E. Spedicato (ed.). Algorithms for Continuous Optimization, 481–500.

DETERMINISTIC GLOBAL OPTIMIZATION[1]

Yu.G. EVTUSHENKO, M.A. POTAPOV
Computing Center
40 Vavilov Str.
117967 Moscow GSP-1
Russia

Revised 14 November 2002

Abstract: Numerical methods for finding global solutions of nonlinear programming and multicriterial optimization problems are proposed. The sequential deterministic approach is used which is based on the non-uniform space covering technique as a general framework. The definitions of the ε-solution for the nonlinear programming problem and the multicriterial optimization problems are given. It is shown that if all the functions, which define these problems, satisfy a Lipschitz condition and the feasible set is compact, then ε-solutions can be found in the process, of only one covering of the feasible set on a non-uniform net with a finite number of function evaluations. Space covering techniques are applied to solving systems of nonlinear equations and minimax problems.

1. Introduction

The purpose of this paper is to describe a unified sequential space covering technique for finding approximate global solutions of various optimization problems. The global optimization is of great importance to all fields of engineering, technology and sciences. In numerous applications the global optimum or an approximation to the global optimum is required.

Numerical methods for seeking global solutions of multiextremal problems, in spite of their practical importance, have been rather poorly developed. This is, no doubt, due to their exceedingly great complexity. We do not detail all the available approaches to this problem. Instead, we shall concentrate on one very promising direction, which is based on the idea of a non-uniform covering of a feasible set. This approach has turned out to be quite universal and, as we shall show, can be used not only for seeking global extrema of functions but also for nonlinear programming problems (NLP), for nonlinear integer programming, for solving systems of nonlinear equations, sequential minimax problems and, most importantly, for multicriterial optimization problems (MOP). For these problems we introduce the notion of ε-solutions and describe numerical methods for finding these solutions. Sequential deterministic algorithms and practical results have been obtained for Lipschizian optimization, where all functions which define the problem satisfy a Lipschitz condition. If moreover a feasible set is compact then an ε-solution of MOP can be found after only one covering of the feasible set on a non-uniform net. This property of the proposed approach simplifies the solution process radically in comparison with the traditional approaches which are based on the use of scalarization techniques (convolution functions, reference point approximation, etc.). In our algorithm the computing time which is needed for finding the approximate solution of a multicriterial optimization problem is close to the time which is needed for the search of the global minima of a single function

[1]Research supported by the grant N 93-012-450 from Russian Scientific fund

on the same feasible set. This approach was proposed in Evtushenko (1985), Evtushenko and Potapov (1984, 1985, 1987), Potapov (1984).

In this paper we briefly present some results in the field of global optimization which were obtained in the Computing Center of Russian Academy of Sciences. Our methods were implemented on a computer and were used in practice. On the basis of the proposed approach we have developed the packages Solvex and Globex, implemented in Fortran and C language. The libraries of algorithms which were included in these packages enable a user to solve the following classes of problems:

1. Unconstrained global and local minimization of a function of several variables.
2. Nonlinear global and local minimization of a function under the equality and inequality constraints.
3. Nonlinear global solution of a multicriterial problem with equality and inequality constraints.

These packages give opportunity to combine the global approach with local methods and this way speeds up the computation considerably. Global nonlinear problems that are solvable in reasonable computer time must be of limited dimension (of order 10 to 20); however, the use of multiprocessors, parallel computing, and distributed processing substantially increases the possibilities of this approach. Our preliminary results of computations on a parallel transputer system are encouraging. The description of the packages for global optimization and the computational experiments will be given in subsequent papers.

2. General concept

We consider the global optimization problem

$$f_* = \text{global} \min_{x \in X} f(x), \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}^1$ is a continuous real valued objective function and $X \subset \mathbb{R}^n$ is a nonempty compact feasible set.

Since maximization can be transformed into minimization by changing the sign of the objective function, we shall consider only the minimization problem here. Subscripts will be used to distinguish values of quantities at a particular point and superscripts will indicate components of vectors.

As a special case, we consider the situation, where X is a right parallelepiped P with sides parallel to the coordinate axes (a box in the sequel):

$$P = \{x \in \mathbb{R}^n : a \le x \le b,\ a \in \mathbb{R}^n,\ b \in \mathbb{R}^n\}. \tag{2}$$

Here and below, the vector inequality $q \le z$, where $q, z \in \mathbb{R}^n$, means that the componentwise inequalities $q^i \le z^i$ hold for all $i = 1, \ldots, n$.

