The Tenth International Symposium on Operations Research and Its Applications (ISORA 2011)
Dunhuang, China, August 28–31, 2011

Multistart Local Search Continuous Global Optimization Method with a Taboo Step and its Condition for Finding the Global Optimum

Hideo Kanemitsu[1]

[1]Course of Computer and Information Sciences Hakodate Campus, Hokkaido University of Education 1-2, Hachimancho, Hakodate 040-8567, Japan

Abstract We introduce a multistart local search-based method with a taboo step for solving continuous global optimization problems with bound constraints. Since this algorithm has a characteristic taboo step[5, 1995] by removing candidate points that converge to the current local optimum in each iteration, the step enables us to avoid repeated convergence to one of an already known optima in a local search. Since a similar step has been proposed by Ursem[13, 1999], known as the *hill-valley* step, we show the difference between the hill-valley step and the taboo step. Finally, we show that the algorithm stops after a finite number of iterations and finds the global optimum under certain conditions.

Keywords global optimization; multistart method; taboo step

1 Introduction

Finding a global minimum $x^{**} = \text{argmin}_{x \in D^n} f(x) : R \rightarrow R^n$ with bound constraints $D^n = \{x = (x_1, x_2, \ldots, x_n) \in R^n \,|\, L_i \leq x_i \leq U_i\}$, is a very well-known global optimization problem, and many methods have been proposed for solving such a problem with multimodal function f. Those methods are mainly classified into methods using the *deterministic approach* and methods using the *stochastic approach*.

Methods using the deterministic approach, such as branch and bound methods[4], find solutions in reduced regions by repeatedly dividing a given region into sub-regions using the Lipschitz constant or interval analysis, and these methods guarantee to deterministically find the global optimum with a given tolerance. However, in these methods, realization of the algorithms are complicated and computational time complexity often exponentially increases with an increased number of variables.

In the stochastic approach[2][9] which includes heuristic methods[3], multistart methods[10] that combine multiple sample points, selected candidates for optimal points and a local optimizer have been proposed. These methods, such as the *multistart method* and *the clustering methods* [2][11][7][8][12], can find the global optimum with a high degree of accuracy and with no special structure and simple feasible regions. Most of these algorithms can be summarized in the following general form:

SL1. Take sample points over the searching region D^n.

SL2. Select (and concentrate) candidates from sampled points according to a certain criterion.
SL3. Apply the local optimizer for each candidate as a starting point to find the local optimum.

However, one problem with these methods is that some candidates repeatedly converge to the same optima that have already been found. In order to avoid this disadvantage, we introduce the following algorithm to eliminate candidates that are expected to converge in duplicate on the same optimum[5].

SLR1. Take sample points over the searching region.
SLR2. Select and concentrate candidates with the lowest function values.
SLR3. Find the local optimum by applying the local optimizer to a starting point that has the lowest function value among the candidates.
SLR4. Remove candidates that are expected to converge to the current local optimum. If no other candidates remain, then terminate, otherwise go to **SLR3**.

Section 2 shows a global optimization problem and gives notations, definitions and properties among some definitions. The details of the above main algorithm and termination property are given in section 3. In section 4, a taboo step that is characterized by **SLR4** is described. In section 5, convergence properties of the main algorithm are discussed. Finally, concluding remarks and discussion are presented.

2 Preliminaries

2.1 Global optimization problem

In this paper, we deal with the following global optimization problem(P):

$$\begin{cases} \text{minimize} & f(x) \equiv f(x_1, x_2, \ldots, x_n)\,, \\ \text{subject to} & L_i \le x_i \le U_i\,,\ \ i = 1, 2, \ldots, n\,, \\ & D^n = \{\, x \in R^n \,|\, L_i \le x_i \le U_i,\ i = 1, 2, \ldots, n \,\}, \end{cases} \tag{P}$$

where D^n is a searching region of such that the region consists of the closed intervals by lower and upper bounds on each variable, and sample points are taken over the region D^n. We assume that the objective function $f(x)$ has a finite number of isolated local minima $x_k^* \in D^n$ $(k = 1, 2, \ldots, M)$. The set X^* of the isolated local minima and the set F^* of its minimal values are written by :

$$X^* = \{\, x_1^*, x_2^*, \ldots, x_M^* \,\}\,, \tag{1}$$

$$F^* = \{\, f(x_1^*), f(x_2^*), \ldots, f(x_M^*) \,\}\,. \tag{2}$$

For simplifying the later description, suppose the function has the unique global minimum x^{**} and its function value f^{**}.

2.2 Notation and definitions

First, we show the notations and definitions.

Notation 1.
Given a set A consisting of finite elements, the j-th element and the size of A are denoted by A_j and $|A|$, respectively.

