Proceedings of the World Congress on Engineering 2007 Vol II
WCE 2007, July 2 - 4, 2007, London, U.K.

A Multidimensional Bisection Method for Minimizing Function over Simplex

A.N. Baushev, E.Y. Morozova

***Abstract*—A new method for minimization problem over simplex, as a generalization of a well-known in one-dimensional optimization bisection method is proposed. The convergence of the method for class of strictly unimodal functions including class of strictly convex functions is proved. The computational results are presented for a set of test problems.**

***Index Terms*— convex set, *n*-dimensional simplex, strictly unimodal function, direct search methods, nonlinear programming, nonlinear optimization.**

I. Introduction

The problem considered is

$$f(x) \to \min, \quad x \in S, \tag{1}$$

where S - *a n-dimensional* simplex in R^n , and f – a continuous function.

For the case $n=1$ one of well-known methods for solving problem (1) is the bisection method which convergence is proved for the case when f is continuous function with the single point of a local minimum (which solves the problem (1)) over set *S*. In this paper we generalize a bisection method for case *n*>1 and demonstrate convergence of our algorithm for the class of strictly unimodal functions.

Definition. Let *D* be a bounded closed convex set in R^n. Function $f: D \to R$ is *strictly unimodal over set D* iff for any segment $\Delta \subset D$ $\# Arg\min\{f(x)|x \in \Delta\}=1$ where we write «$\# A$» for the cardinality of a set *A*.

Notice that a function is strictly unimodal iff for any closed convex subset $D' \subset D$ $\# Arg\min\{f(x)|x \in D'\}=1$. Notice that the class of strictly unimodal functions contains the class of strictly convex functions (over set *D*).

In point 2 we describe some properties of strictly unimodal functions, and also the general structure of the decomposition method. In point 3 we describe the basic algorithm and prove a convergence of this algorithm for the case when f is a continuous strictly unimodal function. In point 4 we illustrate the algorithm by the well-known example of Dennis-Wood test function.

II. PRELIMINARITIES

Let D be a closed bounded set in R^n with nonempty interior, f be a continuous function on *D*, e be some unit vector. Consider for each real t the hyperplane

$$\gamma_t = \left\{x \in R^n \mid \langle e, x\rangle = t\right\}. \tag{2}$$

The family $\{\gamma_t \mid -\infty < t < \infty\}$ covers R^n and whereas *D* is a closed bounded set then there are such values $t_{\min}$ and $t_{\max}$ that for each $t \in [t_{\min}, t_{\max}]$ we have $D_t = D \cap \gamma_t \neq \varnothing$, and for each $t \notin [t_{\min}, t_{\max}]$ - $D_t = \varnothing$. Thus

$$D = \bigcup_{t \in [t_{\min}, t_{\max}]} D_t . \tag{3}$$

The set D_t is *a section* of the set *D* by hyperplane γ_t. Whereas the set D_t lies in γ_t, it is possible to consider D_t as a set in the space of dimensionality $n-1$. Consider the function

$$F(t) = \min\left\{f(x) \mid x \in D_t\right\}. \tag{4}$$

Let

$$\omega_f(\delta) = \sup\{|f(x) - f(y)| : x, y \in D, \|x-y\| < \delta\}$$

be the uniform modulus of continuity of the function f over set *D*,

$$\omega_F(\delta) = \sup\{|F(t) - F(s)| : t, s \in [t_{\min}, t_{\max}], |t-s| < \delta\}.$$

Proposition 1.

1. $F(t)$ is a continuous function on the segment $[t_{\min}, t_{\max}]$.
2. We have the equality
$$\min\left\{f(x) \mid x \in D\right\} = \min\left\{F(t) \mid t \in [t_{\min}, t_{\max}]\right\} \tag{5}$$
3. For any positive δ $\omega_F(\delta) \le \omega_f(\delta)$.

Preposition 2.

1. If f is a convex function on D then the function *F* is convex on the segment $[t_{\min}, t_{\max}]$.
2. If f is a strictly convex function on *D* then the function F is strictly convex on the segment $[t_{\min}, t_{\max}]$.
3. If f is a strictly unimodal function on *D* then the function F is strictly unimodal on the segment $[t_{\min}, t_{\max}]$.

The main idea of our algorithm is motivated by the possibility of using the recursive procedure for the realization of the one-dimensional bisection algorithm.