The set of all global minimum points of the function f (the solutions set) and the set of ε-optimal solutions are defined as follows

$$X_* = \{x_* \in X : f_* = f(x_*) \le f(x),\ \forall x \in X\}, \tag{3}$$
$$X_*^\varepsilon = \{x_\varepsilon \in X : f(x_\varepsilon) \le f_* + \varepsilon\}. \tag{4}$$

For the sake of simplicity of the presentation, we assume throughout this paper that global minimizer sets exist. The existence of an optimal solution in (1) is assured by the well-known

Theorem of Weierstrass. The sets X_*, X_*^ε are nonempty because of the assumed compactness of X and continuity of $f(x)$.

The global optimal value of f is denoted by $f_* = f(x_*)$, $x_* \in X_*$. Our goal is to find at least one point $x_\varepsilon \in X_*^\varepsilon$. Any value $f(x_\varepsilon)$, where $x_\varepsilon \in X_*^\varepsilon$, is called an ε-optimal value of f on X. Let $N_k = \{x_1, x_2, \ldots, x_k\}$ be a finite set of k points in X. After evaluating the objective function values at these points, we define the record value

$$R_k = \min_{1\le i\le k} [f(x_1), f(x_2), \ldots, f(x_k)] = f(x_r), \tag{5}$$

where $r \in [1:k]$; any such point x_r is called a record point.

We say that a numerical algorithm solves the problem (1) after k evaluations if a set N_k is such that $R_k \le f_* + \varepsilon$, or, equivalently, $x_r \in X_*^\varepsilon$. The algorithm is defined by a rule for constructing such a set N_k.

We introduce the Lebesgue level set in X

$$K(\ell) = \{x \in X : \ell - \varepsilon \le f(x),\ \ell \in \mathbb{R}^1,\ \varepsilon \in \mathbb{R}^1\}. \tag{6}$$

Theorem 1. *Let N_k be a set of k feasible points such that*

$$X \subseteq K(R_k), \tag{7}$$

then any record point $x_r \in N_k$ belongs to X_*^ε.

Proof. The set of solutions satisfies $X_* \subseteq X \subseteq K(R_k)$. Let a point x_* belong to X_*, then according to (6) and (7) we have $x_* \in K(R_k)$ and, therefore, $f(x_r) = R_k \le f(x_*) + \varepsilon$. It means that $x_r \in X_*^\varepsilon$. □

If condition (7) holds, then we will say that the set X is covered by $K(R_k)$.

It is very difficult to implement this result as a numerical algorithm because the level set $K(R_k)$ can be very complicated; in general it is not compact and it requires a special program to store it in computer memory. Therefore, we have to weaken the statement of this theorem and impose an additional requirement on the function f. We suppose that for any point $z \in X$ and any level value ℓ, where $\ell \le f(z)$, it is possible to define set $B(z,\ell)$ as follows

$$B(z,\ell) = \{x \in G(z) : \ell - \varepsilon \le f(x)\}, \tag{8}$$

where $G(z) \subset \mathbb{R}^n$ is a closed bounded nonempty convex neighborhood of the point z such that the Lebesgue measure mes $(G(z))$ of a set $G(z)$ is positive and $z \in G(z)$.

Theorem 2 (Main Theorem). *Let N_k be a set of feasible points such that*

$$X \subseteq \bigcup_{x_j \in N_k} B(x_j, R_k), \tag{9}$$

then any record point $x_r \in N_k$ belongs to X_*^ε.

Proof. It is obvious that

$$B(z, f(z)) \subseteq K(f(z)), \quad \bigcup_{x_j \in N_k} B(x_j, R_k) \subseteq K(R_k).$$

Therefore, from condition (9) follows (7), that proves the theorem. □

If the set X is compact, then a finite set N_k, which satisfies the conditions of this Theorem, exists.