Notation 2.
An algorithm is expressed by using the following notation:

$$(r_1, r_2, \ldots, r_q) \leftarrow algo\_name\,(a_1, a_2, \ldots, a_p).$$

This notation means that the algorithm algo_name with input p-arguments $(a_1, a_2, \ldots, a_p)$ *is applied, and then results* $(r_1, r_2, \ldots, r_q)$ *are obtained. Moreover,* $\{*$ descriptions $*\}$ *in algorithms denotes comments.*

Definition 1.
Let the level set $L(\alpha)$ *and the* strict level set $L^s(\alpha)$ *with function level* α *of the problem (P) such that*

$$L(\alpha) = \{\, x \in D^n \,|\, f(x) \leq \alpha,\ \alpha \in R \,\}. \tag{3}$$

$$L^s(\alpha) = \{\, x \in D^n \,|\, f(x) < \alpha,\ \alpha \in R \,\}. \tag{4}$$

Let the connected components of $L(\alpha)$ *and* $L^s(\alpha)$ *that include* $x \in D^n$ *be* $L_c(\alpha; x)$ *and* $L_c^s(\alpha; x)$*, respectively. Then we call the sets the* connected level set *and the* connected strict level set*[6], respectively. Let sets* $L_c^1(\alpha), L_c^2(\alpha), \ldots, L_c^m(\alpha)$ *that are satisfied the equations*

$$\begin{cases} L(\alpha) = L_c^1(\alpha) \cup L_c^2(\alpha) \cup \ldots \cup L_c^m(\alpha), \\ \forall i, \forall j\,(i \neq j) \in \{1, 2, \ldots, m\},\ L_c^i(\alpha) \cap L_c^j(\alpha) = \emptyset \end{cases} \tag{5}$$

be connected components *of level set* $L(\alpha)$.

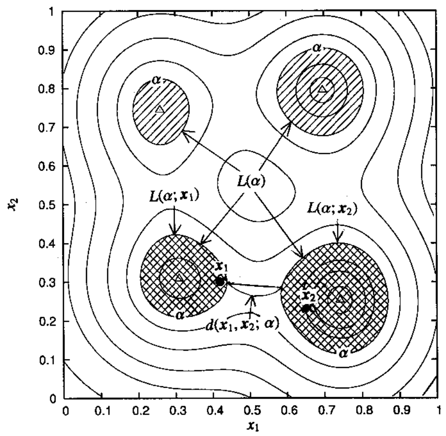


Figure.1 A *level set* $L(\alpha)$, two *connected level sets* $L_c(\alpha; x_1)$ and $L_c(\alpha; x_2)$ have distance $d(x_1, x_2, \alpha)$ between two connected level sets.

Since the searching region D^n is a bounded set, all of these connected components are bounded.

Definition 2.
The distance between two compact sets A and B is defined by

$$d(A, B) = \min\{\, \|x - y\| \,|\, x \in A,\ y \in B \,\}. \tag{6}$$

Typical examples of a *level set*, a *connected level set* of a function defined on D^2 and *distance between two connected level sets* are illustrated in **Figure** 1.

Definition 3.
We define the strongly quasi-convex region $L^s_{qc}(\alpha^*;x^*)$ *at a local minimum* $x^* \in X^*$ *as the largest connected strong level set satisfying the condition of a strongly quasi-convex function[1] as follows:*

$$\begin{cases} L^s_{qc}(\alpha^*;x^*) = \{\, x \,|\, x \in L^s_c(\alpha^*;x^*)\,\}, \\ \qquad \alpha^* = \max\{\ \alpha \mid f((1-\lambda)x_1 + \lambda x_2) < \max\{f(x_1), f(x_2)\} \\ \qquad 0 < \forall\lambda < 1,\ \forall x_1 \neq \forall x_2 \in L^s_c(\alpha;x^*)\ \}. \end{cases} \tag{7}$$

Definition 4.
Let the n-dimensional maximal open ball $B(r^*_i;x^*_i)$ *centered at local minimum* x^*_i *with maximal radius* r^*_i *such that the ball is included in the* strongly quasi-convex region *be*

$$\begin{cases} r^*_i = \max\{\, r \,|\, B(r;x^*_i) \subset L^s_{qc}(\alpha^*_i;x^*_i)\,\} \\ B(r;x^*_i) = \{\ x \mid \|x - x^*_i\| < r\ \}. \end{cases} \tag{8}$$

Moreover, the maximum radius r^{**} *of all open balls is defined as follows:*

$$r^{**} = \min_{i\in[1,M]} r^*_i. \tag{9}$$

Figure 2 shows an example of two *strongly quasi-convex regions* $L^s_{qc}(\alpha^*_1;x^*_1)$ and $L^s_{qc}(\alpha^*_2;x^*_2)$ and two *open balls* $B(r^*_1;x^*_1)$ and $B(r^*_2;x^*_2)$ defined on D^2.

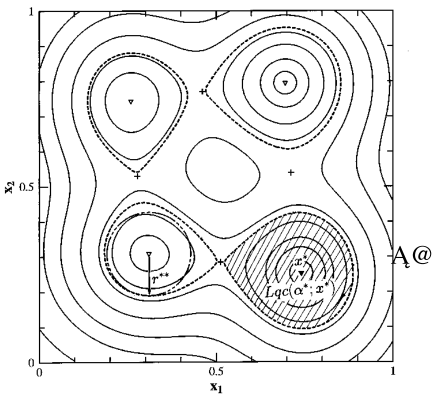


Figure.2 Two *strongly convex regions* $L^s_{qc}(\alpha^*_1;x^*_1)$ and $L^s_{qc}(\alpha^*_2;x^*_2)$ and the *open balls* $B(r^*_1;x^*_1)$ and $B(r^*_2;x^*_2)$.

