Journal of Universal Computer Science, vol. 4, no. 6 (1998), 589-603
submitted: 25/9/97, accepted: 22/4/98, appeared: 28/6/98 © *Springer Pub. Co.*

Enclosure Methods for Multivariate Differentiable Functions and Application to Global Optimization

Frédéric Messine
(Lab. LIA, Département d'Informatique de l'Université de Pau
messine@univ-pau.fr)

Jean-Louis Lagouanelle
(Lab. LIMA, Institut de Recherche en Informatique de Toulouse
lagouane@irit.fr)

Abstract: The efficiency of global optimization methods in connection with interval arithmetic is no more to be demonstrated. They allow to determine the *global optimum* and the corresponding *optimizers*, with certainty and arbitrary accuracy. One of the main features of these algorithms is to deliver a function enclosure defined on a *box* (right parallelepiped). The studied method provides a *lower bound* (or *upper bound*) of a function in that box throughout two different strategies. As we shall see, these algorithms associated with various Branch and Bound methods lead to accelerated convergence and permit to avoid the cluster problem.

Key Words: Global optimization, Interval arithmetic, Multivariate functions, Branch and Bound algorithm, Taylor's expansion, Polyhedral cone.

1 Introduction

In this paper the methods which are proposed, concern the minimization problem in the context of Branch and Bound algorithms and interval arithmetic [6], [10]. We use a first order Taylor's expansion and the mean value theorem to build *affine underestimations* of the function over the box. The intersection of these underestimations is easily computed, then a *lower bound* of the function on the box follows. This *lower bound* is often better than those which are issued from different inclusion functions, in connection with interval arithmetic [9].

First, we shall deal with the building of the affine underestimations and the principle of the algorithm to get the lower bound. Then, its efficiency will be shown in section 5, when used within Branch and Bound algorithm, on simple examples of unconstrained minimization of multivariate functions.

Usually, the minimization problem is described by:

$$\text{Minimize } f(x), \text{subject to } x \in X, \text{ with } X \subseteq \mathbb{R}^n. \tag{1}$$

f is a differentiable n-variables function and X is the right parallelepiped in $\mathbb{R}^n$, defined by $X = \{x = (x_1, \cdots, x_n); x_i^L \leq x_i \leq x_i^U\}, \forall i \in \{1, 2, \cdots, n\}$.

The principle used here, including the underlying mean value theorem, has already been studied in the case of univariate polynomial function by Jaumard, Hansen, and Xiong [5], by Alefeld [1], and by Visweswaran and Floudas [11]. Mladineo gives, for functions that satisfy a Lipchitz condition, a method [8] based on the building of cones with spherical bases, where a mixed linear non-linear system must be solved.

In our method, we get polyhedral cones and we want to solve very simple linear systems to keep the efficiency of the algorithms. The unique solution of each linear system corresponds to the vertex of a polyhedral cone that is meant for a minimizer of f on a box X. This is reached by two suitable choices of vertices of the hypercube as explained in sections 3. Convergence and complexity properties are given in section 4. Numerical experiments and applications to global optimization follow in section 5.

The following section gives the basic tools for the construction of the polyhedral cones.

2 Underestimating affine functions

According to Taylor's expansion of a function f, we get some inclusion function of f in the box X [9]. This inclusion function is called the Taylor form:

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

Where $g(X)$ is an *interval vector* of the gradient inclusion of f; g will be automatically calculated by *automatic differentiation* [3], [6].

Let $X_i = [x_i^L, x_i^U]$ and $G_i(X) = [g_i^L(X), g_i^U(X)]$ for $i = 1, 2, \cdots, n$.

So, for every $x \in X, \dfrac{\partial f}{\partial x_i}(x) \in G_i(X)$.

We assume, for the following, that $g_i^L(X).g_i^U(X) < 0$.

Let $s = (s_1, \cdots, s_n)$ be the coordinates of a vertex S of the box X and

$$g^S(X) = (g_1^S(X), \cdots, g_n^S(X))^T$$

where

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

Then, from the following trivial inequalities:

$$\begin{aligned} (y_i - x_i^L) g_i^L(X) \leq (y_i - x_i^L)\alpha \leq (y_i - x_i^L) g_i^U(X) \\ \text{and } (y_i - x_i^U) g_i^U(X) \leq (y_i - x_i^U)\alpha \leq (y_i - x_i^U) g_i^L(X) \end{aligned}$$

for all $i = 1, 2, \cdots, n$ and any given $\alpha \in G_i(X)$, we can deduce for every vertex S of the box X, an affine underestimation of a function f over X, that is

$$f(y) \geq f(s) + (y-s)^T g^S(X), \forall y \in X \tag{3}$$

So we can obtain 2^n affine underestimations of the function f for the box X. Each one of these hyperplanes is a *maximum supporting hyperplane* at the vertex S, for the box X.