The one-dimensional bisection algorithm solves the problem

$$f(x) \to \min, \quad x \in [a, b], \tag{6}$$

where f is a strictly unimodal function over segment $[a, b]$.

Let $bis(f, a, b, \varepsilon)$ denote the recursive one-dimensional bisection procedure. The inputs for this procedure are: the

A.N. Baushev is with the Applied Mathematics Department, Saint-Petersburg State Transport University, Moskovsky prospect, 9, Saint-Petersburg, 190031, Russia (e-mail: baushev@AB7310.spb.edu)

E.Y. Morozova is with the Applied Mathematics Department, Herzen State Pedagogical University, nab. Moyki, 48, Saint-Petersburg, 191186, Russia (e-mail: melena65@mail.ru)

procedure for calculation values of f, the segment $[a,b]$ and the accuracy ε. The outputs are the estimations x_m for the minimizer x^* and f_m for the value of the minimum of the function f over the segment $[a,b]$.

The iteration of the recursive procedure includes the following steps.

Step 0. If $b-a \ge \varepsilon$, go to step 1, otherwise stop.

Step 1.

$c=\dfrac{a+b}{2}$ $a'=\dfrac{a+c}{2}$, $b'=\dfrac{b+c}{2}$, $f(c)$, $f(a')$, $f(b')$.

Step 2.

If $f(a') \le f(c) \le f(b')$, set $b=b'$.

If $f(a') \ge f(c) \ge f(b')$, set $a=a'$.

If $f(c) \le \min\{f(a'), f(b')\}$, set $a=a'$, $b=b'$.

Step 3. Execute $bis(f,a,b,\varepsilon)$ with new inputs.

III. THE ALGORITHM

Let S be a n-dimensional simplex. Let fix the vertex V^0 and denote by $V^1,...,V^n$ the opposite vertices. Set for each $t \in [0,1]$

$$S_t = conv\left\{V^0 + t(V^1 - V^0),...,V^0 + t(V^n - V^0)\right\}. \qquad (7)$$

The set S_t is the $n-1$-dimensional simplex for $0 < t \le 1$.

Set $x_t \equiv \arg\min\{f(x) \mid x \in S_t\}$. Each simplex S_t for $0 < t < 1$ part the initial simplex S in two sets: $conv\{V^0, S_t\}$ and $conv\{S_1, S_t\}$. Fig. 1 illustrate an example of the partition of the simplex S for the case $n=3$.

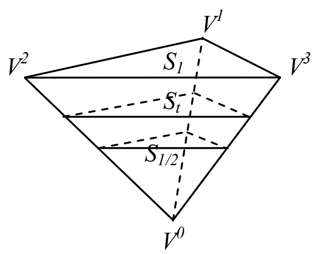


Fig. 1. The partition of the simplex *S.*

Let $bissimpl(f, SS_1, SS_2, d, \varepsilon)$ denote the recursive procedure in case $n > 1$. The inputs for this procedure are: the procedure for calculation values of *f,* boundary simplices SS_1 and SS_2, the current dimension d and the accuracy ε. The outputs are the estimations x_m for the minimizer x^* and f_m for the value of the minimum of the function f over the set $conv\{SS_1, SS_2\}$. Originally d is equal n, then this parameter varies depending on the dimension of the simplex where the point of a minimum is searched. Actually this parameter at first decreases to value 1, and then increases to value $d=n$. Three circles of such calculations we consider as the iteration with number k. Denote by f^k the estimation of a minimum of the function f and by x^k the estimation of the point x^*. The parameter d and the outputs must be declared as global variables and its initial values must be defined before starting procedure $bissimpl(f, SS_1, SS_2, d, \varepsilon)$. More concretely the preliminary step includes the following destinations:

$SS_1 = S_0 = V^0$, $SS_2 = S_1 = conv\left\{V^1, \ldots V^n\right\}$ according to (7), $d=n$;

x^0, x^1, f^0, f^1 we define in a such way that the condition

$$\max\left\{\left|f^k - f^{k-1}\right|, \left\|x^k - x^{k-1}\right\|\right\} < \varepsilon \qquad (8)$$

be failed.

Step 1.

If (8) is hold, stop. Otherwise set $\sigma_1 = SS_1$, $\sigma_2 = SS_2$ and go to step 2.