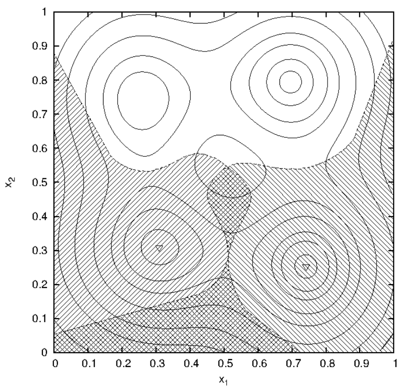


Figure.3 *Two line monotone decreasing regions* $Md(x^*_1)$ to the local minimum x^*_1 and another line monotone decreasing region $Md(x^*_2)$ to the local minimum and x^*_2.

Definition 5.
If the set $Md(x^*)$ *in which the function value* $f(x)$ *is monotonic decreases to the local minimum* x^* *along the line-segment* $[x,x^*]$ *such that*

$$\begin{cases} Md(x^*) = \{\, x \in D^n \,|\, f(x_1) \geq f(x_2)\,, \\ \qquad x_1 = (1-\theta_1)x + \theta_1 x^*,\ x_2 = (1-\theta_2)x + \theta_2 x^*, \\ \qquad 0 \leq \forall\theta_1 < \forall\theta_2 \leq 1\,\}, \end{cases} \tag{10}$$

then we call the region $Md(x^*)$ *the* line monotone decreasing region to the local minimum x^*.

Figure 3 shows an example of the *line monotone decreasing region* $Md(x_1^*)$ and another *line monotonic decreasing region* $Md(x_2^*)$. From **Figure 3**, note that two sets are not disjoint sets in general.

Definition 6.
The region of attraction $A(x^*)$ *of a local minimum* x^* *is the set of starting points such that strictly descent local minimal procedure:Lomin* $(x_s, f_s, f, D^n, \varepsilon_f, \varepsilon_x)$ *of which input parameters are the starting point* x_s*, its function value* f_s*, an objective function* f*, searching region* D^n *and tolerances of* f ε_f *and of* x ε_x *are applied, and then converge to the local minimum* x^* *as follows:*

$$A(x^*) = \{ x_s \in D^n \mid Lomin(x_s, f_s, f, D^n, \varepsilon_f, \varepsilon_x) \rightarrow x^*, \; x^* \in D^n \}. \tag{11}$$

Moreover, we define an adjoining relationship between two connected level sets.

Definition 7.
Let a level set $L(\alpha)$ *consist of connected components* $L_c^1(\alpha), L_c^2(\alpha), \ldots L_c^m(\alpha)$*. Moreover, if there exist two connected components* $L_c^{k1}(\alpha)$ *and* $L_c^{k2}(\alpha)$ $(k1 \neq k2)$ *and there exist two points such that* $x_1 \in L_c^{k1}(\alpha)$ *and* $x_2 \in L_c^{k2}(\alpha)$*, then* $\forall \lambda \in [0,1]$ *and the point* $x = (1-\lambda)x_1 + \lambda x_2$ *on the line segment between* x_1 *and* x_2 *is satisfied as follows:*

$$x \in L_c^{k1}(\alpha) \cup L_c^{k2}(\alpha) \cup (D^n \setminus L(\alpha)), \; x \notin L(\alpha) \setminus (L_c^{k1}(\alpha) \cup L_c^{k2}(\alpha)). \tag{12}$$

In case where, we call these two connected level set $L_c^{k1}(\alpha)$ *and* $L_c^{k2}(\alpha)$ *are* adjoint *on the line segment* $[x_1, x_2]$.

2.3 Relationships among three definitions of regions around each local minimum

Relationships among regions that have been defined in section 2.2. are described.

Proposition 1. *If* $x \in L_{qc}^s(\alpha^*; x^*)$ *and* $x^* \in X^*$*, then* $x \in Md(x^*)$ *(that is, the function value decreases with approach to the local minimum* x^* *on the line segment* $[x, x^*]$*).*

Proof) From **Definition 2** of a *strongly quasi-convex region*, for all $x (\neq x^*) \in L_{qc}^s(\alpha^*; x^*)$, we have

$$\begin{cases} x_1 = (1-\theta_1)x + \theta_1 x^*, \quad 0 < \forall \theta_1 < 1, \\ f(x_1) < \max\{ f(x), f(x^*) \} = f(x). \end{cases} \tag{13}$$

Since $L_{qc}^s(\alpha^*; x^*)$ is a convex set, $x_1 \in L_{qc}^s(\alpha^*; x^*)$. Therefore,

$$\begin{cases} x_2 = (1-\theta)x_1 + \theta x^*, \; 0 < \forall \theta < 1, \\ f(x_2) < \max\{ f(x_1), f(x^*) \} = f(x_1). \end{cases} \tag{14}$$

