arg\min_{\mathbf{C}} f(x)$).

2- Procedure $N_{Passive}(1, a, b)$.

We construct a sequence of points x_j^1 (horizontal grid).

While $x_j^1 \leq b_1 + \frac{\varepsilon}{L\sqrt{2}}$, **do**
If $x_j^1 > b_1$ then $x_j^1 = b_1$
$S = f(x_j)$, ($x_j = (x_j^1, x_1^2)$, with $x_1^2 = a_2 + \frac{\varepsilon}{L\sqrt{2}}$).
If $S < R_1$, $m = S$ and $\arg R_1 = x_j$
$x_{j+1}^1 = x_j^1 + \frac{2\varepsilon}{L\sqrt{2}}$ (see the one-dimensional case on $[a_1, b_1]$).
End While

Otherwise $N_{Passive}(i - 1, a, b)$ (we fix j, (vertical grid))

While $x_j^i \leq b_2 + \frac{\varepsilon}{L\sqrt{2}}$, **do**
If $x_j^i > b_2$ then $x_j^i = b_2$
$x_{j+1}^i = x_j^i + \frac{2\varepsilon}{L\sqrt{2}}$ $i = i + 1$ **End While**.

Theorem 2.3.2. *The number of evaluation points N for a Lipschitz continuous function with constant $L >$ on the rectangle $\mathbf{C} = [a_1, b_1] \times [a_2, b_2]$ in $\mathbb{R}^2$ is given by:*

$$N = \left(\left\lfloor \frac{(b_1 - a_1)L\sqrt{2}}{2\varepsilon} \right\rfloor + 1 \right) \times \left(\left\lfloor \frac{(b_2 - a_2)L\sqrt{2}}{2\varepsilon} \right\rfloor + 1 \right),$$

where $\lfloor . \rfloor$ denotes the floor function of a real number and ε is a given precision.

Proof. If $n = 2$, we have

$$\begin{aligned} x_1^1 &= a_1 + \frac{\varepsilon}{L\sqrt{2}} \\ x_2^1 &= x_1^1 + \frac{2\varepsilon}{L\sqrt{2}} = a_1 + \frac{\varepsilon}{L\sqrt{2}} + \frac{2\varepsilon}{L\sqrt{2}} = a_1 + \frac{(2\times 2-1)\varepsilon}{L\sqrt{2}} \\ x_3^1 &= x_2^1 + \frac{2\varepsilon}{L\sqrt{2}} = a_1 + \frac{(2\times 3-1)\varepsilon}{L\sqrt{2}} \\ x_k &= a_1 + \frac{(2\times k-1)\varepsilon}{L\sqrt{2}} \text{ (this can be shown by induction).} \end{aligned}$$

For the algorithm to converge to the global minimum with the desired precision ε, the number k must satisfy the following condition:

$$x_k^1 \leq b_1 + \frac{\varepsilon}{L\sqrt{2}} < x_{k+1}^1.$$

The number of evaluation points N_1 is given by (see the one-dimensional case on $[a_1, b_1]$)

$$N_1 = \left(\left\lfloor \frac{(b_1 - a_1)L\sqrt{2}}{2\varepsilon} \right\rfloor + 1 \right).$$

So at each call of $N_{Passive}(1, a, b)$, N_1 evaluation points are needed, but the procedure $N_{Passive}(1, a, b)$ will be called $N_2 = \left(\left\lfloor \frac{(b_2-a_2)L\sqrt{2}}{2\varepsilon} \right\rfloor + 1 \right)$ times, hence the total number of evaluation points is $N = N_1 \times N_2$, which was to be demonstrated. □

Exercise 2.3.1. Describe the passive covering algorithm in the general case $n \geq 2$, i.e., $\mathbf{C} = \prod_{i=1}^{n} [a_i, b_i] = [a_1, b_1] \times [a_2, b_2] \times ... \times [a_n, b_n]$.

2.3.2 Evtushenko's iterative covering method

In iterative covering algorithms, the elements $x_1,\ x_2, ...,\ x_k$ of the mesh are denser in the region where the function takes small values and less dense elsewhere. Indeed, at step i of the algorithm, the selection of point x_i $(1 \leq i \leq n)$ depends on the points previously chosen $x_1,\ x_2, ...,\ x_{i-1}$ and their images $f(x_1),\ f(x_2), ...,\ f(x_{i-1})$. The construction of the mesh depends on the function f, the feasible set C, and the chosen metric, in addition to the Lipschitz constant. First and second derivatives of f are sometimes used when they exist.

Evtushenko developed several covering methods, the one we will present is the most well-known and widely discussed in the literature. It is based on excluding cubes whose centers are points obtained through a recursive procedure of one-dimensional global optimization. Suppose the function f is evaluated at points $x_1, x_2, ..., x_k$.
Let $m_k^* = \min\{f(x_1), f(x_2), ..., f(x_k)\}$ and denote by $(B(x_i, r_{ik}))_{1\leq i\leq k}$ the family of balls with centers x_i and radii r_{ik}. Let's demonstrate how this sequence of points is constructed in the method of Evtushenko.

Theorem 2.3.3. *Let f be a Lipschitz function with constant L, defined on a compact set C of the metric space $(\mathbb{R}^n, \mathbf{d})$. Suppose f is evaluated at points $x_1, x_2, ..., x_k$ in C. Define $m_k^* = \min\{f(x_1), f(x_2), ..., f(x_k)\}$. If $C \subset \bigcup_{i=1}^{k} B(x_i, r_{ik})$ with $r_{ik} = \frac{\varepsilon + f(x_i) - m_k^*}{L}$ then m_k^* is the global minimum of f with precision ε on C.*

Proof. According to the Lipschitz condition, we have

$$|f(x) - f(y)| \leq L \|x - y\|, \quad \forall x, y \in C, \tag{1}$$

thus

$$f(y) - L \|x - y\| \leq f(x), \quad \forall x, y \in C \tag{2}$$

which means that for a fixed $y \in C$, if a certain x satisfies:

$$m_k^* - \varepsilon \leq f(y) - L \|x - y\| \tag{3}$$

then $m_k^* - \varepsilon \leq f(x)$. We consider the balls:

$$B(x_i, r_{ik}) = \{x \in \mathbb{R}^n / \quad \|x_i - x\| \leq r_{ik}\};$$

from inequality (3) we have

$$\|x - y\| \leq \frac{\varepsilon + f(y) - m_k^*}{L} \tag{4}$$

so for $y = x_i$, we must have the radii r_{ik} given by:

$$r_{ik} = \frac{\varepsilon + f(x_i) - m_k^*}{L}.$$

It is easy to see that for all $i = 1, ..., k$ and for all $x \in B(x_i, r_{ik})$ we have $m_k^* - \varepsilon \leq f(x)$. Thus if the union of the balls $B(x_i, r_{ik})$ does not cover C, the global minimum can be reached in $C - \bigcup_{i=1}^{k} B(x_i, r_{ik})$. Consequently, the balls $B(x_i, r_{ik})$ can be omitted from the feasible set C; we search for the solution in the remaining part. The problem (**P**) will have a solution when the union of the balls completely covers C.

□

Remark 2.3.4. 1° In the case where the feasible set C is of dimension $n \geq 2$, the covering intervals are replaced by cubes in $\mathbb{R}^n$ whose faces are parallel to the axes of the coordinate system. For this, it suffices to assume that the Lipschitz condition (1) is verified with respect to the Chebyshev norm. In this case, we have

$$|f(x) - f(y)| \leq L \|x - y\|_\infty = L \max_{1 \leq j \leq n} (\left|x^j - y^j\right|) = L\mathbf{d}_{\max}(x, y),$$

where $x = (x^1, x^2, ..., x^n) \in C$ and $y = (y^1, y^2, ..., y^n) \in C$. Therefore, according to inequality (4), for $y = x_i$,

$$\mathbf{d}_{\max}(x, x_i) \leq \frac{f(x_i) - m_k^* + \varepsilon}{L} = r_{ik}.$$

2° If the norm used in (1) is not the Chebyshev norm, for example the Euclidean norm, we obtain

$$|f(x)-f(y)| \leq L\,\|x-y\| = L\sqrt{\sum_{j=1}^{n}|x^j-y^j|^2} \leq L\sqrt{n}\max_{1\leq j\leq n}(|x^j-y^j|) = L\sqrt{n}\,\|x-y\|_\infty\,,$$

thus according to (4), for $y = x_i$,

$$\|x-x_i\| \leq \sqrt{n}\ \mathbf{d}_{\max}(x,x_i) = \sqrt{n}\ \max_{1\leq j\leq n}(|x^j-y^j|) \leq r_{ik}$$

then,

$$\mathbf{d}_{\max}(x,x_i) \leq \frac{r_{ik}}{\sqrt{n}} = \frac{f(x_i)-m_k^*+\varepsilon}{\sqrt{n}L}.$$

The representation given suggests a constructive method for solving the problem (**P**). In summary, the method consists of the following: suppose that for a certain sequence of points $\{x_k\}$ the record m_k^* is determined. The sequence of points x_i and the radii r_{ik} of the ball are stored in memory. If at the new point x_{k+1} we have $f(x_{k+1}) < m_k^*$, we set $m_{k+1}^* = f(x_{k+1})$ and replace the term r_{ik} with r_{ik+1}. If the balls B_{ik+1} cover the set C, then the calculation will be stopped; otherwise, we take a new point x_{k+2} and continue.

3° The set C is covered by balls of different radii. Consider a point x_i for which $m_i^* = f(x_i)$, $1 \leq i \leq k$, obviously we have $f(x_i) = m_i^* \geq m_k^*$ hence,

$$r_{ik} = \frac{f(x_i)-m_k^*+\varepsilon}{L} \geq \frac{\varepsilon}{L},$$

thus the smallest radius is at the point x_i for which $m_i^* = f(x_i)$. (that is $\frac{\varepsilon}{L}$).

Evtushenko's One-Dimensional Method

In the case where $C = [a,\ b]$ in $\mathbb{R}$, according to theorem 3, we can cover the interval $[a,\ b]$ with sub-intervals $\{I_i\}_{1\leq i\leq k}$ centered at $\{x_i\}_{1\leq i\leq k}$ and with radius $r_{ik} = \frac{f(x_i)-m_k^*+\varepsilon}{L}$, such that: $[a,\ b] \subset \bigcup_{i=1}^{k} I_i$.

Since the smallest of the radii r_{ik} is $\frac{\varepsilon}{L}$, we take $x_1 = a + \frac{\varepsilon}{L}$ which is the center of the interval $I_1 = [a, a+\frac{2\varepsilon}{L}]$ with radius $r_{11} = \frac{\varepsilon}{L}$, because at initialization $f(x_1) = m_1^*$.

With this value of x_1, we save time and ensure that we have not overlooked the global minimum in the vicinity of this point.

For the point x_2 which is the center of the interval I_2 with radius $r_{22} = \frac{f(x_2)-m_2^*+\varepsilon}{L}$, we should take $x_2 = x_1 + r_{11} + r_{22}$. However, r_{22} is unknown, but we know that $r_{22} \geq \frac{\varepsilon}{L}$. So we take $x_2 = x_1 + r_{11} + \frac{\varepsilon}{L} = a + \frac{\varepsilon}{L} + \frac{2\varepsilon}{L}$.

In general, for $i \geq 1$, the sequence $\{x_{i+1}\}$ is defined by:

$$\begin{cases} x_1 = a + \frac{\varepsilon}{L} \\ x_{i+1} = x_i + \frac{\varepsilon}{L} + \frac{f(x_i)-m_k^*+\varepsilon}{L} = x_i + \frac{f(x_i)-m_k^*+2\varepsilon}{L}. \end{cases}$$

Remark 2.3.5. - This choice ensures that we do not miss the global minimum of f, as the intervals I_i overlap. We stop when k satisfies $C \subset \bigcup_{i=1}^{k} I_i$.

- If $x_k < b$ and $x_k + r_{kk} \geq b$, then the last point of the sequence is x_k.
- If $x_k + r_{kk} < b$ and $x_k + r_{kk} + \frac{\varepsilon}{L} \geq b$, then the last point of the sequence will be $x_{k+1} = b$.

Evtushenko's Iterative One-Dimensional Algorithm
1. Initialization
Set $k = 1,\ x_1 = a + \frac{\varepsilon}{L},\ x_\varepsilon = x_1,\ f_\varepsilon = f(x_1)$.
2. Steps $k = 2,\ 3, ...$
Set $x_{k+1} = x_k + \frac{\varepsilon}{L} + \frac{f(x_k) - f_\varepsilon + \varepsilon}{L}$.
If $x_{k+1} > b - \frac{\varepsilon}{L}$, then stop.
Otherwise, determine $f(x_{k+1})$.
If $f(x_{k+1}) < f_\varepsilon$, then set $x_\varepsilon = x_{k+1},\ f_\varepsilon = f(x_{k+1})$.
Set $k = k + 1$ and go to step 2.

Theorem 2.3.4. *The number of evaluation points N in Evtushenko's algorithm, in the one-dimensional case, lies between k_1 and $k_2, (k_1 \leq N \leq k_2)$, with:*

$$\begin{aligned} k_1 &= \left\lfloor \log_2\left(\frac{(b-a)L}{2\varepsilon} + 1\right) \right\rfloor + 1 \\ k_2 &= \left\lfloor \frac{(b-a)L}{2\varepsilon} \right\rfloor + 1 \end{aligned}$$

$\lfloor . \rfloor$ *denotes the floor function of a real number.*

Proof. It is clear that N is maximal $\Leftrightarrow$ the radii r_{ik} are maximal (maximum distance), and this is verified when the objective function f exhibits the greatest possible variation between two points x and y in the interval $[a,\ b]$. However, f is Lipschitz, which means

$$\underbrace{|f(x) - f(y)|}_{\text{variation}} \leq L \underbrace{|x - y|}_{\text{distance}},\quad \forall x, y \in [a,\ b].$$

Thus, the most favorable case for the method is the one in which we have

$$\underbrace{|f(x) - f(y)|}_{\text{variation}} = L \underbrace{|x - y|}_{\text{distance}},\quad \forall x, y \in [a,\ b],\ \text{with } f \text{ increasing on } [a,\ b]. \qquad (*)$$

In this case, the Evtushenko algorithm gives us the following sequence of points:

$$\begin{aligned} x_1 &= a + \frac{\varepsilon}{L},\ x_\varepsilon = x_1,\ f_\varepsilon = f(x_1) = m_1^* \text{ (initial record)} \\ x_2 &= x_1 + \frac{1}{L}(f(x_1) - f_\varepsilon + 2\varepsilon) = a + \frac{\varepsilon}{L} + \frac{1}{L}(\underbrace{f(x_1) - m_1^*}_{=0} + 2\varepsilon) = a + \frac{(2^2 - 1)\varepsilon}{L} \\ x_3 &= x_2 + \frac{1}{L}(f(x_2) - f_\varepsilon + 2\varepsilon) = x_2 + \frac{1}{L}(f(x_2) - m_2^* + 2\varepsilon) = \\ &= x_2 + \frac{1}{L}(f(x_2) - \min\{f(x_1), f(x_2)\} + 2\varepsilon) = x_2 + \frac{1}{L}(f(x_2) - f(x_1) + 2\varepsilon) \text{ since } f \text{ is increasing} \\ &= x_2 + \frac{1}{L}(L(x_2 - x_1) + 2\varepsilon) \text{ (according to } (*)). \\ &= 2x_2 - x_1 + \frac{2\varepsilon}{L} = a + \frac{(2^3 - 1)\varepsilon}{L}. \\ &\vdots \\ x_j &= a + \frac{(2^j - 1)\varepsilon}{L}, \text{ for all } j \geq 1 \text{ (by recurrence).} \end{aligned}$$

To ensure that the algorithm converges to the global minimum with the desired precision ε, the number k_1 must satisfy the following condition:

$$x_{k_1} \leq b + \frac{\varepsilon}{L} < x_{k_1+1}.$$

Hence k_1 is the smallest integer $j \in \mathbb{N}^*$ such that

$$\begin{aligned}
k_1 &= \min\left\{j \in \mathbb{N}^* \,/\; b + \frac{\varepsilon}{L} < x_{j+1}\right\} \\
&= \min\left\{j \in \mathbb{N}^* \,/\; b + \frac{\varepsilon}{L} < a + \frac{(2^{j+1}-1)\varepsilon}{L}\right\} \\
&= \min\left\{j \in \mathbb{N}^* \,/\; \frac{(b-a)L}{2\varepsilon} + 1 < 2^j\right\} \\
&= \min\left\{j \in \mathbb{N}^* \,/\; j > \log_2(\frac{(b-a)L}{2\varepsilon} + 1) < j\right\} \\
&= \left\lfloor \log_2(\frac{(b-a)L}{2\varepsilon} + 1)\right\rfloor + 1
\end{aligned}$$

If the function f is constant, meaning $f(x) = c$ for all $x \in [a,\ b]$, then the method coincides with the passive covering method. Indeed:

$$\begin{aligned}
x_1 &= a + \frac{\varepsilon}{L} \\
x_2 &= x_1 + \frac{1}{L}(f(x_1) - f_\varepsilon + 2\varepsilon) = x_1 + \frac{2\varepsilon}{L} = a + \frac{(2\times 2 - 1)\varepsilon}{L} \\
x_3 &= x_2 + \frac{1}{L}\underbrace{(f(x_2) - f_\varepsilon)}_{=0 \text{ since } f \text{ is constant}} + 2\varepsilon) = a + \frac{(2\times 3 - 1)\varepsilon}{L} \\
&\vdots \\
x_k &= a + \frac{(2\times k - 1)\varepsilon}{L}
\end{aligned}$$

To ensure that the algorithm converges to the global minimum with the desired precision ε, the number k_2 must satisfy the following condition:

$$x_{k_2} \leq b + \frac{\varepsilon}{L} < x_{k_2+1}.$$

So $k_2 = \left\lfloor \frac{(b-a)L}{2\varepsilon}\right\rfloor + 1$. In general, we have $k_1 \leq N \leq k_2$. □

Remark 2.3.6. According to Theorem 4, we deduce that the Evtushenko method is not effective for decreasing functions, since the number of evaluation points in this case is equal to the number of evaluation points in the passive algorithm. Indeed:

$$\begin{aligned}
x_3 &= x_2 + \frac{1}{L}(f(x_2) - m_2^* + 2\varepsilon) = x_2 + \frac{1}{L}(f(x_2) - \min\{f(x_1), f(x_2)\} + 2\varepsilon) \\
&= x_2 + \frac{2\varepsilon}{L} \\
&\vdots
\end{aligned}$$

2.4 Global optimization methods using support functions

The main idea behind these methods is to construct convex lower bounding functions (support functions), whether linear or nonlinear, of the objective function over the feasible set in order to relax the non-convex problem. Thus, the global minimizers of these support functions form a sequence that converges to the global minimum of the objective function. In this section, we describe how to construct support functions for a Lipschitz continuous objective function or a function that is twice continuously differentiable.

2.4.1 Piyavskii-Shubert method

Problem Statement. Consider the following one-dimensional global minimization problem:

$$\min_{x\in[a,b]} f(x) \tag{\textbf{P}}$$

which we aim to solve with a required precision of $\varepsilon > 0$, where f is a Lipschitz continuous function with constant $L > 0$, defined on an interval $[a,b]$ in $\mathbb{R}$ and taking values in $\mathbb{R}$. In other words, f satisfies

$$\forall\ x, y \in [a, b]\ , \qquad |f(x) - f(y)| \leq L\,|x - y| \tag{2.1}$$

The Piyavskii-Shubert method [11] involves constructing increasingly refined coverings of the objective function over the optimization domain $C = [a, b] \subset \mathbb{R}$. It is applicable to one-dimensional Lipschitz continuous functions with Lipschitz constant $L > 0$ over $[a, b]$.

