Journal of Global Optimization **10:** 257–281, 1997.

A Global Optimization Algorithm for Multivariate Functions with Lipschitzian First Derivatives*

VICTOR P. GERGEL
Software Department, University of Nizhni Novgorod, Gagarin Ave., 23, Nizhni Novgorod, 603600, Russia

(Received: 16 March 1995; accepted: 1 June 1996)

Abstract. In this paper we propose a new multi-dimensional method to solve unconstrained global optimization problems with Lipschitzian first derivatives. The method is based on a partition scheme that subdivides the search domain into a set of hypercubes in the course of optimization. This partitioning is regulated by the decision rule that provides evaluation of the "importance" of each generated hypercube and selection of some partition element to perform the next iteration. Sufficient conditions of global convergence for the new method are investigated. Results of numerical experiments are also presented.

Key words: Global optimization, multiextremal algorithms, Lipschitzian first derivatives, convergence, numerical experiments.

1. Introduction

In this paper we consider the problem of finding the global (or absolute) minimizer for a multivariate function $f(y)$, i.e.,

$$min f(y),\ y \in D, \tag{1.1}$$

where the domain of search D is a hypercube,

$$D = \{y \in R^N : a_i \leqslant y_i \leqslant b_i, 1 \leqslant i \leqslant N\}, \tag{1.2}$$

R^N is the N-dimensional Euclidean space and the objective function $f(y)$ to be minimized may be multiextremal. These problems are of substantial interest – see, for example, Archetti and Schoen (1984), Dixon and Szegö (1978), Horst and Pardalos (1995), Horst and Tuy (1990), Pardalos and Rosen (1990), Pintér (1996), Rinnooy Kan and Timmer (1989), Strongin (1978), Törn and Žilinskas (1989).

Generally any estimates of the global minimizer of the function $f(y)$ are based on some assumptions about the properties of the function behaviour. One of the fruitful approaches is using the idea that the function $f(y)$ satisfies the Lipschitz condition

$$|f(y_1) - f(y_2)| \leqslant K||y_1 - y_2|| \text{ for any } y_1, y_2 \in R^N, \tag{1.3}$$

* This work was supported by RFBR Grant N 95-01-01073.

where K is a constant. In the framework of this approach a number of well-known global optimization methods are proposed (see, for example, Evtushenko (1971), Hansen *et al.* (1992), Piyavskii (1972), Shubert (1972), Strongin (1978)).

In recent years the above mentioned approach has been extended to take into account some additional information about the functions to be minimized. For instance, Hansen *et al.* (1989), Gergel (1992), Breiman and Cutler (1993), Baritompa (1994) considered the functions with Lipschitzian first derivatives, i.e. satisfying the condition

$$|f'(y_1) - f'(y_2)| \leqslant L||y_1 - y_2|| \text{ for any } y_1, y_2 \in R^N, \tag{1.4}$$

where $f'(y)$ is the first partial derivative of $f(y)$ with respect to the direction y_1y_2. Choosing any iteration point in the optimization process, these methods use the values of the first derivatives of the objective function in addition to the function values. This additional search information is used to increase the efficiency of finding the global minimizer.

In this paper we discuss a new method for minimizing multiextremal functions for which the first derivatives satisfy the Lipschitz condition (1.4). This method generalizes the univariate global optimization technique given in Gergel (1992) to a multi-dimensional case. The extension of the original algorithm is based on the idea of reducing the initial multi-dimensional optimization problems to equivalent one-dimensional problems using the scheme given in Pintér (1986) (some alternative schemes for such reduction are reviewed, for example, in Butz (1968), Evtushenko and Potapov (1994), Meewella and Mayne (1989), Strongin (1978)).

Optimization problems considered in this paper are rather specific. In contrast with various known methods (see, e.g., Breiman and Cutler (1993), Baritompa (1994)), we assume that the value of the Lipschitz constant may be unknown *a priori*. Similar problems (in the case when the condition (1.3) is used instead of (1.4)) have been investigated in the framework of the information approach to global optimization (see Strongin (1978)). Such problems are also studied in Sergeyev (1995) where the major attention was dedicated to estimating local Lipschitz constants.

If the problem belongs to this type then any guaranteed bounds for the function to be minimized can not be determined. As a result, the general approaches developed in Horst and Tuy (1990), Pintér (1996) cannot be used in this case to establish convergence properties of global optimization techniques. That is why to analyse the proposed method we have to carry out some extended investigations.

For the sake of simplicity we present our approach in two successive steps. Initially the one-dimensional case is considered and Section 2 contains the general scheme of the method, a detailed description of the univariate version of the method and some results of numerical experiments. Then we apply this scheme to the multi-dimensional case (Section 3). First we briefly discuss how to extend the univariate optimization technique to a multivariate one and then provide the algorithmic scheme of the method. Finally, we investigate convergence properties of the method (Section 4) and describe the results of numerical experiments (Section 5).

2. Algorithm for the One-Dimensional Case

In this Section the objective function is a univariate differentiable function $f(x)$, the domain of search D is an interval $[a, b]$ and the condition (1.4) is satisfied with an unknown constant L.

APPROACH. Let us start with a brief description of the univariate global optimization method proposed in Gergel (1992). This method belongs to the class of characteristical optimization algorithms (see Grishagin (1979), the related approach presented in Pintér (1986)) and can be described as follows.

Iteration points generated by the method subdivide the initial search domain into a set of subintervals. At any iteration, to make the next optimization step, the method examines the "importance" of each subinterval and selects a point for the next function value evaluation within the most "important" subinterval. To solve the optimization problem successfully the "importance" of the subinterval must correspond to the possibility of the global minimizer being within the subinterval (the value of importance expressed numerically is called *the characteristic*).

To assign characteristics for subintervals of the search domain the following approach has been used.

As can be proved from the condition (1.4), taking into account the Taylor expansion of $f(x)$, limited to the second order term, the next inequality is satisfied

$$f(\bar{x}) \geqslant f(x) + f'(x)(\bar{x} - x) - 0.5L(\bar{x} - x)^2 \tag{2.1}$$

with $x, \bar{x} \in [a, b]$. It means that if for any subinterval $[x_1, x_2] \in [a, b]$ the function values are evaluated at the interval endpoints then

$$f(x) \geqslant max \begin{pmatrix} f(x_1) + f'(x_1)(x - x_1) - 0.5L(x - x_1)^2 \\ f(x_2) - f'(x_2)(x_2 - x) - 0.5L(x_2 - x)^2 \end{pmatrix}, \tag{2.2}$$

where $x \in [x_1, x_2]$. From (2.1), (2.2) we can estimate the least value of the objective function $f(x)$ within the interval $[x_1, x_2]$ (see Figure 1)

$$f(x) \geqslant R[x_1, x_2] = f(x_1) + f'(x_1)(\hat{x} - x_1) - 0.5L(\hat{x} - x_1)^2, \tag{2.3}$$

where

$$\hat{x} = \frac{-(f(x_2) - f(x_1)) + (f'(x_2)x_2 - f'(x_1)x_1) + 0.5L(x_2^2 - x_1^2)}{L(x_2 - x_1) + (f'(x_2) - f'(x_1))}. \tag{2.4}$$

Below the expressions (2.3), (2.4) are used to calculate the values of the interval characteristics, but various numerical estimates of the Lipschitz constant are used instead of L.

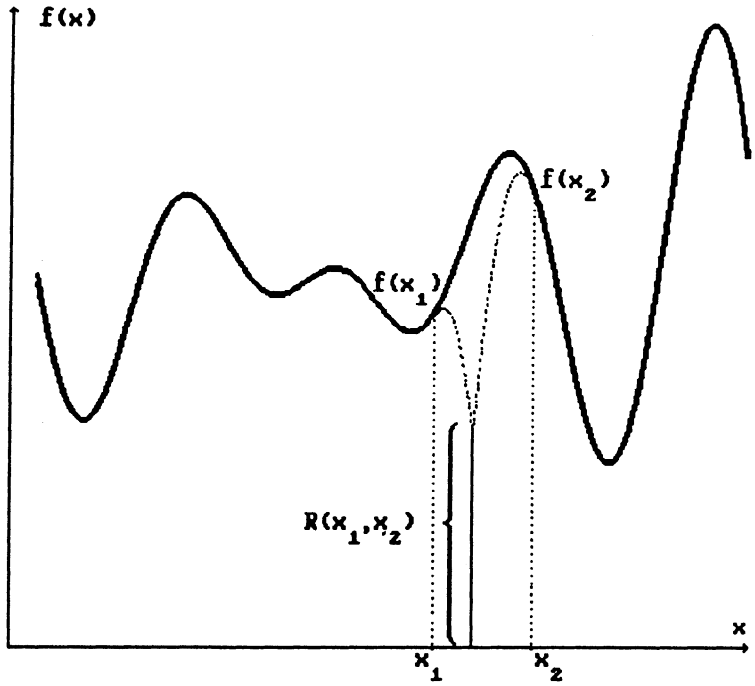

Figure 1. The least value estimate of the function $f(x)$ within the interval $[x_1, x_2]$ using the Lipschitz condition for the first derivatives.