The construction of numerical methods is split in two parts: 1) the definition of a set $B(z, \ell)$, 2) the definition of covering rule. Let us consider the first part. Assume that the function f satisfies a Lipschitz condition on $\mathbb{R}^n$ with constant L. It means that for any x and $z \in \mathbb{R}^n$, we have

$$|f(x) - f(z)| \leq L\|x - z\|. \tag{10}$$

In this case, we can write that

$$B(z, \ell) = \{x \in \mathbb{R}^n : \|x - z\| \leq r = [\varepsilon + f(z) - \ell]/L\}, \tag{11}$$

i.e. $B(z, \ell)$ is a ball of radius r and a center z. If $x \in B(z, \ell)$ and $\ell \leq f(z)$, then from (10) we obtain that condition $\ell - \varepsilon \leq f(x)$ holds. If $f(z) = \ell$, then the ball $B(z, f(z))$ has minimal radius $\rho = \varepsilon/L$. The smallest edge of a hypercube inscribed into ball $B(z, f(z))$ is $2\rho/\sqrt{n}$.

Suppose the function f is differentiable and for any x and z of the convex compact set X we have

$$\|f_x(x) - f_x(z)\| \leq M\|x - z\|,$$

where M is a constant. In this case, $B(z, \ell)$ can be constructed as a ball centered at $\bar{z}$ with radius $\bar{r}$:

$$\begin{aligned} B(\bar{z}, \ell) &= \{x \in \mathbb{R}^n : \|x - \bar{z}\| \leq \bar{r}\}, \ \bar{z} = z - f_x(z)/M, \\ \bar{r}^2 M^2 &= [\|f_x(z)\|^2 + 2M[f(z) + \varepsilon - R_k]]. \end{aligned}$$

These formulas were given in Evtushenko (1971, 1985). Other more complicated cases were considered in Evtushenko and Potapov (1984, 1987).

If $f(x)$ satisfies (10), then, evaluating the function on a regular orthogonal grid of points, $2\rho/\sqrt{n}$ apart in each direction, and choosing the smallest function value solves the problem. But this is generally impractical due to the large number of function evaluations that would be required. Many authors proposed various improvements to the grid method. The main idea of non-uniform grid methods is the following: a set of spheres is constructed such that the minimum of the function over the union of the spheres differs by at most ε from the minimum of the function at their centers. When a sufficient number of spheres is utilized (such that X is a subset of their union), the problem is solved. Such an approach is more efficient than the use of a simple grid.

In recent years a rapidly-growing number of deterministic methods has been published for solving various multiextremal global optimization problems. Many of them can be viewed as realization of a basic covering concept. Theorem 2 suggests a constructive approach for solving problem (1). If inclusion (9) holds, i.e. the union of sets $B(x_j, R_k)$ covers the feasible set, then the set N_k solves the problem (1). As a rule we construct sample sequence of feasible points $N_j = \{x_1, x_2, \ldots, x_j\}$ where the function $f(x)$ has been evaluated and such that the set $W = \bigcup_{x_j \in N_k} B(x_j, R_j)$ covers X. According to (5) $R_j \geq R_k$ for any $1 \leq j \leq k$, hence $B(x_j, R_j) \subseteq B(x_j, R_k)$. If $X \subseteq W$, then condition (9) holds. A family of numerical methods based on such an approach is called a space covering technique. Many optimization methods have been developed on the basis of the covering idea.

The volume of the current covering set $B(z, R_i)$ essentially depends on the current record value R_i and it is greatest for $R_i = f_*$, but the value f_* is usually not known. Hence, to extend this set it is desirable that the current record value be as close as possible to f_*. For this purpose we use the auxiliary procedures of finding a local minimum in the problem (1). If in the computation process we obtain that $f(x_{i+1}) < R_i$, then we use a local search algorithm and if we find a point $\bar{x} \in X$ at which $f(\bar{x}) < f(x_{i+1})$ then we take the quantity $f(\bar{x})$ as a current record value and the vector $\bar{x}$ as a current record point. After this we continue the global covering process. Therefore, the optimization package which we developed for global minimization includes a library of well-known algorithms for local minimization. The coherent utilization of both these techniques substantially accelerates the computation.

All results given in this paper can be extended straightforwardly to integer global optimization problems and mixed-integer programming. In this case all functions which define a problem must satisfy the Lipschitz condition (10) for any feasible vectors x, y whose components are integers. When solving practical problems, we often take the accuracy $\varepsilon = 0.1$ and the Lipschitz constant $L = 10$, therefore, the minimal radius of the covering ball is equal to $\rho = \varepsilon/L = 0.01$. Assume that we compute the value of a function f at a point $x \in X$ with integer coordinates, then we can exclude all points which are inside of the hypercube centered at x and having the edge lengths equal to two. Therefore, the minimal radius of the covering balls is greater or equal to one. We take the covering set B into account and exclude the union of the hypercube and B. It is possible to consider another common case where only a part of the variables are integers. For this case we use different covering sets in the spaces of integer and continuous variables. All methods presented here can be used for solving nonlinear integer programming problems and integer multicriterial optimization problems. Moreover, the integer assumption accelerates and simplifies greatly the covering process.