Definition 2.4.1. We say that a function g is an underestimator (or support function) of the function f on $[a, b]$ if

$$\forall x \in [a, b], \quad g(x) \leq f(x).$$

According to (2.1), we have

$$f(y) - L\,|x - y| \leq f(x) \leq f(y) + L\,|x - y|\,, \quad \forall\ x, y \in [a, b] \tag{2.2}$$

Let's fix $y \in [a, b]$. It follows from (2.2) that the concave function: $F(x) = f(y) - L\,|x - y|$ is an underestimator (or support function) of f on the interval $[a, b]$. We will construct a sequence of points $\{x_k\}$ and a family of sets $\{E_k\}$ satisfying $[a, b] \subset \underset{k}{\cup} E_k$ as follows: Consider the function,

$$F_1(x) = f(x_1) - L\,|x - x_1|\,, \tag{2.3}$$

where $x_1 = y$ is any point in $[a, b]$. The next point of the iteration is defined by:

$$x_2 = \arg\min_{x\in D} f(x).$$

Replacing x_1 with x_2 in (2.3), we obtain an underestimator of $f(x)$ on $[a, b]$, but

$$\max(f(x_1) - L\,|x - x_1|\,, f(x_2) - L\,|x - x_2|)$$

is a new underestimator of f and it is closer to $f(x)$. Therefore, we take

$$F_2(x) = \max_{1\le i\le 2}(f(x_i) - L\,|x - x_i|) \quad \text{and} \quad x_3 = \arg\min_{x\in[a,b]} F_2(x).$$

At step k, we take

$$F_k(x) = \max_{1\le i\le k}(f(x_i) - L\,|x - x_i|), \tag{2.5}$$

and

$$x_{k+1} = \arg\min_{x\in[a,b]} F_k(x) \tag{2.6}$$

In the Piyavskii method, we take the following family of sets as the covering of $[a, b]$:

$$\begin{aligned} E_k &= \left\{x \in \mathbb{R} : \max_{1\le i\le k}(f(x_i) - L\,|x - x_i|) \ge f_k^* - \epsilon\right\} \\ &= \left\{x \in \mathbb{R} : F_k(x) \ge f_k^* - \epsilon\right\}, \end{aligned}$$

where f_k^* is an approximation of the global minimum of f. In the Piyavskii-Shubert method, the sequence of points $\{x_k\}$ is defined by the general term:

$$x_{k+1} = \arg\min_{x\in[a,b]} F_k(x), \text{ with } x_1 = \frac{a+b}{2} \text{ and } F_k(x) = \max_{1\le i\le k}(f(x_i) - L\,|x - x_i|),$$

in this case $F_k(x)$ is a piecewise linear underestimator (support function).

Piyavskii-Shubert one-dimensional algorithm

1. **Initialization.**

$k \leftarrow 1, \quad x_1 \leftarrow \frac{a+b}{2}, \quad x_{opt} \leftarrow x_1, \quad f_{opt} \leftarrow f(x_{opt})$.
$F_{opt} \leftarrow f_{opt} - \frac{L}{2}(b-a),\ F_1(x) = f(x_1) - L\,|x - x_1|$.

2. **While** $f_{opt} - F_{opt} > \varepsilon$ **do**

$x_{k+1} \leftarrow \min_{[a,b]} F_k(x)$.
If $f(x_{k+1}) < f_{opt}$; $f_{opt} \leftarrow f(x_{k+1})$; $x_{opt} \leftarrow x_{k+1}$;
End If
$F_{k+1}(x) \leftarrow \max_{1\le i\le k+1}(f(x_i) - L\,|x - x_i|)$
$F_{opt} \leftarrow \min_{[a,b]} F_{k+1}(x)$.
$k \leftarrow k + 1$ **End While**

Remark 2.4.1. 1) This method is simple and very effective in the one-dimensional case. In the multidimensional case where C is a box $\prod_{i=1}^{n}[a^i, b^i]$ in $\mathbb{R}^n$, the algorithm loses its efficiency. Some extensions of the Piyavskii method in both one-dimensional and multi-dimensional cases will be given eventually.
2) The restriction of the function $F_k(x)$ on each subinterval $[x_{i-1}, x_i]$ is a piecewise linear function (see figure 2). Its global minimum is attained at the intersection point of two lines with equations:

$$\begin{aligned} y_1(x) &= f(x_i) - L(x_i - x) \\ y_2(x) &= f(x_{i-1}) - L(x - x_{i-1}) \end{aligned}$$

y_1 is increasing and y_2 is decreasing. The intersection point of y_1 and y_2 is explicitly given for $i = 2, ..., k$ by

$$\begin{aligned} x_k &= \frac{1}{2L}(f(x_{i-1}) - f(x_i)) + \frac{1}{2}(x_i + x_{i-1}) \\ \text{and } F_k(x_k) &= \min_{[x_{i-1}, x_i]} F_k(x) = \frac{1}{2}(f(x_{i-1}) + f(x_i)) - \frac{L}{2}(x_i - x_{i-1}). \end{aligned}$$

Convergence of the Piyavskii-Shubert Method

Theorem 2.4.1. *The Piyavskii-Shubert method applied to Lipschitz continuous functions converges in a finite number of iterations to the global minimum with precision less than or equal to ε.*

Proof. Let m denote the exact global minimum of f on $[a, b]$, and let f_ε be the approximate global minimum of f obtained by the Piyavskii method. We will show that $f_\varepsilon - m \leq \varepsilon$.

Suppose the algorithm has reached the global minimum f_ε with precision ε after k point evaluations $x_1, x_2, \ldots, x_k$.

Then, we have:

$$f_\varepsilon = \min\{f(x_1), f(x_2), \ldots, f(x_k)\} \text{ (the k-th record).}$$

And $f_\varepsilon - \min_{[a,b]} F_k(x) \leq \varepsilon$ (since $F_k(x)$ is an underestimator and f_ε are evaluations of f), where $F_k(x) = \max_{1 \leq i \leq k}(f(x_i) - L|x - x_i|)$.

Since $F_k(x) \leq f(x)$ for all $x \in [a, b]$, then

$$\min_{[a,b]} F_k(x) \leq m \leq f_\varepsilon$$

and consequently

$$f_\varepsilon - f_\varepsilon \leq f_\varepsilon - m \leq f_\varepsilon - \min_{[a,b]} F_k(x) \leq \varepsilon.$$

This completes the proof. □

2.4.2 Brent's method

Problem Statement. Consider the following one-dimensional global minimization problem:

$$\min_{x \in [a,b]} f(x) \qquad (\mathbf{P})$$

which we want to solve with a required precision $\varepsilon > 0$, where f is a C^2 function defined on the interval $[a, b]$ of $\mathbb{R}$ and takes values in $\mathbb{R}$.

Brent's method is a one-dimensional method developed by Brent. It applies to twice continuously differentiable functions whose second derivative is bounded. In other words, there exists a constant $M > 0$ such that:

$$\left| f''(x) \right| \leq M, \text{ for all } x \in [a, b].$$

The idea is to construct an increasing sequence of piecewise parabolic functions $(\varphi_k)_{k\in\mathbb{N}}$ on $[a,b]$ such that:

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

Let's recall that a function F is piecewise parabolic on $[a,b]$ if there exists a finite family of disjoint intervals $(\mathbb{I}_i)_{1\leq i\leq k}$ such that: $[a,b] = \underset{1\leq i\leq k}{\cup} \mathbb{I}_i$ and for all $i = 1, ..., k$, the restriction of F to the interval $\mathbb{I}_i$ is a parabolic function. This allows us to generate a sequence of points $x_k^* = \underset{x\in[a,b]}{\arg\min}\varphi_k$, which converges to an optimal solution of problem $(\mathbf{P})$.
Brent's method relies on the following result:

Theorem 2.4.2. *Let $f : [a,b] \to \mathbb{R}$ be a function twice continuously differentiable and satisfies $\left|f''(x)\right| \leq M$ for all $x \in [a,b]$. Then for any x_1, $x_2 \in [a,b]$ such that $x_1 \leq x_2$, the parabola $\varphi(x)$ defined by:*

$$\varphi''(x) = M, \ \forall x \in [a,b] \quad and \quad \left\{ \begin{array}{c} \varphi(x_1) = f(x_1) \\ and\ \varphi(x_2) = f(x_2) \end{array} \right.$$

is an underestimate (or support function) of f on $[x_1, x_2]$.

Proof. Let's define $g(x) = \varphi(x) - f(x)$ for $x \in [a,b]$, then
$g''(x) = \varphi''(x) - f''(x) = M - f''(x) \geq 0$, thus g is convex and we have: $g(x_1) = g(x_2) = 0$, hence we necessarily have

$$g(x) \leq 0, \forall x \in [x_1, x_2], \text{ where } \varphi(x) \leq f(x), \forall x \in [x_1, x_2].$$

□

Remark 2.4.2. If the function f is evaluated at points $a = x_0 \leq x_1 \leq ... \leq x_k \leq x_{k+1} = b$ within $[a,b]$, then we can construct a parabola $\varphi_{[x_j,x_{j+1}]}$ on each interval $[x_j, x_{j+1}]$ for $j = 0, ..., k$, satisfying:

$$\left\{ \begin{array}{l} \varphi_{[x_j,x_{j+1}]}(x_j) = f(x_j) \\ \varphi_{[x_j,x_{j+1}]}(x_{j+1}) = f(x_{j+1}) \\ \text{and } \varphi''_{[x_j,x_{j+1}]}(x) = M \text{ for } x \in [x_j, x_{j+1}]. \end{array} \right.$$

This results in a sequence of piecewise parabolic functions, where the general term is denoted by:

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

Now, let's find a parabola $\varphi(x)$ satisfying $\varphi''(x) = M$ and passing through the two points $(x_1, f(x_1))$ and $(x_2, f(x_2))$, where $x_1 \neq x_2$. Assume $\varphi(x) = \frac{M}{2}x^2 + \alpha x + \beta$. We have

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

This system has the solution:

$$\begin{aligned} \alpha &= \frac{f(x_1) - f(x_2)}{x_1 - x_2} - \frac{M}{2}(x_1 + x_2) \\ \beta &= \frac{x_1 f(x_2) - x_2 f(x_1)}{x_1 - x_2} + \frac{M}{2}x_1 x_2. \end{aligned}$$

The minimum of φ necessarily occurs at the point x_0, which satisfies $\varphi^{'}(x_0) = 0$, hence

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

Example 2.4.1. Figure 2 below shows an example of a function f of class $C^2([0,1])$ and its minorant function $\varphi(x)$ (parabolic) on the interval $[0,1]$, where f is given by

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

We have:

$$\begin{aligned} f^{'}(x) &= 4\pi\cos((x - \frac{1}{2})4\pi) + 12x, \forall x \in [0,1]\,, \\ f^{''}(x) &= -16\pi^2\cos((x - \frac{1}{2})4\pi) + 12, \forall x \in [0,1]\,, \end{aligned}$$

thus we have

$$\left|f^{''}(x)\right| \leq 16\pi^2 + 12, \forall x \in [0,1]\,.$$

We set $M = 16\pi^2 + 12$. There exists then a parabolic minorant function $\varphi(x)$ such that $\varphi(x) \leq f(x), \forall x \in [0,1]$ and $\varphi(0) = f(0)$ and $\varphi(1) = f(1)$.

The parabolic function $\varphi(x)$ is defined by $\varphi(x) = \frac{M}{2}x^2 + \alpha x + \beta$ with $\alpha = \frac{f(0)-f(1)}{0-1} - \frac{M}{2}(0+1) = -8\pi^2$ and $\beta = \frac{0f(1)-1f(0)}{0-1} + \frac{M}{2}0 \times 1 = 2$, whence $\varphi(x) = \frac{M}{2}x^2 + \alpha x + \beta = (8\pi^2 + 6)\,x^2 + (-8\pi^2)x + 2$
and $x_0 = -\frac{\alpha}{M} = \frac{f(1)-f(0)}{M(0-1)} + \frac{0+1}{2} = \frac{8\pi^2}{16\pi^2+12} = 0.464\,69 \in [0,1]$

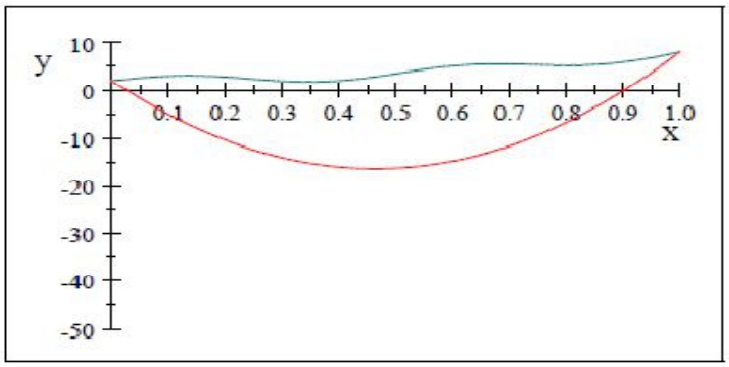

Figure 2.2: The parabolic minorant function (in red) of the objective function (in blue)

Brent's Algorithm.

1. Initialization.
Set $k = 1,\ x_1 = \frac{f(a)-f(b)}{M(b-a)} + \frac{a+b}{2},\ x_\epsilon = \arg\min\{f(a),\ f(x_1), f(b)\}\,,\ f_\epsilon = f(x_\epsilon)$.
$F_\epsilon = \varphi_{[a,b]}(x_1),\ F_1 = \left[\varphi_{[a,x_1]}, \varphi_{[x_1,b]}\right]$.

2. Step $k = 2, 3, ...$

If $f_\epsilon - F_\epsilon \leq \epsilon$, then stop.

Otherwise, determine $x_{k+1} = \arg\min F_k([a,b])$.
If $f(x_{k+1}) \leq f_\epsilon$, then set $f_\epsilon = f(x_{k+1})$, $x_\epsilon = x_{k+1}$.
Set $F_{k+1}(x) = \left[\varphi_{[a,x_1]}, \varphi_{[x_1,x_2]}, ..., \varphi_{[x_k,b]}\right]$ and $F_\epsilon = \min F_{k+1}([a,b])$.
Set $k = k+1$ and go to 2.

2.4.3 Conclusion

The covering methods mentioned in this chapter are simple to program and very effective in the one-dimensional case. Brent's method generates a smaller number of evaluation points than the methods of Evtushenko and Piyavskii-Shubert, but it is applicable only to the class of functions that are twice differentiable with bounded second derivatives, whereas the other algorithms apply to a larger class (Lipschitz functions, Hölder functions, etc.). Additionally, its disadvantage is that it does not generalize to the multidimensional case, and it loses its efficiency if the objective function is fairly flat near the global minimum.

2.5 Exercises

Exercise 2.5.1. Let $P \subset \mathbb{R}^n$ be a convex set and let f be a continuously differentiable function on an open set Ω containing P with the gradient of f bounded on P. Then show that f is a Lipschitz function on P with Lipschitz constant L given by

$$L = \sup\left\{\|\nabla f(x)\| \ , x \in P\right\}.$$

Example 2.5.1. Let $f(x) = \sin\left((x - \frac{1}{2})4\pi\right) + 6x^2 + 2, \ \ x \in [0,1]$. Calculate L. [5]

Solution 2.5.1. Let $x, y \in P$, since P is convex, there exists $z \in P$ such that $z = \lambda x + (1-\lambda)y, \lambda \in [0,1]$. Using the Mean Value Theorem,

$$\begin{aligned} f(x) - f(y) &= (x-y)^t \cdot \nabla f(z) \\ &\Rightarrow \ |f(x) - f(y)| = \left|(x-y)^t \cdot \nabla f(z)\right|, \end{aligned}$$

since ∇f is bounded on P,

$$\begin{aligned} |f(x) - f(y)| &= \left|(x-y)^t \cdot \nabla f(z)\right| \leq \|\nabla f(z)\| \, \|x-y\| \\ &\leq \sup_P \|\nabla f(z)\| \, \|x-y\|. \end{aligned}$$

Let $L = \sup\left\{\|\nabla f(x)\| \ , x \in P\right\}$ then

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

Thus, f is Lipschitz on P with Lipschitz constant $L = \sup_P \|\nabla f(z)\|$.

Remark 2.5.1. 1) If the set P is a compact subset of $\mathbb{R}^n$, we can apply the previous result, i.e., f is Lipschitz with constant
$L = \max_P \|\nabla f(z)\|$. (because $\nabla f(x)$ is continuous on the compact P, thus $\nabla f(x)$ is bounded).

2) According to the expression of the Lipschitz constant, we note that L is the smallest constant that satisfies the Lipschitz condition.

Example 2.5.2. Let $f(x) = \sin((x - \frac{1}{2})4\pi) + 6x^2 + 2$, $x \in [0, 1]$. We will show that f is Lipschitz and calculate L.
f is differentiable on the compact $[0, 1]$ of $\mathbb{R}$, according to the previous result, f is Lipschitz on $[0, 1]$ and the Lipschitz constant is given by the expression

$$L = \max_{[0,1]} \left| f\,'(x) \right|.$$

Calculating $f\,'(x) = 4\pi \cos((x - \frac{1}{2})4\pi) + 12x$ and $\left|f\,'(x)\right| = \left|4\pi \cos((x - \frac{1}{2})4\pi) + 12x\right| \leq 4\pi + 12, \forall x \in [0, 1]$.
Here, $4\pi + 12 = L = \max\limits_{[0,1]} \left|f\,'(x)\right|$ (the smallest Lipschitz constant).

Exercise 2.5.2. Let f, g, f_i $(i = 1, ..., m)$ be Lipschitz functions on a compact set C in $\mathbb{R}^n$. Show that the following propositions hold:
i) Each linear combination of the functions f_i $(i = 1, ..., m)$ is a Lipschitz function on C.
ii) The functions $\max\limits_{i=1,...,m} f_i$ and $\min\limits_{i=1,...,m} f_i$ are Lipschitz functions on the compact set C.
iii) The product function $f \cdot g$ is Lipschitz on C.
iv) Let f be a Lipschitz function on an interval C_1 with constant L_1 and g be Lipschitz on an interval C_2 with constant L_2 such that $f(C_1) \subset C_2$. Then the composite function $g \circ f$ is Lipschitz on C_1 with constant $L_1 L_2$.

Example 2.5.3. The function $h(x) = (2x-1)^4$. This function is Lipschitz on every bounded interval. Let $f(x) = 2x - 1$ and $g(x) = x^4$. Consider the interval $C_1 = \left[-\frac{1}{2}, \frac{3}{2}\right]$, then $f(C_1) \subset [-2, 2] = C_2$. Since g is Lipschitz on C_2 with constant 32, and the Lipschitz constant of f is 2, hence h is Lipschitz with constant $32 \times 2 = 64$ on C_1. (We can determine the Lipschitz constant of h using the result from Exercise 1).

Exercise 2.5.3. Let f be a function defined on a bounded interval $[a, b]$ of $\mathbb{R}$ by:

$$f(x) = \begin{cases} f_1(x) \text{ if } x \in [a, c] \\ f_2(x) \text{ if } x \in [c, b] \end{cases}$$

where $c \in [a, b]$ and such that $f_1(c) = f_2(c)$.

Show that if f_1 is Lipschitz with Lipschitz constant L_1 on $[a, c]$ and f_2 is Lipschitz with Lipschitz constant L_2 on $[c, b]$, then f is Lipschitz with constant $L = L_1 + L_2$ on $[a, b]$. Provide an example.

Example 2.5.4. Consider the function $f(x) = \begin{cases} |x| & \text{if } x \in [-1, 1] \\ x^2 & \text{if } x \in [1, 2] \end{cases}$. Here, we have $f_1(x) = |x|$ and $f_2(x) = x^2$, and $f_1(1) = f_2(1)$.
f_1 is Lipschitz with constant 1 on $[-1, 1]$ and f_2 is Lipschitz on $[1, 2]$ with constant 4, thus f is Lipschitz on $[-1, 2]$ with constant $L = 1 + 4 = 5$.

Note: f is not differentiable on $[-1, 2]$. We cannot use the result from Exercise 1, but f is Lipschitz.

Exercise 2.5.4. Let $f : [0, c] \to \mathbb{R}$ be a Lipschitz function with constant $L > 0$ satisfying $f(0) = f(c)$. Consider the function $g : [0, 2c] \to \mathbb{R}$ defined by:

$$g(x) = \begin{cases} f(x) & \text{if } x \in [0, c] \\ f(x - c) & \text{if } x \in [c, 2c] \end{cases},$$

then show that g is Lipschitz on $[0, 2c]$. Determine the Lipschitz constant of g.

Exercise 2.5.5. Differentiable Lipschitz problems with a single variable [6] .