SCHEME. The global Method using Derivative values (MD) proposed in Gergel (1992) can be described as follows.

To start with MD, the values of the function $f(x)$ and its derivative $f'(x)$ are calculated at the points $x^0 = a$ and $x^1 = b$ (such calculations will be referred to as *trials* at corresponding points). The point $x^{(s+1)}$, $s \geqslant 1$ of the next $(s+1)$-th iteration is selected in the following way:

1. Order the points $x^0, ..., x^s$ of previous trials by increasing their coordinates, i.e.

$$a = x_0 < x_1 < ... < x_i < ... < x_s = b; \tag{2.5}$$

2. Compute absolute values of the slopes of the function to be minimized over each interval (x_{i-1}, x_i), $1 \leqslant i \leqslant s$,

$$M_i = max \begin{cases} \dfrac{|z'_i - z'_{i-1}|}{x_i - x_{i-1}} \\ \dfrac{2[-(z_i - z_{i-1}) + z'_{i-1}(x_i - x_{i-1})]}{(x_i - x_{i-1})^2}, \; 1 \leqslant i \leqslant s, \\ \dfrac{2[(z_i - z_{i-1}) - z'_i(x_i - x_{i-1})]}{(x_i - x_{i-1})^2}, \end{cases} \tag{2.6}$$

where $z_i = f(x_i)$, $z'_i = f(x_i)$, $1 \leqslant i \leqslant s$, and x_i, $1 \leqslant i \leqslant s$, are from (2.5);

3. Compute an estimate of the Lipschitz constant

$$M = \max_{1 \leqslant i \leqslant s} M_i, \quad m = \begin{cases} 1 & \text{if } M = 0, \\ rM & \text{if } M > 0, \end{cases} \tag{2.7}$$

   where $r > 1$ is *the parameter* of the method and M_i, $1 \leqslant i \leqslant s$, are from (2.6);
4. Calculate for each interval (x_{i-1}, x_i), $1 \leqslant i \leqslant s$, the value

$$R(i) = z_{i-1} + z'_{i-1}(\hat{x}_i - x_{i-1}) - 0.5m_i(\hat{x}_i - x_{i-1})^2, \tag{2.8}$$

   where

$$m_i = 0.5m(1 + (M_i/m)^2), \tag{2.9}$$

$$\hat{x}_i = \frac{-(z_i - z_{i-1}) + (z'_i x_i - z'_{i-1} x_{i-1}) + 0.5m_i(x_i^2 - x_{i-1}^2)}{m_i(x_i - x_{i-1}) + (z'_i - z'_{i-1})}. \tag{2.10}$$

   We shall call $R(i)$ *the characteristic* of the interval (x_{i-1}, x_i);
5. Find among the intervals (x_{i-1}, x_i), $1 \leqslant i \leqslant s$ an interval with the minimal characteristic

$$R(t) = \min_{1 \leqslant i \leqslant s} R(i); \tag{2.11}$$

6. Execute a new trial at the point $x^{s+1} = \hat{x}_t$, where t is from (2.11) and $\hat{x}_t$ is calculated according to (2.10).

The algorithm stops when

$$(x_t - x_{t-1}) < \varepsilon, \tag{2.12}$$

where t is from (2.11), and $\varepsilon > 0$ is a preset accuracy. The values

$$z_s^\star = \min_{0 \leqslant i \leqslant s} f(x^i), \quad x_s^\star = arg \min_{0 \leqslant i \leqslant s} f(x^i) \tag{2.13}$$

may be taken as *an estimate of the global solution of the problem (1.1), (1.2).*

REMARK 1. The value of m in (2.7) presents a numerical estimate of the Lipschitz constant. Increasing the method parameter r improves the reliability of the method implementation (because an adequate estimate of the Lipschitz constant of the objective function may be provided). But, on the other hand, large values of r can be a reason for increasing the number of method iterations implemented by the method until the stopping condition (2.12) is satisfied.

It should be noted that the value of m may overestimate the Lipschitz constant obtained for the objective function over subintervals of the search domain (see the related discussion in Hansen *et al.* (1992), Sergeyev (1995), Pintér (1996)). That is why the method uses a "local" estimation of the Lipschitz constant for each search subinterval (x_{i-1}, x_i), $1 \leqslant i \leqslant s$, in accordance with (2.9).

REMARK 2. The characteristic $R(i)$ from (2.8) is a lower estimate for values of the objective function $f(x)$ within the interval (x_{i-1}, x_i) (see Figure 1). Taking account of (2.11) we can conclude that the point x^{s+1} of a new trial is taken within the interval (x_{t-1}, x_t) containing the estimate of the smallest possible value of the objective function $f(x)$ over the interval $[a, b]$.

CONVERGENCE. The convergence properties of the method are given by the following Theorem (see Gergel (1992)).

THEOREM 1. *If MD is used to solve the problem (1.1)–(1.2) and the condition (1.4) is true, then for any accumulation point $\bar{x}$ of the minimizing sequence $\{x^s\}$ generated by this algorithm the following statements are valid:*

(1) the point $\bar{x}$ is a local minimizer if the function $f(x)$ has a finite number of local extrema within the interval $[a, b]$;

(2) $z^s = f(x^s) \geqslant f(\bar{x}), s \geqslant 1$, i.e. the algorithm does not converge to points where the function value exceeds the result of some realized trial;

(3) if there is another accumulation point $\hat{x}$ of the sequence $\{x^s\}$, then $f(\bar{x}) = f(\hat{x})$, i.e. a simultaneous convergence to points with different function values is impossible. Hence, if the minimized function is not a constant then the method will produce a nonuniform net in the interval $[a, b]$;

(4) if at some step we obtain $m > 2L$, where m is from (2.7), then $\bar{x}$ will be a global minimizer of $f(x)$ and, moreover, the set of all accumulation points of the sequence $\{x^s\}$ will coincide with the set of global minimizers of $f(x)$.

NUMERICAL EXAMPLE. To demonstrate the method potentialities we present the results of minimizing the objective function (see Strongin (1978), Hansen *et al.* (1992))

$$f(x) = \sin x + \sin 10x/3 + \ln x - 0.84x + 3, \quad x \in [2.7, 7.5]. \tag{2.14}$$

The parameter r from (2.7) and the accuracy parameter ε from (2.12) have been taken as follows: $r = 1.1, \quad \varepsilon = 0.0001(b - a)$.

The function to be minimized is shown in Figure 2a. In Figure 2b the series of vertical strokes indicates the trial points x_i, $0 \leqslant i \leqslant 20$, from (2.5) in which the values of the function $f(x)$ and its first derivative $f'(x)$ have been calculated (the dark rectangle denotes the group of closely located points). In Figure 2c the segments of the broken line connect successively (from bottom to top) the points corresponding to pairs $(x^s, s), (x^{s+1}, s + 1)$, where x^s is the coordinate, s is the number of trials. The total number of iterations is 21 until the stopping condition from (2.12) is satisfied.