Let

$$\theta_2 = (1-\theta)\theta_1 + \theta > \theta_1. \tag{15}$$

Then we obtain the following equation by substituting equation (13) for x_1 of into equation (14)

$$\begin{aligned} x_2 &= (1-\theta)x_1 + \theta x^* \\ &= (1-\theta)\{(1-\theta_1)x + \theta_1 x^*\} + \theta x^* \\ &= (1-\theta)(1-\theta_1)x + \{(1-\theta)\theta_1 + \theta\}x^* \\ &= (1-\theta_2)x + \theta_2 x^*, \quad 0 < \forall\theta_1 < \forall\theta_2 < 1\,. \end{aligned} \tag{16}$$

From equations (13) to (16), for all $x(\neq x^*) \in L^s_{qc}(\alpha^*;x^*)$, we have

$$\begin{cases} f(x_1) > f(x_2)\,, \\ \qquad \text{where } x_1 = (1-\theta_1)x + \theta_1 x^*,\ \ x_2 = (1-\theta_2)x + \theta_2 x^*, \\ \qquad\qquad 0 < \forall\theta_1 < \forall\theta_2 < 1\,. \end{cases} \tag{17}$$

The above condition is included in definition (10) of the line monotone decreasing region, and thus $x \in Md(x^*)$. ■

From the above property and equation (10), the following relationship holds:

$$B(r^*;x^*) \subset L^s_{qc}(\alpha^*;x^*) \subset Md(x^*). \tag{18}$$

3 Main algorithm

3.1 Details of the algorithm

In this section, we show the main algorithm Mg for finding the global minimum x^{**} and the corresponding function value f^{**} for an objective function f over a searching region D^n. The steps of the algorithm are as follows.

Algorithm Mg $\quad (x^{**}, f^{**}) \leftarrow Mg\,(f, D^n, N, \gamma, Ns, Nc, h, r^{**}, \varepsilon_f, \varepsilon_x)\,;$

$\{*$ N is the number of samples, γ is ratio of random sampling, $Ns(<N)$ is the number of selecting candidates, $Nc(<Ns)$ is the number of concentrating candidates, $h>0$ is the step size, r^{**} is the radius of open ball, ε_x is tolerance for x-coordinates, and ε_f is tolerance for the function value. $*\}$

M1. [Initialize]

$\{*$ Initialize $f^{**}, \overline{X}^*, \overline{F}^*$, the set of candidate $X^{(0)}$, $F^{(0)}$, and iteration counter k. $*\}$

(S3.1) $\quad \overline{X}^* \leftarrow \emptyset;\ \overline{F}^* \leftarrow \emptyset;\ f^{**} \leftarrow \infty;\ X^{(0)} \leftarrow \emptyset;\ F^{(0)} \leftarrow \emptyset;\ k \leftarrow 0\,.$

M2. [Take sample points]

$\{*$ Sample uniformly N-points by applying procedure *U_Samples*, obtain the set of sample points $X^{(0)} = \{x_1, x_2, \ldots, x_N\}$ and set of these function values $F^{(0)} = \{f(x_1), f(x_2), \ldots, f(x_N)\}$. $*\}$

(S3.2) $\quad (X^{(0)}, F^{(0)}) \leftarrow U\_Samples(f, D^n, N, \gamma);$

M3. [Select candidates]

$\{*$ Sort sample points ascent order with respect to function values and select Ns-candidates with smallest function values in the sets $X^{(0)}$, $F^{(0)}$. $*\}$

(S3.3) $\quad X^{(1)} = \{x_i \in X^{(0)} \mid f(x_1) \leq \cdots \leq f(x_i) \leq \cdots \leq f(x_{Ns})\}\,;$

$\qquad\quad F^{(1)} = \{f(x_i) \in F^{(0)} \mid f(x_1) \leq \cdots \leq f(x_i) \leq \cdots \leq f(x_{Ns})\}\,.$

M4. [Concentrate candidates by line search]

$\{*$ Concentrate Nc-candidates $X_i^{(1)}$, $(i = Ns-Nc+1, Ns-Nc+2, \ldots, Ns)$ with Nc-largest function values around local minima using *line search* along steepest descent direction $-\nabla f(X_i^{(1)})$. $*\}$

(S3.4) **for** $i \leftarrow Ns-Nc+1$ **to** Ns **do**

$$X_i^{(1)} \leftarrow \underset{\lambda}{\operatorname{argmin}}\{\, f(X_i^{(1)} + \lambda(-\nabla f(X_i^{(1)}))\,\};$$

$$F_i^{(1)} \leftarrow \min_{\lambda}\{\, f(X_i^{(1)} + \lambda(-\nabla f(X_i^{(1)}))\,\};$$

od.

M5. [Apply local search procedure]

$\{*$ Increment iteration counter k by one, and set x_s, f_s as the smallest function value in the set $F^{(k)}$. $*\}$

(S3.5) $k \leftarrow k+1$.

(S3.6) $x_s \leftarrow \operatorname{argmin}\{\, f(x_i) \in F^{(k)}\,\}$; $f_s \leftarrow \min\{\, f(x_i) \in F^{(k)}\,\}$.