Determine the Lipschitz constants of the following functions on each indicated domain.
$(16x^2 - 24x + 5)\, e^{-x}, [1.9, 3.9]$, Ans. 3>2.937
$(-3x + 1.4) \sin 18x, [0, 1.2]$, Ans. 36>35.455
$(x + \sin x)\, e^{-x^2}, [-10, 10]$, Ans. 2.5
$x^{2/3} - (x^2 - 1)^{1/3}$, $[0.001, 0.99]$, Ans. 8.5
$\frac{-x^2+5x-6}{x^2+1}, [-5, 5]$, Ans. 6.5
$\sum_{k=1}^{5} k \cos\left[(k+1)x + k\right], [-10, 10]$, Ans. 67

Exercise 2.5.6. **Non-differentiable Lipschitz problems with a single variable.** Determine the Lipschitz constants of the following functions on each indicated domain [6].

1°
$$\min_{x\in[0,4]} f(x) = \frac{1}{4}\left(\left|x - \frac{3}{2}\right| - |\sin 10x| + 3\right)$$

2°
$$f(x) = \begin{cases} -4x^2 + \frac{89}{25} & x \le \frac{4}{5} \\ \sin(5x-4) + x + \frac{1}{5} & x > \frac{4}{5} \end{cases} \qquad x \in [0, 5]$$

3°
$$\min_{x\in[0,4]} f(x) = \begin{cases} 6(x - \frac{1}{2})^2 - \frac{1}{2} & x \le \frac{1}{2} \\ \frac{1}{4}(x - \frac{5}{2}) & x > \frac{1}{2} \end{cases}$$

4°
$$\max_{[0,5]} \left\{ \left(x - \frac{3}{4}\right)\left(x - \frac{21}{5}\right), \ -\left(x - \frac{11}{5}\right)(x-3) \right\}$$

5°
$$\min_{x\in[0,10]} f(x) = \begin{cases} x - 4 & x \le 3 \\ \frac{8}{9}x^2 - 8x + 15 & 3 < x \le 6 \\ -x + 5 & x > 6 \end{cases}$$

6°
$$f(x) = -5\left|\sin(\frac{1}{2}(x - \frac{1}{2})\right| + 1, \qquad x \in [0, 10]$$

7°
$$\min_{x\in[0,4]} f(x) = -\cos(3x)\,|x \sin x| + 8$$

8°
$$f(x) = \frac{8}{25} - e^{-x}\,|\sin(3\pi x)|\ , \quad x \in [0, 4]$$

9°
$$f(x) = \begin{cases} -(x-2)^2 & \text{if } x \le 3 \\ -2\ln(x-2) - 1 & \text{else} \end{cases} \qquad x \in [0, 6]$$

Exercise 2.5.7. Let $f : [a, b] \to \mathbb{R}$ be a Lipschitz function with constant $L > 0$, and let F be a piecewise linear underestimator of f on each subinterval $[x_{i-1}, x_i]$ of $[a, b]$. Show that F attains its global minimum at the intersection point of two lines with equations:

$$y = f(x_i) - L(x_i - x) \quad \text{and} \quad y = f(x_{i-1}) - L(x - x_{i-1}),$$

which is explicitly given by:

$$\begin{aligned} x^* &= \frac{1}{2L}\left(f(x_{i-1}) - f(x_i)\right) + \frac{1}{2}\left(x_i + x_{i-1}\right) \\ \text{and} \quad F(x^*) &= \frac{1}{2}\left(f(x_{i-1}) + f(x_i)\right) - \frac{L}{2}\left(x_i - x_{i-1}\right). \end{aligned}$$

Exercise 2.5.8. Consider the following global minimization problem:

$$\min_{x\in[0,4]} f(x) = \begin{cases} 6\left(x - \frac{1}{2}\right)^2 - \frac{1}{2} & \text{if } x \leq \frac{1}{2} \\ \frac{1}{4}\left(x - \frac{5}{2}\right) & \text{if } x > \frac{1}{2} \end{cases} \qquad (P)$$

1. Determine the Lipschitz constant $L > 0$ of f on the interval $[0, 4]$.
2. Show that problem (P) has at least one solution in $[0, 4]$.
3. Determine the number of passes N_{Pass} and iterations N_{Iter} necessary to achieve the global minimum of f on $[0, 4]$, with a given precision $\varepsilon > 0$.
4. Calculate three approximations x_1, x_2, and x_3 in terms of ε and L, using the methods of **Piyavskii** and **Evtushenko.**
5. Using the **Brent** algorithm, compute the global minimum of f on $\left[0, \frac{1}{2}\right]$ with precision $\varepsilon = 0.1$.
6. Determine the global minimum of f on $\left[\frac{1}{2}, 4\right]$, and deduce $\min\limits_{[0,4]} f(x)$.

Exercise 2.5.9. Inspired by the **Piyavskii** method, describe an algorithm to solve the following global optimization problem with precision $\varepsilon > 0$:

$$\max_{x\in[a,b]} f(x) \qquad (P)$$

where f is a Lipschitz function on $[a, b]$ with constant L.

Note: Without transforming problem (P) into the problem $\min\limits_{x\in[a,b]} (-f(x))$.

Chapter 3

Global Multidimensional Optimization Methods

Contents

3.1 Introduction

The authors have utilized the generalization of the Piyavskii algorithm to the single-variable case. For the multivariable case, a direct generalization of the Piyavskii algorithm is very challenging, as determining the local minima of the surrogate function overestimating the objective function requires determining the intersection of several parabolic hypersurfaces. They proposed a procedure to partition and eliminate uninteresting hyperrectangles (Branch-and-Bound) by constructing piecewise constant lower bound functions.

The Branch-and-Bound method [10] is intended for solving a class of multidimensional global optimization problems. It is an iterative method that divides a given hyperrectangle $\Omega \subset \mathbb{R}^n$ into smaller and smaller sub-boxes. At each sub-box of Ω, a lower bound of the objective function f is constructed in order to eliminate parts that do not contain the global optimum and to select the sub-box that needs to be divided.
This method is widely used for solving a large number of mathematical problems with various structures using tools from convex analysis.

3.2 Branch-and-Bound Separation and Evaluation Method

3.2.1 General Branch and Bound Algorithm

Consider the global minimization problem:

$$\min_{x\in\Omega} f(x) \qquad (P)$$

where Ω is a compact subset of $\mathbb{R}^n$ and $f : \Omega \to \mathbb{R}$ is a continuous and non-convex function on Ω.
The $B\&B$ algorithm consists of generating two convergent sequences $\{BS_k\}$ and $\{BI_k\}$ of upper and lower bounds respectively for the minimum value of the function f of problem (P).
We denote by BS : Upper Bound and BI : Lower Bound.
An initial relaxation D of the feasible set Ω will be defined such that: $\Omega \subset D$.
D is convex and can be a rectangle, a cone, a cube, etc. At each iteration k, the lower and upper bound problems will be solved on a finite number of subsets of D. We denote these subsets as $D_{k_i} \in I_k$ where I_k is the set of active subsets at iteration k. On each subset D_{k_i}, the lower and upper bounds BI_k and BS_k will be calculated by relaxing f over $D_{k_i} \cap \Omega$ respectively. This method uses the "best-first" strategy. Indeed, the final lower and upper bounds for iteration k will be given by:

$$\begin{cases} BS_k = \min\limits_i BS_{k_i} \\ BI_k = \min\limits_i BI_{k_i}, \end{cases}$$

respectively, and any subset on which the lower bound exceeds BS_k will be eliminated, as $\min f$ cannot be attained on such a subset. Let x_ε be the optimal solution of problem (P) for the following.

The Branch-and-Bound Algorithm

1. Build the set D such that: $\Omega \subset D$.
2. Let's set: $k = 1$, $I_k = D$, and $\varepsilon > 0$, a given precision.
3. Construct the lower and upper bound problems of $\min f(x)$ on D. Let BI_k and BS_k be the respective solutions.
4. If $BS_k - BI_k \leq \varepsilon$, then we stop and set: $\min f(x) = BS_k$ and $x_\varepsilon = x^k \in \{x : f(x) = BS_k, \quad x \in \Omega \cap D\}$.

5. Otherwise, subdivide I_k into two (or a finite number of) subsets D_{k_1} and D_{k_2} such that:

$$\bigcup_{i=1}^{2} D_{k_i} = D; \quad \text{et} \quad D_{k_1} \cap D_{k_2} = \varnothing$$

6. Construct the lower and upper bound problems of $\min f(x)$ on $\Omega \cap D_{k_i}$, $i = 1, 2$. Let $BI_1,\ BS_1$ and $BI_2,\ BS_2$ be the respective solutions.
7. Let's set:

$$\begin{aligned} BS_{k+1} &= \min\{BS_{k_1},\ BS_{k_2},\ BS_k\} \\ BI_{k+1} &= \min\{BI_{k_1},\ BI_{k_2}\}. \end{aligned}$$

8. Let's set: $I_k = \{D_{k_1},\ D_{k_2}\}$.
9. Eliminate from I_k any subset $D_{k_i}, \qquad i = 1, 2.$ such that:

$$BI_{k_i} > BS_{k+1}$$

10. Let's set: $k = k + 1$ and go to 2.

D_k : The current subset.
BI_k and BS_k : The lower and upper bounds respectively found at the k^{th} iteration.
x_k : The solution obtained at the k^{th} iteration.

3.2.2 Branch and Bound Algorithm on the Main Diagonal of an Hyper-rectangle in $\mathbb{R}^n$

Description of the method

Consider the following global optimization problem:

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

where f is a Hölder continuous function defined on the hyperrectangle $\Omega = \{x \in \mathbb{R}^n : a \leq x \leq b\}$ where $a, b \in \mathbb{R}^n$, and takes values in $\mathbb{R}$.
It has already been noted that the Piyavskii method is numerically inefficient for $n \geq 2$. For this reason, we prefer to present the extension of Branch-and-Bound by essentially applying the one-dimensional version of the Piyavskii-Shubert method to the main diagonal of the rectangular sets in the partition of Ω [14].
Let a_M, $b_M \in \mathbb{R}^n$, denote the lower and upper bounds of the hyperrectangle M, i.e.,

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

The B&B-Piyavskii-Diagonal HOL$_2$ Algorithm

1) Initialization:
$M_0 = \Omega, \quad \mathcal{M}_0 = \{M_0\}$.
$\alpha_0 = \min\{f(a), f(b)\}, \quad x^0 \in \{a, b\}$ such that $f(x^0) = \alpha_0$.
$\beta_0 = \max\{f(a), f(b)\} - h\|b - a\|^{\alpha}$.
Go to step k.

2) Step $k = 1, 2, ...$

At the beginning of step k, we have the current partition $\mathcal{M}_{k-1}$ of a subset of $M_0 = \Omega$ which is still of interest and for all $M \in \mathcal{M}_{k-1}$, there exist bounds $\beta(M)$, $\alpha(M)$ such that:

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

moreover, we have the bounds β_{k-1}, α_{k-1} such that:

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

with $\alpha_{k-1} = f(x^{k-1})$, where x^{k-1} is the record obtained in step $(k-1)$.

k.1. Exclude any subset $M \in \mathcal{M}_{k-1}$ satisfying:

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

Consider the class $\Re_k$ of the remaining subsets of $\mathcal{M}_{k-1}$.

k.2. Choose a class $\Im_k \subset \Re_k$, such that:

$$\Im_k \cap \arg\min \{\beta(M) : M \in \Re_k\} \neq \emptyset.$$

Then, in each $\widetilde{M} \in \Im_k$, consider the point $x_{\widetilde{M}}$ given by:

$$x_{\widetilde{M}} = \frac{1}{2}(a_{\widetilde{M}} + b_{\widetilde{M}}) + \frac{f(a_{\widetilde{M}}) - f(b_{\widetilde{M}})}{2h \left\| b_{\widetilde{M}} - a_{\widetilde{M}} \right\|^{\alpha}}(b_{\widetilde{M}} - a_{\widetilde{M}})$$

and subdivide $\widetilde{M}$ into two, by the hyperplane containing $x_{\widetilde{M}}$ and orthogonal to the longest face of $\widetilde{M}$.

Let $\mathcal{M}'_k$ be the set of all new elements of the partition.

For each $M' \in \mathcal{M}'_k$, denote by $\widetilde{M'} \in \Im_k$, the hyperrectangle whose subdivision generates M'.

k.3. For each $M' \in \mathcal{M}'_k$, set:

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

k.4. Determine

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

k.5. If $\alpha_k - \beta_k \leq \varepsilon$, then stop.

Otherwise, go to step $k+1$.

3.3 The reducing transformation method

3.3.1 Presentation of the Method

This is a multidimensional global optimization method that was developed in the early 1980s at the MEDIMAT Laboratory. It was invented by Y. Cherruault and collaborators [1], [21] [18]. However, it is in the works of Y. Cherruault that the method received its theoretical justification, complemented by the results of A. Ziadi, G. Mora, and Y. Cherruault.
The main idea is to transform an optimization problem with n variables into an optimization problem with a single variable. It is a reductive transformation using approximations of a compact set S in $\mathbb{R}^n$ by α-dense curves. Let $f(x_1, x_2, \ldots, x_n)$ be a function of n variables $x_1, x_2, \ldots, x_n$. The method consists of expressing these variables using a single variable by densifying the compact set S with the help of a simple curve of class C^∞. We will highlight what we call the reductive transformation:

$$x_i = h_i(t), \quad t > 0, \ i = 1, 2, \ldots, n$$

where t is a real variable and the h_i are very regular functions.
This transformation allows approximating the compact set S by a curve with the desired precision, meaning it reduces a multivariable function to a single-variable function $f^*(t)$ [1].
We denote by:

μ : the Lebesgue measure.
A : a closed and bounded interval of $\mathbb{R}$, (generally $A = [0, T]$, with $T > 0$).
α : a real number, strictly positive and assumed to be very small compared to the dimensions of the hyperrectangle $S = \prod_{i=1}^{n} [a_i, b_i]$, ($\alpha << \min_{1 \le i \le n} (b_i - a_i)$), where $n \geq 2$, is an integer.

The Aliénor Method for Multidimensional Global Optimization

Consider the following multidimensional global optimization problem:

$$\min_{x \in S \subset \mathbb{R}^n} f(x) \tag{Pb}$$

where f is a continuous function with constant $L(\varepsilon)$ defined on a compact set S, and taking values in $\mathbb{R}$.
The Aliénor transformation

$$x_i = h_i(t), \ i = 1, \ldots, n$$

is used to approximate the function f by the single-variable function f^* defined by:

$$f^*(t) = f(h_1(t), h_2(t), ..., h_n(t)).$$

The minimization problem (Pb) is then approximated by the minimization problem

$$\min_{t \in [0,T]} f^*(t) \tag{Pb*}$$

which we can solve simply using unidimensional methods (e.g., modified Piyavskii method).

3.3.2 Construction of α-dense curves in a box in $\mathbb{R}^n$

Definition 3.3.1. A parameterized curve in $\mathbb{R}^n$ defined by:

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

is said to be α-dense in S, or that it α-densifies S, if for every $x \in S$, there exists $t \in A$ such that

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

where d is the Euclidean distance in $\mathbb{R}^n$.

Consider the function $h(t) = (h_1(t), h_2(t), ..., h_n(t))$, defined on an interval A of $\mathbb{R}$ and taking values in a box $S \subset \mathbb{R}^n$. We study the relationship between the component functions $h_1, h_2, ..., h_n$ so that the function h generates an α-dense curve in the box S [16].

Theorem 3.3.1. *Let $h = (h_1, h_2, ..., h_n) : A \longrightarrow S$ be a continuous function and $t_2, t_3, ..., t_n, \alpha$ be strictly positive numbers such that:*

1. *h_1 is surjective.*
2. *For all $i = 2, ..., n$, h_i reaches the bounds a_i and b_i in every closed interval of length t_i.*
3. *For all $i = 2, ..., n$, and for every interval I of A we have:*

$$\mu(I) < t_i \Longrightarrow \mu(h_{i-1}(I)) < \alpha.$$

Then the parameterized curve defined by $h(t) = (h_1(t), h_2(t), ..., h_n(t))$, for $t \in A$, is $\sqrt{n-1}\alpha$-dense in $S = \prod_{i=1}^{n} [a_i, b_i]$.

Proof. (By induction)
i) For $n = 2$, $S = [a_1, b_1] \times [a_2, b_2]$.
Let $x = (x_1, x_2) \in S$. Consider the interval $[x_1 - \alpha, x_1 + \alpha]$.
Since h_1 is surjective, there exists a closed interval $I \subset A$ such that:

$$h_1(I) = [x_1 - \alpha, x_1 + \alpha] \cap [a_1, b_1].$$

But $\mu(h_1(I)) \geq \alpha$, so $\mu(I) \geq t_2$ and consequently $h_2(I) = [a_2, b_2]$. Moreover, $x_2 \in [a_2, b_2]$ so $\exists t' \in I$ such that $x_2 = h_2(t')$, and since $h_1(t') \in [x_1 - \alpha, x_1 + \alpha] \cap [a_1, b_1]$ then $|x_1 - h_1(t')| \leq \alpha$.
We deduce that

$$\begin{aligned} \|x - h(t')\| &= \|(x_1, x_2) - (h_1(t'), h_2(t'))\| \\ &\leq \sqrt{(x_1 - h_1(t'))^2 + (x_2 - h_2(t'))^2} \\ &\leq \alpha. \end{aligned}$$

ii) Assume that the theorem is verified for $(n-1)$.
Let $x = (x_1, x_2, ..., x_n) \in \prod_{i=1}^{n} [a_i, b_i]$. Consider the interval $[x_{n-1} - \alpha, x_{n-1} + \alpha]$. Given that h_{n-1} is surjective, there exists a closed interval $I \subset A$ such that:

$$h_{n-1}(I) = [x_{n-1} - \alpha, x_{n-1} + \alpha] \cap [a_{n-1}, b_{n-1}],$$

but $\mu(h_{n-1}(I)) \geq \alpha$, so $\mu(I) \geq t_n$. It follows that $h_n(I) = [a_n, b_n]$. Let $x_n \in [a_n, b_n]$, so $\exists t'' \in I$ such that $x_n = h_n(t'')$. Moreover, the sequence $t_2, t_3, ..., t_n$ is increasing, so the functions $h_1, h_2, ..., h_n$ restricted to the interval I are surjective and satisfy the hypotheses of the theorem.
Since $h_{n-1}(t'') \in [x_{n-1} - \alpha, x_{n-1} + \alpha] \cap [a_{n-1}, b_{n-1}]$, then

$$|x_{n-1} - h_{n-1}(t'')| \leq \alpha,$$

we deduce that:

$$\begin{aligned} \|x - h(t'')\| &= \|(x_1, x_2, ..., x_{n-1}, x_n) - (h_1(t''), h_2(t''), ..., h_{n-1}(t''), h_n(t''))\| \\ &\leq \sqrt{n-1}\alpha. \end{aligned}$$

□

Example of α-dense Curves

Let $h = (h_1, h_2, ..., h_n) : [0, \frac{\pi}{\alpha_1}] \to \prod_{i=1}^{n} [a_i, b_i]$ (with $n \geq 2$) be a function defined by:

$$\begin{cases} h_1(t) &= \frac{a_1 - b_1}{2} \cos(\alpha_1 t) + \frac{a_1 + b_1}{2} \\ h_2(t) &= \frac{a_2 - b_2}{2} \cos(\alpha_2 t) + \frac{a_2 + b_2}{2} \\ &\vdots \\ h_n(t) &= \frac{a_n - b_n}{2} \cos(\alpha_n t) + \frac{a_n + b_n}{2}, \end{cases}$$

where $\prod_{i=1}^{n} [a_i, b_i]$ is a box in $\mathbb{R}^n$ and $\alpha, \alpha_1, \alpha_2, ..., \alpha_n$ are given strictly positive constants satisfying the relation:

$$\alpha_i = \frac{\pi(b_{i-1} - a_{i-1})\sqrt{n-1}}{2\alpha} \alpha_{i-1}, \text{ for } \forall i = 2, ..., n.$$

Then the curve defined by $h(t) = (h_1(t), h_2(t), ..., h_n(t))$ for $t \in [0, \frac{\pi}{\alpha_1}]$ is α-dense in $\prod_{i=1}^{n} [a_i, b_i]$.

Proof. Indeed

1) h_1 is surjective.