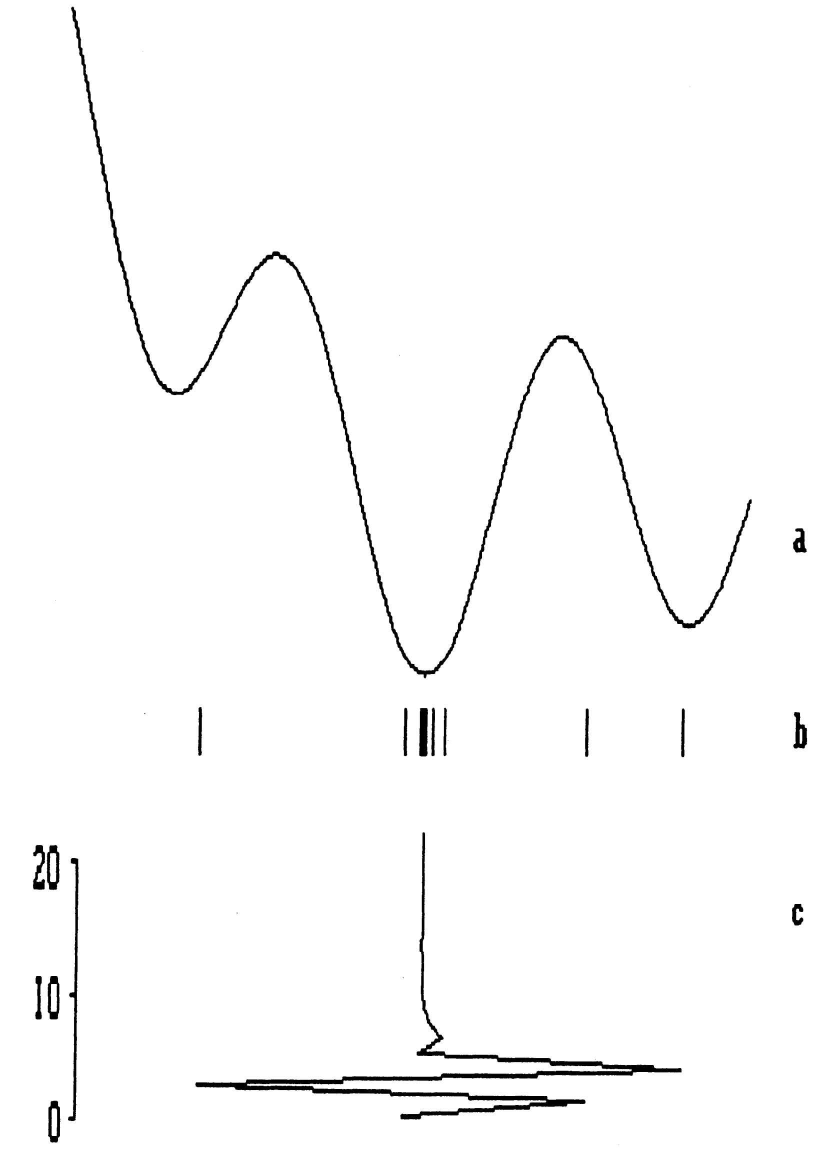

Figure 2. An example of implementation of the Method using Derivative values (a) the view of the function to be minimized, (b) the points generated by the method, (c) optimization dynamics.

Results of more extended numerical experiments for this method are given in Gergel and Sergeyev (1994) where some comparison with other methods are also presented.

3. Algorithm for the Multi-Dimensional Case

Let us consider the N-dimensional ($N > 1$) case. Assume that the objective function $f(y)$ from (1.1) is differentiable and the condition (1.4) is satisfied with an unknown constant L.

APPROACH. We propose a method for solving multi-dimensional global optimization problems that is formulated as a diagonal extension of the univariate optimization technique described above (the approach is used for such extension has been suggested in Pintér (1986)). Some necessary concepts and notation are given below.

In accordance with this approach at any iteration the search domain D is subdivided into several hypercubes. Let D_i, $1 \leqslant i \leqslant s$, denote these hypercubes, α^i and β^i denote the lower left and the upper right vertices of D_i, $1 \leqslant i \leqslant s$, accordingly, i.e.

$$D_i = \{y \in D : \alpha^i_j \leqslant y_j \leqslant \beta^i_j,\ \ 1 \leqslant j \leqslant N\},\ \ 1 \leqslant i \leqslant s, \tag{3.1}$$

where s is the number of the current iteration. To split the search domain, the hypercubes have to be taken in such a way that

$$\cup^s_{i=1} D_i = D,\ \ < D_i > \cap < D_j > = \emptyset,\ \ i \neq j,$$

where $< D >$ denotes the interior of D (the pair D_i and D_j may have a common face). It is also required that the values of the objective function and its first partial derivatives at the points α^i, β^i, $1 \leqslant i \leqslant s$, of the current partition of D should be calculated at previous iterations. An example of such a partition of D is depicted in Figure 3a (trial points are denoted by bold dots).

On the basis of some partition of the search domain a multi-dimensional method can be formulated as an univariate one. To perform an iteration, the method has to calculate characteristics for the hypercubes $D_i, 1 \leqslant i \leqslant s$, then to choose the hypercube which has the best (maximal or minimal) characteristic. Finally, the chosen hypercube is used for selecting some new iteration points and then it is split into several parts.

To describe a multi-dimensional extension of our univariate method let us denote

$$v_i = f(\alpha^i),\ \ u_i = f(\beta^i),\ 1 \leqslant i \leqslant s, \tag{3.2}$$

$$v'_i = \left(\sum^N_{j=1} f'_j(\alpha^i) d^i_j \right) / \Delta^i,\ 1 \leqslant i \leqslant s, \tag{3.3}$$

$$u'_i = \left(\sum^N_{j=1} f'_j(\beta^i) d^i_j \right) / \Delta^i,\ 1 \leqslant i \leqslant s, \tag{3.4}$$

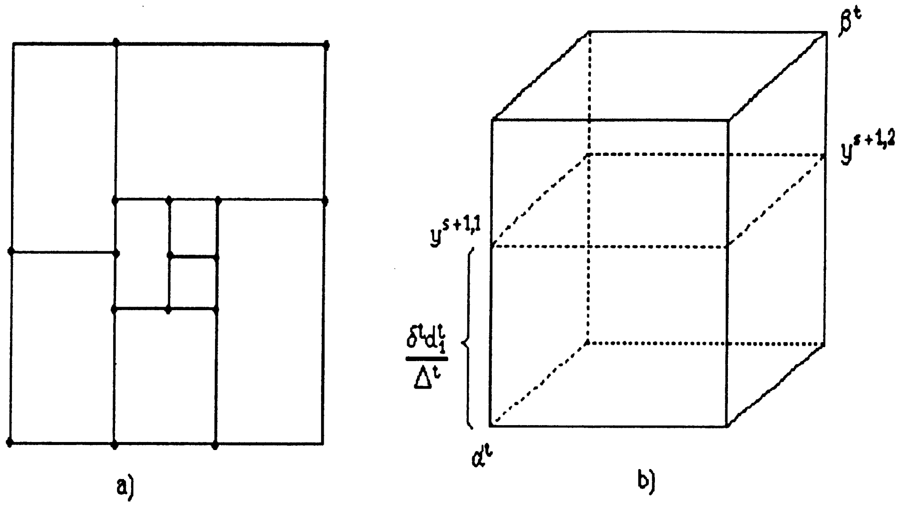

Figure 3. The method using derivative values for the multi-dimensional case (MDM): a) the structure of the search domain partition ($N = 2$), b) the scheme of splitting a partition element ($N = 3$).

where $f'_j(y)$ is the first partial derivative of $f(y)$ with respect to the j-th coordinate of y, d^i_j is the length of the j-th edge of D_i, $1 \leqslant i \leqslant s$, i.e.

$$d^i_j = \beta^i_j - \alpha^i_j,\ 1 \leqslant j \leqslant N,\ 1 \leqslant i \leqslant s, \tag{3.5}$$

and Δ^i is the length of the main diagonal of D_i, $1 \leqslant i \leqslant s$, i.e.

$$\Delta^i = \rho(\alpha^i, \beta^i) = \left[\sum_{j=1}^{N}(\beta^i_j - \alpha^i_j)^2\right]^{1/2},\quad 1 \leqslant i \leqslant s. \tag{3.6}$$

As it can be noted, $v_i(u_i)$ and $v'_i(u'_i)$ denote the values of the objective function $f(y)$ and its first directional derivative with respect to the main diagonal of D_i at the points α^i, β^i, $1 \leqslant i \leqslant s$, respectively.

SCHEME. The global Method using Derivative values for the Multi-dimensional case (MDM) can be described as follows.

First, two initial trials are calculated at the points $y^{1,1} = a$, $y^{1,2} = b$, i.e.

$$D_1 = D,\ \alpha^1 = a,\ \beta^1 = b, \tag{3.7}$$

where the vectors a and b are from (1.2). The points $y^{s+1,1}$, $y^{s+1,2}$, $s \geqslant 1$ of the next $(s+1)$-th iteration are selected in the following way:

1. Compute the absolute values of the slopes of the function $f(y)$ over the hypercubes $D_i, 1 \leqslant i \leqslant s$

$$M_i = max \begin{cases} \dfrac{|u_i' - v_i'|}{\Delta^i} \\ \dfrac{2[-(u_i - v_i) + v_i'\Delta^i}{(\Delta^i)^2}, \ 1 \leqslant i \leqslant s, \\ \dfrac{2[(u_i - v_i) - u_i'\Delta^i}{(\Delta^i)^2}, \end{cases} \tag{3.8}$$

where v_i, v_i', u_i, u_i', Δ^i, $1 \leqslant i \leqslant s$, are from (3.2)–(3.6);

2. Compute an estimate of the Lipschitz constant

$$M = \max_{1 \leqslant i \leqslant s} M_i, \quad m = \begin{cases} 1 & \text{if } M = 0, \\ rM & \text{if } M > 0, \end{cases} \tag{3.9}$$

where $r > 1$ is *the parameter* of the method;

3. Calculate for each hypercube D_i, $1 \leqslant i \leqslant s$, the characteristic

$$R(i) = v_i + v_i'\delta^i - 0.5m_i(\delta^i)^2, \tag{3.10}$$

where

$$m_i = 0.5m(1 + (M_i/m)^2), \tag{3.11}$$

$$\delta^i = \frac{-(u_i - v_i) + u_i'\Delta^i + 0.5m_i(\Delta^i)^2}{m_i\Delta^i + (u_i' - v_i')}. \tag{3.12}$$