3 Global minimum enclosure methods

The first aim of the following algorithm is to get a lower bound of a function f on a box X. This will be achieved by the intersection of $n+1$ hyperplanes chosen amongst the 2^n hyperplanes that we can build with the properties given in section 2; but that does not give automatically a lower bound of f on X; except in the case of an univariate function [5], [11].

For $k = 1, \cdots, 2^n$ let Π_k be a maximum supporting hyperplane, u_k its affine representation and E_k^+ the restriction to the box X of the associated half-space defined by

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

classically we have the

Property 1 *If* $\bigcap_{k \in K} E_k^+$ *is a polyhedral cone with vertex C which contains the graph of the function f then for $|K| = n+1$, $\{C\} = \bigcap_{k \in K} \Pi_k$ exists and is unique.*

Remark. The choice of $|K| = n+1$ is minimal, it is possible to set $|K| > n+1$ with an analogous convexity property for a polyhedrical set but then the computation of the lower bound becomes too expensive.

One way to obtain existence and uniqueness of such a point C is given by the following choice of the hyperplanes Π_k.

First of all the $n+1$ vertices S_k selected are not contained in the same hyperplane of $\mathbb{R}^n$.

Then each affine function u_k satisfies the relations

$$u_k(s_k) = f(s_k) \tag{4}$$

$$u_k(x) \leq f(x), \forall x \in X \tag{5}$$

$$u_k(s_k) \geq u_k(x), \forall x \in X \tag{6}$$

Conditions (4) and (5) follow directly from the fact that Π_k is a maximum supporting hyperplane.

Condition (6) supposes, here, that according to section 2, $g_i^L(X) g_i^U(X) < 0$ for all i; this is obtained by application of a monotonicity test to f on X.

Let $(x_c, z_c) \in \mathbb{R}^n \times \mathbb{R}$ the coordinates of the point C and $u_k(x) = f(s_k) + (x - s_k)^T g^{S_k}(X)$, we introduce now the

Definition 1. A set of $n+1$ vertices of X is said *admissible* if the vertex C of the polyhedral cone defined by $n+1$ hyperplanes, built as above, satisfies $z_c \leq f(x), \forall x \in X$ and $x_c \in X$. For commodity the simplex, convex hull of the vertices $S_0 S_1 \cdots S_n$ is said admissible.

Such a point C gives a lower bound z_c of f and x_c is destined to be a current minimizer for problem (1). Let us see on a short example that any arbitrary set of $n+1$ vertices of a box X in $\mathbb{R}^n$ may not be admissible.

Example 1.

$$\begin{aligned} f : X \subset I\!R^2 &\longrightarrow I\!R \\ (x_1, x_2) &\longmapsto f(x_1, x_2) = x_1^2 - 2x_1x_2 + 3x_1 - 5x_2 \\ X = &\quad [-5, 5] \times [-15, 10] \end{aligned}$$

then $G(X) = ([-27, 43], [-15, 5])^T$

Let $s_1 = (-5, -15)^T, s_2 = (5, -15)^T, s_3 = (5, 10)^T$. For these vertices and u_1, u_2, u_3 the corresponding hyperplanes

$$\begin{aligned} u_1(x_1, x_2) &= f(-5, -15) - 27(x_1 + 5) - 15(x_2 + 15) \\ u_2(x_1, x_2) &= f(5, -15) + 43(x_1 - 5) - 15(x_2 + 15) \\ u_3(x_1, x_2) &= f(-5, 10) - 27(x_1 + 5) + 5(x_2 - 10) \end{aligned}$$

we get $(x_c, z_c) = (-3.57, -15, -103.6)$ but $f(5, 10) = -110$, hence z_c is not a lower bound; $\{s_1, s_2, s_3\}$ is not admissible.

In sections 3.1 and 3.2 we deal with two kinds of admissible sets of vertices; in section 3.2 we show that such sets always exist and can be easily identified.

3.1 Admissible right simplexes

The first idea to get straightforward formulae from linear systems is to consider the $n + 1$ vertices of a right simplex generated by a reference vertex S_0 and by the n adjacent vertices of a box X.

So doing the linear system is easily solved because from the vertex S_0 to one of the others S_k, only one variable is modified.

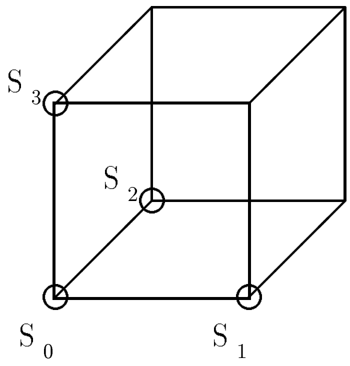

Figure 1: Right Simplex

The following result gives a *sufficient condition* for a right simplex to be admissible.

Theorem 2. *Let S_0 be a vertex of a box X in $\mathbb{R}^n$, let $S_1, S_2, \cdots, S_n$ be the vertices adjacent to S_0. If $g^{S_k}(X)$ is defined by (2) for all k. Then the corresponding right simplex T_0 is admissible if and only if*

$$\sum_{i=1}^{n} \frac{g_i^{S_0}(X)}{g_i^{S_i}(X) - g_i^{S_0}(X)} \geq -1 \tag{7}$$

where $g_i^{S_k}(X) = g_i^L(X)$ if $(s_k)_i = x_i^L$ and $g_i^{S_k}(X) = g_i^U(X)$ if $(s_k)_i = x_i^U$.