2) For all $i = 2, 3, ..., n$, h_i reaches the bounds a_i and b_i in any closed interval of length $t_i = \frac{\pi}{\alpha_i}$, there exist $t_i^{'} = \frac{2\pi k}{\alpha_i}$ and $t_i^{''} = \frac{\pi}{\alpha_i} + \frac{2\pi k}{\alpha_i} \in I$, $k \in \mathbb{N}$.

3) For all $i = 2, 3, ..., n$, and for any interval I such that $\mu(I) < \frac{\pi}{\alpha_i}$ we have $\mu(h_{i-1}(I)) < \frac{\alpha}{\sqrt{n-1}}$, this results from the fact that for an interval I of length strictly less than $\frac{\pi}{\alpha_i}$, we

have:

$\forall t^{'}, t^{''} \in I$,

$$\begin{aligned}
\left|h_{i-1}(t^{'}) - h_{i-1}(t^{''})\right| &= \left|\frac{a_{i-1}-b_{i-1}}{2}\right| \left|\cos(\alpha_{i-1}t^{'}) - \cos(\alpha_{i-1}t^{''})\right| \\
&\leq \left|\frac{a_{i-1}-b_{i-1}}{2}\right| \left|2\sin(\alpha_{i-1}\frac{t^{'}-t^{''}}{2})\sin(\alpha_{i-1}\frac{t^{'}+t^{''}}{2})\right| \\
&\leq |b_{i-1}-a_{i-1}| \left|\sin(\alpha_{i-1}\frac{t^{'}-t^{''}}{2})\right| \\
&\leq (b_{i-1}-a_{i-1}) \left|\alpha_{i-1}\frac{t^{'}-t^{''}}{2}\right| \\
&\leq (b_{i-1}-a_{i-1})\alpha_{i-1} \left|\frac{t^{'}-t^{''}}{2}\right| \\
&\leq \frac{(b_{i-1}-a_{i-1})}{2}\alpha_{i-1}\frac{\pi}{\alpha_i} \leq \frac{\alpha}{\sqrt{n-1}}.
\end{aligned}$$

According to theorem (2.3.1) the parameterized curve defined by the function h is α-dense in $\prod_{i=1}^{n}[a_i, b_i]$. □

3.3.3 Constrained global optimization of multivariate functions based on dimensionality reduction through α-dense curves

This subsection deals with the reducing transformation method to study multi-dimensional constrained global optimization problems where the objective function is Hölderian and non-differentiable over a general compact set D of $\mathbb{R}^n$. The fundamental principle is to give explicitly a parametric representation $x_i = \ell_i(t), 1 \leq i \leq n$ of α-dense curve ℓ_α in the compact D, for t in an interval $\mathbb{I}$ of $\mathbb{R}$, which allows to convert the initial problem to a one-dimensional Hölder unconstrained one. Thus, we can solve the problem by using an efficient algorithm available in the case of functions depending on a single variable. A relation between the densification parameter α of the curve ℓ_α and the accuracy of attaining the optimal solution is given. Some concrete α-dense curves in the non-convex feasible region D are constructed.

Let us consider the following constrained global optimization problem

$$\begin{cases} \min F(x) \\ \text{subject to } g_i(x) \leq 0,\ i \in I \\ x \in \Omega \end{cases} \tag{1.1}$$

where $x = (x_1, \cdots, x_n)^T$ is the real vector of $\mathbb{R}^n$ represents the n variables, I is a finite index set and Ω is a compact hyper-rectangle given by

$$\Omega = \left[x_1^l, x_1^u\right] \times \cdots \times \left[x_n^l, x_n^u\right] \subset \mathbb{R}^n,$$

and $x_i^l, x_i^u, i = 1, \cdots, n$ are real constant specifying the hyper-rectangle Ω. The functions $F, g_i\ (i \in I) : \Omega \to \mathbb{R}$ are assumed to be Hölderian and non-differentiable over Ω. Denote

the feasible region of solutions for problem (1.1) associated to inequality constraints and explicit bounds by

$$D = \{x \in \Omega \ , \ g_i(x) \leq 0, \ \text{for all } i \in I\},$$

with the assumption that the feasible region D is bounded and robust, i.e. it is the closure of a non-empty open set.

Definition 1. [4]*A function $F : \Omega \to \mathbb{R}$, is said to be Hölder continuous on a hyper-rectangle Ω, if there exists a pair of finite positive constants H and $\beta < 1$, satisfying*

$$|F(x) - F(y)| \leq H \left\|x - y\right\|^{\beta}, \quad \text{for every } x, y \in \Omega, \tag{1.2}$$

where $\|.\|$ denotes the Euclidean vector norm in $\mathbb{R}^n$, other norms could be also considered.

The two real constants H and β are respectively called Hölder constant and Hölder exponent. If F is a Hölder function with constant H and exponent β on Ω, then it is also Hölderian with constant $H^{'} \geq H$ or exponent $\beta^{'} \leq \beta$. Note that a Hölder function need not be differentiable. A Hölder function with a low value of β is much more irregular than a Hölder function with a large value of β.

In this work, we shall generalize the modified Piyavskii's algorithm MPA which we have introduced in [12] for minimizing multivariate Hölderian functions over a hyper-rectangle of $\mathbb{R}^n$ to minimization problems of strictly Hölder functions of several variables undergoing a finite set of constraints which are also strictly Hölderian.

As the problem (1.1) is difficult to solve, there are only a few works in the literature (see [4]). In this paper, we develop an approach which does not use penalty functions or Lagrange multipliers. It is based on the reduction of the dimension of problem (1.1) by means of a transformation allowing to write the n variables $x_i, 1 \leq i \leq n$ in functions of a single variable t, given by $x_i = \ell_i(t), 1 \leq i \leq n$. In this way, becomes on one-dimensional problem without constraints. Recently, the reduction approach, coupled with some one-dimensional optimization methods has proved its effectiveness in solving various types of global optimization problems where the objective function is Lipschitzian and the feasible set is a relatively small hyper-rectangle [1, 3, 12, 19]. In many cases, the feasible region is a non-convex compact set whose boundaries are non-smooth functions. We treat the problem (1.1) in the case where the feasible set D can be formulated, with a finite number of inequalities, as follows:

$$D = \left\{ x \in \Omega, \quad \begin{array}{c} a \leq x_1 \leq b \\ \vdots \\ \varphi_i(x_1, ..., x_{i-1}) \leq x_i \leq \psi_i(x_1, ..., x_{i-1}), \ 2 \leq i \leq n \end{array} \right\}, \tag{1.3}$$

where $a, b \in \mathbb{R}$ such that $a \leq x_1^l$, $b \leq x_1^u$ and the functions φ_i and ψ_i $(i \geq 2)$ are all Hölderian and non-differentiable over a hyper-rectangles Ω_i given by

$$\Omega_i = \left[x_1^l, x_1^u\right] \times ... \times \left[x_{i-1}^l, x_{i-1}^u\right],$$

respectively with the constants $h_i > 0, \ \beta_i < 1$ and $H_i > 0, \ \gamma_i < 1$, that is,

$$\begin{aligned} |\varphi_i(x) - \varphi_i(y)| &\leq h_i \|x - y\|^{\beta_i}, \quad \text{for all } x, y \in \Omega_i, \\ |\psi_i(x) - \psi_i(y)| &\leq H_i \|x - y\|^{\gamma_i}, \quad \text{for all } x, y \in \Omega_i. \end{aligned}$$

In our approach, the objective function $F(x)$ of the initial problem (1.1) is approximated by a function of a single variable defined by $f(t) = F(\ell_1(t), ..., \ell_n(t))$. A computable parametric α-dense curve $\ell_\alpha(t) = (\ell_1(t), ..., \ell_n(t))$ in the feasible region D should be used and ensured the passage from the constrained multi-dimensional problem (1.1) to the unconstrained one-dimensional simpler one

$$\min_{t \in \mathbb{I}} f(t) \tag{1.4}$$

where the univariate objective function $f(t)$ of the problem (1.4) is Hölderian and non-differentiable over the interval $\mathbb{I} = [0, T]$. The constant T is the upper bound of the function $\ell_\alpha(t)$ on which the global minimum of $f(t)$ in $\mathbb{I}$ will be located. It can be proved that the global minimum of the function $F(x)$ on the compact D is approximated by the global minimum of the function $f(t)$ on the interval $\mathbb{I}$ [12]. This due to the fact that $f(t)$ is the restriction of $F(x)$ to the α-dense curve in D. Thus, we can solve the problem (1.4) by applying a more efficient and performing algorithm proposed for minimizing functions depending on a single variable and satisfying the Hölder condition (1.2). We will devote section 4 to this type of problem.

Illustrative example. The following two-variables function $F(x, y) = \frac{1}{2}\sin(\sqrt{|x - y|}) - \frac{1}{2}\cos(\sqrt{|x + y|})$ is Hölderian on the square $[-5, 5]^2$. This example is inspired by the Egg Hölder function when $n = 2$, see [14] . The function F has on the square, several local minima, -0.5 as global minimum value and three global minimizers $(0, 0); (-5, 5); (5, -5)$. There are marked on the contour plot of $F(x, y)$ (see Fig.1 below), at which $F(x, y)$ is not differentiable and not Lipschitzian.

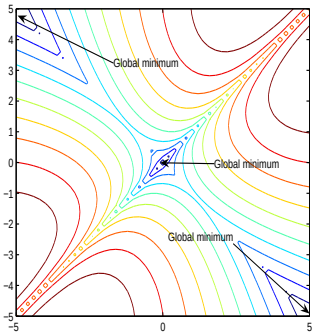

Figure 3.1: A contour plot in 2D of the two-variables Egg Hölder function

The paper is organized as follows. In section 2, we give some preliminaries and notations. The reducing transformation method is presented with concrete α-dense curves in section 3. In section 4 we give an algorithm for a one-dimensional unconstrained problem. Section 5 presents some numerical experiments and section 6 concludes the work.

3.3.4 Lower and Upper bounds

In the preliminaries, we start by recalling that the diameter of the hyper-rectangle Ω_i, is, for each $i = 1, ..., n$,

$$d_i = \max\{\|x - \widetilde{x}\|,\ x, \widetilde{x} \in \Omega_i\} = \sqrt{\left(x_1^l - x_1^u\right)^2 + ... + \left(x_{i-1}^l - x_{i-1}^u\right)^2},$$

and the center of Ω_i, is

$$c_i = \frac{1}{2}\left(x_1^l + x_1^u, ..., x_{i-1}^l + x_{i-1}^u\right).$$

Let the following numbers:

$$\begin{cases} L_i = \varphi_i(c_i) - h_i\left(\frac{d_i}{2}\right)^{\beta_i} \\ U_i = \psi_i(c_i) + H_i\left(\frac{d_i}{2}\right)^{\gamma_i}. \end{cases}$$

Since the functions φ_i and ψ_i $(i \geq 2)$ are Hölderian on Ω_i then by the Hölder condition (1.2) we have

$$\begin{array}{rcl} \varphi_i(y) - h_i\|x - y\|^{\beta_i} & \leq & \varphi_i(x), \text{ for all } x, y \in \Omega_i \\ \psi_i(y) - H_i\|x - y\|^{\gamma_i} & \leq & \psi_i(x), \text{ for all } x, y \in \Omega_i. \end{array}$$

In the last two inequalities, if a value of y is fixed by the center c_i of Ω_i, then, the constants L_i and U_i given above, are respectively lower and upper bounds of the functions φ_i and ψ_i on the hyper-rectangle Ω_i, for $i \geq 2$, with $L_1 = a$ and $U_1 = b$, [4], [12].

Notation and sequence of parameters.

In all this work, the constants h_i, H_i, β_i, γ_i, are assumed to be known a priori and selected such $(h_i)_{2\leq i\leq n}$,$(H_i)_{2\leq i\leq n}$ form two increasing sequences and $(\beta_i)_{2\leq i\leq n}$,$(\gamma_i)_{2\leq i\leq n}$ two decreasing one. Let $\alpha > 0$ and very small compared to the feasible region D and let, for $i \geq 2$,

$$\begin{cases} \delta_1 = 1, \delta_i = 1 + \max\left(h_i\alpha^{\beta_i - 1}\delta_{i-1}^{\beta_i}, H_i\alpha^{\gamma_i - 1}\delta_{i-1}^{\gamma_i}\right) \\ \lambda_i = \max\left(1, h_i(\alpha\delta_{i-1})^{\beta_i - 1}, H_i(\alpha\delta_{i-1})^{\gamma_i - 1}\right). \end{cases}$$

The sequence of numbers $(\delta_i)_{1\leq i\leq n}$ is strictly increasing. Indeed, for $i = 2$, since $h_2 > 0$, $\beta_2 < 1$ and $H_2 > 0$, $\gamma_2 < 1$, then $\delta_2 > \delta_1$. We suppose that $\delta_i < \delta_{i+1}$ for all i, we have

$$h_{i+1} < h_{i+2} \text{ and } \beta_{i+2} < \beta_{i+1} \text{ then for } \alpha << 1 \text{ we have } \alpha^{\beta_{i+1}} < \alpha^{\beta_{i+2}}.$$

Anologously,

$$H_{i+1} < H_{i+2} \text{ and } \gamma_{i+2} < \gamma_{i+1} \text{ then } \alpha^{\gamma_{i+1}} < \alpha^{\gamma_{i+2}}.$$

By suitably choosing the sequence δ_i such that $\delta_i < \delta_{i+1}$ we have

$$\delta_i^{\beta_{i+1}} < \delta_{i+1}^{\beta_{i+2}} \text{ and } \delta_i^{\gamma_{i+1}} < \delta_{i+1}^{\gamma_{i+2}},$$

however,

$$\begin{array}{rcl} h_{i+1}\alpha^{\beta_{i+1}-1}\delta_i^{\beta_{i+1}} & < & h_{i+2}\alpha^{\beta_{i+2}-1}\delta_{i+1}^{\beta_{i+2}} \\ \text{and } H_{i+1}\alpha^{\gamma_{i+1}-1}\delta_i^{\gamma_{i+1}} & < & H_{i+2}\alpha^{\gamma_{i+2}-1}\delta_{i+1}^{\gamma_{i+2}}, \end{array}$$

whence $\delta_{i+1} < \delta_{i+2}$.

In the literature, there exist several methods based on the reduction of the dimension of a multi-dimensional global optimization problem to a univariate case. The method based on filling the feasible region by a curves was studied by numerous authors, see e.g.Butz [2]. They consider the approximation of Peano type curves for this purpose. These curves were first introduced by Peano (1890) then Hilbert (1891) and have the property of passing through all points of a hyper-rectangle of $\mathbb{R}^n$ and are known as *space-filling-curves* [13]. The method based on the approximation by Peano type curves to explore the feasible region have been systematically developed and improved by Strongin, Sergeyev and their collaborators [23]. The monograph [23] treats in detail the theoretical and numerical aspects of this approach. The last works involving this approach have shown that it is competitive in practice and particularly when the feasible set is a hyper-rectangle relatively small of dimension ≤ 4 [23]. Sergeyev and al. have solved the global minimization problem when the objective function $F(x), x \in \mathbb{R}^n$, satisfies the Lipschitz condition over a hyper-rectangle of $\mathbb{R}^n$ using computable approximations of Peano functions to reduce the original problem to a univariate one satisfying the Hölder condition over an interval of $\mathbb{R}$. They also proposed some techniques for estimating locally and globally the Lipschitz and the Hölder constants when these constants are not a priori known [23].
The reduction approach that we will present in our work, is different with the approach given above. The reducing transformation method called Alienor (RT), which has been proposed and developed by Cherruault et al.[1], [18], 191](1980), is based on the approximation of an n variables of the objective function by a single variable function by using a continuous parametric curve ℓ_α, defined from a real interval $\mathbb{I}$ onto the n-dimensional compact D and satisfying a densification property described by the definition given below. Such a function ℓ_α is said to be a *space-densifying-curve* in D [16, 17]. However, the introduction of these α-dense curves has allowed analytical argumentation of certain dimension reduction methods in global optimization techniques [3], [19]].

Definition 3.3.2. Let S be a subset of the metric space E and $\alpha > 0$. We say that a continuous curve ℓ_α of E is α-dense in S, if $\ell_\alpha \subset S$ and $\mathbf{d}(x, \ell_\alpha) \leq \alpha$ for all $x \in S$, where $\mathbf{d}(x, \ell_\alpha) = \inf_{y \in \ell_\alpha} \mathbf{d}(x, y)$.

Definition 3.3.3. Let $\alpha > 0$. We say that a subset S of the metric space E is α-densifiable if there exists a continuous curve of E, rectifiable and α-dense in S.

Remark 3.3.1. Designate by $\mathcal{L}_{\alpha,\ell}$ the length of a continuous curve ℓ_α parametrically represented by the function $\ell_\alpha(t),\ t \in \mathbb{I}$, where $\mathbb{I}$ is an interval of $\mathbb{R}$. ($\mathcal{L}_{\alpha,\ell}$ is independent of the chosen parametrization). We say that the curve ℓ_α is rectifiable if $\mathcal{L}_{\alpha,\ell}$ is finite.

The curves used in the dimensionality reduction method belong to two different classes. Butz, Strongin and Sergeyev et al.[2], [23], use numerical approximations of Peano type curves whereas Cherruault and their collaborators [15], use α-dense ones. The approximations of Peano curves are obviously α-dense but the curves used by Cherruault et al. do not necessarily converge to Peano type curves. Although both classes of curves have the common property to approach as much as wanted all the points of an hyper-rectangle of $\mathbb{R}^n$, there are important differences: they do not possess the same regularity, they differ also concerning repartition uniformity (on the hyper-rectangle) and asymptotic behavior. In this work, we are interested by compact sets of type (1.3) which are α-densifiable and

note that if the functions φ_i and ψ_i are only Hölderian we can construct compacts which are not α-densifiable, in this case we can transform our problem by defining "a Lagrange penalty function" by choosing a huge hyper-rectangle $C \subset \mathbb{R}^n$ which comprises the feasible set of the problem (1.1) and so we solve the transformed problem where the objective function is Hölderian on the hyper-rectangle C (for the detail see [20]).

3.3.5 A constructive method for generating α-dense curves in the feasible region

Some interesting results concerning the existence of α-dense curves with minimal length of an α-densifiable compact in general metric space were given by Ziadi et al. [17]. In the case where the feasible set is a hyper-rectangle Ω of $\mathbb{R}^n$, some ways of constructing the α-dense curves in Ω have been given, see [10, 18]. In this subsection we shall give the relationship existing between the components ℓ_i $(i = 1, ..., n)$ of the parametrization in order to generate a new class of α-dense curves in a compact $D \subset \mathbb{R}^n$ whose boundary is defined by a non-smooth functions. Denote by $\mathbf{m}$ the Lebesgue measure, the number $\alpha > 0$ is supposed to be less than the length of the intervals $\left[\ell_i\left(t_i^{'}\right), \ell_i\left(t_i^{''}\right)\right], t_i^{'}, t_i^{''} \in \mathbb{I}$ for all $i = 2, ..., n$.
In the rest of this section, the sup-norm $\|x\| := \max\limits_{1 \leq i \leq n} \{|x_i|\}$ is used in order to simplify calculations, but other norms can also be considered. Next, keeping the same notations introduced above, we give a general main result for generating α-dense curves in D.

Theorem 3.3.2. *Let $\ell_\alpha = (\ell_1, ..., \ell_n) : \mathbb{I} \to D$ be a continuous function and $t_2, ..., t_n, \alpha$ strictly positive numbers such that:*
1) ℓ_1 is surjective.
2) For all $i = 2, ..., n$, and for all interval $\mathbb{I}_{i-1}$ with length t_i, there exist $t_i^{'}$ and $t_i^{''} \in \mathbb{I}_{i-1}$ such that:

$$\begin{cases} \ell_i(t_i^{'}) = \varphi_i(\ell_1(t_i^{'}), ..., \ell_{i-1}(t_i^{'})) \\ \ell_i(t_i^{''}) = \psi_i(\ell_1(t_i^{''}), ..., \ell_{i-1}(t_i^{''})) \end{cases}$$

3) For all $i = 2, .., n$, and for all closed interval $\mathbb{J}$, we have:

$$\mathbf{m}(\mathbb{J}) < t_i \Rightarrow \mathbf{m}(\ell_{i-1}(\mathbb{J})) < \alpha.$$

Then, the parameterized curve defined by $\ell_\alpha(t) = (\ell_1(t), ..., \ell_n(t))$ for $t \in \mathbb{I}$, is $\delta_{n-1}\lambda_n\alpha$-dense in D where δ_{n-1}, λ_n are the parameters designated in section 2.