$\{*$ Find the local minimum x^* and the local minimal value f^* by applying the local search procedure *Lomin* from a starting point x_s. $*\}$

(S3.7) $(x^*, f^*) \leftarrow Lomin(x_s, f_s, f, D^n, \varepsilon_f, \varepsilon_x)$.

$\{*$ Remove starting point x_s and its function value f_s from the sets $X^{(k)}$ and $F^{(k)}$. $*\}$

(S3.8) $X^{(k+1)} \leftarrow X^{(k)} - \{x_s\}$; $F^{(k+1)} \leftarrow F^{(k)} - \{f_s\}$.

M6. [Check whether the current local minimum was firstly found]

$\{*$ Check the current local minimum x^* can be the same within the tolerance ε_x as one in the set of local minima $\overline{X}^{*(k)}$ ($isol$= $false$) or not ($isol = true$). $*\}$

(S3.9) $isol = true$;

for $i \leftarrow 1$ **to** $|\overline{X}^*|$ **do**

if $\|x^* - \overline{X}_i^*\| > \varepsilon_x$ **then** $isol = false$; **break**; **fi**.

od.

$\{*$ If the local minimum x^* is firstly found ($isol = true$), then add the point x^* and it's function value f^* to the related two sets $\overline{X}^*$ and $\overline{F}^*$. If the set of candidates $X^{(k+1)}$ is empty set, then terminate the algorithm. If the local minimum has been already found, then go to step **M5** $*\}$

(S3.10) **if** $isol = true$ **then**

$\overline{X}^* \leftarrow \overline{X}^* + \{x^*\}$; $F^* \leftarrow F^* + \{f^*\}$;

if $f^* > f^{**}$ **then** $x^{**} \leftarrow x^*$; **fi**;

fi.

(S3.11) **if** $X^{(k+1)} = \emptyset$ **then return**.

(S3.12) **if** $isol = false$ **then goto M5**.

M7. [Reduce candidates]

$\{*$ Remove candidates from $X^{(k+1)}$ that are expected to converge to the current local minimum x^* by applying procedure Rc. $*\}$

(S3.13) $(X^{(k+1)}, F^{(k+1)}) \leftarrow Rc\,(X^{(k+1)}, F^{(k+1)}, x^*, f^*, h, r^{**})$.

$\{*$ If the set of candidate points is an empty set, then stop the algorithm. Otherwise, go to **M5**. $*\}$

(S3.14) **if** $X^{(k+1)} = \emptyset$ **then return**.

(S3.15) **else** $X^{(k)} \leftarrow X^{(k+1)}$; $F^{(k)} \leftarrow F^{(k+1)}$; **goto M5**.

3.2 Termination property

Firstly, we show following the termination property of the algorithm Mg.

Proposition 2. *The algorithm* Mg *always stops at a finite number of iterations* $k \leq Ns$.

Proof) The step (S3.8) is always executed at every iteration k and the step (S3.13) is executed in the case where $isol = true$. $|X^{(k+1)}| = |X^{(k)}| - 1$ *holds at (S3.8) and the number of obtained candidates* $|X^{(k+1)}|$ *by the algorithm* Rc *is smaller than or equal to* $|X^{(k+1)}|$ *at step (S3.8). Therefore,* $|X^{(k+1)}| \leq |X^{(k)}| - 1$ *is satisfied for every iteration.*

Since the initial number of elements of the candidates set $|X^{(1)}|$ *is* Ns*, there exists some values of* k *such that* $k \leq Ns$ *and* $X^{(k)} = \emptyset$*. Thus, from the stop condition, the algorithm* Mg *terminates at a finite number of iterations.* ■

4 Taboo step for avoiding multiple convergence to one of already known local optima

In this section, we introduce the taboo step **M7** of our algorithm [5, 1995] that reduces the candidate points, and we show the differences between Ursem's [13, 1999] *hill-valley* step and our taboo step.

4.1 Details of the taboo step for reducing candidates

At step **M7** of the main algorithm in section 3, the number of candidates is reduced. This step plays an important role in that it removes candidates that can converge to the current local minimum. By this step, the point with the smallest function value in remaining candidates becomes an effective starting point for the local search procedure *Lomin* in the next iteration. If the main algorithm does not use this step, e.g., *the multistart method*, the algorithm becomes very inefficient, since many of the starting points for the local search converge to one of the already found local minima.

The algorithm Rc removes the j-th point $X_j^{(k)}$ from the set $X^{(k+1)}$ such that the function f can monotonically decrease to the open ball $B(r^{**};x^*)$ along the line segment $[X_j^{(k+1)}, x^*]$. As a result, the set of candidates $X^{(k+1)}$ and the set of function values $F^{(k+1)}$ are obtained. The steps of the algorithm are as follows:

$$Algorithm\ Rc \quad (F^{(k+1)}, X^{(k+1)}) \leftarrow Rc(F^{(k+1)}, X^{(k+1)}, x^*, f^*, h, r^{**});$$

R1. [Loop for reducing candidates]

(S4.1) **for** $j \leftarrow 1$ **to** $|X^{(k+1)}|$ **do**

R2. [Check the function monotonic decrease with approaching the open ball.]