Step 2.

If d=1, execute $bis(f,a,b,\varepsilon)$ with $a= SS_1$, $b= SS_2$ Otherwise, go to step 3.

Step 3.

Two cases are possible.

1) SS_1 and SS_2 are d-dimensional simplices. Let $V^0_{SS_1}, V^1_{SS_1},..., V^d_{SS_1}$ and $V^0_{SS_2}, V^1_{SS_2},..., V^d_{SS_2}$ be vertices of simplices SS_1 and SS_2 accordingly. Then we define $S_{\frac{1}{2}}$ $S_{\frac{1}{4}}$ $S_{\frac{3}{4}}$ by

$$\begin{aligned} S_t = conv\Big\{&V^0_{SS_1} + t\left(V^0_{SS_2} - V^0_{SS_1}\right), V^1_{SS_1} + t\left(V^1_{SS_2} - V^1_{SS_1}\right), \\ &..., V^d_{SS_1} + t\left(V^d_{SS_2} - V^d_{SS_1}\right)\Big\}. \end{aligned} \qquad (9)$$

2) One of the sets SS_1, SS_2 is a vertex, another is d-dimensional simplex. In this case $S_{\frac{1}{2}}$ $S_{\frac{1}{4}}$ $S_{\frac{3}{4}}$ are defined by (7). Set $d = d-1$.

Step 4.

For each of simplices $S_{\frac{1}{2}}$ $S_{\frac{1}{4}}$ $S_{\frac{3}{4}}$ the following actions must be done:

1) Fix a vertex V^0 in the simplex S_t and let $V^1,...,V^n$ be an opposite vertices.

2) Execute $bissimpl(f, SS_1, SS_2, d, \varepsilon)$ with new values $SS_1 = V^0$ and $SS_2 = conv\left\{V^1,...,V^n\right\}$.

Step 5.

Let x^1_m x^2_m, x^3_m and f^1_m f^2_m, f^3_m be results of the previous step (for $S_{\frac{1}{2}}$ $S_{\frac{1}{4}}$ $S_{\frac{3}{4}}$ accordingly).

If $f^2_m \le f^1_m \le f^3_m$, set $SS_1 = \sigma_1$ $SS_2 = S_{\frac{3}{4}}$.

If $f^2_m \ge f^1_m \ge f^3_m$, set $SS_1 = S_{\frac{1}{4}}$ $SS_2 = \sigma_2$.

If $f^1_m \le \min\left\{f^2_m, f^3_m\right\}$, set $SS_1 = S_{\frac{1}{4}}$, $SS_2 = S_{\frac{3}{4}}$.

Set $d = d+1$.

Step 6. Execute $bissimpl(f, SS_1, SS_2, d, \varepsilon)$ with current inputs.

Theorem. Let $x(\varepsilon)$ be the final estimation of the minimizer x^* for the function f where f is a continuous strictly unimodal function over *n-dimensional* simplex S then $\lim\limits_{\varepsilon \to 0} x(\varepsilon) = x^*$.

Sketch of proof. Let $\{\varepsilon_k\} \downarrow 0$ is chosen arbitrary, $x_k \equiv x(\varepsilon_k)$, $k = 1,2,...$ Set $\{f_k\}$ corresponding sequence of estimations of the minimum f over S. The condition

$$\max\{|f_k - f_{k-1}|, \|x_k - x_{k-1}\|\} < \varepsilon_k$$

imply the existence of the limits $\lim_{k\to\infty} x_k = x'$, $\lim_{k\to\infty} f_k = f'$. One can show that

$$x' = x^*, \ f' = f^*. \tag{10}$$

We shall use the induction on $n = \dim(S)$. For $n = 1$ our algorithm coincides with the well-known one-dimensional bisection method. So the assertion of the theorem is valid.

We shall assume the assertion of the theorem is valid for the dimensionalities $\le n-1$ and show that it is valid for the dimensionality *n*.

Assume that the relations (10) are failed. According to the induction assumption the points x' and x^* must belong to different simplices, i.e. $x' \in S_{t'}$, $x^* \in S_{t^*}$, $t' \neq t^*$. For the definiteness let $t' < t^*$. From the strictly unimodality of f follows the relation $\{x_k\} \subset conv\{S_{t'}, S_0\}$ that contradicts to the inequalities of step 5 of the algorithm (see previous point). So, the equations (10) and the theorem are valid.