Proof. The proof is obtained by induction in two parts.

Part 1. $n = 2$.

Consider the interval $[x_1 - \alpha, x_1 + \alpha]$ where $(x_1, x_2) \in D$. Since the function ℓ_1 is surjective, there exists a closed interval $\mathbb{I}_1 \subset \mathbb{I}$ such that:

$$\ell_1(\mathbb{I}_1) = [x_1 - \alpha, x_1 + \alpha] \cap [a, b].$$

But $\mathbf{m}(\ell_1(\mathbb{I}_1)) \geq \alpha$, so $\mathbf{m}(\mathbb{I}_1) \geq t_2$.

Thus, we have for all $t \in \mathbb{I}_1$: $|x_1 - \ell_1(t)| \leq \alpha$ moreover, there exists $t_2^{'}, t_2^{''} \in \mathbb{I}_1$ such that:

$$\begin{cases} \ell_2(t_2^{'}) = \varphi_2(\ell_1(t_2^{'})) \\ \ell_2(t_2^{''}) = \psi_2(\ell_1(t_2^{''})). \end{cases}$$

On the other hand, according to the definition of the feasible region D, we have:

$$\varphi_2\left(x_1\right) \leq x_2 \leq \psi_2\left(x_1\right).$$

As the functions φ_2, ψ_2 are Hölderian on Ω_2, respectively with the constants $h_2, \beta_2, H_2, \gamma_2$, i.e.,

$$\begin{cases} \left|\varphi_2\left(x_1\right) - \varphi_2(\ell_1(t_2^{'}))\right| \leq h_2 \left|x_1 - \ell_1(t_2^{'})\right|^{\beta_2} \leq h_2\alpha^{\beta_2} \\ \left|\psi_2\left(x_1\right) - \psi_2(\ell_1(t_2^{''}))\right| \leq H_2 \left|x_1 - \ell_1(t_2^{''})\right|^{\gamma_2} \leq H_2\alpha^{\gamma_2}, \end{cases}$$

then

$$\ell_2(t_2^{'}) - h_2\alpha^{\beta_2} \leq x_2 \leq \ell_2(t_2^{''}) + H_2\alpha^{\gamma_2}.$$

We will study the following three cases:

$$\begin{cases} (1)\ \ell_2(t_2^{'}) - h_2\alpha^{\beta_2} \leq x_2 < \ell_2(t_2^{'}) \\ (2)\ \ell_2(t_2^{'}) \leq x_2 \leq \ell_2(t_2^{''}) \\ (3)\ \ell_2(t_2^{''}) < x_2 \leq \ell_2(t_2^{''}) + H_2\alpha^{\gamma_2}. \end{cases}$$

If (1) holds, we have:

$$\left|x_2 - \ell_2(t_2^{'})\right| \leq h_2\alpha^{\beta_2},$$

and therefore

$$\left\|x - \ell_\alpha(t_2^{'})\right\| = \max\left\{\left|x_1 - \ell_1(t_2^{'})\right|, \left|x_2 - \ell_2(t_2^{'})\right|\right\} \leq \max\left\{\alpha, h_2\alpha^{\beta_2}\right\}.$$

If (2) that satisfied, there exists $t_0 \in \left[t_2^{'}, t_2^{''}\right]$ such that:

$$x_2 = \ell_2\left(t_0\right),$$

leading to

$$\|x - \ell_\alpha(t_0)\| \leq \alpha.$$

If the last case (3) holds, we have:

$$\left|x_2 - \ell_2(t_2^{''})\right| \leq H_2\alpha^{\gamma_2},$$

and therefore

$$\left\|x - \ell_\alpha(t_2^{''})\right\| \leq \max\left\{\alpha, H_2\alpha^{\gamma_2}\right\}.$$

We deduce that the parameterized curve $\ell_\alpha\left(t\right) = \left(\ell_1\left(t\right), \ell_2\left(t\right)\right)$ for $t \in \mathbb{I}$, is $\lambda_2\alpha$-dense in D with $\lambda_2 = \max\left(1, h_2\alpha^{\beta_2 - 1}, H_2\alpha^{\gamma_2 - 1}\right)$.

Part 2. For $n > 2$.

Let $x = (x_1, ..., x_n) \in D$ and consider the interval $[x_1 - \alpha, x_1 + \alpha]$, since ℓ_1 is surjective, there exists a closed interval $\mathbb{I}_1 \subset \mathbb{I}$ such that:

$$\ell_1\left(\mathbb{I}_1\right) = [x_1 - \alpha, x_1 + \alpha] \cap [a, b],$$

But $\mathbf{m}\left(\ell_1\left(\mathbb{I}_1\right)\right) \geq \alpha$, then $\mathbf{m}\left(\mathbb{I}_1\right) \geq t_2$.

It follows that $|x_1 - \ell_1(t)| \leq \alpha$ for all $t \in \mathbb{I}_1$. In addition, there exist $t_2^{'}$, $t_2^{''} \in \mathbb{I}_1$ such that:

$$\begin{cases} \ell_2(t_2^{'}) = \varphi_2\left(\ell_1(t_2^{'})\right) \\ \ell_2(t_2^{''}) = \psi_2\left(\ell_1(t_2^{''})\right). \end{cases}$$

Since $\left|\ell_2(t_2^{'}) - \ell_2(t_2^{''})\right| \geq \alpha$, then,

$$\mathbf{m}\left(\ell_2\left(\left[t_2^{'}, t_2^{''}\right]\right)\right) \geq \alpha,$$

and hence

$$\mathbf{m}\left(\left[t_2^{'}, t_2^{''}\right]\right) \geq t_3.$$

On the other hand, we have $\varphi_2(x_1) \leq x_2 \leq \psi_2(x_1)$, and as φ_2 and ψ_2 are Hölderian on Ω_2, respectively with the constants $h_2, \beta_2, H_2, \gamma_2$, then

$$\ell_2(t_2^{'}) - h_2\alpha^{\beta_2} \leq x_2 \leq \ell_2(t_2^{''}) + H_2\alpha^{\gamma_2}.$$

We still have to study the three cases:

$$\begin{cases} (1)\ \ell_2(t_2^{'}) - h_2\alpha^{\beta_2} \leq x_2 \leq \ell_2(t_2^{'}) \\ (2)\ \ell_2(t_2^{'}) \leq x_2 \leq \ell_2(t_2^{''}) \\ (3)\ \ell_2(t_2^{''}) \leq x_2 \leq \ell_2(t_2^{''}) + H_2\alpha^{\gamma_2}. \end{cases}$$

If (1) holds, we take $\mathbb{I}_2 = \left[t_2^{'}, t_2^{'} + t_3\right]$.

If (2) is valid, there exists $\mathbb{I}_2 \subset \left[t_2^{'}, t_2^{''}\right]$ such that:

$$\ell_2\left(\mathbb{I}_2\right) = [x_2 - \alpha, x_2 + \alpha] \cap \left[\ell_2(t_2^{'}), \ell_2(t_2^{''})\right].$$

If the case (3) holds, we take $\mathbb{I}_2 = \left[t_2^{''} - t_3, t_2^{''}\right]$.

It is clear that $\mathbb{I}_2 \subset \left[t_2^{'}, t_2^{''}\right]$ and $\mathbf{m}\left(\mathbb{I}_2\right) \geq t_3$ in the three cases. Indeed, for (2) we have:

$$\mathbf{m}\left(\ell_2\left(\mathbb{I}_2\right)\right) \geq \alpha,$$

and hence $\mathbf{m}\left(\mathbb{I}_2\right) \geq t_3$ from the assumption *3)* of Theorem 1.

In addition, for all $t \in \mathbb{I}_2$, we have:

If (1) holds, then

$$|x_2 - \ell_2(t)| \leq \left|x_2 - \ell_2(t_2^{'})\right| + \left|\ell_2(t_2^{'}) - \ell_2(t)\right| \leq \alpha + h_2\alpha^{\beta_2}.$$

If (2) holds, then

$$|x_2 - \ell_2(t)| \leq \alpha.$$

If the condition (3) holds, then

$$|x_2 - \ell_2(t)| \leq \left|x_2 - \ell_2(t_2^{''})\right| + \left|\ell_2(t_2^{''}) - \ell_2(t)\right| \leq \alpha + H_2\alpha^{\gamma_2}.$$

Consequently, for all $t \in \mathbb{I}_2$, we have:

$$|x_2 - \ell_2(t)| \leq \alpha \left(1 + \max\left(h_2\alpha^{\beta_2 - 1}, H_2\alpha^{\gamma_2 - 1}\right)\right) = \alpha\delta_2.$$

We have $\mathbf{m}(\mathbb{I}_2) \geq t_3$, then there exist $t_3^{'}, t_3^{''} \in \mathbb{I}_2$ such that:

$$\begin{cases} \ell(t_3^{'}) = \varphi_3(\ell_1(t_3^{'}), \ell_2(t_3^{'})) \\ \ell_3(t_3^{''}) = \psi_3(\ell_1(t_3^{''}), \ell_2(t_3^{''})), \end{cases}$$

hence

$$\left|\ell_3(t_3^{'}) - \ell_3(t_3^{''})\right| \geq \alpha,$$

consequently

$$\mathbf{m}\left(\ell_3\left(\left[t_3^{'}, t_3^{''}\right]\right)\right) \geq \alpha.$$

and therefore

$$\mathbf{m}\left(\left[t_3^{'}, t_3^{''}\right]\right) \geq t_4.$$

On the other hand, we have: $\varphi_3(x_1, x_2) \leq x_3 \leq \psi_3(x_1, x_2)$ and φ_3, ψ_3 are two Hölder functions on Ω_3,, i.e.,

$$\begin{aligned} \left|\varphi_3(x_1, x_2) - \varphi_3(\ell_1(t_3^{'}), \ell_2(t_3^{'}))\right| &\leq h_3\left[\max\left\{\left|x_1 - \ell_1(t_3^{'})\right|, \left|x_2 - \ell_2(t_3^{'})\right|\right\}\right]^{\beta_3} \\ &\leq h_3\left[\max\{\alpha, \alpha\delta_2\}\right]^{\beta_3} = h_3(\alpha\delta_2)^{\beta_3}. \end{aligned}$$

$$\begin{aligned} \left|\psi_3(x_1, x_2) - \psi_3(\ell_1(t_3^{''}), \ell_2(t_3^{''}))\right| &\leq H_3\left[\max\left\{\left|x_1 - \ell_1(t_3^{''})\right|, \left|x_2 - \ell_2(t_3^{''})\right|\right\}\right]^{\gamma_3} \\ &\leq H_3\left[\max\{\alpha, \alpha\delta_2\}\right]^{\gamma_3} = H_3(\alpha\delta_2)^{\gamma_3}. \end{aligned}$$

Whence

$$\ell_3(t_3^{'}) - h_3(\alpha\delta_2)^{\beta_3} \leq x_3 \leq \ell_3(t_3^{''}) + H_3(\alpha\delta_2)^{\gamma_3}.$$

Obviously, we will study the following cases:

$$\begin{cases} (4)\ \ell_3(t_3^{'}) - h_3(\alpha\delta_2)^{\beta_3} \leq x_3 \leq \ell_3(t_3^{'}) \\ (5)\ \ell_3(t_3^{'}) \leq x_3 \leq \ell_3(t_3^{''}) \\ (6)\ \ell_3(t_3^{''}) \leq x_3 \leq \ell_3(t_3^{''}) + H_3(\alpha\delta_2)^{\gamma_3}. \end{cases}$$

If (4) holds, we take $\mathbb{I}_3 = \left[t_3^{'}, t_3^{'} + t_4\right]$.

If (5) is satisfied, there exists $\mathbb{I}_3 \subset \left[t_3^{'}, t_3^{''}\right]$ such that:

$$\ell_3(I_3) = [x_3 - \alpha, x_3 + \alpha] \cap \left[\ell_3(t_3^{'}), \ell_3(t_3^{''})\right].$$

If the condition (6) is valid, we take $\mathbb{I}_3 = \left[t_3^{''} - t_4, t_3^{''}\right]$.

Hence, we have $\mathbb{I}_3 \subset \left[t_3^{'}, t_3^{''}\right]$ and $\mathbf{m}(\mathbb{I}_3) \geq t_4$ in all cases. In addition, for all $t \in \mathbb{I}_3$, we have:

If (4) holds:

$$|x_3 - \ell_3(t)| \leq \left|x_3 - \ell_3(t_3^{'})\right| + \left|\ell_3(t_3^{'}) - \ell_3(t)\right| \leq \alpha + h_3(\alpha\delta_2)^{\beta_3}.$$

If (5) is satisfied, then:

$$|x_3 - \ell_3(t)| \leq \alpha.$$

If the condition (6) holds:

$$|x_3 - \ell_3(t)| \leq \left|x_3 - \ell_3(t_3^{''})\right| + \left|\ell_3(t_3^{''}) - \ell_3(t)\right| \leq \alpha + H_3(\alpha\delta_2)^{\gamma_3}.$$

Then in all cases, we have for every $t \in \mathbb{I}_3$

$$|x_3 - \ell_3(t)| \leq \alpha\left\{1 + \max\left(h_3\alpha^{\beta_3 - 1}\delta_2^{\beta_3}, H_3\alpha^{\gamma_3 - 1}\delta_2^{\gamma_3}\right)\right\},$$

i.e.,

$$|x_3 - \ell_3(t)| \leq \alpha\delta_3.$$

With the same process given previously until the step $(n-1)$, we obtain by induction for all $t \in \mathbb{I}_{n-1}$,

$$|x_{n-1} - \ell_{n-1}(t)| \leq \alpha\delta_{n-1},$$

with $\mathbf{m}(\mathbb{I}_{n-1}) \geq t_n$ and $\mathbb{I}_{n-1} \subset \mathbb{I}_{n-2} \subset \cdots \subset \mathbb{I}_1 \subset \mathbb{I}$.
Therefore, there exist $t_n^{'}, t_n^{''} \in \mathbb{I}_{n-1}$ such that:

$$\begin{cases} \ell_n(t_n^{'}) = \varphi_n\left(\ell_1(t_n^{'}), ..., \ell_{n-1}(t_n^{'})\right) \\ \ell_n(t_n^{''}) = \psi_n\left(\ell_1(t_n^{''}), ..., \ell_{n-1}(t_n^{''})\right). \end{cases}$$

On the other hand, according to the definition of the feasible region, we have:

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

Moreover, the two functions φ_n, ψ_n are Hölderian on the hyper-rectangle Ω_n, and as shown in section 2, the sequence of numbers $(\delta_i)_{1\leq i\leq n}$ is increasing, we have

$$\begin{aligned} &\left|\varphi_n(x_1, ..., x_{n-1}) - \varphi_n\left(\ell_1(t_n^{'}), ..., \ell_{n-1}(t_n^{'})\right)\right| \\ &\leq h_n\left[\max\left\{\left|x_1 - \ell_1(t_n^{'})\right|, ..., \left|x_{n-1} - \ell_{n-1}(t_n^{'})\right|\right\}\right]^{\beta_n} \\ &\leq h_n\left[\max\left\{\alpha, \alpha\delta_2, ..., \alpha\delta_{n-1}\right\}\right]^{\beta_n} = h_n(\alpha\delta_{n-1})^{\beta_n}. \end{aligned}$$

$$\begin{aligned} &\left|\psi_n(x_1, ..., x_{n-1}) - \psi_n\left(\ell_1(t_n^{''}), ..., \ell_{n-1}(t_n^{''})\right)\right| \\ &\leq H_n\left[\max\left\{\alpha, \alpha\delta_2, ..., \alpha\delta_{n-1}\right\}\right]^{\gamma_n} = H_n(\alpha\delta_{n-1})^{\gamma_n}. \end{aligned}$$

Hence

$$\ell_n(t_n^{'}) - h_n(\alpha\delta_{n-1})^{\beta_n} \leq x_n \leq \ell_n(t_n^{''}) + H_n(\alpha\delta_{n-1})^{\gamma_n}.$$

Again, we have three cases to check:

$$\begin{cases} (7)\ \ell_n(t_n^{'}) - h_n(\alpha\delta_{n-1})^{\beta_n} \leq x_n \leq \ell_n(t_n^{'}) \\ (8)\ \ell_n(t_n^{'}) \leq x_n \leq \ell_n(t_n^{''}) \\ (9)\ \ell_n(t_n^{''}) \leq x_n \leq \ell_n\left(t_n^{''}\right) + H_n(\alpha\delta_{n-1})^{\gamma_n}. \end{cases}$$

If (7) holds, we have:

$$\left|x_n - \ell_n(t_n^{'})\right| \leq h_n(\alpha\delta_{n-1})^{\beta_n},$$

and

$$\begin{aligned}\left\|x-\ell_\alpha(t_n^{'})\right\| &= \max\left\{\left|x_1-\ell_1(t_n^{'})\right|,...,\left|x_{n-1}-\ell_{n-1}(t_n^{'})\right|,\left|x_n-\ell_n(t_n^{'})\right|\right\}\\ &\leq \max\left\{\alpha,\alpha\delta_2,...,\alpha\delta_{n-1},h_n\left(\alpha\delta_{n-1}\right)^{\beta_n}\right\}\\ &\leq \max\left\{\alpha\delta_{n-1},h_n\left(\alpha\delta_{n-1}\right)^{\beta_n}\right\}.\end{aligned}$$

If (8) is satisfied, there exists $t_0\in\mathbb{I}_{n-1}$ such that $x_n=\ell_n\left(t_0\right)$, consequently

$$\begin{aligned}\left\|x-\ell_\alpha\left(t_0\right)\right\| &= \max\left\{\left|x_1-\ell_1\left(t_0\right)\right|,...,\left|x_{n-1}-\ell_{n-1}\left(t_0\right)\right|,\left|x_n-\ell_n\left(t_0\right)\right|\right\}\\ &\leq \max\left\{\alpha,\alpha\delta_2,...,\alpha\delta_{n-1},\alpha\right\}\\ &\leq \alpha\delta_{n-1}.\end{aligned}$$

If the last condition (9) holds, we have:

$$\left|x_n-\ell_n(t_n^{''})\right|\leq H_n\left(\alpha\delta_{n-1}\right)^{\gamma_n}.$$

And therefore

$$\begin{aligned}\left\|x-\ell_\alpha(t_n^{''})\right\| &= \max\left\{\left|x_1-\ell_1(t_n^{''})\right|,...,\left|x_{n-1}-\ell_{n-1}(t_n^{''})\right|,\left|x_n-\ell_n(t_n^{''})\right|\right\}\\ &\leq \max\left\{\alpha,\alpha\delta_2,...,\alpha\delta_{n-1},H_n\left(\alpha\delta_{n-1}\right)^{\gamma_n}\right\}\\ &\leq \max\left\{\alpha\delta_{n-1},H_n\left(\alpha\delta_{n-1}\right)^{\gamma_n}\right\}.\end{aligned}$$

We deduce that the parameterized curve $\ell_\alpha\left(t\right)=\left(\ell_1\left(t\right),...,\ell_n\left(t\right)\right)$ for $t\in\mathbb{I}$, is $\delta_{n-1}\lambda_n\alpha$-dense in D. □

Corollary 1. *Let $\ell_\alpha=(\ell_1,...,\ell_n):\mathbb{I}\to D$, be a continuous function, whose components ℓ_i ,$(i=1,...,n)$ are Hölder functions with constants $c_i>0$ and exponents $\mu_i<1,(i=1,...,n)$ (resp.). Let $t_2,t_3,...,t_n,\alpha$ be strictly positive numbers such that:*

1) ℓ_1 is surjective,

2) For all $i=2,...,n$ and for all interval $\mathbb{I}_{i-1}$ with length t_i, there exist $t_i^{'}$ and $t_i^{''}\in\mathbb{I}_{i-1}$ such that:

$$\begin{cases}\ell_i(t_i^{'})=\varphi_i\left(\ell_1(t_i^{'}),...,\ell_{i-1}(t_i^{'})\right)\\ \ell_i(t_i^{''})=\psi_i\left(\ell_1(t_i^{''}),...,\ell_{i-1}(t_i^{''})\right)\end{cases}$$

3) For all $i=2,...,n$

$$c_{i-1}<\frac{\alpha}{t_i^{\mu_{i-1}}}.$$

Then, the parameterized curve defined by $\ell_\alpha\left(t\right)=\left(\ell_1\left(t\right),...,\ell_n\left(t\right)\right)$ for $t\in\mathbb{I}$, is $\delta_{n-1}\lambda_n\alpha$-dense in D.