(S4.2) {∗ Check whether the function f monotonic decrease with approaching the open ball $B(r^{**},x^*)$ on the line segment $[X_j^{(k+1)},x^*]$($m\_decrs = true$) or not($m\_decrs = false$) . ∗}

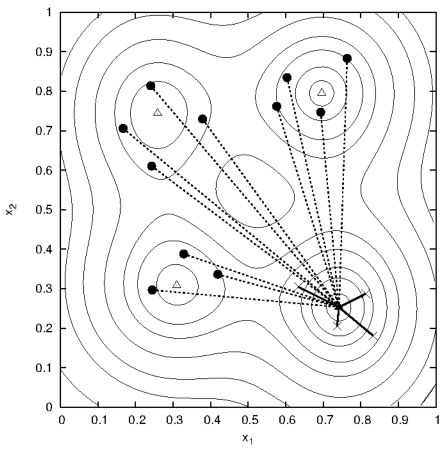
(S4.2.1) $f^{(0)} \leftarrow F_j^{(k+1)}$; $x^{(0)} \leftarrow X_j^{(k+1)}$;

(S4.2.2) $e \leftarrow (x^* - x^{(0)}) / \| x^* - x^{(0)} \|$;
(S4.2.3) $m\_decrs \leftarrow true$; $i \leftarrow 1$;
(S4.2.4) **while** $x^{(i-1)} \notin B(r^{**}, x^*)$ **and** $ih / \|x^* - x^{(0)}\| \leq 1$ **do**
(S4.2.5) $x^{(i)} \leftarrow x^{(0)} + ihe$; $f^{(i)} \leftarrow f(x^{(i)})$;
(S4.2.6) **if** $f^{(i)} \leq f^{(i-1)}$ **then** $i \leftarrow i+1$;
(S4.2.7) **else** $m\_decrs \leftarrow false$; **break**; **fi**.
od.

R3. [Remove the candidate such that $m\_decrs = true$ **.]**

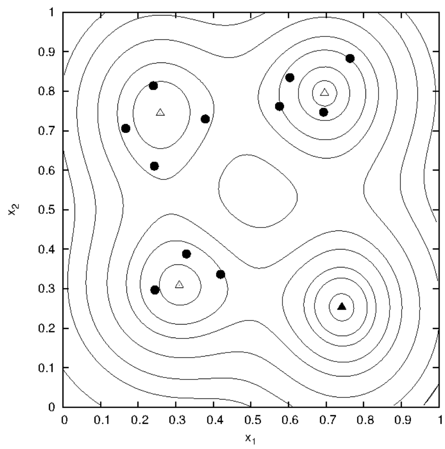
(S4.3) {$*$ If $m\_decrs = true$, then remove the candidate $X_j^{(k+1)}$ and it's function value $F_j^{(k+1)}$ from $X^{(k+1)}$ and $F^{(k+1)}$. $*$}
(S4.3.1) **if** $m\_decrs = true$ **then**
(S4.3.2) $X^{(k+1)} \leftarrow X^{(k+1)} - \{X_j^{(k+1)}\}$; $F^{(k+1)} \leftarrow F^{(k+1)} - \{F_j^{(k+1)}\}$; **fi**.
od.

The actions of the algorithm *Rc* are shown in Figure 4.1 and Figure 5.



●: a remaining candidate
×: a removing candidate

Figure.4 Algorithm *Rc* being operated



●: a remained candidate

Figure.5 Algorithm *Rc* is finished

Proposition 1.
If there exist $j1$ *such that* $X_{j1}^{(k+1)} \in Md(x^*) \cup B(r^{**}, x^*)$*, then* $m\_decrs = true$ *at step (S4.3.1).*

Proof) In the case where $x^{(0)} \equiv X_{j1}^{(k+1)} \in B(r^{**}, x^*)$, then $m\_decrs = true$ at step (S4.3) from steps (S4.2.3) and (S4.2.4). In the case where $x^{(0)} \equiv X_{j1}^{(k+1)} \notin B(r^{**}, x^*)$, for $x^{(i-1)}$

such that $x^{(i-1)} \notin B(r^{**}, x^*)$, from step (S4.2.5), the equations:

$$\begin{aligned} x^{(i-1)} &= x^{(0)} + (i-1)he \\ &= x^{(0)} + (i-1)h\frac{x^* - x^{(0)}}{\|x^* - x^{(0)}\|} \\ &= \left(1 - \frac{(i-1)h}{\|x^* - x^{(0)}\|}\right)x^{(0)} + \frac{(i-1)h}{\|x^* - x^{(0)}\|}x^* \end{aligned} \tag{19}$$

$$\begin{aligned} x^{(i)} &= x^{(0)} + ihe \\ &= x^{(0)} + ih\frac{x^* - x^{(0)}}{\|x^* - x^{(0)}\|} \\ &= \left(1 - \frac{ih}{\|x^* - x^{(0)}\|}\right)x^{(0)} + \frac{ih}{\|x^* - x^{(0)}\|}x^* \end{aligned} \tag{20}$$

hold. Here, let $\theta_1 \equiv (i-1)h/\|x^* - x^{(0)}\|$ and $\theta_2 \equiv ih/\|x^* - x^{(0)}\|$, $i \geq 1$ by steps (S4.2.3) and (S4.2.6) and $ih/\|x^* - x^{(0)}\| \leq 1$ at step (S4.2.4), thus $0 \leq \theta_1 < \theta_2 \leq 1$. Since $x^{(0)} \in Md(x^*)$, $f(x^{(i-1)}) < f(x^{(i)})$ by **Definition 5**, that is, $f^{(i)} < f^{(i-1)}$, step (S4.2.5) is always executed. Therefore, $m\_decrs = true$. ■

From this proposition, if $m\_decrs = true$, then (S4.3.2) is executed and $X_{j1}^{(k+1)} \notin X^{(k+1)}$ at the end of (S4.3).