Proof. Let $s^k = (s_1^k, s_2^k, \cdots, s_n^k)^T$ be the coordinates of the vertex S_k for all k and u_k the affine function

$$u_k(x) = f(s^k) + (x - s^k)^T g^{S_k}(X)$$

First, note that S_k differs from S_0 only by one coordinate, supposed to be x_k.

Consider the auxiliary problem

$$\begin{cases} \text{Minimize} & z \\ \text{subject to} & (x, z) \in E_k^+, \forall k \in K \\ & x \in X \end{cases} \tag{8}$$

If x^* is a solution of this linear program, then (x^*, z^*) satisfies property 1; z^* is a lower bound of f on the box X and $x^* \in X$.

Standard form of (8) is

$$\begin{cases} \text{Minimize} & z \\ \text{subject to} & z = e_k + f(s^k) + (x - s^k)^T g^{S_k}(X), \\ & 0 \leq e_k, k = 0, 1, 2, \cdots, n (x, z) \in E_k^+, \forall k \in K \\ & x \in X \end{cases} \tag{9}$$

e_k is a slack variable.

The optimal solution of (9) is reached at the unique extremal point C^*, vertex of the polyhedral cone:

$$C^* = \bigcap_{k=0}^{n} \Pi_k$$

If $x_1, x_2, \cdots, x_n$ are the "basis" variables and $e_k = 0, k = 0, 1, \cdots, n$, $x^* = (x_1^*, x_2^*, \cdots, x_n^*)^T$ is solution of the linear system

$$z = f(s^k) + (x - s^k)^T g^{S_k}(X), k = 0, 1, 2, \cdots, n$$

Moreover S_k is adjacent to S_0 for all k and g^{S_k} is built following (2). Then by subtraction of equation (k) from equation (0) one gets directly

$$\begin{cases} x_k^* = \dfrac{f(s^0) - f(s^k)}{g_k^{S_k}(X) - g_k^{S_0}(X)} + \dfrac{s_k^k g_k^{S_k}(X) - s_k^0 g_k^{S_0}(X)}{g_k^{S_k}(X) - g_k^{S_0}(X)} \\ k = 1, 2, \cdots, n \end{cases} \tag{10}$$

$x^* \in X$ that is proved using (3) and the eventually optimal value z^* for the problem (8) is

$$z^* = f(s^0) + (x^* - s^0)^T g^{S_0}(X) \tag{11}$$

But this solution is optimal if and only if the marginal costs are non negative.

From the standard form (9) we get

$$z = z^* + \sum_{k=0}^{n} cm_k e_k$$

$$\text{with } cm_0 = 1 + \sum_{k=1}^{n} \frac{g_k^{S_0}(X)}{g_k^{S_k}(X) - g_k^{S_0}(X)}$$

$$\text{and } cm_k = -\frac{g_k^{S_0}(X)}{g_k^{S_k}(X) - g_k^{S_0}(X)}, k = 1, \cdots, n$$

but the assumption $g_k^L(X) g_k^U(X) < 0$ induces that $cm_k > 0$ for any $k \neq 0$ and optimality of (x^*, z^*) implies that $cm_0 > 0$.

Let $w(G_i(X)) := g_i^U(X) - g_i^L(X)$ and $< G_i(X) > := \min\{g_i^U(X), |g_i^L(X)|\}$ be respectively the width and the mignitude of the interval $G_i(X)$, for $i = 1, 2, \cdots, n$.

The following condition of existence of an admissible right simplex is deduced from the worst case in the previous inequality (7).

Theorem 3. *For $n \geq 3$ admissible right simplex exists on the box X if and only if*

$$\sum_{i=1}^{n} \frac{< G_i(X) >}{w(G_i(X))} \leq 1 \qquad (12)$$

Remark. As it can be seen in Algorithm 1 when no admissible right simplex can be found on the box, this condition is weakened and a lower bound will be found again, but with a loss of accuracy.

For univariate or bivariate functions, inequality (12) is always satisfied and a right simplex can be found.

This leads to the

Algorithm 1-

- Application of a monotony test (elimination of components).
- Choose a right simplex satisfying (12); determination of the set K.
- For $n \geq 3$ if no such simplex exists then apply one of these local strategies :
  - relaxation of the gradient; then modify Π_k until the condition is satisfied,
  - Return the intersection of the polyhedral cone with the unbounded cylinder which has for base the box X: return the "lower point" so obtained.
- Compute x^*, z^* by (10) and (11).

Proof.
- if the optimality condition (7) is satisfied then a lower bound is directly obtained by computation of $\{C\} = \bigcap_{k \in K} \Pi_k$,
- Relaxation: one moves some hyperplanes Π_k until the condition (7) is satisfied,
- Intersection is equivalent to compute directly the first order Taylor inclusion function [10].