4. Find among the hypercubes D_i, $1 \leqslant i \leqslant s$ a hypercube with the minimal characteristic

$$R(t) = \min_{1 \leqslant i \leqslant s} R(i); \tag{3.13}$$

5. Execute new trials at the points

$$y^{s+1,1} = (y_1^{s+1,1}, \ldots, y_N^{s+1,1}), \tag{3.14}$$

$$y^{s+1,2} = (y_1^{s+1,2}, \ldots, y_N^{s+1,2}), \tag{3.15}$$

where

$$y^{s+1,1} = \begin{cases} \beta_j^t, & j \neq l, \\ \alpha_l^t + \delta^t d_l^t/\Delta^t, & j = l, \end{cases}$$

$$y^{s+1,2} = \begin{cases} \alpha_j^t, & j \neq l, \\ \alpha_l^t + \delta^t d_l^t/\Delta^t, & j = l, \end{cases}$$

and l is the number of the longest edge of D_t, i.e.

$$d_l^t = \max_{1 \leqslant j \leqslant N} d_j^t, \quad d_j^t, \ 1 \leqslant j \leqslant N, \text{ from (3.5).} \tag{3.16}$$

6. Subdivide the hypercube D_t into two parts

$$\bar{D}_t = \{y \in D_t : \bar{\alpha}_j^t = \alpha_j^t \leqslant y_j \leqslant \bar{\beta}_j^t = y_j^{s+1,1}, 1 \leqslant j \leqslant N\}, \tag{3.17}$$

$$\hat{D}_t = \{y \in D_t : \hat{\alpha}_j^t = y_j^{s+1,2} \leqslant y_j \leqslant \hat{\beta}_j^t = \beta_j^t, 1 \leqslant j \leqslant N\}, \tag{3.18}$$

where t is from (3.13) and $y^{s+1,1}$, $y^{s+1,2}$ are from (3.14)–(3.15). Then replace the hypercube D_t in the search domain partition by its subhypercubes $\bar{D}_t$, $\hat{D}_t$ (the scheme of such subdivision is shown in Figure 3b).

The algorithm stops when

$$\Delta^t \leqslant \varepsilon, \tag{3.19}$$

where t is from (3.13), and $\varepsilon > 0$ is a preset accuracy. The values

$$z_s^\star = \min_{1 \leqslant i \leqslant s} \min_{1 \leqslant j \leqslant 2} f(y^{i,j}), \quad x_s^\star = arg \min_{1 \leqslant i \leqslant s} \min_{1 \leqslant j \leqslant 2} f(y^{i,j}) \tag{3.20}$$

may be taken as *an estimate of the global solution of the problem (1.1), (1.2).*

REMARK 3. The hypercube characteristics can be computed according to the equivalent formula based upon the upper right boundary point of the hypercube

$$R(i) = u_i - u_i'(\Delta^i - \delta^i) - 0.5 m_i (\Delta^i - \delta^i)^2, \tag{3.21}$$

where δ^i is from (3.12) and Δ^i from (3.6).

The convergence conditions of MDM are considered in the next section.

4. Conditions of Convergence

It should be noted that though our method has been formulated in the framework of the partition algorithm scheme (see Pintér (1986)) but the assumption that the Lipschitz constant of the function to be minimized is unknown prevents us from using general convergence results obtained in Pintér (1986). In particular, in this case one cannot propose any characteristics (a partition operator) which satisfy the monotonicity condition (see for details Pintér (1996)). That is why to analyse the proposed method we have to carry out some extended investigations.

LEMMA 1. *For the value of* δ^i, $1 \leqslant i \leqslant s$, *from (3.12) the inequality*

$$max[\delta^i, \Delta^i - \delta^i] \leqslant \frac{1}{2}\left(1 + \frac{2}{r+1}\right)\Delta^i \tag{4.1}$$

is satisfied, where Δ^i is the length of the main diagonal of the hypercube D^i, $1 \leqslant i \leqslant s$, from (3.1).

Proof. From (3.8) it follows that

$$-(u_i - v_i) \leqslant -v_i' \Delta^i + 0.5 M_i (\Delta^i)^2.$$

Taking into account (3.12), we shall get

$$\begin{aligned} \delta^i &\leqslant \frac{(u_i' - v_i')\Delta^i + 0.5(m_i + M_i)(\Delta^i)^2}{m_i \Delta^i + (u_i' - v_i')} = \\ &= \Delta^i \left(1 - \frac{0.5(m_i - M_i)}{m_i + (u_i' - v_i')/\Delta^i}\right) \leqslant \Delta^i \left(1 - \frac{0.5(m_i - M_i)}{m_i + M_i}\right) = \\ &= \Delta^i \left(1 - \frac{0.5(1 - M_i/m_i)}{1 + M_i/m_i}\right). \end{aligned}$$

From (3.9), (3.11),

$$\delta^i \leqslant \Delta^i \left(1 - \frac{0.5(1 - 1/r)}{1 + 1/r}\right) = \frac{1}{2}(1 + \frac{2}{r+1})\Delta^i.$$

The inequality,

$$\Delta^i - \delta^i \leqslant \frac{1}{2}(1 + \frac{2}{r+1})\Delta^i$$

can be shown similarly.

LEMMA 2. *Let $\bar{y}$ be a limit point (a point of accumulation) of the sequence $\{y^s\}$. Then for any other limit point $\hat{y}$, it follows that*

$$f(\bar{y}) = f(\hat{y}). \tag{4.2}$$

Proof. As $\bar{y}$ is a limit point and Lemma 1 is true then a sequence of hypercubes $D_{j(s)}$ should exist so that

$$\bar{y} \in D_{j(s)} \text{ and } \lim_{s \to \infty} \Delta^j = 0, \tag{4.3}$$

where Δ^j is from (3.6). As a result

$$\lim_{s \to \infty} R(j(s)) = f(\bar{y}), \tag{4.4}$$

Then for any $\theta > 0$ there exists an iteration $p > 1$ so that the inequality

$$f(\bar{y}) - \theta \leqslant R(j(s)) \leqslant f(\bar{y}) + \theta \tag{4.5}$$

is valid if $s > p$.

Now let us assume the contrary, i.e. that the condition (4.2) is not true. Without loss of generality, let

$$f(\bar{y}) < f(\hat{y}) \tag{4.6}$$

In this case any hypercube D_l with

$$\hat{y} \in D_l \text{ and } R(l) > f(\hat{y}) - \theta, \tag{4.7}$$

where θ is from (4.5) and

$$\theta = (f(\hat{y}) - f(\bar{y})/2 \tag{4.8}$$

cannot be chosen by MDM as the hypercube with the minimal characteristic (see (4.5)–(4.8)). From (3.13)–(3.16) it follows that if $s > p$ then new trial points will not fall within such hypercubes. As a result (4.3)–(4.4) cannot be obtained for $\hat{y}$, but this contradicts the fact that $\hat{y}$ is a limit point.

Thus the condition (4.2) must be true.

THEOREM 2. *Let $\{y^s\}$ be the sequence generated by MDM in the course of solving the problem (1.1)–(1.2) and assume that the condition (1.4) is satisfied. If at some iteration $p > 1$ for the value m from (3.9) the inequality*

$$m > \gamma L, \tag{4.9}$$

where

$$\gamma = \left[\frac{2(r+1)}{r-1}\right]^2 \left[\frac{K}{L\varepsilon}\left(\frac{3r+5}{r+1}\right) + 1\right]$$

(K is the Lipschitz constant from (1.3), ε is from (3.19), r is from (3.9)) is true, then any point of the absolute minimum $y^\star$ is a limit point of the sequence $\{y^s\}$. Moreover, any limit point $\bar{y}$ of this sequence will be a global minimizer of $f(y)$.

Proof. Initially we shall prove that if (4.9) is true then

$$R(i) \leqslant f(y) \tag{4.10}$$

for any $y \in D_i$, $1 \leqslant i \leqslant s$. Let σ^i be the distance between α^i and y, i.e.

$$\sigma^i = \rho(\alpha^i, y) = \left[\sum_{j=1}^{N}(y_j - \alpha_j^i)^2\right]^{1/2}.$$

From (3.10)

$$R(i) = v_i + v_i'\delta^i - 0.5m_i(\delta^i)^2 = v_i - K\sigma^i - 0.5L(\sigma^i)^2 + Q_i,$$

where

$$Q_i = v_i'\delta^i + K\sigma^i - 0.5m_i(\delta^i)^2 + 0.5L(\sigma^i)^2.$$

Taking into account the Taylor expansion of $f(y)$ at α^i it follows that (4.10) is true if $Q_i \leqslant 0$. Let us consider a case when this inequality is valid.