Sometimes we can take advantage of special knowledge about the problem and suggest modifications customized to use the additional information. For example, suppose that we have some prior knowledge about the upper bound Θ for optimal value f_*. We define positive numbers $0 < \delta_1 < \delta_2 \leq \varepsilon$ such that $\varepsilon \geq \Theta - f_* + \delta_2$, where $f_* \leq \Theta$. Let introduce a covering ball

$$H_j = \{x \in \mathbb{R}^n : \|x - x_j\| \leq r_j\}, \; r_j = (f(x_j) - \Theta - \delta_1)/L.$$

Now we cover the set X by balls H_j. For a sequence of feasible points $x_1, x_2, x_3, \ldots$ we evaluate the objective function and using (5) we determine the record value and the corresponding record point. Suppose that for all $1 \leq j < s$ we have $f(x_j) > \Theta + \delta_2$, $r_j > (\delta_2 - \delta_1)/L > 0$. If x_s is the first point such that $f(x_s) \leq \Theta + \delta_2$, then we conclude that $R_s = f(x_s) \leq f_* + \varepsilon$, $x_s \in X_*^\varepsilon$. We say that the global approximate solution of Problem (1) is found after s evaluations and we terminate computations.

Theorem 3. *Let N_k be a set of k feasible points such that $X \subseteq \bigcup_{j=1}^{k} H_j$. Assume that $f(z)$ satisfies the Lipschitz condition* (10) *and X_* is nonempty. Suppose that an upper estimation Θ of f_* is known. Then there exists an index i such that $x_i \in X_*^\varepsilon$, $f(x_i) \leq \Theta + \delta_2$, where $1 \leq i \leq k$.*

Proof. By contradiction. Assume that $R_k > \Theta + \delta_2$. According to (10) for any point x, which belongs to H_j, we have:

$$f(x) \geq f(x_j) - L\|x - x_j\| \geq f(x_j) - r_j L = \theta + \delta_1 > f_*.$$

Therefore, $f(x) > f_*$ on H_j. Because of arbitrariness of j, $1 \leq j \leq k$, the same property holds everywhere on X. It means that X_* is empty. Hence there exists a point $x_i \in N_k$ such that $x_i \in X_*^\varepsilon$. $\square$

We can use Theorem 2 and 3 simultaneously taking sequentially the following covering radius

$$\bar{r}_j = \frac{f(x_j) - \min[R_j - \varepsilon, \; \Theta + \delta_1]}{L}.$$

All covering algorithms can easily be modified and still retain their basic form.

We will not describe here all non-uniform covering techniques, instead we only mention some directions: the layerwise covering algorithm (Evtushenko (1971, 1985)), the bisection algorithm (Evtushenko and Ratkin (1987), Ratschek et. al. (1988)), the branch and bound approach (Volkov (1974), Evtushenko and Potapov (1987), Potapov (1984), Horst and Tuy (1989)), the chain covering algorithm (Evtushenko et. al. (1992)). The covering rules are developed mainly for the case where X is a right parallelepiped. In most of these papers the feasible set X is covered by hypercubes inscribed into covering balls B.

Other non-determenistic approaches for global optimization can be found in the recent survey by Betro (1991), and in the books by Horst and Tuy (1990), Torn and Zilinskas (1991).

3. Solution of nonlinear programming problems

The approach described in the preceding section carries over to solving nonlinear programming problems. The feasible set X can be nonconvex and nonsimply connected. Therefore, very often, it is difficult to realize algorithmically a covering of X by balls or boxes. It is easier to cover a rather simple set P that contains the set X. For example, if all components of a vector x are bounded, then P can be the "box" defined by (2). Suppose the global minimum is sought:

$$f_* = \text{global} \min_{x \in P \cap X} f(x), \tag{12}$$

$$X = \{x \in \mathbb{R}^n : \Psi(x) = 0\} \tag{13}$$

where $\Psi : \mathbb{R}^n \to \mathbb{R}^1$ and the intersection $P \cap X$ is nonempty.