Proof. It is a consequence of Theorem 1. The conditions *1)* and *2)* of the Theorem 1 are satisfied. It remains to check the condition *3)*. Indeed, let $i=2,...,n$ and $\mathbb{I}_{i-1}$ an interval of $\mathbb{I}$ satisfying $\mathbf{m}\left(\mathbb{I}_{i-1}\right)<t_i$. For $t^{'},t^{''}\in\mathbb{I}_{i-1}$, we have:

$$\left|\ell_{i-1}(t^{'})-\ell_{i-1}(t^{''})\right|\leq c_{i-1}\left|t^{'}-t^{''}\right|^{\mu_{i-1}}\leq c_{i-1}\mathbf{m}\left(I\right)^{\mu_{i-1}};$$

involving

$$\mathbf{m}\left(\ell_{i-1}\left(\mathbb{I}_{i-1}\right)\right)\leq c_{i-1}\mathbf{m}\left(\mathbb{I}_{i-1}\right)^{\mu_{i-1}}<c_{i-1}t_i^{\mu_{i-1}}<\alpha.$$

3.3.6 Constructing a concrete α-dense curves

In this subsection, we will use the following notations:

$$\begin{aligned}\Phi_1(t) &= a \text{ and } \Phi_i(t) = \varphi_i\left(\ell_1(t), ..., \ell_{i-1}(t)\right) \text{ for } i \geq 2\\ \Psi_1(t) &= b \text{ and } \Psi_i(t) = \psi_i\left(\ell_1(t), ..., \ell_{i-1}(t)\right) \text{ for } i \geq 2,\end{aligned}$$

where φ_i and ψ_i for $i \geq 2$, are the functions defined in the feasible region (1.3).

Theorem 2. *Consider the compact D given in the problem (1.1) and let the parameterized curve $\ell_\alpha = (\ell_1, ..., \ell_n) : \mathbb{I} \to D$ be defined by*

$$\ell_i(t) = \Phi_i(t)\cos^2\left(\frac{\alpha_i t}{2}\right) + \Psi_i(t)\sin^2\left(\frac{\alpha_i t}{2}\right) \text{ for } i = 1, ..., n,$$

where $\alpha_1, ..., \alpha_n, \alpha$ are strictly positive numbers satisfying the relationship

$$(\alpha_i/\pi)^{\beta_1 ... \beta_{i-1}} \geq (1/2\alpha)\left(\left(U_{i-1} - L_{i-1}\right)\alpha_{i-1} + 2\left(h_{i-1}z_{i-2}^{\beta_{i-1}} + H_{i-1}z_{i-2}^{\gamma_{i-1}}\right)\right),$$

with

$$z_i = \max_{1 \leq k \leq i}\left(\frac{1}{2}\left(U_k - L_k\right)\alpha_k + \left(h_k z_{k-1}^{\beta_k} + H_k z_{k-1}^{\gamma_k}\right)\right), \quad z_0 = 0,$$

and L_i, U_i are respectively the lower and upper bounds of the functions φ_i and ψ_i on the hyper-rectangles Ω_i(see section 2). Then, the parameterized curve $\ell_\alpha(t) = (\ell_1(t), ..., \ell_n(t))$ for $t \in [0, \pi/\alpha_1]$ is $\delta_{n-1}\lambda_n\alpha$-dense in D.

Proof. This result is a consequence of Theorem 1. Indeed, the proof is obtained in a recurrent way:

(i) For $i = 2$.

1) ℓ_1 is surjective by definition.

2) If we take $t_2 = \dfrac{\pi}{\alpha_2}$, then for all interval $\mathbb{I}_1$ with length t_2, there exist $t', t'' \in \mathbb{I}_1$ such that:

$$\begin{cases} \ell_2(t') = \varphi_2(\ell_1(t')) \\ \ell_2(t'') = \psi_2(\ell_1(t'')) \end{cases} \Rightarrow \begin{cases} \cos^2(\alpha_2 t'/2) = 1 \\ \sin^2(\alpha_2 t''/2) = 1 \end{cases}$$
$$\Rightarrow \begin{cases} t' = 2\pi k/\alpha_2 \\ t'' = (2k+1)\pi/\alpha_2 \end{cases}, \ k \in \mathbb{N}.$$

3) For any interval $\mathbb{I}_1$ satisfying $\mathbf{m}(\mathbb{I}_1) < \dfrac{\pi}{\alpha_2}$, we have $\mathbf{m}(\ell_1(\mathbb{I}_1)) < \alpha$.

$\forall t', t'' \in \mathbb{I}_1$:

$$\begin{aligned}\left|\ell_1(t') - \ell_1(t'')\right| &= \left|a(\cos^2(\alpha_1 t'/2) - \cos^2(\alpha_1 t''/2)) + b(\sin^2\left(\alpha_1 t'/2\right) - \sin^2(\alpha_1 t''/2))\right| \\ &\leq (b-a)\left|\sin\frac{\alpha_1 t' - \alpha_1 t''}{2}\right| \leq \tfrac{1}{2}(b-a)\alpha_1\left|t' - t''\right| \leq \tfrac{1}{2}(b-a)\alpha_1\frac{\pi}{\alpha_2} \leq \alpha.\end{aligned}$$

By Theorem 1, the parametrized curve $\ell(t) = (\ell_1(t), \ell_2(t))$ defined on $[0, \pi/\alpha_1]$ is $\lambda_2\alpha$-dense in D with $\frac{\alpha_2}{\pi} \geq \frac{(b-a)\alpha_1}{2\alpha}$.

(ii) We suppose that the theorem is true for i.

2) For all $i = 2, ..., n$ and for all interval $\mathbb{I}_{i-1}$ with length $t_i = \dfrac{\pi}{\alpha_i}$, there exist $t_i^{'} = \frac{2\pi k}{\alpha_i}$ and $t_i^{''} = \frac{\pi}{\alpha_i} + \frac{2\pi k}{\alpha_i} \in \mathbb{I}_{i-1}$; $k \in \mathbb{N}$, such that:

$$\begin{cases} \ell_i(t^{'}) = \varphi_i(\ell_1(t^{'}), ..., \ell_{i-1}(t^{'})) \\ \ell_i(t^{''}) = \psi_i(\ell_1(t^{''}), ..., \ell_{i-1}(t^{''})). \end{cases}$$

3) We suppose that the inequality holds for $i = 2, ..., n$, i.e., for any closed interval $\mathbb{J}$ of length less than $\frac{\pi}{\alpha_i}$; we have:

$$\mathbf{m}(\mathbb{J}) < \frac{\pi}{\alpha_i} \Rightarrow \mathbf{m}(\ell_{i-1}(\mathbb{J})) < \alpha.$$

Consequently

$\forall t^{'}, t^{''} \in \mathbb{J} : \left|\ell_{i-1}(t^{'}) - \ell_{i-1}(t^{''})\right| \leq$

$\left[\frac{1}{2}(U_{i-1} - L_{i-1})\alpha_{i-1} + \left(h_{i-1} z_{i-1}^{\beta_{i-1}} + H_{i-1} z_{i-1}^{\gamma_{i-1}}\right)\right] \left|t^{'} - t^{''}\right|^{\beta_1 ... \beta_{i-1}} \leq \alpha.$

Consider the closed interval $\mathbb{J}$, such that: $\mathbf{m}(\mathbb{J}) < \dfrac{\pi}{\alpha_{i+1}}$, show that for $t^{'}, t^{''} \in \mathbb{J}$, we have:

$$\left|\ell_i(t^{'}) - \ell_i(t^{''})\right| = \left| \begin{array}{l} \Phi_i(t^{'})\cos^2(\alpha_i t^{'}/2) + \Psi_i(t^{'})\sin^2(\alpha_i t^{'}/2) - \\ \Phi_i(t^{''})\cos^2(\alpha_i t^{''}/2) - \Psi_i(t^{''})\sin^2(\alpha_i t^{''}/2) \end{array} \right.$$

$$\left. + (\Phi_i(t^{'}) - \Psi_i(t^{'}))\cos^2(\alpha_i t^{''}/2) - (\Phi_i(t^{'}) - \Psi_i(t^{'}))\cos^2(\alpha_i t^{''}/2) \right|$$

$$= \Big|(\Phi_i(t^{'}) - \Psi_i(t^{'}))(\cos^2(\alpha_i t^{'}/2) - \cos^2(\alpha_i t^{''}/2)) + (\Phi_i(t^{'}) - \Phi_i(t^{''}))(\cos^2\left(\alpha_i t^{''}/2\right)$$

$$(\Psi_i(t^{'}) - \Psi_i(t^{''})\sin^2(\alpha_i t^{''}/2)\Big|$$

$$\leq \left|\Phi_i(t^{'}) - \Psi_i(t^{'})\right| \left|\cos^2(\alpha_i t^{'}/2) - \cos^2(\alpha_i t^{''}/2)\right| + \left|\Phi_i(t^{'}) - \Phi_i(t^{''})\right| + \left|\Psi_i(t^{'}) - \Psi_i(t^{''})\right|.$$

Let us take into account that the constants L_i, U_i are respectively the lower and upper bounds of the functions φ_i, ψ_i on the hyper-rectangles Ω_i(as shown in section 2). If we suppose that $\gamma_i \geq \beta_i$ (or $\gamma_i < \beta_i$) for each i, then we have

$$\left|\ell_i(t^{'}) - \ell_i(t^{''})\right| \leq \tfrac{1}{2}(U_i - L_i)\alpha_i \left|t^{'} - t^{''}\right| + h_i \left(\max\left\{\left|\ell_1(t^{'}) - \ell_1(t^{''})\right|, ..., \left|\ell_{i-1}(t^{'}) - \ell_{i-1}(t^{''})\right|\right\}\right)^{\beta_i} +$$

$$+ H_i \left(\max\left\{\left|\ell_1(t^{'}) - \ell_1(t^{''})\right|, ..., \left|\ell_{i-1}(t^{'}) - \ell_{i-1}(t^{''})\right|\right\}\right)^{\gamma_i}$$

$$\leq \tfrac{1}{2}(U_i - L_i)\alpha_i \left|t^{'} - t^{''}\right| +$$

$$h_i \left(\max\left\{\tfrac{1}{2}(b-a)\alpha_1, ..., \tfrac{1}{2}(U_{i-1} - L_{i-1})\alpha_{i-1} + \left(h_{i-1} z_{i-2}^{\beta_{i-1}} + H_{i-1} z_{i-2}^{\gamma_{i-1}}\right)\right\} \left|t^{'} - t^{''}\right|^{\beta_1 ... \beta_{i-1}}\right)^{\beta_i}$$

$$+ H_i \left(\max\left\{\tfrac{1}{2}(b-a)\alpha_1, ..., \tfrac{1}{2}(U_{i-1} - L_{i-1})\alpha_{i-1} + \left(h_{i-1} z_{i-2}^{\beta_{i-1}} + H_{i-1} z_{i-2}^{\gamma_{i-1}}\right)\right\} \left|t^{'} - t^{''}\right|^{\beta_1 ... \beta_{i-1}}\right)^{\gamma_i}$$

$$\leq \left[\tfrac{1}{2}(U_i - L_i)\alpha_i + (h_i z_{i-1}^{\beta_i} + H_i z_{i-1}^{\gamma_i})\right] \left|t^{'} - t^{''}\right|^{\beta_1 ... \beta_i}$$

$$< \left[\tfrac{1}{2}(U_i - L_i)\alpha_i + (h_i z_{i-1}^{\beta_i} + H_i z_{i-1}^{\gamma_i})\right] \left(\frac{\pi}{\alpha_{i+1}}\right)^{\beta_1 ... \beta_i} < \alpha.$$

Hence the inequality holds for $i + 1$.

Remark 2. *All the functions $\ell_i(t)$ defined in Theorem 2 are Hölderian with constants $c_i > 0$ and exponents $\mu_i < 1$, for each $i = 1, ..., n$, given by*

$$\begin{cases} c_i = \frac{1}{2}(U_i - L_i)\alpha_i + (h_i z_{i-1}^{\beta_i} + H_i z_{i-1}^{\gamma_i}) \\ \mu_i = \beta_1 ... \beta_i. \end{cases}$$

It is well known that if a function with several variables $F(x)$ is Hölderian of the constant H and the exponent β' on the compact Ω, then the function of a single variable $f(t) = F(\ell_\alpha(t))$ is Hölderian with the constant $h = H\left(\max_{1\le i\le n} c_i\right)^{\beta'}$ and exponent $\beta = \beta'\left(\min_{1\le i\le n} \mu_i\right)$ on the interval $[0, \pi/\alpha_1]$.

Examples of a compact sets which are α-densifiable

In this subsection and according to Theorem 2, the illustrations figures 2-6 given below, represent a densification of some non-convex sets and hyper-rectangles (for $n = 2$ and $n = 3$), by the supports of the different α-dense curves $\ell_\alpha(t)$ on $[0, \pi]$ (with $\alpha_1 = 1$). Note that the parametrized curves obtained in the examples D_1, D_3 and D_4 are regular of class $\mathcal{C}^\infty$ whereas the curve in D_2, is only continuous.

a) Examples with non-convex sets

Let $D_1 = \left\{(x_1, x_2) \in \mathbb{R}^2 / \ -1 \le x_1 \le 1, -\frac{1}{4}x_1^2 + \frac{1}{2}x_1 - 1 \le x_2 \le \frac{1}{4}x_1^2 + \frac{1}{2}x_1 + 1\right\}$, here, we have $\beta_2 = \gamma_2 = 1$.

$D_2 = \left\{(x_1, x_2) \in \mathbb{R}^2 / \quad x_1^2 + x_2^2 \le 1, \ \ |x_1| - (x_2 + 1)^2 \le 0\right\}$,
$\varphi_2(x_1) = \sqrt{|x_1|} - 1$ and $\psi_2(x_1) = \sqrt{1 - x_1^2}$ and $\beta_2 = \gamma_2 = 1/2$.

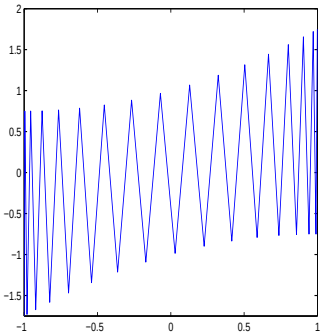

Figure 3.2: The α-dense curve ℓ_α in D_1 with $\alpha = 0.2$.

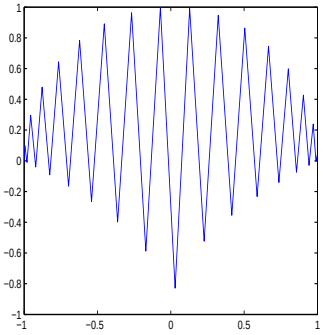

Figure 3.3: The α-dense curve ℓ_α in D_2 with $\alpha = 0.2$.

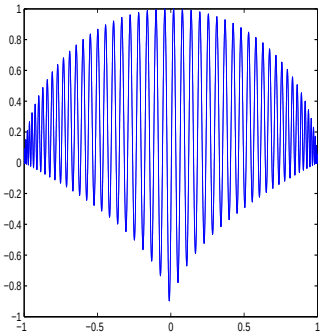

Figure 3.4: The α-dense curve ℓ_α in D_2 with $\alpha = 0.01$.

Remark 3. In particular, if the functions φ_i and ψ_i $(i \geq 2)$ in the definition of the feasible set D, are all constants and if $D = \Omega$, i.e.,

$$a = x_1^l \text{ and } \varphi_i(x_1, ..., x_{i-1}) = x_i^l,\ 2 \leq i \leq n,$$
$$b = x_1^u \text{ and } \psi_i(x_1, ..., x_{i-1}) = x_i^u,\ 2 \leq i \leq n.$$

Then the feasible set D becomes an hyper-rectangle of $\mathbb{R}^n$. According to Theorem 2, the hyper-rectangle D is α-densified by the parametrized curve $\ell_\alpha = (\ell_1, ..., \ell_n)$ defined on

$[0, \pi/\alpha_1]$, by the following expression:

$$\ell_i(t) = \left(\frac{x_i^l - x_i^u}{2}\right)\cos^2\left(\frac{\alpha_i t}{2}\right) + x_i^u,\ 1 \leq i \leq n,$$

where the constants $\alpha_1, ..., \alpha_n, \alpha$ satisfying the relationship:

$$\left(\frac{\alpha_i}{\pi}\right) \geq \left(\frac{\alpha_{i-1}}{2\alpha}\right)(x_{i-1}^u - x_{i-1}^l),\ 2 \leq i \leq n.$$

But in the implementation of the reducing transformation, one can choose these last constants as:

$$\alpha_1 = 1, \alpha_i = \left(\frac{\pi}{\alpha}\right)^{i-1}\prod_{k=1}^{i-1}(x_k^u - x_k^l),\ 2 \leq i \leq n.$$

b) Examples with the hyper-rectangles

Let $D_3 = \{(x_1, x_2) \in \mathbb{R}^2 / \ -1 \leq x_1 \leq 1, -1 \leq x_2 \leq 1\}$, here, we have $\varphi_2(x_1) = -1$ and $\psi_2(x_1) = 1$.

$D_4 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 / \ -1 \leq x_i \leq 1,\ i = 1, 2, 3\}$, here, we have $\varphi_2(x_1) = -1$,$\psi_2(x_1) = 1$, $\varphi_3(x_1, x_2) = -1, \psi_3(x_1, x_2) = 1$.

Remark 4. At the end of this section, it should be mentioned that, in the reducing transformation method (RT), the number of evaluation points necessary for reaching the global minimum, with the accuracy ε, of the univariate objective function $f(t) = F(\ell_\alpha(t))$, depends on the length $\mathcal{L}_{\alpha,\ell}$ of the α-dense curve ℓ_α in the feasible set, which allows to convert the constrained multi-dimensional problem (1.1) to the one-dimensional unconstrained one. As for all $\alpha > 0$, the set of α-dense curves in a compact D of $\mathbb{R}^n$, contains a curve with minimal length, see Ziadi et al.[17], it is therefore, very interest to find other curves with small length in order to improve the reducing transformation method.

3.3.7 One-dimensional unconstrained problem

The univariate global minimization problem of Hölder functions over a closed and bounded interval of $\mathbb{R}$ appear frequently in several life applications, for instance, the simple plant location problem under a uniform delivered price policy, see [4]. The authors gave a transformation that reduces the problem of the infinite horizon into an equivalent one-dimensional optimization problem of Hölder function over an interval of $\mathbb{R}$.

Let us consider $\ell_\alpha(t) = (\ell_1(t), ..., \ell_n(t))$ the $\delta_{n-1}\lambda_n\alpha$-dense curve in D, for $t \in [0, \pi/\alpha_1]$ which is obtained from Theorem 2. Then by using the reducing transformation $x_i = \ell_i(t)$, for $i = 1, ..., n$, the initial problem (1.1) is approached by the one-dimensional unconstrained one,

$$\min_{t\in[0,\pi/\alpha_1]} f(t), \tag{4.1}$$

where $f(t) = F(\ell_\alpha(t))$ is a Hölder non-differentiable function on the interval $[0, \pi/\alpha_1]$ with constants h and the exponent $\beta < 1$. In the rest of this section, we focus on the

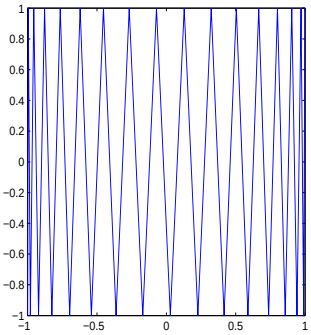

Figure 3.5: The α-dense curve ℓ_α in the square D_3 with $\alpha = 0.2$.

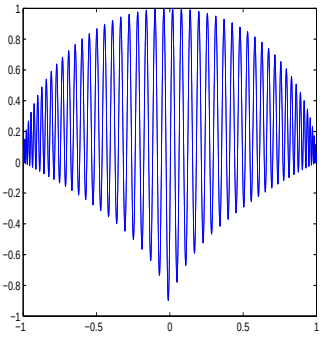

Figure 3.6: The α-dense curve ℓ_α in the cube D_4 with $\alpha = 0.3$.

one-dimensional global minimization problem (4.1) .