From (4.1) we can conclude

$$Q_i \leqslant K(\delta^i + \Delta^i) - 0.5 m_i(\delta^i)^2 + 0.5L(\Delta^i)^2 \leqslant$$
$$\leqslant K\Delta^i \left[\frac{r+3}{2(r+1)} + 1\right] - 0.5\left[m_i\left(\frac{r-1}{2(r+1)}\right)^2 + L\right](\Delta^i)^2. \qquad (4.11)$$

From (4.11) it can be shown that $Q_i \leqslant 0$ if

$$m_i \geqslant 2\left[\frac{2(r+1)}{r-1}\right]^2\left(\frac{1}{\Delta^i}\right)^2\left[0.5K\Delta^i\frac{3r+5}{r+1} + 0.5L(\Delta^i)^2\right] \geqslant$$
$$\geqslant \left[\frac{2(r+1)}{r-1}\right]^2\left[\frac{K}{L\varepsilon}\frac{3r+5}{r+1} + 1\right]L$$

(the last inequality follows from $\Delta^i \geqslant \varepsilon$).

Now let us prove that $y^\star$ has to be a limit point of the sequence $\{y^s\}$.

Assume the contrary, i.e. $y^\star$ is not a point of accumulation. Then there will exist a number $q > 0$ such that for all iterations $l \geqslant q$ trials will not fall in the hypercubes D_i such that $y^\star \in D_i$. But in accordance with (4.9), (4.10)

$$R(i) \leqslant f(y^\star) \qquad (4.12)$$

if $s \geqslant \max(p,q)$. On the other hand, for any limit point $\bar{y}$ it follows (4.3), (4.4). But $f(\bar{y}) \geqslant f(y^\star)$ and (4.12) contradicts to (4.4). Then our assumption is not true and the point $y^\star$ of absolute minimum of $f(y)$ is a limit point of the sequence $\{y^s\}$.

The fact that any limit point $\bar{y}$ has to be a global minimizer follows immediately from Lemma 2.

REMARK 4. The coefficient γ from (4.9) can be written in a simple form for the range of large values of the method parameter r

$$\gamma \geqslant \gamma^+ = 4(1 + 3K/L\varepsilon).$$

MODIFICATIONS. To provide additional convergence properties for MDM we shall consider some modifications of the rules (3.8), (3.10)–(3.12).

Let us extend the expressions (3.2)–(3.6) for any two points $p, q \in D$:

- the distance between p and q is equal to

$$\Delta(p,q) = \left[\sum_{j=1}^{N} d_j^2(p,q)\right]^{1/2} \qquad (4.13)$$

where $d_j(p,q)$ is the difference between the j-th coordinates of p and q, i.e.,

$$d_j(p,q) = q_j - p_j, \quad 1 \leqslant j \leqslant N. \qquad (4.14)$$

Instead of (3.3)–(3.4) now we calculate the values

$$\bar{v}' = \begin{cases} \left[\sum_{i \in I} f_i'(p) d_i(p,q) \right] / \Delta(p,q), & \text{if } |I| > 1, \\ \min_{1 \leqslant i \leqslant N} [f_i'(p) \mathtt{sign}(d_i(p,q))], & \text{if } |I| \leqslant 1, \end{cases} \tag{4.15}$$

$$\bar{u}' = \begin{cases} \left[\sum_{j \in J} f_j'(q) d_j(p,q) \right] / \Delta(p,q), & \text{if } |J| > 1, \\ \max_{1 \leqslant j \leqslant N} [f_j'(q) \mathtt{sign}(d_j(p,q))], & \text{if } |J| \leqslant 1, \end{cases} \tag{4.16}$$

where

$$I = \{i \ : \ 1 \leqslant i \leqslant N, \ f_i'(p) d_i(p,q) < 0\},$$
$$J = \{j \ : \ 1 \leqslant j \leqslant N, \ f_j'(p) d_j(p,q) > 0\},$$

and $|I|$ $(|J|)$ denotes the number of elements of the set I (J), $\mathtt{sign}(x)$ denotes the sign of x. As a result of calculations, $\bar{v}'$ $(\bar{u}')$ in (4.15)–(4.16) contains the smallest (the largest) value of the first directional derivative of $f(y)$ with respect to any direction

$$\alpha \boldsymbol{py} \text{ such that } (y_j - p_j) d_j(p,q) > 0, \ 1 \leqslant j \leqslant N,$$
$$(\alpha \boldsymbol{qy} \text{ such that } (y_j - q_j) d_j(p,q) > 0, \ 1 \leqslant j \leqslant N)$$

at the point p (q), where $d_j(p,q)$, $1 \leqslant j \leqslant N$, are from (4.14).

The modification the MDM is formed by the following rules.

Modification 1. The formula

$$M_i = \ max \begin{cases} M(\alpha^i, \beta^i), \\ M(\alpha^i, y_s^\star), \ y_s^\star \neq \alpha^i, \\ M(y_s^\star, \beta^i), \ y_s^\star \neq \beta^i, \end{cases} \tag{4.17}$$

where $y_s^\star$ is from (3.20) and

$$M(p,q) = \ max \begin{cases} |\bar{u}' - \bar{v}'| / \Delta(p,q), \\ 2[-(u-v) + \bar{v}' \Delta(p,q)] / \Delta^2(p,q), \\ 2[\ (u-v) - \bar{u}' \Delta(p,q)] / \Delta^2(p,q), \end{cases} \tag{4.18}$$

has to be used instead of (3.8) for

$$i \in I_s^\star = \{i, \ 1 \leqslant i \leqslant s \ : \ y_s^\star \in D_i\}. \tag{4.19}$$

Modification 2. The formula

$$R(i) = \ min \begin{cases} R(\alpha^i, \beta^i), \\ R(\alpha^i, y_s^\star), \ y_s^\star \neq \alpha^i, \\ R(y_s^\star, \beta^i), \ y_s^\star \neq \beta^i, \end{cases} \tag{4.20}$$

where m_i is from (3.11) and

$$R(p,q) \;=\; v + \bar{v}'\hat{\delta} - 0.5m_i\hat{\delta}^2, \tag{4.21}$$

$$\hat{\delta} \;=\; \frac{-(u-v) + \bar{u}'\Delta(p,q) + 0.5m_i\Delta^2(p,q)}{m_i\Delta(p,q) + (\bar{u}' - \bar{v}')}, \tag{4.22}$$

has to be used instead of (3.10) for the case (4.19).

REMARK 5. The hypercube characteristics $R(p,q)$ from (4.21) can be computed according to the equivalent formula based upon the upper right boundary point of the hypercube

$$R(i) = u - \bar{u}'(\Delta(p,q) - \hat{\delta}) - 0.5m_i(\Delta(p,q) - \hat{\delta})^2, \tag{4.23}$$

where $\hat{\delta}$ is from (4.22) and $\Delta(p,q)$ is from (4.13).

We shall denote the multi-dimensional method combined with (4.17)–(4.22) as MDM1.

THEOREM 3. *Let the point $\bar{y}$ be a limit point of the sequence $\{y^s\}$ generated by MDM1 in the course of solving the problem (1.1), (1.2). Then the function values at any iteration point cannot be less than $f(\bar{y})$, i.e.*

$$f(y^{s,j}) \geqslant f(\bar{y}),\; s \geqslant 1,\; 1 \leqslant j \leqslant 2.$$

Proof. Assume the contrary, i.e. the value

$$f(y^{p,j}) < f(\bar{y})$$

has been obtained at some iteration $p > 1$. As a result the minimum function value $f(y_s^\star)$ from (3.20) is also less than $f(\bar{y})$, i.e.

$$z_s^\star = f(y_s^\star) < f(\bar{y}) \tag{4.24}$$

If $s > p$, then the rules (4.15)–(4.16) provide that there exists a hypercube D_l, $1 \leqslant l \leqslant s$, that contains the point $y_s^\star$ and

$$\bar{v}'(\alpha^l, y_s^\star) \leqslant 0 \text{ or } \bar{u}'(y_s^\star, \beta^l) \leqslant 0.$$

In both cases the characteristic of D_l calculated in accordance with (4.21) or (4.23) is less than $z_s^\star$, i.e.

$$R(l) < z_s^\star \tag{4.25}$$

Thus from some iteration $s > p$ there exists a hypercube which has its characteristic less than the characteristics of the hypercubes containing the point $\bar{y}$ (see (4.4), (4.24), (4.25)). As a result $\bar{y}$ cannot be a limit point. This fact contradicts the assumption of the Theorem.

THEOREM 4. *Let the point $\bar{y}$ be any limit point (a point of accumulation) of the sequence $\{y^s\}$ generated by MDM1 in the course of minimization of the bounded function $f(y)$, $y \in D$. If the point $\bar{y}$ belongs to the interior of D then $\bar{y}$ will be a local minimizer of the function $f(y)$.*

Proof. Let us assume the opposite, i.e. that the limit point $\bar{y}$ is not a local minimizer. Then the first partial derivative of $f(y)$ with respect to some coordinate of y is not equal to zero ($f_i'(\bar{y}) \neq 0$ for some i, $1 \leqslant i \leqslant N$). Otherwise, the point $\bar{y}$ will be a local maximizer, or a saddle point. Let us consider all possible situations.