The scalar function $\Psi(x)$ is equal to zero everywhere on X, and greater than zero outside X. As before, we denote by X_*, the global solutions set of problem (12), which is assumed to be nonempty, $X_* \subseteq P \cap X$. It is obvious that $X_* \subseteq P$. We extend the feasible set X by introducing the ε-feasible set X_ε, define the set of approximate global solutions X_*^ε of the problem (12) and two Lebesgue level sets in P:

$$X_\varepsilon = \{x \in \mathbb{R}^n : \Psi(x) \le \varepsilon\}, \tag{14}$$

$$\begin{aligned} X_*^\varepsilon &= \{x \in P \cap X_\varepsilon : f(x) - \varepsilon \le f_*\}, \\ T(\varepsilon) &= \{x \in P : 0 < \varepsilon < \Psi(x), \ \ \varepsilon \in \mathbb{R}^1\}, \\ K(\ell) &= \{x \in P : \ell - \varepsilon \le f(x), \ \ \ell \in \mathbb{R}^1\}. \end{aligned} \tag{15}$$

Let $N_k = \{x_1, \ldots, x_k\}$ be a set of k points from the set P, where the functions $f(x)$ and $\Psi(x)$ have been evaluated. The record point x_r and the record value R_k are defined similarly to (5):

$$R_k = \min_{x_i \in N_k \cap X_\varepsilon} f(x_i) = f(x_r).$$

If the intersection $N_k \cap X_\varepsilon$ is empty, then R_k is not defined and the set $K(R_k)$ is assumed to be empty.

Theorem 4. *Assume that the set of global solutions X_* of problem (12) is nonempty. Let N_k be a set of points from P such that*

$$P \subseteq K(R_k) \cup T(\varepsilon), \tag{16}$$

then there exists at least one record point $x_r \in N_k$ which belongs to X_*^ε.

Proof. The intersection $X_* \cap T(\varepsilon)$ is empty because $\Psi(x) > \varepsilon > 0$ for any point $x \in T(\varepsilon)$. Therefore, all points from $T(\varepsilon)$ can not belong to the set X_* of global solutions of problem (12). Hence using (16), we obtain $X_* \subseteq P \subseteq K(R_k)$. It means that $K(R_k)$ is nonempty and the intersection $N_k \cap X_\varepsilon$ is also nonempty. If $x_* \in X_*$, then there exists a record point x_r such that $x_r \in N_k$, $x_r \in P \cap X_\varepsilon$ and $f(x_r) - \varepsilon \leq f(x_*)$. Taking into account the definition (15), we conclude that $x_r \in X_*^\varepsilon$. $\square$

If functions $f(x)$ and $\Psi(x)$ satisfy the Lipschitz condition with the same constant L, then with each point $x_s \in N_k$ we associate a ball B_{sk} centered in x_s and with radius r_{sk}:

$$\begin{array}{lll} B_{sk} & = & \{x : \|x - x_s\| \leq r_{sk}\}, \qquad r_{sk} = \max[\hat{r}_{sk}, \tilde{r}_{sk}], \\ \hat{r}_{sk} & = & (f(x_s) - R_k + \varepsilon)/L, \qquad \tilde{r}_{sk} = (\Psi(x_s) - \delta)_+/L, \end{array} \tag{17}$$

where $a_+ = \max[a, 0]$, $0 < \delta < \varepsilon$.

Theorem 5. *Assume that the set of global solutions X_* of problem* (12) *is nonempty. Suppose that the functions f and Ψ satisfy the Lipschitz condition* (10) *and the set N_k of the points from P is such that $P \subseteq \bigcup_{i=1}^{k} B_{ik}$. Then any record point x_r belongs to X_*^ε.*

Proof. The solution set satisfies $X_* \subseteq P \subseteq \bigcup_{i=1}^{k} B_{ik}$. Consider a point $x_* \in X_*$. Then there exists at least one covering ball B_{sk} such that $x_* \in B_{sk}$. Hence, according to the definition of a ball B_{sk}, we have $\|x_* - x_s\| \leq r_{sk}$. We prove that $r_{sk} > 0$. Suppose that the radius of this ball is equal to $\tilde{r}_{sk}$, then we have

$$\|x_* - x_s\| \leq \tilde{r}_{sk} = (\Psi(x_s) - \delta)_+/L.$$

If $\Psi(x_s) \geq \delta$, then $\tilde{r}_{sk} = (\Psi(x_s) - \delta)Ł$ and, taking into account the Lipschitz condition, we obtain

$$\Psi(x_*) \geq \Psi(x_s) - L\|x_* - x_s\| \geq \delta > 0.$$

The above inequality is impossible because $x_* \in X_*$ and $\Psi(x_*) = 0$.