Proposition 2.
If there exist j2 and connected components $L_c^1(F_{j2}^{(k+1)}), L_c^2(F_{j2}^{(k+1)}), \ldots, L_c^{\bar{l}}(F_{j2}^{(k+1)})$ $\left(= L_c(F_{j2}^{(k+1)}; x^*)\right), \ldots, L_c^m(F_{j2}^{(k+1)})$ *of* $L(F_{j2}^{(k+1)})$ *such that*

$$\begin{cases} d\left(L_c(F_{j2}^{(k+1)}; x^*), L_c^l(F_{j2}^{(k+1)})\right) > h, & (l \neq \bar{l},\ l = 1, \ldots, m), \\ d\left(B(r^{**}; x^*), L_c^l(F_{j2}^{(k+1)})\right) > h, & (l \neq \bar{l},\ l = 1, \ldots, m), \end{cases} \tag{21}$$

then $m\_decrs = false$ *at step (S4.3.1) of Algorithm Rc.*

Proof) At step (S4.2.1), we set $x^{(0)} \equiv X_{j2}^{(k+1)}$ and $f^{(0)} \equiv F_{j2}^{(k+1)}$. Then $L(f^{(0)})$ can be represented by m-number of connected components as follows:

$$L(f^{(0)}) = L_c^1(f^{(0)}) \cup L_c^2(f^{(0)}) \cup \cdots \cup L_c^m(f^{(0)}). \tag{22}$$

If $m\_decrs = true$, then

$$x^{(i)} \in L(f^{(0)}) = L_c^1(f^{(0)}) \cup L_c^2(f^{(0)}) \cup \cdots \cup L_c^m(f^{(0)}), \qquad (i = 0, 1, \ldots) \tag{23}$$

because $f^{(0)} \geq f^{(1)} \geq f^{(2)} \geq \cdots$ while the condition of step (S4.2.5) is *true*. Let a point that is included in $L_c(f^{(0)}; x^*)$ or $B(r^{**}, x^*)$ with smallest i be $x^{(l)}$, then $x^{(l)} \in L_c(f^{(0)}; x^*)$ or $x^{(l)} \in B(r^{**}, x^*)$, and $x^{(l-1)} \notin L_c(f^{(0)}; x^*)$ or $x^{(l-1)} \notin B(r^{**}, x^*)$. Moreover, there exist $L_c^{l3}(f^{(0)})$ such that $x^{(l-1)} \in L_c^{l3}(f^{(0)})$ and $L_c^{l3}(f^{(0)}) \subset L(f^{(0)}) \setminus L_c(f^{(0)}; x^*)$ or $L_c^{l3}(f^{(0)}) \subset L(f^{(0)}) \setminus B(r^{**}, x^*)$. From $\|x^{(l)} - x^{(l-1)}\| = h$,

$$\begin{cases} d\left(L_c(f^{(0)}; x^*), L_c^{l3}(f^{(0)})\right) \leq h & \text{or} \\ d\left(B(r^{**}, x^*), L_c^{l3}(f^{(0)})\right) \leq h \end{cases} \tag{24}$$

holds, and this contradicts equation (22). Thus, $m\_decrs = false$. ■

From this proposition, if $m\_decrs = false$, then no (S4.3.2) is executed and $X_{j1}^{(k+1)} \in X^{(k+1)}$ at end of (S4.3).

4.2 Differences between Ursem's *hill-valley* step and our taboo step

Ursem's hill-valley step and our taboo step are the same for searching on a line segment consisting of two different points.

However, our step is different from Ursem's step as follows.

(1). Ursem's step allows all two different point in population, but one of the two points is always a local optimum x^*.

(2). Step h and inner point $\bar{x}$ is determined the base ratio $\gamma = 0.25$ of distance between two different points x_i and x_j as follows

$$\begin{cases} h = \gamma \|x_i - x_j\|, & 1 \le i, j(i \ne j) \le N_p \\ \bar{x} = x_i + k \cdot h(x_i - x_j), & \gamma = 0.25,\ k = 1,2,3, \end{cases} \tag{25}$$

but our step length h is always constant.

From the above two different matters and section **4.2**, the convergence property cannot hold at Ursem's step.

5 Conditions for finding the global minimum

We show the conditions for finding the global minimum.