Constructing piecewise affine under-estimator functions.

Since the objective function $f(t)$ of problem (4.1) is Hölderian on $[0, \pi/\alpha_1]$, then we have

$$\left|f(t) - f(t^{'})\right| \leq h \left|t - t^{'}\right|^{\beta}, \ \forall t, t^{'} \in [0, \pi/\alpha_1] \,. \tag{4.2}$$

The one-dimensional deterministic method that we will use for solving this problem, is that of Piyavskii [11], which is the best known and more discussed in the literature among the covering global optimization methods. It applies to Lipschitz functions on a closed and bounded interval of $\mathbb{R}$ whose Lipschitz constant is known. Recall that an under-estimator of a function f: $[a, b] \to \mathbb{R}$, is a function U: $[a, b] \to \mathbb{R}$, such that for all $x \in [a, b]$, $U(x) \leq f(x)$. In the case of Hölder functions, it follows from the Hölder condition (4.2) that, for all $t, t^{'} \in [0, \pi/\alpha_1]$ we have

$$f(t^{'}) - h \left|t - t^{'}\right|^{\beta} \leq f(t).$$

If a point $t^{'} \in [0, \pi/\alpha_1]$ is fixed, then the function

$$U(t) = f(t^{'}) - h \left|t - t^{'}\right|^{\beta},$$

is an under-estimator function for f on $[0, \pi/\alpha_1]$, i.e.,

$$U(t) \leq f(t), \text{for all } t \in [0, \pi/\alpha_1].$$

A sequence of points, $t_1, t_2, ..., t_k$, converging to the global minimizer of f on $[0, \pi/\alpha_1]$, is generated by constructing a sequence of an under-estimator functions of f over $[0, \pi/\alpha_1]$. In the starting step, a sampling point t_1 is chosen at the mid-point of the interval $[0, \pi/\alpha_1]$ as $t_1 = \pi/2\alpha_1$, and at t_1 we have the under-estimator $U_1(t) = f(t_1) - h|t - t_1|^\beta$, which represented by two parabolic curves as shown in Fig.7. Let's set $t_2 = \arg\min U_1([0, \pi/\alpha_1])$, we have obtain a new under-estimator of f, which is a piecewise parabolic function and is closer to f, as

$$U_2(t) = \max_{i=1,2} \left\{ f(t_i) - h|t - t_i|^\beta \right\}.$$

At the iteration k, the function given by

$$U_k(t) = \max_{i=1,...,k} \left\{ f(t_i) - h|t - t_i|^\beta \right\},$$

which satisfies

$$U_k(t) \leq U_{k+1}(t) \leq f(t), t \in [0, \pi/\alpha_1],$$

is also an under-estimator of f on $[0, \pi/\alpha_1]$. In fact, in this algorithm, we construct an increasing sequence of a piecewise parabolic functions which are the under-estimators of the objective function f. The restriction of the under-estimator U_k on each sub-interval $[t_{i-1}, t_i]$, $i = 2, ..., k$ of $[0, \pi/\alpha_1]$, can be given as

$$U_i(t) = \max_i \{q_i(t), r_i(t)\},$$

where

$$q_i(t) = f(t_{i-1}) - h(t - t_{i-1})^\beta \quad \text{and } r_i(t) = f(t_i) - h(t_i - t)^\beta.$$

The stopping criterion for this algorithm is when one has $f(t_{opt}) - U(t_{opt}) \leq \varepsilon$, where t_{opt} is the optimal global minimizer of a piecewise under-estimator function U of the objective function f on $[0, \pi/\alpha_1]$ and ε is the required accuracy. On each sub-interval $[t_{i-1}, t_i]$, $i = 2, ..., k$ of $[0, \pi/\alpha_1]$, the function $U_i(t)$ is strictly convex and non-differentiable and its global minimum can be calculated by determining the intersection point of the two parabolic curves q_i and r_i (see Fig.7. below). Then at each iteration we must solve a non-linear algebraic equation:

$$q_i(t) - r_i(t) = 0. \tag{4.3}$$

Usually, it is not easy to find the explicit solution of this last equation, particularly in the case where the parameter β is sufficiently small or where it is not as $\beta = \frac{1}{m}, m \in \mathbb{N}^*, m \neq 1$. We can even say that the equation (4.3) presents the same difficulty as a local optimization problem. The technique used by Gourdin et al.[4] is based on resolution at each step of the algorithm, the non-linear equation (4.3) in the case when $\beta = \frac{1}{m}$ and m is the integers $2, 3, 4$ and h is given, because it is possible to obtain the analytical expression of the intersection point of $q_i(t)$ and $r_i(t)$. Lera and Sergeyev [11] have replaced this intersection point by an intersection point of two linked segments of the parabolas $q_i(t)$ and $r_i(t)$. They even consider this technique when the constant h is not given a priori. The

previous techniques become very difficult to realise in the case of multivariate Hölderian functions, since the under-estimators of the objective function are also multivariate and the determination of their local minima at each iteration is the intersection of several parabolic hyper-surfaces on a hyper-rectangles of $\mathbb{R}^n$. In this case, the equation (4.3) is still non-linear and with several variables. To overcome this situation, a deterministic approach is proposed. It is composed of two parts, the first one, is to approximate both the multivariate Hölderian function F of the initial problem (1.1) to a univariate one and the feasible domain D, by using a new computable parameterized α-dense curves $\ell_\alpha(t)$ in D as previously described. The second one, is to apply the modification of the Piyavskii's algorithm to the one-dimensional unconstrained problem (4.1). The technique used for the Hölder univariate minimization does not need to solve at iteration the equation (4.3). It consists on constructing a sequence of piecewise affine under-estimators of f on $[0, \pi/\alpha_1]$ similar to the previous procedure, which is essentially based on the following result.

Theorem 3 *A real function F defined on a compact set D of $\mathbb{R}^n$ is continuous if and only if for any $\varepsilon > 0$ there exists a constant $L > 0$ such that for all $x, y \in D$,*

$$|F(x) - F(y)| \leq L \|x - y\| + \varepsilon. \tag{4.4}$$

A version of the proof of this result is found in [15], but in this last one, the author assumes that the set D is a convex of $\mathbb{R}^n$, whereas this condition is not necessary if D is a compact. We have gave a simpler proof of this result in [12]. It is based on the fact that if D is a compact of $\mathbb{R}^n$, then the set of Lipschitz functions on D defined by

$$\mathcal{L}(F, D) = \{F : D \to \mathbb{R}, \exists L > 0, |F(x) - F(y)| \leq L\|x - y\|, \forall x, y \in D\},$$

is dense in the space $\mathcal{C}(F, D)$ of continuous functions on D, on which the sup-norm $\|F - G\| = \sup_{x \in D} |F(x) - G(x)|$, is defined. The above result has a very important theoretical advantage, but in the numerical implementation of the majority of covering global optimization methods, we need to know the value of the constant L in (4.4), unfortunately, it is usually difficult to calculate it. The algorithm of Piyavskii used in this paper is applied to a class of Hölderian functions. Since all the functions of this class defined on a compact D of $\mathbb{R}^n$, with the knowledge of the constants h and β, are continuous, then the condition (4.4) of the previous theorem is satisfied. However, in most cases, it is not easy to evaluate explicitly these constants. We can then, proceed to the numerical evaluation. Since theorem 3 only ensures the existence of the constant L, without giving us how to evaluate it. Due to this reason, several procedures can be envisaged, with the aim of applying the previous algorithm and improving the convergence. From the inequality (4.4), one can directly estimate the constant L, instead of estimating h and β. The estimates of the constant L have an influence on the convergence to the global minimum. Among the procedures allowing for estimating L, there is one, such that the constant L is locally estimated in different sub-intervals of the search domain, the other procedure, is to estimate globally the constant L, i.e., valid for the entire search domain, until the optimal solution is obtained. More details are discussed in [12, 14]. Another idea that can be applied for some covering global optimization methods which use the constant L, is to construct an appropriate increasing sequence of a real positive constants $(L_k), k \in \mathbb{N}$ which satisfies the following conditions: $L_k \to \infty$ for $k \to \infty$ and for a given accuracy

$\varepsilon > 0$, there exists $j \in \mathbb{N}$, such that $L \leq L_j$. Consequently, the inequality (4.4) is holds with the constant L_j for which the convergence to the global minimum is guaranteed [19]. As mentioned earlier, after using the α-dense curve $\ell_\alpha(t)$ in the search region D, the univariate function $f(t) = F(\ell_\alpha(t))$ is Hölderian on the interval $[0, \pi/\alpha_1]$, however, f is certainly continuous and hence condition (4.4) is satisfied. In this case, there is no need to estimate the constant L, because here, the value of the constant L is given explicitly with respect to h, β and the accuracy ε. Let us give the following result and for more details of the proof, one can see [12].

Theorem 4 *Let $f(t) = F(\ell_\alpha(t))$ be a real univariate Hölder function with constant h and exponent $\beta < 1$ defined on the interval $[0, \pi/\alpha_1]$, then for all $\varepsilon > 0$, we have:*

$$\left| f(t) - f(t') \right| \leq L_\varepsilon \left| t - t' \right| + \varepsilon, \ \forall t, t' \in [0, \pi/\alpha_1] \tag{4.5}$$

with

$$L_\varepsilon = \begin{cases} \frac{\beta h^{1/\beta}}{\varepsilon^{(1/\beta)-1}} & \text{for} \quad \beta \in \left\{ \frac{1}{m}, m \in \mathbb{N}^*, m \neq 1 \right\} \\ \\ \frac{\beta h^{(1/\beta)+1}}{\varepsilon^{1/\beta}} & \text{for} \quad \beta \in]0,1[\setminus \left\{ \frac{1}{m}, m \in \mathbb{N}^* \right\}. \end{cases}$$

For a given $\varepsilon > 0$ and for a known constants h and $\beta < 1$, the set of all such constant L_ε is closed in $\mathbb{R}$. Then the smallest value theoretically always exists. In [12] we have suggested another technique for calculating a different constant than L_ε that satisfies (4.5). Now, from the inequality (4.5), we construct an increasing sequence of piecewise affine under-estimators of the objective function $f(t) = F(\ell_\alpha(t))$ on the interval $[0, \pi/\alpha_1]$, instead parabolic:

$$L_k(t) = \max_{i=1,\ldots,k} \left\{ f(t_i) - L_\varepsilon \left| t - t_i \right| - \varepsilon \right\},$$

and on each sub-interval $[t_{i-1}, t_i]$, $i = 2, ..., k$ of $[0, \pi/\alpha_1]$, the restriction of L_k can be given as

$$L_i(t) = \max_i \left\{ v_i(t), w_i(t) \right\},$$

such that v_i and w_i are two linear functions which are drawn with slope L_ε, as shown in Fig.7., where their expressions are given by

$$\begin{aligned} v_i(t) &= f(t_{i-1}) - L_\varepsilon (t - t_{i-1}) - \varepsilon, \quad t \in \left[t_{i-1}, \widetilde{t_i} \right] \\ w_i(t) &= f(t_i) - L_\varepsilon (t_i - t) - \varepsilon, \ t \in \left[\widetilde{t_i}, t_i \right]. \end{aligned}$$

Note that the global minimum of the functions L_i on the interval $[t_{i-1}, t_i]$, is calculated at $\widetilde{t_i}$ the intersection point of two lines v_i and w_i (see Fig. 7. above) which is easily found by solving over the interval $[t_{i-1}, t_i]$, a simple linear equation with the single variable t even β is not an integer or is large:

$$v_i(t) - w_i(t) = 0.$$

The explicit solution $\widetilde{t_i}$ of the last equation, and the value $L_i\left(\widetilde{t_i}\right)$ are given by

$$\begin{cases} \widetilde{t_i} = \frac{1}{2}(\frac{1}{L_\varepsilon}(y_{i-1} - y_i) + t_{i-1} + t_i), \ i = 2, ..., k \\ L_i\left(\widetilde{t_i}\right) = \frac{1}{2}(y_{i-1} + y_i - L_\varepsilon (t_i - t_{i-1}) - 2\varepsilon), \end{cases}$$

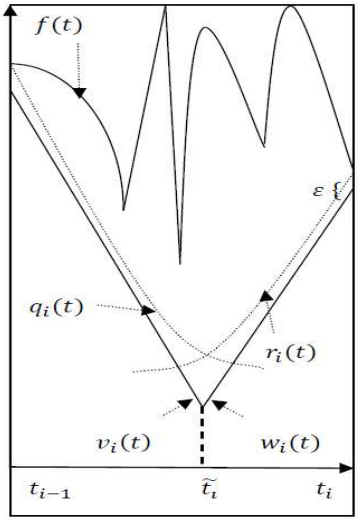

Figure 3.7: The piecewise affine under-estimator of $f(t) = F(\ell_\alpha(t))$ on $[t_{i-1}, t_i]$.

where $y_j = F(\ell_\alpha(t_j))$, $j = 1, ..., k$. Then we have

$$L_i\left(\widetilde{t}_i\right) = \min_{t\in[t_{i-1},t_i]} L_i(t) \leq f(t), \forall t \in [t_{i-1}, t_i].$$

The modified algorithm of Piyavskii MPA proceeds then to evaluate the global minimum of the affine under-estimator of $f(t) = F(\ell_\alpha(t))$ on the search domain $[0, \pi/\alpha_1]$ through the construction of an increasing sequence of piecewise affine under-estimators. This algorithm can be extended to the case when the constant L_ε is a priori unknown. Let's now show the combined method of reducing transformation and the modified Piyavskii's method, which is called RT-MPA, as follows:

`Algorithm` RT-MPA The reducing transformation combined with the modified Piyavskii's method.

Input :

· Define $\ell_\alpha(t)$, $t \in [0, \pi/\alpha_1]$, a parametrized curve α-dense in a nonconvex D with $\alpha_1 = 1$.

· Choose $\varepsilon > 0$, the given small search accuracy.

· For a given $\varepsilon > 0$, and for the known constants h and β of the one-dimensional function f, compute the constant L_ε.

· For each subinterval $[t_{i-1}, t_i]$, $i = 2, ..., k$, compute the point $\widetilde{t}_i$ (The trial points) and the value $L_i\left(\widetilde{t}_i\right)$ (the lower bound of the function $f(t)$ on the curve $\ell_\alpha(t)$).

Output :

· t_{opt} , f_{opt}, the optimal global minimizer and global minimum of $f(t)$ on $[0, \pi/\alpha_1]$.

· x_{opt}, F_{opt}, the optimal global minimizer and global minimum of $F(x)$ over D.

`Step 1. Initialisation`

$k \leftarrow 1, t_1 \leftarrow \pi/2\alpha_1, t_{opt} \leftarrow t_1$

$f_{opt} \leftarrow f(t_{opt}), L_{opt} \leftarrow f_{opt} - \pi L_\varepsilon/2\alpha_1 - \frac{\varepsilon}{2}, L_1(t) \leftarrow f(t_1) - L_\varepsilon |t - t_1| - \frac{\varepsilon}{2}$

`Step 2.`

`While` $f_{opt} - L_{opt} > \varepsilon$ `do`

$t_{k+1} \leftarrow \arg\min_{t \in [0, \pi/\alpha_1]} L_k(t)$

`if` $f(t_{k+1}) < f_{opt}$ `then` $f_{opt} \leftarrow f(t_{k+1}); t_{opt} \leftarrow t_{k+1}$

`end if`

$L_{k+1}(t) \leftarrow \min_{1 \le i \le k+1} \left\{ f(t_i) - L_\varepsilon |t - t_i| - \frac{\varepsilon}{2} \right\}$

$L_{opt} \leftarrow \min_{t \in [0, \pi/\alpha_1]} L_{k+1}(t)$

$k \leftarrow k + 1$

`End While`

Relationship between α density and ε accuracy

Theorem 5. *Let ℓ_α a curve α-dense in D and ε an accuracy with which the global minimum of the problem* (1.1) *on the compact D is to be searched, then we have*

$$\begin{cases} \varepsilon(\alpha) = 2H(\delta_{n-1}\lambda_n\alpha)^\beta \\ s.t. \quad \underset{\alpha \to 0}{\varepsilon(\alpha)} \to 0. \end{cases}$$

Proof. Suppose that x^* is an exact global minimizer of the objective function $F(x)$ on the compact D and t^* an exact global minimizer of the function $f(t) = F(\ell_\alpha(t))$ on the interval $\mathbb{I} = [0, \pi/\alpha_1]$. Since the parametrized curve $\ell_\alpha(t)$ is $\delta_{n-1}\lambda_n\alpha$-dense in D, then there exists $t' \in \mathbb{I}$ such that

$$\left\| x^* - \ell_\alpha(t') \right\| \le \delta_{n-1}\lambda_n\alpha.$$

The modified algorithm of Piyavskii MPA gives us a minimizer $t''$of the function $f(t)$ on the curve $\ell_\alpha(t)$ for $t \in \mathbb{I}$, satisfying

$$\left| f(t^*) - f(t'') \right| \le \frac{\varepsilon}{2}.$$

Then we have

$$\left| F(x^*) - f(t'') \right| \le \underbrace{\left| F(x^*) - f(t') \right|}_{\le \frac{\varepsilon}{2}} + \underbrace{\left| f(t') - f(t'') \right|}_{\le \frac{\varepsilon}{2}} \le \varepsilon,$$

and as the function $F(x)$ is Hölderian on Ω, it follows that

$$\left|F(x^*) - f(t')\right| = \left|F(x^*) - F(\ell_\alpha(t'))\right| \leq H \left\|x^* - \ell_\alpha(t')\right\|^\beta.$$

So to have the desired accuracy ε, it suffices to choose:

$$\alpha = (\frac{\varepsilon}{2H})^{1/\beta} \frac{1}{\delta_{n-1}\lambda_n}.$$

Consequently, the densification parameter α is calculated, depending on the required search accuracy.

Problem 1

$$\min F_1(x_1, x_2) = -\left|\cos x_1 \cos x_2 \; e^{\left|1 - \frac{\left(x_1^2 - x_2^2\right)^{\frac{1}{2}}}{\pi}\right|}\right|$$

on the non-convex set $D = \{(x_1, x_2) \in \mathbb{R}^2 / \; g_1(x_1, x_2) \leq 0 \text{ and } g_2(x_1, x_2) \leq 0 \; \}$
where $g_1(x_1, x_2) = -1 + x_2 - \frac{1}{4}\left|x_1\right|^{1/3}$ and $g_2(x_1, x_2) = -1 - x_2 + \frac{1}{4}\left|x_1\right|^{1/3}$
the set D is contained in the rectangle $\Omega = [-1, 1] \times [-1, 2]$.

In this problem we take: $\varphi_2(x_1) = \frac{1}{4}\left|x_1\right|^{1/3} - 1$ and $\psi_2(x_1) = \frac{1}{4}\left|x_1\right|^{1/3} + 1$.The global minimizer of F_1 is $x^* = (0, 0)^T$ at which both the objective and the constraint functions $\varphi_2(x_1)$ and $\psi_2(x_1)$ are Hölderian and non-smooth.