If $\Psi(x_s) < \delta$, then $\tilde{r}_{sk} = 0$ and $\Psi(x_s) < \varepsilon$. Therefore, $x_s \in X_\varepsilon$ and there exist x_r and R_k such that

$$\hat{r}_{sk} = \frac{f(x_s) - R_k + \varepsilon}{L} \geq \frac{\varepsilon}{L} > \tilde{r}_{sk} = 0.$$

This contradicts the definition (17) of r_{sk}. Hence we have $r_{sk} = \hat{r}_{sk} = (f(x_s) - R_k + \varepsilon)/L$. Using the Lipschitz condition (10), we obtain

$$f(x_*) \geq f(x_s) - L\|x_* - x_s\| \geq R_k - \varepsilon = f(x_r) - \varepsilon.$$

Taking into account the definition of X_*^ε, we conclude that $x_r \in X_*^\varepsilon$. $\square$

If a current point $x_s \in N_k$, where $s \leq k$, is such that $x_s \in X_\varepsilon$ then $\Psi(x_s) \leq \varepsilon$, $f(x_s) \geq$ $\geq R_s \geq R_k$ and the radius of a covering ball satisfies $r_{sk} \geq \varepsilon/L$. If $x_s \notin X_\varepsilon$, then $\Psi(x_s) > \varepsilon$ and $r_{sk} \geq (\varepsilon - \delta)/L$. Consequently, if the set P is compact, then a finite set N_k, which satisfies the conditions of Theorem 5, exists.

As a rule in a practical computation a right parallelepiped (2) plays the role of the set P, which is used in the statement (12). Due to it all covering methods mentioned in previous section can be used for solving (12). During the computational process the sequence R_i monotonically decreases, therefore, $B_{ii} \subset B_{ik}$ for any $1 \leq i \leq k$. Problem (12) will be solved, if we find the sequence N_k such that the union of balls B_{ik} or B_{ii}, $1 \leq i \leq k$ covers the set P. Such a finite set exists.

Consider the particular case of problem (1) where

$$f_* = \text{global} \min_{x \in P} f(x). \tag{18}$$

It is worthwhile to compare problem (18) of finding the minimum of $f(x)$ on P with the problem (12) under the additional constraint $x \in X$. At first glance, it seems paradoxical (although it is true), that finding the global solution of problem (12) is simpler than solving problem (18). The constraint $x \in X$ provides an additional possibility to increase the radii of the covering balls on $P \backslash X$. Hence, the additional constraints merely simplify the problem of finding global solutions. If we know some properties of the problem we should add them to the definition of the set X. We illustrate this idea using a simple version of problem (18). Suppose that $f(x)$ is differentiable on a box P, then the necessary conditions of the minimum can be written in the form

$$\varphi^i(x) = (x^i - a^i)(b^i - x^i)\frac{\partial}{\partial x^i} f(x) = 0, \qquad 1 \leq i \leq n.$$

We introduce the feasible set as follows

$$X = \{x \in \mathbb{R}^n : \sum_{i=1}^{n} (\varphi^i(x))^2 = 0\}.$$

Now we solve the problem (12) instead of (18) and simplify the covering process in this way.

The function $\Psi(x)$ which defines the feasible set can be found on the basis of penalty functions. Consider the case where the feasible set is defined by equality and inequality type constraints

$$X = \{x \in \mathbb{R}^n : h(x) \leq 0,\ g(x) = 0\}, \qquad h : \mathbb{R}^n \to \mathbb{R}^c, \quad g : \mathbb{R}^n \to \mathbb{R}^m. \tag{19}$$

Introduce the vector-valued function $h_+(x) = [h_+^1(x), \ldots, h_+^c(x)]$, $h_+^i = \max[0, h^i]$. Let $\|z\|_p$ denote a Hölder vector norm of a vector $z \in \mathbb{R}^s$:

$$\|z\|_p = \left(\sum_{i=1}^{s} |z^i|^p \right)^{1/p}, \qquad 1 \leq p \leq \infty.$$