This algorithm runs only for right simplexes; the aim of the next section is to answer positively the question : for a given box $X \in I\!R^n$, is it always possible to find an admissible set of $n+1$ vertices?

3.2 Construction of admissible simplexes

Let $x^L = (x_1^L, x_2^L, \cdots, x_n^L)$ and $x^U = (x_1^U, x_2^U, \cdots, x_n^U)$ be the vertices of X with extremal coordinates. $\bar{S}$ is called the opposite of a given vertex S on the box X when $\bar{s} = x^L + x^U - s$.

We naturally associate to the box X, the symmetric directed graph constructed from its vertices and edges. Then the following result shows that, for any given vertex S and the opposite one $\overline{S}$ on the box X, it is always possible to find a path of length n : $SS^{k_1} \cdots S^{k_{n-1}}\overline{S}$ such that the corresponding simplex is admissible.

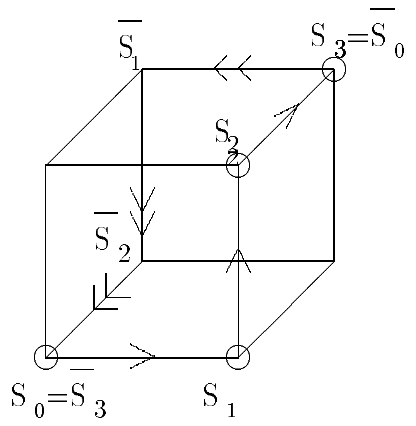

Figure 2: Admissible Simplex

$g_i^S(X)$ being defined by (2), we denote

$$K_i = -\frac{|g_i^S(X)|}{w(G_i(X))}, i = \{1, \cdots, n\}$$

Theorem 4. *Let S be an arbitrary vertex of the box X, $\overline{S}$ its opposite. Let the real numbers K_i be sorted in increasing order so that*

$$-1 < K_{k_1} \leq K_{k_2} \leq \cdots \leq K_{k_n} < 0$$

Then the simplex defined by vertices $SS^{k_1} \cdots S^{k_{n-1}} S^{k_n} (= \overline{S})$ is admissible *when the path from S to $\overline{S}$ is defined for $j = 0, 1, 2, \cdots, n-1$ by*

$$\begin{cases} s_i^{k_{j+1}} := s_i^{k_j}, i \in \{1, 2, \cdots, n\} \setminus k_{j+1} \\ s_{k_{j+1}}^{k_{j+1}} := x_{k_{j+1}}^L + x_{k_{j+1}}^U - s_{k_{j+1}}^{k_j} \end{cases} \quad (13)$$

where $S^{k_0} = S$ *and* $S^{k_n} = \bar{S}$.

This path is unique for strict inequalities.

Proof. One proceeds analogously to the previous theorem. The only difference lies on the fact that in this case, two consecutive vertices have $n-1$ identical coordinates. Consequently x^* optimal solution of the associated linear program is obtained directly by subtraction of two consecutive equations; and then, we get

$$x^*_{k_j} = \frac{f(s^{k_{j-1}}) - f(s^{k_j}) + g_{k_j}^{S_{k_j}}(X) s_{k_j}^{k_j} - g_{k_j}^{S_{k_{j-1}}}(X) s_{k_j}^{k_{j-1}}}{g_{k_j}^{S_{k_j}}(X) - g_{k_j}^{S_{k_{j-1}}}(X)} \qquad (14)$$

for $k_j = 1, 2, \cdots, n$

and

$$z^* = f(s) + (x^* - s)^T g^S(X) \qquad (15)$$

one must show that $x^* \in X$ and that z^* is a lower bound of f.

$x^* \in X$ is straightforward when one writes $x^L_{k_j} \le x^*_{k_j} \le x^U_{k_j}$, for all k_j under the hypothesis $g^L_{k_j} g^U_{k_j} < 0$ and (3).

Otherwise z^* is the minimal value of auxiliary problem (8) if and only if the marginal costs are non-negative because

$$z = z^* + e_{k_0} + \sum_{k_j=1}^{n} \frac{e_{k_{j-1}} - e_{k_j}}{g_{k_j}^{S_{k_j}}(X) - g_{k_j}^{S_{k_{j-1}}}(X)} g_{k_j}^{S_{k_0}}(X)$$

which implies that

$$0 < 1 + \frac{g_{k_1}^{S_{k_0}}(X)}{g_{k_1}^{S_{k_1}}(X) - g_{k_1}^{S_{k_0}}(X)}$$

and

$$\frac{g_{k_j}^{S_{k_0}}(X)}{g_{k_j}^{S_{k_j}}(X) - g_{k_j}^{S_{k_{j-1}}}(X)} \le \frac{g_{k_{j+1}}^{S_{k_0}}(X)}{g_{k_{j+1}}^{S_{k_{j+1}}}(X) - g_{k_{j+1}}^{S_{k_j}}(X)}$$