(1). Consider the case $f_i'(\bar{y}) \neq 0$ for some i, $1 \leqslant i \leqslant N$. Without loss of generality, let $f_i'(\bar{y}) < 0$. Then the two following situations are possible.

In the first case at some iteration q a new trial is carried out precisely at the point $\bar{y}$. Therefore for all iterations $s > q$ there exists a hypercube D_l, $l = l(s)$, containing the point $\bar{y}$ and

$$\beta_i^l - \bar{y}_i > 0.$$

From Theorem 2, (4.15) and taking into account our assumption it follows that $\bar{v}_l(\bar{y}, \beta^l) < 0$. As a result,

$$R(l) < f(\bar{y})$$

and from (4.3)–(4.4)

$$\lim_{s \to \infty} \Delta^{s(l)} \to 0.$$

Since, in accordance with (3.16), the coordinates y_i, $1 \leqslant i \leqslant N$, for subdividing the hypercube D_t from (3.13) are used consecutively, so at some iteration $p > 0$ the point $\hat{y}$ of a new trial is such that

$$f(\hat{y}) < f(\bar{y}). \tag{4.26}$$

This inequality contradicts Theorem 3.

Now it is necessary to consider the second situation when we have $\bar{y} \neq y^{s,\nu}$, $s \geqslant 1,\ 1 \leqslant \nu \leqslant 2$, i.e. trial points do not coincide with the limit point. Let the point $\bar{y}$ belong to a hypercube D_j, $j = j(s)$, at the s-th iteration. If there exists convergence to the point $\bar{y}$ then (4.3)–(4.4) are true. Due to the fact that the coordinates y_i, $1 \leqslant i \leqslant N$, for subdividing the hypercube D_t from (3.13) are used consecutively and taking account of our assumption, at some iteration $q > 0$ a new trial is carried out at some point $\hat{y}$ in which (4.26) holds. But this situation also contradicts Theorem 3.

Now we can conclude that, if the point $\bar{y}$ is not a local minimizer of $f(y)$, it should have at least $f_i'(\bar{y}) = 0$ for any $i,\ 1 \leqslant i \leqslant N$.

(2). Consider the case when $f_i'(\bar{y}) = 0$ for any i, but the limit point $\bar{y}$ is a local maximizer. As $\bar{y}$ is a limit point then there exists a subsequence of hypercubes such that (4.3)–(4.4) are true. Therefore, at some iteration $q > 0$ a new trial is carried out at some point $\hat{y}$ such that (4.26) holds.

As it contradicts Theorem 3, the limit point cannot be a local maximizer.

(3). Similarly, it can be shown that the situation when the point $\bar{y}$ is a saddle point is also impossible.

Having considered all this, we can conclude that $\bar{y}$ must be a point of local minimum of $f(y)$ and the Theorem has been proved.

5. Numerical Results

In this Section we present some numerical experiments to study computational behaviour of the proposed optimization techniques.

In the first series of experiments we use test optimization problems given in Dixon and Szegö (1978). The first of these functions is the Goldstein and Price (GP) function stated as follows:

$$\begin{aligned} F(y) \;=\; & [1+(y_1+y_2+1)^2(19-14y_1+3y_1^2-14y_2+6y_1y_2+3y_2^2)]\,* \\ & *[30+(2y_1-3y_2)^2(18-32y_1+12y_1^2+48y_2-36y_1y_2+27y_2^2)], \end{aligned}$$

where $-2 \leqslant y_1, y_2 \leqslant 2$. Level curves of this function are shown in Figure 4. As it can be noted, $f(y)$ has four local minima. The global minimum is at the point $(0, -1)$ with the value $F(0, -1) = 3$.

Initially to solve this problem we have set the method parameter $r = 1.4$, where r is from (3.9), and the accuracy $\varepsilon = 0.1||b - a||$, where ε is from (3.19) and a, b is from (1.2).

The Method using Derivative values for the Multi-dimensional case (MDM, see Section 3) has made 66 iterations, that corresponds to 132 function and first derivative value evaluations. The global minimum estimate is

$$z_s^\star = 17.63, \;\; y_s^\star = (-0.21, -1.11).$$

To estimate the global minimum more precisely the proposed method can be applied in *the mixed mode* when purely global iterations are combined with local steps around the current estimate of the optimum. According to this scheme (see Strongin *et al*. (1988)) for each odd iteration the value of $R(i)$ in (3.10) and (4.20) are replaced with

$$\hat{R}(i) = (R(i) - R^-(i))/[((u_i - z_s^\star)(v_i - z_s^\star))^2 + 10^{-6}m_i], \tag{5.1}$$

where

$$R^-(i) = \min_{1 \leqslant i \leqslant s} R(i),$$

and $z_s^\star$ is from (3.20), m_i is from (3.11).

Making the second experiment we have set $\varepsilon = 0.01||b - a||$ and $r = 1.4$ and the mixed scheme has been employed after 50 iteration MDM. In this case the total number of iterations is 68 that corresponds to 136 function and

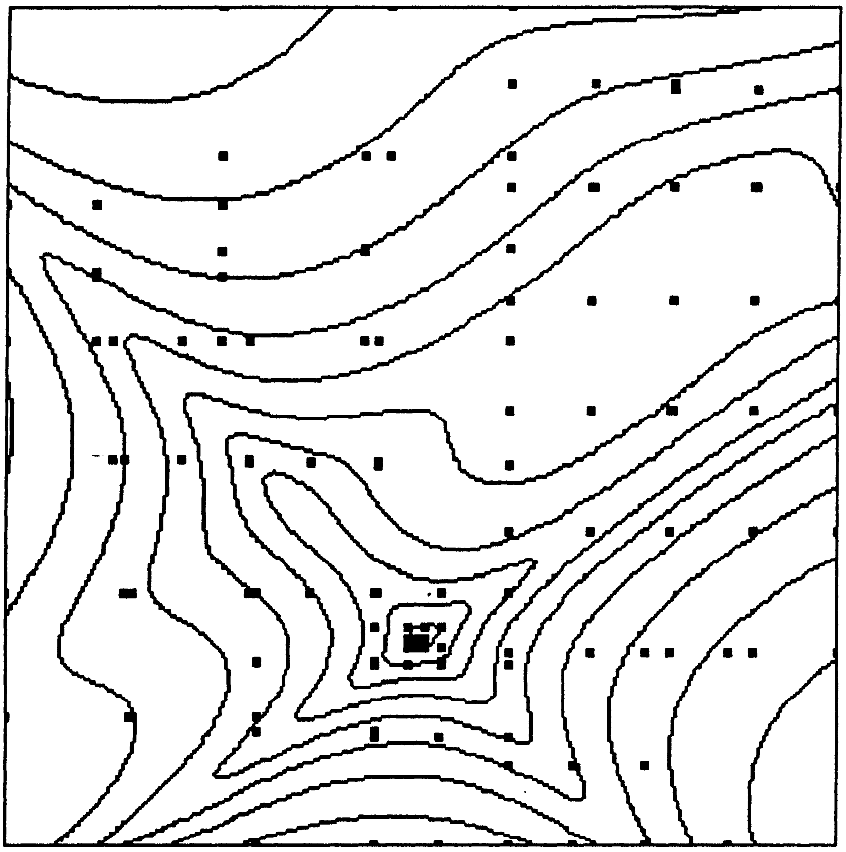

Figure 4. Minimizing the Goldstein and Price function by MDM.

first derivative value evaluations. Trial points marked by bold dots are shown in Figure 4. The global minimum estimate is

$$z_s^\star = 3.77,\ \ y_s^\star = (-0.045, -1.037).$$

Let us discuss how to set a value of the method parameter r from (3.9). As it can be noted, the parameter is used to increase a numerical estimate of the Lipschitz constant of the function to be minimized (see (3.9)). In this connection the parameter can be regarded as *a reliability coefficient* of the method. Increasing the value of the parameter can improve the probability of finding the global minimum (because an adequate estimate of the Lipschitz constant is provided). On the other hand, large values of r can be a reason for increasing the number of method iterations implemented by the method until the stop condition is satisfied. To illustrate the influence of the method parameter, the experiment with $r = 2.0$ and $\varepsilon = 0.01||b - a||$ has been carried out (the mixed scheme has been employed similarly). In this case the total number of iterations is 78, that corresponds to 156 function and first derivative value evaluations. The global minimum estimate is