The parametrized curve $\ell_\alpha(t) = (\ell_1(t), \ell_2(t))$ defined on the interval $[0, \pi/\alpha_1]$ by

$$\begin{cases} \ell_1(t) = -\cos \alpha_1 t \\ \ell_2(t) = \frac{1}{4}\left|\cos \alpha_1 t\right|^{1/3} - \cos \alpha_2 t \end{cases}$$

with

$$\begin{cases} \alpha_2 \geq \frac{2\pi\alpha_1}{\alpha} \\ \text{and } \lambda_2 = \max\left(1, 1/(4\alpha^{2/3})\right) \end{cases}$$

is $\lambda_2\alpha$-dense in D. The number α is calculated according to the accuracy ε with which this problem is numerically solved. The problem 1 can be reduced to the following Hölder one-dimensional problem:

$$\min_{t \in [0, \pi/\alpha_1]} f(t) = F_1(\ell_\alpha(t)).$$

Problem 2

$$\min F_3(x_1, x_2, x_3) = \left|\sin\left(0.5 + 9.5\sqrt{x_1^2 + x_2^2 + x_3^2}\right)\right|$$

$$\begin{aligned} \text{subject to} \quad & \sqrt{|x_1|} - 1 \leq x_2 \leq \sqrt{1 - x_1^2} \\ & \left|\sin\sqrt{x_1^2 + x_2^2}\right| \leq x_3 \leq \left|3 + \cos\sqrt{x_1^2 + x_2^2}\right| \\ & -1 \leq x_i \leq 1, i = 1, 2 \text{ and } 0 \leq x_3 \leq 4 \end{aligned}$$

In this problem we take:
$$\begin{cases} \varphi_2(x_1) = \sqrt{|x_1|} - 1 \quad , \ \psi_2(x_1) = \sqrt{1 - x_1^2} \\ \varphi_3(x_1, x_2) = \left|\sin\sqrt{x_1^2 + x_2^2}\right|, \ \psi_3(x_1, x_2) = \left|3 + \cos\sqrt{x_1^2 + x_2^2}\right|. \end{cases}$$

Then, the parametrized curve $\ell_\alpha(t) = (\ell_1(t), \ell_2(t), \ell_3(t))$ defined on $[0, \pi/\alpha_1]$ by

$$\begin{cases} \ell_1(t) = -\cos\alpha_1 t \\ \ell_2(t) = \frac{1}{2}\varphi_2(\ell_1(t))(1 + \cos\alpha_2 t) + \frac{1}{2}\psi_2(\ell_1(t))(1 - \cos\alpha_2 t) \\ \ell_3(t) = \frac{1}{2}\varphi_3(\ell_1(t), \ell_2(t))(1 + \cos\alpha_3 t) + \frac{1}{2}\psi_3(\ell_1(t), \ell_2(t))(1 - \cos\alpha_3 t) \end{cases}$$

with
$$\begin{cases} \sqrt{\alpha_3/\pi} \geq \frac{1}{\alpha}\left(\alpha_2 + \alpha_1^{-1/2} + \sqrt{2}\alpha_1^{-2/3}\right), \\ \lambda_3 = \max\left(1, \sqrt{2/\alpha\delta_2}, \sqrt{3}/\left(\alpha\delta_2\right)^{2/3}\right) \\ \text{and } \delta_2 = 1 + (2/\alpha)^{0.5} \end{cases}$$

is $\delta_2\lambda_3\alpha$-dense in the feasible region.

Problem 3. Find the global minimum of non-smooth Hölder function subject to an inequality constraint. This problem com from [12].

$$\begin{aligned} &\min F_2(x_1, x_2) = \tfrac{1}{2}\sin(\sqrt{|x_1 - x_2|}) - \tfrac{1}{2}\cos(\sqrt{|x_1 + x_2|}) \\ &\text{subject to} \quad \tfrac{1}{2}x_1^{1/3} \leq x_2 \leq \exp\left(\sqrt{x_1}\right) \\ &\qquad\qquad 0 \leq x_1 \leq 1, \ 0 \leq x_2 \leq 2.72 \end{aligned}$$

In this problem we take $\varphi_2(x_1) = \frac{1}{2}x_1^{1/3}$ and $\psi_2(x_1) = \exp\left(\sqrt{x_1}\right)$.

Then, the parametrized curve $\ell_\alpha(t) = (\ell_1(t), \ell_2(t))$ defined by

$$\begin{cases} \ell_1(t) = \frac{1}{2}(1 - \cos\alpha_1 t) \qquad \text{for} \ \ t \in [0, \pi/\alpha_1] \\ \ell_2(t) = \frac{1}{2}\varphi_2(\ell_1(t))(1 + \cos\alpha_2 t) + \frac{1}{2}\psi_2(\ell_1(t))(1 - \cos\alpha_2 t) \end{cases}$$

with
$$\begin{cases} \alpha_2 \geq \frac{2\pi\alpha_1}{\alpha} \\ \text{and } \lambda_2 = \max\left(1, 1/2\alpha^{2/3}, 2.72/\sqrt{\alpha}\right) \end{cases}$$

is $\lambda_2\alpha$-dense in the feasible region.

Problem 4.

$$\begin{aligned} &\min F_4(x_1, x_2) = -20\exp(-0.2\sqrt{\tfrac{|x_1| + |x_2|}{2}}) \\ &\text{subject to} \quad x_1^2 + x_2^2 \leq 16 \\ &\qquad\qquad -4 \ \leq x_i \leq 4, \ \ i = 1, 2. \end{aligned}$$

In this problem we take:
$$\begin{cases} a = -4, \quad b = 4 \\ \varphi_2(x_1) = -\sqrt{16 - x_1^2}, \ \psi_2(x_1) = \sqrt{16 - x_1^2}. \end{cases}$$

Then, the parametrized curve $\ell_\alpha(t) = (\ell_1(t), \ell_2(t))$ defined by

$$\begin{cases} \ell_1(t) = -4\cos\alpha_1 t \ , \qquad \text{for} \ \ t \in [0, \pi/\alpha_1] \\ \ell_2(t) = -4\sin\alpha_1 t\cos\alpha_2 t, \end{cases}$$

with

$$\begin{cases} \alpha_2 \geq \frac{2\pi\alpha_1}{\alpha} \\ \text{and } \lambda_2 = \max(1, 4/\sqrt{\alpha}) \end{cases}$$

is $\lambda_2\alpha$-dense in the feasible region.

Problem 5.

$$\begin{aligned} &\min F_5(x_1, x_2) = \sum_{k=1}^{3} \tfrac{1}{2k} \left|\cos\left(\left(\tfrac{3}{2k} + 1\right) x_1 + \tfrac{1}{2k}\right)\right| |x_1 - x_2|^{1/3} \\ &\text{subject to} \quad (2x_1 - 3)^2 + (2x_2 - 3)^2 - 9 \leq 0 \\ &\qquad\qquad 0 \leq x_i \leq 3, \ i = 1, 2. \end{aligned}$$

In this problem we take:

$$\begin{cases} a = 0, \quad b = 3 \\ \varphi_2(x_1) = \frac{3}{2} - \sqrt{\frac{9}{4} - \left(x_1 - \frac{3}{2}\right)^2}, \ \psi_2(x_1) = \frac{3}{2} + \sqrt{\frac{9}{4} - \left(x_1 - \frac{3}{2}\right)^2}. \end{cases}$$

Then, the parametrized curve $\ell_\alpha(t) = (\ell_1(t), \ell_2(t))$ defined by

$$\begin{cases} \ell_1(t) = \frac{3}{2}(1 - \cos\alpha_1 t) \qquad \text{for} \ \ t \in [0, \pi/\alpha_1] \\ \ell_2(t) = \frac{1}{2}\varphi_2(\ell_1(t))(1 + \cos\alpha_2 t) + \frac{1}{2}\psi_2(\ell_1(t))(1 - \cos\alpha_2 t) \end{cases}$$

with

$$\begin{cases} \alpha_2 \geq \frac{2\pi\alpha_1}{\alpha} \\ \text{and } \lambda_2 = \max(1, 3/\sqrt{\alpha}) \end{cases}$$

is $\lambda_2\alpha$-dense in the feasible region.

Problem 6.

$$\begin{aligned} &\min F_6(x_1, x_2) = \left|\cos(0.5 + 9.5\sqrt{x_1^2 + x_2^2})\right| \\ &\text{subject to} \quad |x_1| - (x_2 + 1)^2 \leq 0 \\ &\qquad\qquad x_1^2 + x_1^2 \leq 1 \end{aligned}$$

The objective function F_6 has 0 as global minimum value and has a continuum of global minimizers on its non-convex feasible region. In this problem we take

$$\begin{cases} a = -1, \quad b = 1 \\ \varphi_2(x_1) = \sqrt{|x_1|} - 1, \ \psi_2(x_1) = \sqrt{1 - x_1^2}. \end{cases}$$

Then, the parametrized curve $\ell_\alpha(t) = (\ell_1(t), \ell_2(t))$ defined by

$$\begin{cases} \ell_1(t) = -\cos\alpha_1 t \qquad \text{for} \ \ t \in [0, \pi/\alpha_1] \\ \ell_2(t) = (-1 + \sin\alpha_1 t + \sqrt{|-\cos\alpha_1 t\ |})\cos^2\frac{\alpha_2 t}{2} + \sin\alpha_1 t, \end{cases}$$

with

$$\begin{cases} \alpha_2 \geq \frac{2\pi\alpha_1}{\alpha} \\ \text{and } \lambda_2 = \max\left(1, \sqrt{2/\alpha}\right), \end{cases}$$

is $\lambda_2\alpha$-dense in the feasible region.

3.4 Conclusion

A new coupled method is proposed called RT-MPA method, for solving non-convex, non-smooth constrained global optimization problem. We have consider a class of problems when the objective and the constraint functions are all Hölderian and non-differentiable, and we have shown how to solve such a problem without using the Lagrangian and the dual problem. Indeed, this class of problems is very difficult to treat, because the non-differentiability of the objective and the constraint functions, with the existence of a large number of local and global extrema. This method is a combination of two methods. In fact, by means of the reducing transformation method, we converted the multivariate initial problem to a univariate Hölder non-differentiable one. Before starting the minimization procedure, we approach the non-convex feasible area with an α-dense curve by constructing explicitly a parametric representation $x_i = \ell_i(t), 1 \leq i \leq n$ defined on the interval $\mathbb{I}$ of $\mathbb{R}$. For minimizing the univariate Hölder function on the interval $\mathbb{I}$, we have proposed a modification of Piyavskii's algorithm MPA which we have introduced in [12] and which does not require any other assumptions but only the knowledge of the constants h and β. We have used the Piyavskii's algorithm because it is the best known and most studied in the literature. The reducing transformation method is rather effective for dealing with difficult problems and its numerical implementation is very simple and does not require a memory space. Although this method has its own advantages, the disadvantage is that it generates an excessive number of evaluation points of the objective function, particularly when the dimension of the feasible set is greater then four or when the objective function is subject to great variations.

3.5 Exercises

Exercise 3.5.1. [6] **Lipschitz problems in multiple variables.**
Determine the Lipschitz constants $L_i (i = 1, 2, \infty)$ with respect to the norms $\|\cdot\|_i$ for the following functions:

$f(x, y) = 4xy \sin(4\pi y), \quad (x, y) \in [0, 1] \times [0, 1]$
$f(x, y) = \sin(2x + 1) + 2 \sin(3y + 2), \ (x, y) \in [0, 1] \times [0, 1]$
$f(x, y) = -\max\left\{\sqrt{3}x + y, -2y, y - \sqrt{3}x\right\}, (x, y) \in [-1, 1] \times [-1, 1]$
$f(x, y) = e^{-x^2} \sin x - |y|, (x, y) \in [0, 10]^2$
$f(x, y) = -100(y - x^2)^2 - (x - 1)^2, (x, y) \in [-3, 3] \times [-1.5, 4.5]$
$f(x, y, z) = \sin x \sin xy \sin xyz \ , (x, y, z) \in [0, 4]^3$.

Hint. We have:

$$\begin{aligned} \|\cdot\|_\infty &\leq \|\cdot\|_1 \leq n \|\cdot\|_\infty \\ \|\cdot\|_\infty &\leq \|\cdot\|_2 \leq \sqrt{n} \|\cdot\|_\infty \\ \|.\|_2 &\leq \|.\|_1 \leq \sqrt{n} \|.\|_2 \\ n^{-1/2} \|.\|_1 &\leq \|.\|_2 \leq \|.\|_1 \end{aligned}$$

Exercise 3.5.2. Let $X = (x, y) \in \mathbb{R}^2$ and $f : \mathbb{R}^2 \to \mathbb{R}$ be a function defined by:

$$f(X) = \frac{-1}{\|X - (1, 2)\|_2^2 + 3} \ .$$

1- Let $g : \mathbb{R} \to \mathbb{R}$ be defined by:

$$g(t) = 2\,|t-\alpha|\,.\,\left((t-\alpha)^2+3\right)^{-2}, \qquad \alpha \in \mathbb{R}.$$

a) Show that g reaches its maximum at the points: $t_1 = \alpha + 1$ and $t_2 = \alpha - 1$.
b) Calculate $\max\limits_{\mathbb{R}} g(t)$.

Hint. Compute $g'(t)$ by studying the two cases $t \geq \alpha$ and $t \leq \alpha$.

2- Show that for all $x, y \in \mathbb{R}$, we have

$$\begin{aligned}\left|\frac{\partial f(x,y)}{\partial x}\right| &\leq 2\,|x-1|\,.\,\left((x-1)^2+3\right)^{-2}\\ \left|\frac{\partial f(x,y)}{\partial y}\right| &\leq 2\,|y-2|\,.\,\left((y-2)^2+3\right)^{-2},\end{aligned}$$

and deduce the Lipschitz constant of f over $\mathbb{R}^2$.

Exercise 3.5.3. Consider the following global minimization problem:

$$\underset{(x,y)\in\Omega}{Min}\ f(x,y)$$

where $f(x,y) = \max\left\{-2y, \sqrt{3}x + y, -\sqrt{3}x + y\right\}$ and $\Omega = \left\{(x,y) \in \mathbb{R}^2 / \begin{array}{l} -1 \leq x \leq 1 \\ -1 \leq y \leq 1 \end{array}\right\}$.

1- Show that f is Lipschitz over Ω and determine the Lipschitz constant L.

2- Show that f is continuous over Ω. Deduce the existence of $(x_0, y_0) \in \Omega$ such that $f(x_0, y_0) \leq f(x,y), \forall (x,y) \in \Omega$.

3- What is the minimum number of iterations N_{pass} required to achieve the global minimum of f over Ω with precision $\varepsilon = 10^{-2}$ using the passive covering algorithm? Calculate three approximations within Ω using this algorithm.

4- Does f admit an upper bounding function over Ω? If yes, determine one.

Exercise 3.5.4. Consider the following global maximization problem:

$$\left\{\begin{array}{l}\max f(x_1, x_2)\\ (x_1, x_2) \in \mathbf{P}\end{array}\right.$$

where $f(x_1, x_2) = x_1^2 + 2x_1x_2 + x_2^2 - |x_1|$ and $\mathbf{P} = [-1, 2] \times [1, 2]$.

1° Show that there exists a constant $L > 0$ such that:

$$|f(x_1,x_2) - f(y_1,y_2)| \leq L\,\|(x_1,x_2) - (y_1,y_2)\|_1\,, \quad \forall (x_1,x_2),(y_1,y_2) \in \mathbf{P}.$$

Hint. In $\mathbb{R}^2$, $\left\|.\right\|_2 \leq \|.\|_1 \leq \sqrt{2}\,\|.\|_2$ and $\left\|\,(x_1,x_2)\right\|_1 = \max\{|x_1|\,,|x_2|\}$, $\|(x_1,x_2)\|_2 = \sqrt{x_1^2+x_2^2}$.

2° Deduce the existence of a point $(x_1^*, x_2^*) \in \mathbf{P}$, such that

$$f(x_1,x_2) \leq f(x_1^*, x_2^*), \forall (x_1,x_2) \in \mathbf{P}.$$

3° Let the function $g(x_1, x_2) = 1 + L \max\{|x_1|, |x_2 - 1|\}$, show that

$$f(x_1, x_2) \leq g(x_1, x_2), \forall (x_1, x_2) \in \mathbf{P}.$$

Does there exist a function h such that $h(x_1, x_2) \leq f(x_1, x_2), \forall (x_1, x_2) \in \mathbf{P}$?. If yes, provide one.

4° Define the recursive sequence of the Evtushenko algorithm starting from $X^\circ = (-1, 1)^t$ and $\varepsilon = 10^{-1}$.

Exercise 3.5.5. Consider the following global minimization problem:

$$(P) \begin{cases} \min \ f(x_1, x_2, x_3) = x_1 + 4x_2 + |x_3| \\ (x_1, x_2, x_3) \in [0, 1]^3 \end{cases}$$

1- Does f have a global minimum $x^* \in [0, 1]$? Justify.

2- Determine the number of function evaluations N required to reach x^* using the passive covering method with precision $\varepsilon = 10^{-1}$.

3- How can we improve the number N?

Hint.
$\|x\|_2 = (\sum x_i^2)^{1/2}$, $\|x\|_1 = \sum |x_i|$, $\|x\|_\infty = \max(|x_i|)$,
$x = (x_1, ..., x_n) \in \mathbb{R}^n$, $\|x\|_1 \leq n \|x\|_\infty$, $\|x\|_2 \leq \sqrt{n} \|x\|_\infty$.

Exercise 3.5.6. Let f be a Lipschitz function with Lipschitz constant L_f over the square $\mathbf{P} = [1, 2]^2$ defined by:

$$f(x, y) = (x + y)^2 - \sqrt{x},$$

and let $g(x, y)$ be a function defined on $\mathbf{P}$ by $g(x, y) = f(1, 1) - L_f \|(x, y) - (1, 1)\|_\infty$.

1- Show that g is Lipschitz with constant L_g to be determined.

2- Determine the constant L_f.

3- Prove the following inequality:

$$3 - 8(|x - 1| + |y - 1|) \leq f(x, y), \forall (x, y) \in \mathbf{P}.$$

Bibliography

[1] H. Ammar and Y. Cherruault, Approximation of several variables function by a one variable function and application to global optimization, Mathematical and computer Modeling 18(2)(1993) 17-21.

[2] A. R., Butz, Space filling curves and mathematical programming, Information and control 12(4)(1968) 313-330.

[3] D. Guettal and A. Ziadi, Reducing transformation and global optimization, Applied Mathematics and Computation 218(2012) 5848-5860.

[4] E. Gourdin, B. Jaumard, and R. Ellaia, Global Optimization of Hölder function, J. of Global Optimization 8(1996) 323-348.

[5] R. Horst and H. Tuy, Global Optimization, Deterministic Approach, Springer-Verlag, Berlin, 1993.

[6] R. Horst and P. M. Pardalos, Handbook of Global Optimization, Kluwer Academic Publishers, Dordrecht, 1995.

[7] S. K. Mishra, Some test functions for global optimization and performance of repulsive particle swarn method, MPRA Paper, http://mpra.ub.uni-muenchen.de/2718/ (2007).

[8] G. Mora and Y. Cherruault, Characterization and generation of α-dense curves. Computers and Mathematics with Applications 33(9)(1997) 83-91.

[9] G. Mora, Y. Cherruault and A. Ziadi, Functional equations generating space densifying curves. Computers and Mathematics with Applications 39(2000) 45-55.

[10] J. Pinter, Global optimization in action, Continuous and Lipschitz optimization: Algorithm, Implementations and Applications, Kluwert, Dordrecht, 1996.

[11] S. A. Piyavskii, An algorithm for finding the absolute minimum for a function, Theory of Optimal Solution 2(1967) 13-24.

[12] M. Rahal, Extension de certaines méthodes de recouvrement en optimisation globale, Thèse de Doctorat en Sciences, Université Ferhat Abbas, Sétif 1, Algeria, 2009.

[13] H. Sagan, Space-filling curves, Springer-Verlag, New York, 1994.

[14] A. Törn and A. Zilinska, Global optimization, Lecture Notes in Computer Sciences, Springer-Verlag, 1998.

[15] R. J. Vanderbei, Extension of Piyavskii's algorithm to continuous global optimization, J. Glob. Opt. 14 (1999) 205-216.

[16] A. Ziadi and Y. Cherruault, Generation of α-dense curves in a cube of $\mathbb{R}^n$, Kybernetes 27(4)(1998) 416-425.

[17] A. Ziadi, Y. Cherruault and G. Mora, The existence of α-dense curves with minimal lenght in a metric space. Kybernetes 29(2)(2000) 219-230.

[18] A. Ziadi and Y. Cherruault, Generation of α-dense curves and application to global optimization, Kybernetes 27(1)(2000) 71-82.

[19] A. Ziadi, D. Guettal and Y. Cherruault Global optimization: the Alienor mixed method with Piyavskii-Shubert technique, Kybernetes 34(7/8)(2005) 1049-1058.

[20] Jean-Christophe CULIOLI, Introduction à l'optimisation, Ellipeses, (1994).

[21] Yves Cherruault and Gaspar Mora, Optimisation Globale Theorie des courbes α-dense, ECONOMICA, (2005).

[22] Gerard Berthiau et Patrick Siarry, Etat de l'art des méthodes d'optimisation globale, Rairo Operations Research 35 (2001) 329-365.

[23] Ya. D. Sergeyev, R. G. Strongin and D. Lera, Introduction to Global Optimization Exploiting Space-Filling-Curves, NewYork, Springer, 2013.