for $j = 1, 2, \cdots, n-1$

But considering (13) and (2) we get

$$g_{k_j}^{S_{k_{j-1}}}(X) = g_{k_j}^{S_{k_{j-2}}}(X) = \cdots = g_{k_j}^{S_{k_0}}(X)$$

and

$$g_{k_j}^{S_{k_j}}(X) = \frac{g_{k_j}^{S_{k_{j-1}}}(X) - g^L_{k_j}(X)}{g^U_{k_j}(X) - g^L_{k_j}(X)} g^L_{k_j}(X) + \frac{g^U_{k_j}(X) - g_{k_j}^{S_{k_{j-1}}}(X)}{g^U_{k_j}(X) - g^L_{k_j}(X)} g^U_{k_j}(X)$$

for $j = 1, 2, \cdots, n$. Therefore, in any case,

$$\frac{g_{k_j}^{S_{k_0}}(X)}{g_{k_j}^{S_{k_j}}(X) - g_{k_j}^{S_{k_{j-1}}}(X)} = -\frac{|g_{k_j}^{S_{k_0}}(X)|}{w(G_{k_j}(X))}$$

because $g_j^L(X) < 0 < g_j^U(X)$, for $j = 1, 2, \cdots, n$.

Hence for, any initial vertex S, at least one admissible simplex, can be found applying the rule given in theorem (4) that gives a path from S to $\bar{S}$.

If the initial vertex is $\bar{S}$ instead of S, it is easily seen that one finds the same path in reverse order towards S. Then, for a box X, it exists at least 2^{n-1} admissible simplexes and as much lower bounds for f. We do not study here the best choice amongst all these possibilities, a choice that must take into account the values of f and its derivative.

For illustration, we apply this method to example 1

$s = (5, -15)^T$ is the initial vertex, $g^S = (43, -15)^T$ and $K_1 = -\frac{43}{70}$, $K_2 = -\frac{15}{20}$ which gives

$$-1 < K_2 < K_1 < 0$$

then an admissible path from $s = (5, -15)^T$ to $\bar{s} = (-5, 10)^T$ is obtained by modifying first the coordinate x_2 which leads to the vertex $(5, 10)$ and secondly the coordinate x_1 which leads to $(-5, 10)$ the terminal vertex.

One gets with the corresponding hyperplanes $x^* = (\frac{8}{7}, -5)$ and $z^* = -\frac{2211}{7} \simeq -315.9$.

Now we see why the right simplex $\{s_1, s_2, s_3\}$ was not admissible in this example; generally for a bivariate function on a box X, two simplexes are admissible; which, in this case, coincides with the right ones.

To get a *lower bound* of a function f from theorem (4), we apply

Algorithm 2-

- Choose a vertex S; determine $g^S(X)$.
- Application of monotony test (elimination of components).
- Computation of

$$K_k = -\frac{|g_k^S(X)|}{g_k^U(X) - g_k^L(X)}, k = \{1, \cdots, n\}$$

- Classify the associated numbers K_k in increasing order.
- Determination of an admissible simplex following theorem (4)
- Compute $\{x^*, z^*\}$ with (14) and (15).

Remark. Several very efficient improvements may be used to get tighter lower bounds: Slope matrices will advantageously take place of derivatives when their computation is possible or narrower intervals for the components of g may also result from Taylor expansion of E. Hansen [3], where some interval arguments are replaced by real quantities.

If an upper bound is required instead of a lower bound, Algorithm 2 is modified as follows

For a given vertex S of the box X, let $H_i = |K_i|$, for $i = 1, 2, \cdots, n$.

$$g_i^S(X) = \frac{s_i - x_i^U}{x_i^L - x_i^U} g_i^U(X) + \frac{s_i - x_i^L}{x_i^U - x_i^L} g_i^L(X) \tag{16}$$

The real numbers H_i are classified in increasing order

$$0 < H_{i_n} \leq H_{i_{n-1}} \leq \cdots \leq H_{i_1} < 1 \tag{17}$$

Then an admissible path from S to $\bar{S}$ is found by modifying the coordinates of S in the order $x_{i_1}, x_{i_2}, \cdots, x_{i_n}$.

Unfortunately, the paths for a lower bound and an upper bound of a function f on a box X are distinct excepted in extreme cases for example if $g_i^L(X) = C_1$ and $g_i^U(X) = C_2$ for all i.

When inequalities are strict in (17), the two paths are opposite which means that S and $\bar{S}$ are respectively the initial vertices for lower and upper bounds then $\bar{S}_{k_i}$ is the opposite vertex of S_{k_i} i.e. $\bar{s}_{k_i} = x^L + x^U - s_{k_i}$ for all i, see figure (2).

4 Optimization algorithm using these enclosure processes

The methods previously described are integrated, by the mean of Algorithm 1 and Algorithm 2 in interval Branch and Bound algorithms [10] to solve the global minimization problem (1).