If $p = 1, 2, \infty$, then we have the Manhatten, Euclidean and Chebyshev norms, respectively:

$$\|z\|_1 = \sum_{i=1}^{s} |z^i|, \quad \|z\|_2 = \left(\sum_{i=1}^{s} (z^i)^2 \right)^{1/2}, \quad \|z\|_\infty = \max_{1 \leq i \leq s} |z^i|.$$

Suppose that each component of h and g satisfies the Lipschitz condition with constant L:

$$|h^i(x) - h^i(y)| \leq L\|x - y\|, \qquad |g^i(x) - g^i(y)| \leq L\|x - y\|.$$

Then using Theorem 1.5.2 from Evtushenko (1985), it is easy to show that $h_+(x)$ satisfies a Lipschitz condition with the same constant L. If we use the well-known inequality

$$|\|a\| - \|b\|| \leq \|a - b\|,$$

then we obtain that the function

$$\Psi(x) = \left\| \begin{array}{c} h_+(x) \\ g(x) \end{array} \right\|_p = \left[\sum_{i=1}^{c} (h_+^i(x))^p + \sum_{j=1}^{m} |g^j(x)|^p \right]^{1/p}, \tag{20}$$

also satisfies the Lipschitz condition with constant L.

The great advantage of the proposed approach lies in the tact that the constraints need not be dealt with separately and that the classical and modern local numerical methods can be used as auxiliary procedures to improve the record values and in this way to accelerate the computations.

4. Numerical solution of global multicriterial minimization problem

The multicriterial minimization problem has numerous applications in diverse fields of science and technology and plays a key role in many computer-based decision support systems. Various complex engineering problems and practical design require multicriterial optimization. The non-uniform covering technique developed above can be extended for the solution of multicriterial minimization problems.

In minimizing a number of objective functions $F^1(x), F^2(x), \dots, F^m(x)$ it can not be expected in general that all of the objective functions attain their minimum value simultaneously. The objectives usually conflict with each other in that any improvement of one objective can be achieved only at the expense of another. For such multiobjective optimization the so-called Pareto optimality is introduced.

Define the vector-valued function $F^\top(x) = [F^1(x), F^2(x), \dots, F^m(x)]$, $F : \mathbb{R}^n \to \mathbb{R}^m$. Let $Y = F(X)$ be the image of X under the continuous mapping $F(x)$. The problem of global multicriterial minimization of the vector-valued function $F(x)$ on an admissible set X is denoted by

$$\text{global} \min_{x \in X} F(x). \tag{21}$$

The set X_* of solutions of this problem is defined as follows:

$$X_* = \{x \in X : \text{if it exists } w \in X \text{such that } F(w) \le F(x), \text{then } F(w) = F(x)\}. \tag{22}$$

To solve problem (21) means to find the set X_*. In papers on multicriterial optimization, X_* is usually called the set of effective solutions, and its image $Y_* = F(X_*)$ is called the Pareto set. The sets X_* and Y_* are assumed to be nonempty.

In many practical problems x is a vector of decisions and $y = F(x)$ is a criterion vector or outcome of the decisions. Therefore, we say that x belongs to the decision space $\mathbb{R}^n$ and that y belongs to the criteria space $\mathbb{R}^m$. We can rewrite the definition (22) in the criteria space:

$$Y_* = \{y \in Y : \text{if it exists } q \in Y \text{such that } q \le y, \text{then } q = y\}. \tag{23}$$

We say that a vector y_1 is better than or preferred to y_2 if $y_1 \le y_2$ and $y_1 \ne y_2$. From definition (23) it follows that a point y_1 belongs to the Pareto set, if there are no points in Y, which are better than the point y_1. In the same way we can compare the points in decision space. We say that $x_1 \in X$ is better than $x_2 \in X$ if $F(x_1) \le F(x_2)$, $F(x_1) \ne F(x_2)$. In the last case we can say also that x_2 is worse than x_1.

A Pareto-optimal solution is efficient in the sense that none of the multiple objective functions can be downgraded without any other being upgraded. Any satisfactory design for the multiple objectives must be a Pareto-optimal solution.

The structure of X_* turns out to be very complicated even for the simplest problems. It often happens that this set is not convex and not simply connected, and every attempt to describe it with the help of approximation formulas is extremely difficult to realize. Therefore, in Evtushenko and Potapov (1984,1987) and Potapov (1984) we defined the new concept of an ε-optimal solution and gave a rule for finding it.