Remark. If the function f satisfies the Lipschitz condition with exponent $\alpha > 0$ and with constant C that is

$$|f(x) - f(y)| \le C\|x - y\|^{\alpha} \tag{11}$$

for all $x, y \in S$, one can change the stopping rule (8) by the condition

$$\|x^k - x^{k-1}\| \le \tilde{\varepsilon}(\alpha, C). \tag{12}$$

In this case it is not difficult to see that the number of iterations $N_n(\tilde{\varepsilon})$ needed to hold (12) is about $(N_1(\tilde{\varepsilon}))^n$. So, the working time of the algorithm increases exponentially with the dimensionality of the problem (1).

IV. THE EXAMPLE

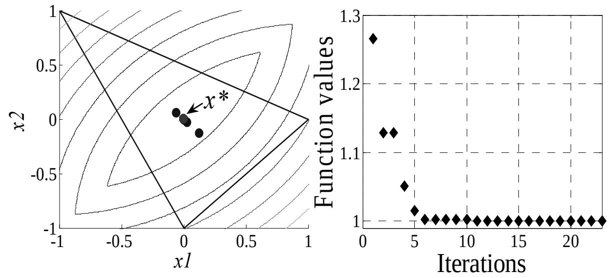
Consider the following variant of Dennis-Wood function [1]:

$$f(x) = \frac{1}{2}\max\{\|x - c_1\|^2, \|x - c_2\|^2\}, \tag{13}$$

where $c_1 = (1,-1)$ $c_2 = -c_1$. This function is continuous and strictly convex, but its gradient is discontinuous everywhere on the line $x_1 = x_2$.

In fact, numerical tests show [2] that compass search applied to (13) frequently converge to a point of the form (a,a), with $a \neq 0$. This sort of failure was observed in [3] for the multidirectional search algorithm. Besides in [1] it is shown that a modification of the Nelder-Mead algorithm [4] can fail to converge to the minimizer of the function (13).

The level sets of this function are shown in fig. 2 (a). The sequence of the points x^k generated by our algorithm converge to the point $x^*(0,0)$ is marked. The simplex $S = conv\{(-1,1),(0,-1),(1,0)\}$ was chosen as an initial simplex. The accuracy $\varepsilon = 10^{-6}$ was achieved after 23 iterations (fig. 2 (b)).



(a) (b)

Fig. 2. (a) The level sets of the function (13), the initial simplex and the sequence $\{x^k\}$; (b) the sequence $\{f^k\}$.

V. CONCLUSION

We have exposed our algorithm for the class of strictly unimodal functions only. However one can show that the algorithm can be applied for a wider class of functions, namely, we consider the class of functions Φ_S, where S - the *n*-dimensional simplex, defined as follows: $f \in \Phi_S$ iff for any segment $\Delta \subseteq S$ each local minimum of f over this segment is also a global minimum of the function f over this segment. The class Φ_S contains a subclass of strictly unimodal functions over set *S.*

Notice also this algorithm can serve as a basis element for solving the unconstrained minimization problem for the function $f \in \Phi_{R^n}$.

We realized our algorithm in MatLab. The program was tested for different examples of minimization of nonsmooth functions, some of them can be found in [5].

REFERENCES

[1] J. E. Dennis, Jr. Woods and J. Daniel, "Optimization on microcomputers: The Nelder-Mead simplex algorithm", *in New Computing Environments: Microcomputers in Large-Scale Computing*, A. Wouk, ed., SIAM.-Philadelphia, 1987, pp. 116-122.

[2] T. J. Kolda, R. M. Lewis, V. Torczon, "Optimization by Direct Search: New Perspectives on Some Classical and Modern Methods", *SIAM Review*, 2003, vol. 45, No. 3, pp. 385-482.

[3] V. Torczon, "On the convergence of the multidirectional search algorithm", *SIAM J. Optim.*, 1991, vol. 1, pp. 123-145.

[4] J.A. Nelder and R. Mead, "A simplex method for function minimization", *Computer Journal*, 1965, 7, pp. 308-313.

[5] E.Y. Morozova, "A direct search recursive algorithm for minimization of function of several variables", *The Review of applied and industrial mathematics*, 2006, vol. 13, No. 5, pp. 783-796. (In Russian).