This is the topic of

Algorithm 3-

- Initial enclosure of the minimum on X, say $[F^L, F^U]$
- Insert (X, F^L) in a list L
- $\tilde{f} \longleftarrow +\infty$
- While the minimizers are not found with a given tolerance
- Do
  - Extract the first element from L, write it as follows: $(Y, \tilde{y})$
  - Divide Y into two parts (along its larger edge): we get $V^{(1)}$ and $V^{(2)}$.
  - For $i = 1, 2$
  - Do
    - Enclosure the global minimum on $V^{(i)} = [F^L, F^U]$
    - if $F^L > \tilde{f}$ then
      - Go to the next step
    - else
      - if $F^U < \tilde{f}$ then
      - . $\tilde{f} \longleftarrow F^U$
      - . Remove all the elements of L which lower bound is greater than the new $\tilde{f}$.
      - end if
      - Insert $(V^{(i)}, F^L)$ in L following the increasing order of the F^L.
    - end if
  - end For
- End While
- Return $\tilde{f}$ and L.

Remark. – The upper bound F^U may also be progressively computed in the algorithm

– We have three possibilities to enclose the global minimum:

1. right simplex and, if necessary, intersection with unbounded cylinder which has for base the box X.
2. right simplex and, if necessary, gradient relaxation.
3. research of an admissible simplex with two opposite vertices of X.

– When the size of a box X is sufficiently small, it is not necessary to compute for each sub-box of X the gradient according to inclusion monotonicity but this induces a slower convergence.

For every admissible simplex defined by the vertices $S_{k_0}, S_{k_1}, \cdots, S_{k_n}$ of X, the following inequalities hold:

$$\min_{j=1,2,\cdots,n}\{f(S_{k_j})\} \geq f_{\min} \geq z^* \tag{18}$$

with $f_{\min} = \inf_{x\in X}\{f(x)\}$.

4.1 Complexity and convergence

Each underestimation of f requires

– an interval inclusion function of grad f
– $n+1$ point evaluations of the function f

and finally the computation of $\{x^*, z^*\}$ requires $O(3n)$ elementary operations for a given box X.

This estimate is relative to a single box; but considering, the splitting of a box into adjacent sub-boxes, , which is done in global optimization problems, the number of evaluations of f may be perceptibly reduced. Indeed if n vertices belong to the interface of two sub-boxes, then one more vertex is necessary on each sub-box.

Moreover changing a single vertex does not affect all the coordinates of x^* and brings only a partial contribution in the computation of z^*.

Let z_L^* and z_U^* be respectively the computed lower and upper bounds.

Let $w(X)$ and $w(G)$ be the width vectors, $w(X) = x^U - x^L$ and

$$w(G) = (w(G_1(X)), w(G_2(X)), \cdots, w(G_n(X)))^T$$

and

$$|g^S(X)| = (|g_1^S(X)|, |g_2^S(X)|, \cdots, |g_n^S(X)|)^T$$

where $g^S(X)$ is defined by (2), then properties of convergence of affine estimates run from

Theorem 5. *For a differentiable function $f : X \subset I\!R^n \to I\!R$; if an admissible set $\{S_{k_j}\}_{j=0,1,\cdots,n}$ is used to get z_L^* and the opposite one $\{\bar{S}_{k_j}\}_{j=0,1,\cdots,n}$ to get z_U^* for z_U^* on the box X, then, assuming $g_k^U(X)g_k^L(X) < 0, \forall k$, the following inequalities are valid:*

$$|z_L^* - f_{\min}| \leq \min_{j=0,1,\cdots,n}(w(X), |g^{S_{k_j}}(X)|)_{I\!R^n} \tag{19}$$

$$z_U^* - z_L^* \leq (w(X), w(G))_{I\!R^n} \tag{20}$$

And the excess width of affine estimation is at least of order 2 for a regular function.

Proof. The proof is quite straightforward drawn from the arguments :

– For all j,

$$|z_L^* - f_{\min}| \leq |z_L^* - f(s_{k_j})| = (x_L^* - s_{k_j})^T g^{S_{k_j}}(X)$$

and on equivalent inequality for z_U^*.

– For any $x \in X$,

$$\frac{\partial f}{\partial x_i}(x) = \Theta g_i^L(X) + (1-\Theta) g_i^U(X), \Theta \in]0,1[$$

– Finally, EW excess width of affine estimate for a function f is given by

$$EW = (z_U^* - z_L^*) - (f_{\max} - f_{\min}) = z_U^* - f_{\max} - (f_{\min} - z_L^*) \leq (w(X), w(G))_{\mathbb{R}^n}$$

which induces a convergence of order 2 as soon as, for example, the first derivative of f satisfies a Lipschitz condition.