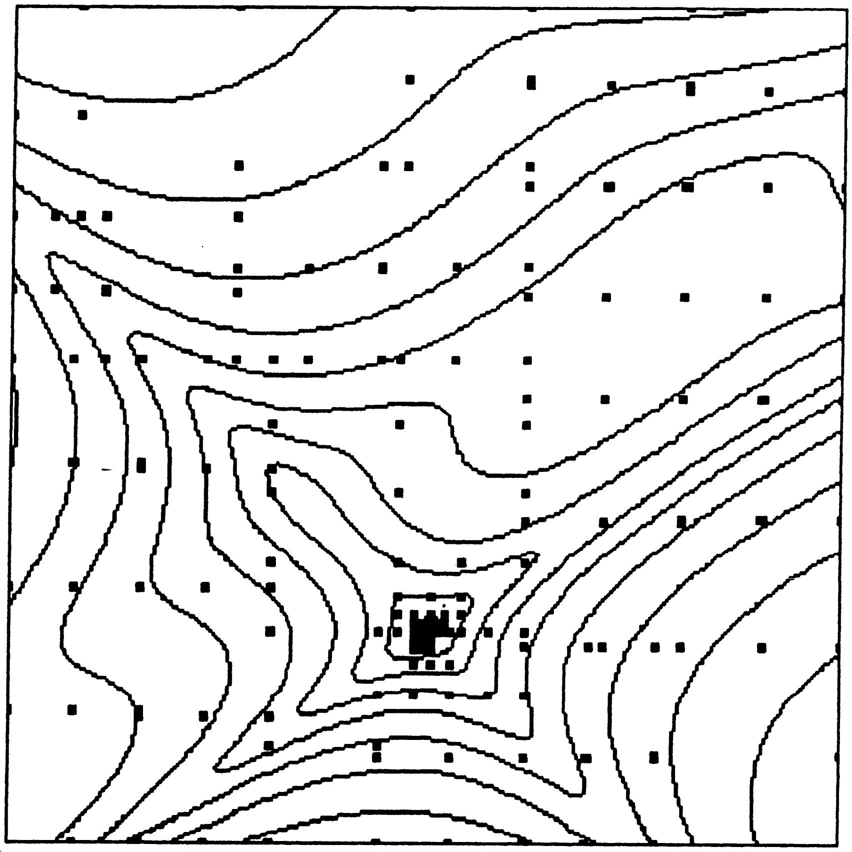

Figure 5. Minimizing the Goldstein and Price function by the modified version of MDM.

$$z_s^\star = 3.0011, \;\; y_s^\star = (-0.00031, -0.99851).$$

In practice the range from 1.5 to 3.0 can be recommended for the parameter values.

To solve the same problem the modified method MDM1 (see Section 4) with $r = 1.4$ and $\varepsilon = 0.01||b-a||$ has made 96 iterations. The global minimum estimate is

$$z_s^\star = 3.03, \;\; y_s^\star = (0.0042, -1.0067).$$

To compare both variants of the proposed method the trial points of MDM1 are given in Figure 5.

Table 1 contains the results of minimizing the Goldstein and Price function by other methods (these examples are taken from Breiman and Cutler (1993)). Besides, Table 1 also gives the results of minimizing the Branin (RCOS) function:

$$f(y) = (y_2 - (5.1/4\pi^2)y_1^2 + (5/\pi)y_1 - 6)^2 + (10 - 1/8\pi)\cos y_1 + 10,$$

where $-5 \leqslant y_1 \leqslant 10$, $0 \leqslant y_2 \leqslant 15$.

Table 1. Results of Minimizing Test Problems by some well-known methods

Method	Number of function evaluations	
	Function	
	GP	RCOS
Törn	2499	1558
de Biase and Frontini	378	597
Price	2500	1600
Bremermann	300	160
Rinnooy Kan and Timmer	148	206
Snyman and Fatti	474	–
Vanderbilt and Louie	1186	557
Breiman and Cutler	–	269
MDM	136	124
MDM1	192	148

The methodology of the computational experiments to compare numerical methods is discussed widely (see, for example, Jackson *et al.* (1991), Horst and Pardalos (1995), etc.). The scheme used in the paper has been proposed in Grishagin (1978), Strongin (1978). In the framework of this approach the method to be compared has to be applied for solving a wide variety of test optimization problems. In performing these computations, the values of method parameters have to be fixed and the problems have to be selected in a random way. As a result of these calculations a number of pairs $\{(s, p_s)\}$ can be obtained , where s is the number of function and first derivative evaluations and p_s is the fraction of functions for which the global minimum has been found with the given accuracy after this amount of optimization iterations. These pairs are referred to as *operational characteristics* of the method and they can be plotted as a broken line graph.

To estimate operational characteristics of the proposed method we used a set of multiextremal functions, each of them stated as follows (see Strongin (1978)):

$$f(y) = \left\{ \left(\sum_{i=1}^{7} \sum_{j=1}^{7} [A_{ij} a_{ij}(y) + B_{ij} b_{ij}(y)] \right)^2 + \right.$$
$$\left. + \left(\sum_{i=1}^{7} \sum_{j=1}^{7} [C_{ij} a_{ij}(y) - D_{ij} b_{ij}(y)] \right)^2 \right\}^{1/2}, \qquad (5.2)$$

where

$$a_{ij}(y) = \sin \pi i y_1 \sin \pi j y_2,$$
$$b_{ij}(y) = \cos \pi i y_1 \cos \pi j y_2,$$

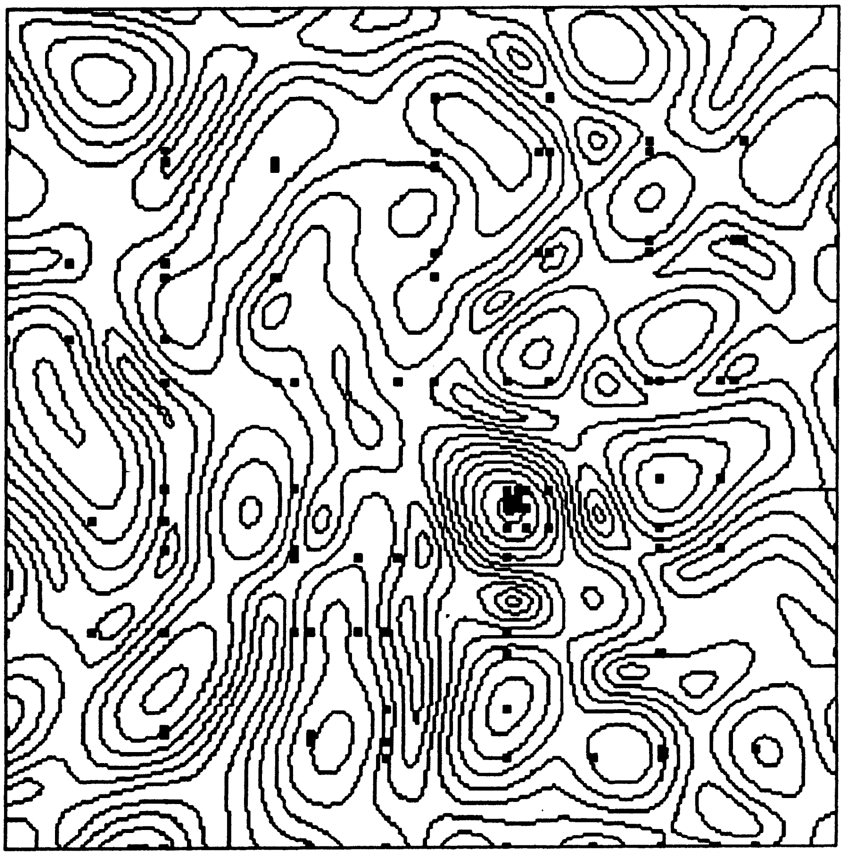

Figure 6. Minimizing the function from the second series of numerical experiments.

and $0 \leqslant y_1, y_2 \leqslant 1$. To explain the nature of these functions it should be noted that the expressions (5.2) describe the stresses of a thin elastic plate under transverse load. Level curves for one such function are shown in Figure 6.

The values $A_{ij},\ B_{ij},\ C_{ij},\ D_{ij}$ for any function to be minimized are generated by a random mechanism. A total of 50 functions has been chosen.

Within these experiments we use MDM with the method parameter $r = 1.4$ and the accuracy $\varepsilon = 0.01||b - a||$. The results of minimizing (trial points) for one of the functions (5.2) are shown in Figure 6.