Theorem 1.
If in the main algorithm Mg, the following two conditions are satisfied, then the global minimum x^{**} can always be found without duplication.
(C1) *There exists $i1$ such that $X_{i1}^{(1)} = \mathrm{argmin}\{ f(x) \,|\, x_i \in L_c(F_{N_s}^{(1)}; x^{**}) \subset Md(x^*)\}$ at step (S3.5).*
(C2) *There exists an iteration $k1$, and the following sub-conditions hold.*
(C2-1) *For all $k < k1$ at (S3.13), The conditions of* **Proposition 2** *hold at point $X_{i1}^{(1)}$ and its function value $F_{i1}^{(1)}$ in procedure Rc. That is, there exists $j2$ such that $X_{j2}^{(k+1)} = X_{i1}^{(1)}$ and $F_{j2}^{(k+1)} = F_{i1}^{(1)}$, and there exists connected components $L_c^1(F_{j2}^{(k+1)})$, $L_c^2(F_{j2}^{(k+1)})$, $\ldots, L_c^{\bar{l}}(F_{j2}^{(k+1)})$ $\left(= L_c(F_{j2}^{(k+1)}; x^*)\right), \ldots, L_c^m(F_{j2}^{(k+1)})$ of $L(F_{j2}^{(k+1)})$ such that*

$$\begin{cases} d\left(L_c(F_{j2}^{(k+1)}; x^*), L_c^l(F_{j2}^{(k+1)})\right) > h, & (l \ne \bar{l},\ l = 1, \ldots, m), \\ d\left(B(r^{**}; x^*), L_c^l(F_{j2}^{(k+1)})\right) > h, & (l \ne \bar{l},\ l = 1, \ldots, m). \end{cases} \tag{26}$$

(C2-2) *If there exists $X_{i1}^{(1)}$ in the iteration $k1$, then $X_{i1}^{(1)} = x_s = \mathrm{argmin}\{ f(x_i) \in F^{(k1)}\}$, $F_{i1}^{(1)} = f_s = \min\{ f(x_i) \in F^{(k1)}\}$ holds at (S3.6) of Step M5. Moreover, if $(x^*, f^*) \leftarrow Lomin(x_s, f_s, f, D^n, \varepsilon_f, \varepsilon_x)$, then $x^* = x^{**}$ at (S3.7).*

Proof) If sub-condition **(C2-1)** holds, then $m\_decrs = false$ from **Proposition 2**. Thus, in iteration 1 to $k1-1$, $X_{j2}^{(k+1)} = X_{i1}^{(1)} \in X^{(k+1)}$ at the end of (S4.3), that is, the end of Rc . Then, at (S3.4) in the next iteration $k1$, if the sub-condition **(C2-2)** holds, the global minimum x^{**} can be found. ∎

References

[1] Bazaraa, M. S., Sherali, H. D. and Shetty, C, M., *Nonlinear Programming —Theory and Algorithms (3rd ed.)*, John Wiley & Sons, 2006.

[2] Boender, C.G.E., Rinnooy Kan, A.H.G. and Timmer G.T. A Stochastic Method for Global Optimization, *Math. Program.* 22 (1982) 125–140.

[3] Dzemyda, G., Šaltenis, V. and Žilinskas, A., *Stochastic and Global Optimization,* Kluwer Acad. Publ., Dordrecht, 2002.

[4] Horst, R., Tuy, H. , *Global Optimization, Deterministic Approaches,* Springer-Verlag, Berlin, 1996.

[5] Kanemitsu, H., Murai, T. and Shimbo, M.: "A global minimization algorithm by removing points converging on the already known local minima," *Proc. of the International Symposium on Operations Research with Applications in Engineering, Technology and Management(ISORA'95)*, World Pub. Co., Beijing(China), (1995) 1–9.

[6] Kanemitsu, H., Miyakosi, M. and Shimbo, M.: "Properties of Unimodal and Multimodal Functions Defined by Local Minimal Values Set," *Electronics and Communications*, Part3, Vol.81, No.1, (1998) 42–51.

[7] Rinnooy Kan, A.H.G. and Timmer, G.T. , Stochastic Global Optimization Methods. Part I: Clustering Methods, *Math. Program.* 39 (1987) 27–56.

[8] Rinnooy Kan, A.H.G. and Timmer, G.T. , Stochastic Global Optimization Methods. Part II: Multi Level Methods, *Math. Program.* 39 (1987) 57–78.

[9] Schöen, F. , Stochastic Techniques for Global Optimization: A Survey of Recent Advances, *J. of Global Optim.* 1 (1990) 207–228.

[10] Snyman, J.A. and Fatti, L.P. , "A multi-start global minimization algorithm with dynamic search trajectories," *J. of Optim. Theory Appl.*, 54 (1987) 121-141.

[11] Timmer, G.T. , *Global optimization: A stochastic approach,* Ph.D. Dissertation, Econometric Institute, Erasmus University, Rotterdam, Holland, (1984).

[12] Törn, A.A., Antanas A. Žilinskas , *Global Optimization, Lecture Notes in Computer Science 350*, Springer-Verlag, Berlin, 1989.

[13] Ursem, R.K.: "Multinational evolutionary algorithms," *Proc. Congress of Evolutionary Computation (CEC-99)*, Vol.3, Washington, DC.(USA), pp.1633-1644, 1999.