5 Numerical experiments

First we search the global minimum of polynomial functions of 2, 3, or 4 variables with different degrees and of functions involving trigonometric functions.

$$
\begin{aligned}
f_1(x) &= 1 + (x_1^2 + 2)x_2 + x_1 x_2^2, \text{ with } X = [1,2] \times [-10,10] \\
f_2(x) &= 2x_1^2 - 1.05x_1^4 + x_2^2 - x_1 x_2 + \frac{1}{6}x_2^6, \text{ with } X = [-2,4]^2 \\
f_3(x) &= (x_1 - 2x_2 - 7)^2 + (2x_1 + x_2 - 5), \text{ with } X = [-2.5, 3.5] \times [-1.5, 4.5] \\
f_4(x) &= [1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)] \\
&\quad \times [30 + (2x_1 - 3x_2)^2(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)] \\
&\quad \text{with } X = [-2,2]^2 \\
f_5(x) &= (x_1 - 1)(x_1 + 2)(x_2 + 1)(x_2 - 2)x_3^2, \text{ with } X = [-2,2]^3 \\
f_6(x) &= 4x_1^2 - 2x_1x_2 + 4x_2^2 - 2x_2x_3 + 4x_3^2 - 2x_3x_4 + 4x_4^2 + 2x_1 - x_2 + 3x_3 \\
&\quad + 5x_4, \text{ with } X = [-1,3] \times [-10,10] \times [1,4] \times [-1,5] \\
f_7(x) &= x_1^2 + x_2^2 - \cos 18x_1 + x_1 \sin 18x_2 + x_3 \cos x_3 + x_1x_2x_3 \\
&\quad \text{with } X = [1,500]^3
\end{aligned}
$$

Pb	f_1			f_2			f_3			f_4			$-f_2$		
	N	T(s)	C	N	T(s)	C	N	T(s)	C	N	T(s)	C	N	T(s)	C
NE	1550	16,80	286	649	7,73	150	218	1,08	36	-	-	-	456	4,02	83
T1	114	2,29	9	158	3,76	21	174	1,64	5	5373	356,73	42	331	8,16	26
T1+M	27	0,32	5	18	0,27	6	23	0,23	3	4769	129,4	38	25	0,53	8
RS	64	0,65	8	87	1,33	25	55	0,51	3	3379	84,83	24	131	1,76	32
RS+M	23	0,29	5	18	0,29	6	19	0,16	3	3318	83,5	24	22	0,50	9

. N: Number of iterations,
. T(s): CPU time in seconds,

. C: Number of elements enclosing the optimizers in the list L,
. -: the algorithm does not give any solution after 15 minutes,
. NE: Classical Ichida Fujii Algorithm [4] using natural inclusion extension,
. T1: Ichida Fujii Algorithm using Taylor inclusion function [10],
. T1+M: Ichida Fujii Algorithm using Taylor inclusion function [10] with elimination of components by monotony,
. RS: Optimization algorithm using the "best" right simplex,
. RS+M: Method RS with elimination of components by monotony.

Pb	f_5			$-f_5$			f_6			f_7		
	N	T(s)	C	N	T(s)	C	N	T(s)	C	N	T(s)	C
NE	2832	30,99	312	322	1,43	8	-	-	-	-	-	-
T1	731	17,31	24	421	11,88	8	3959	201,86	216	-	-	-
T1+M	185	1,81	24	108	1,11	2	468	9,07	77	3058	71,30	22
RS1	329	3,58	24	228	2,36	8	861	19,31	137	6483	351,11	2934
RS2	271	3,00	18	25	0,28	2	564	12,18	101	6370	311,61	2934
AS	319	3,09	22	228	3,05	8	878	22,25	121	10221	601,42	3012
RS1+M	101	1,06	24	210	2,05	8	302	6,73	86	2093	47,50	65
RS2+M	59	0,62	12	25	0,28	2	205	4,48	67	1786	34,16	17
AS+M	110	1,71	22	34	0,72	2	309	10,05	78	2211	48,01	117

. RS1: Optimization algorithm using right simplexes and intersection of the polyhedral cone with the unbounded cylinder,
. RS2: Optimization algorithm using right simplexes and gradient relaxation,
. AS: Optimization algorithm using the research of an admissible simplex,
. RS1+M: Method RS1 with elimination of components by monotony,
. RS2+M: Method RS2 with elimination of components by monotony,
. AS+M: Method AS with elimination of components by monotony.

(18) is not taken into account, the different methods are compared uniquely by *lower bounds*. These examples are sufficient to see the improvement provided by the algorithms 1, and 2 compared to the classical Branch and Bound algorithm.

Moreover the problem of clusters [2] is strongly reduced.

The three different methods exposed here seem to have quite the same efficiency with perhaps a thin advantage for the algorithm using gradient relaxation. Admissible simplexes have always the same initial vertex $s = (x_1^L, \cdots, x_n^L)$ which is surely not the best strategy.