Each function has been minimized several times with different numbers of purely global steps when the formula (5.1) was not applied. Operational characteristics that have been obtained are given in Figure 7. To compare the results, we present in Figure 7 operational characteristics for some other methods (see Strongin (1978)), viz.:

– the uniform grid (UG) algorithm;,

– the random or Monte-Carlo (MC) method;

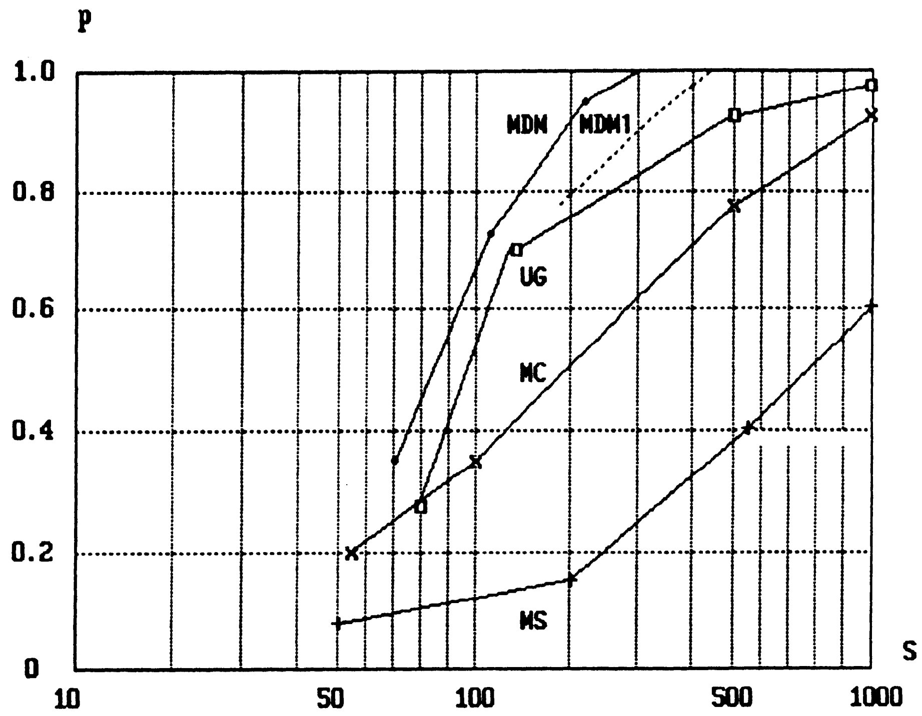

Figure 7. Operational characteristics for the global optimization methods.

– the multi-start (MS) local scheme of using the Nelder–Mead (1964) method from various start points selected randomly in the search domain.

To mark the operational characteristics of the presented optimization techniques we use the signs "+", "x" and "□" respectively.

As it can be expected, operational characteristics obtained for the method with various parameter values are different. For instance, when we use $r = 2.0$ instead of $r = 1.4$ the average number of optimization iterations is increased by approximately 10%.

Operational characteristics of MDM1 with $\varepsilon = 0.01||b - a||$ and $r = 1.4$ are also presented in Figure 7 (the dashed broken line).

6. Conclusions

A new multi-dimensional method for solving unconstrained global optimization problems with Lipschitzian first derivatives has been proposed in this paper. In accordance with this method, the search domain is subdivided into a set of hypercubes in the course of optimization. This partitioning is regulated by the decision rule that provides evaluation of the "importance" of each generated hypercube and selection of some partition element to perform the next iteration. Sufficient conditions of global convergence have been established for two variants of the algorithm.

The second modification has additional theoretical properties (see Theorem 4) but the first one demonstrates a better numerical behaviour.

To examine computational behaviour, operational characteristics have been calculated by solving a wide variety of test global optimization problems. As we can conclude, the proposed method is quite competitive in comparison with the other optimization techniques.

Acknowlegments

The author thanks R.G. Strongin for many useful discussions and the anonymous referees for their helpful suggestions.

The revised version of the paper has been prepared when the author was a guest at József Attila University (Szeged, Hungary). This visit was undertaken in accordance with the joint Hungarian–Russian project and was supported by OMFB Grant OF-17/94. The author thanks T. Csendes for his friendly attention to this work.

References

Archetti, F. and Shoen, F. (1984), A Survey on the Global Optimization Problems: General Theory and Computational Approaches, *Annals of Operation Research* **1**, 87–110.

Baritompa, W. (1994), Accelerations for a Variety of Global Optimization Methods, *J. of Global Optimization* **4**, 37–45.

Breiman, L. and Cutler, A. (1993), A Deterministic Algorithm for Global Optimization, *Math. Programming* **58**, 179–199.

Butz, A.R. (1968), Space Filling Curves and Mathematical Programming, *Information and Control* **12**, 319–330.

Dixon, L.C.W. and Szegö, G.P., eds. (1978), *Towards Global Optimization*, **2**, North-Holland, Amsterdam.

Evtushenko, Yu.G. (1971), Algorithm for Finding the Global Extremum of a Function (Case of a Nonuniform Mesh), *USSR Computational Mathematics and Mathematical Physics* **11**, 1390–1403.

Evtushenko, Yu.G. and Potapov, M.A. (1994), Deterministic Global Optimization, In: Spedicato E., ed. *Algorithms for Continuous Optimization*, 481–500, Kluwer, Dordrecht.

Galperin, E.A. (1985), The Cubic Algorithm, *J. Math. Analyses and Appl.* **112**, 635–640.

Gergel, V.P. (1992), A Global Search Algorithm Using Derivatives, In: Neymark Yu. I., ed., *Systems Dynamics and Optimization*, 161–178, N.Novgorod University Press, N.Novgorod (in Russian).

Gergel, V.P. and Sergeyev, Ya. D. (1994), *Sequential and Parallel Global Optimization Algorithms Using Derivatives*, University of Calabria, Department of Mathematics, Report n. 7/94.

Grishagin, V.A. (1978), Operational Characteristics for Some Algorithms of Global Search, In: *Problems of Random Search*, **7**, 198–206, Zinatne, Riga (in Russian).

Grishagin, V.A. (1979), On Convergence Conditions for One Class of Global Search Algorithms, In: *Numerical Methods on Nonlinear Programming*, Proceedings of the All-Union Seminar, 82–84, Kharkov, Russia (in Russian).

Hansen, P., Jaumard B. and Lu, S.-H. (1989), Global Optimization of Univariate Functions by Sequential Polynomial Approximation, *Int. J. of Computer Mathematics* **28**, 183–193

Hansen, P., Jaumard, B. and Lu, S.-H. (1992), Global Optimization of Univariate Lipschitz Functions: 2. New Algorithms and Computational Comparison, *Math. Programming* **55**, 273–292.

Horst, R. and Tuy, H. (1990), *Global Optimization – Deterministic Approaches*, Springer, Berlin (2nd edn., 1993).

Horst, R. and Pardalos, P.M., eds. (1995), *Handbook of Global Optimization*, Kluwer, Dordrecht.

Jackson, R.H.F., Boggs, P.T., Nash, S.G. and Powell, S. (1991), Guidelines for Reporting Results of Computational Experiments. Report of the ad hoc committee, *Math. Programming* **49**, 413–425.

Meewella, C.C. and Mayne, D.Q. (1989), Efficient Domain Partitioning Algorithms for Global Optimization of Rational and Lipschitzian Continuous Functions, *J. Optim. Theory Appl.* **61**, 247–270.

Nelder, J.A. and Mead, R. (1964), A Simplex Method for Function Minimization, *Computer J.* **7**, 308–313.

Pardalos, P.M. and Rosen, J.B., eds. (1990), *Computational Methods in Global Optimization*, *Annals of Operations Research* **25**.

Piyavskii, S.A. (1972), An Algorithm for Finding the Absolute Extremum of a Function, *USSR Computational Mathematics and Mathematical Physics* **12**, 57–67.

Pintér, J. (1986), Extended Univariate Algorithms for N-Dimensional Global Optimization, *Computing* **36**, 91–103.

Pintér, J. (1996), *Global Optimization in Action*, Kluwer, Dordrecht.

Rinnooy, A.H.G. and Timmer, G.H. (1989), Global Optimization: A Servey, In: Nemhauser G.L. *et al*., eds., *Handbooks of Operations Research* **1**, 631–662, North-Holland, Elsevier Science Publishers.

Sergeyev, Ya.D. (1995), An One-Dimensional Deterministic Global Minimization Algorithm, *Russian J. Computational Mathematics and Mathematical Physics* **35**, 705–717.

Shubert, B.O. (1972), A Sequential Method Seeking the Global Maximum of a Function, *SIAM J. of Numerical Analysis* **9**, 379–388.

Strongin, R.G. (1978), *Numerical Methods for Multiextremal Problems*, Nauka, Moskow (in Russian).

Strongin, R.G., Gergel, V.P., Markin, D.L. (1988), Multicriterion Multiextreme Optimization with Nonlinear Constraints, In: Lewandovski, A. and Volkovich, V., eds., *Lecture Notes in Economics and Mathematical Systems* **351**, 120–127. Proceedings 1988. Springer. IIASA, Laxenburg/Austria 1991.

Törn, A.A. and Žilinskas, A. (1989), *Global Optimization, Lecture Notes in Computer Sciences*, **350**, Springer, Berlin.