Secondly to test the efficiency of our enclosure methods, one computes the bounds of the function f_6 with decreasing width

$$
\begin{aligned}
X_1 &= [-1,1] \times [-1,1] \times [-1,1] \times [-1,1] \\
X_2 &= [0,1] \times [0,1] \times [0,1] \times [-1,1] \\
X_3 &= [0,1] \times [0,1] \times [0,1] \times [0,1] \\
X_4 &= [0,0.5] \times [0,1] \times [0,1] \times [-1,1] \\
X_5 &= [0,0.5] \times [0,0.5] \times [0,1] \times [-1,1] \\
X_6 &= [0.5,1] \times [0.5,1] \times [0.5,1] \times [0,1]
\end{aligned}
$$

In this example $0 \in G_i(X_j), i = 1,2,3,4$ and $j = 1,2,4$, in the others cases $0 \notin G_4(X_3)$, $0 \notin G_1(X_5)$ and $0 \notin G_i(X_6), i = 1,2,3$.

Boxes / Methods	X_1	X_2	X_3	X_4	X_5	X_6
$f(X_i)$	$[-3, 29]$	$[-1.6, 20]$	$[-0.1, 20]$	$[-1.6, 16]$	$[-1.6, 16]$	$[4, 19]$
NE	$[-17, 33]$	$[-12, 28]$	$[-7, 26]$	$[-11, 24]$	$[-9, 21]$	$[-1.5, 24.5]$
T1	$[-55, 55]$	$[-24, 32]$	$[-13.5, 27.5]$	$[-21, 27]$	$[-19.1, 24.9]$	$[-1.75, 22.3]$
RS1	$[-44, 82]$	$[-7.4, 28.6]$	$[-2, 23]$	$[-7.2, 24.8]$	$[-6, 22]$	$[3.8, 19.3]$
RS2	$[-45.3, 77.4]$	$[-7.4, 28.6]$	$[-2, 23]$	$[-7.2, 24.8]$	$[-1.6, 19]$	$[4, 19]$
AS	$[-32.2, 55.7]$	$[-7.8, 28.1]$	$[-2.2, 22.8]$	$[-7.6, 24.4]$	$[-6, 22]$	$[3.8, 19.3]$

Methods used to compute upper and lower bounds are:

. $f(X_i)$: range of f over $X_i, i = 1, \cdots, 6$,
. NE: natural inclusion extension function,
. T1: Taylor inclusion function,
. RS1: right simplex and intersection of the polyhedral cone with the unbounded cylinder,
. RS2: right simplex and gradient relaxation,
. AS: research of an admissible simplex.

First, we can see that our enclosure methods are not automatically better than classical natural inclusion extension of the function. However, when the width of the boxes decreases, our algorithms are more efficient and in some cases, they are very accurate and they give the true bounds (method RS2 over box X_6).

6 Conclusion

The methods developed in this paper, give a lower bound and/or an upper bound over a box X for a differentiable multivariate function, and therefore may be used as an evaluation function, as well as enclosure methods of a global optimum. For a function defined over a box X of $\mathbb{R}^n$, one estimation of a minimum, for example, requires point evaluations in $n+1$ vertices of the box X and an interval evaluation of the partial derivatives.

Efficiency of algorithms is obtained by suitable choices of these vertices.

Numerical results are very satisfactory on unconstrained problems; the first results on optimization problems with constraints are promising.

References

1. G. ALEFELD *Bounding the Slope of Polynomial Operators and some Applications*, Computing : 26 (227-237) - (1980).
2. K. DU, R. B. KEARFOTT *The Cluster Problem in Multivariate Global Optimization*, Journal of Global Optimization : 10 (27-32) - (1996).
3. E. HANSEN *Global optimization Using Interval Analysis*, MARCEL DEKKER, Inc. 270 Madison Avenue, New York, New York 10016 - (1992).
4. K. ICHIDA, Y. FUJII *An Interval Arithmetic Method for Global Optimization*, Computing : 23 (85-97) - (1979).
5. P. HANSEN, B. JAUMARD, J. XIONG *Decomposition and Interval Arithmetic Applied to Global Minimization of Polynomial and Rational Functions*, Journal of Global Optimization : 3 (421-437) - (1992).
6. R. B. KEARFOTT *Rigorous Global Search: Continuous Problems*, Kluwer Academic Publishers, Dordrecht, Boston, London - (1996).

7. J.L. LAGOUANELLE, F. MESSINE *Algorithme d'encadrement de l'optimum global d'une fonction différentiable*, Compte Rendu à l'Académie des Sciences, Numerical Analysis, T. 326, Série I, (1998).
8. R.H. MLADINEO *An Algorithm for Finding the Global Maximum of a Multimodal, Multivariate Function*, Mathematical Programming : 34 (188-200) - (1986).
9. R. E. MOORE *Interval Analysis*, Prentice Hall, Inc. Englewood Cliffs, N.J. - (1966).
10. H. RATSCHEK, J. ROKNE *New Computer Methods for Global optimization*, ELLIS HORWOOD LIMITED Market Cross House, Cooper Street, Chichester, West Sussex, PO19 1EB, England - (1988).
11. V. VISWESWARAN, C.A. FLOUDAS *Uncontrained and Constrained Global Optimization of Polynomial Function in one Variable*, Journal of Global Optimization: 2 (73-99) - (1992).