Definition 1. *A set $A \subseteq X$ is called an ε-optimal solution of problem* (21) *if*

1) *for each point $x_* \in X_*$ there exists a point $z \in A$ such that $F(z) - \varepsilon e \le F(x_*)$, e means the m-dimensional vector of ones;*

2) *the set A does not contain two distinct points x and z such that $F(x) \le F(z)$.*

Numerical methods for finding global extrema of functions of several variables can be used for constructing the ε-optimal solution of problem (21). Let $N_k = \{x_1, \dots, x_k\}$ be a set of k points in X. We shall define a sequence of sets $A_k \subseteq N_k$ as k increases, while trying in the final analysis to find ε-optimal solutions.

RULE 1 FOR CONSTRUCTING A_k. The set A_1 consists of the single point $x_1 \in N_1 \subseteq X$. Suppose that N_k, N_{k+1}, and A_k are known. We compute the vector $F(x_{k+1})$ at a point $x_{k+1} \in \in N_{k+1} \subseteq X$. Three cases are possible:

1. If it turns out that among the elements $x_i \in A_k$ there are some such that $F(x_{k+1}) \le F(x_i)$, $F(x_{k+1}) \ne F(x_i)$, then they are all removed from A_k, the point x_{k+1} is included in A_k and this new set is denoted by A_{k+1}.

2. If it turns out that there exists at least one element $x_i \in A_k$ such that $F(x_i) \le F(x_{k+1})$, then x_{k+1} is not included in A_k, and the set A_k is denoted by A_{k+1}.

3. If the conditions of the two preceding cases do not hold, then the point x_{k+1} is included in the set A_k, which is denoted now by A_{k+1}.

In the first case we exclude from A_k all points which are worse than the new point x_{k+1}. In the second case the new point x_i is worse than at least one point from A_k, therefore, this point is not included in the set A_k. In the last case the point x_{k+1} is equivalent (or equally preferred) to all points which belong to the set A_k.

The definition of the Lebesgue set (6) is replaced now by

$$K(\ell) = \{x \in X : \ell - \varepsilon e \le F(x), \;\; \ell \in \mathbb{R}^m, \;\; e \in \mathbb{R}^m, \;\; \varepsilon \in \mathbb{R}^1\}. \tag{24}$$

Theorem 6. *Let the finite set A_k of admissible points be such that*

$$X \subseteq \bigcup_{x_j \in A_k} K(F(x_j)). \tag{25}$$

Then the set A_k determined by the first rule for constructing A_k forms an ε-optimal solution of the multicriterial problem (21).

Proof. From the covering condition (25) it follows that the set X_* is also covered, i.e. for any point $x_* \in X_*$ there exists a point $x_s \in A_k$ such that $x_* \in K(F(x_s))$. Then from definition (24) we obtain $F(x_s) - \varepsilon e \le F(x_*)$. Because of arbitrariness of the point $x_* \in X_*$ we conclude that the set A_k forms an ε-optimal solution set of problem (21). □

In a manner similar to the second section we suppose that for any point $z \in X$ and level vector $\ell \in \mathbb{R}^m$, where $\ell \le F(z)$, it is possible to define the set

$$B(z, \ell) = \{x \in G(z) : \ell - \varepsilon e \le F(x)\}.$$

From the inclusion $B(z, \ell) \subseteq K(F(z))$ it follows that the set A_k is an ε-optimal solution if

$$X \subseteq \bigcup_{x_s \in A_k} B(x_s, F(x_s)).$$

We assume now that each component of the vector-valued function $F(x)$ satisfies a Lipschitz condition on X with one and the same constant L. Therefore, for any x and z in X, we have the vector condition

$$F(z) - eL\|x - z\| \le F(x).$$

In this case we can use the following covering balls

$$B(x_j, F(x_i)) = \{x \in \mathbb{R}^n : \|x - x_j\| \le r^i_{jk}\},$$

where $r^i_{jk} = [\varepsilon + h^i_{jk}]/L$, the index i is a function of indexes j, k and it is found as a solution of the following maximin problem

$$h^i_{jk} = \max_{x_c \in A_k} \min_{s \in [1:m]} (F^s(x_j) - F^s(x_c)) = \min_{s \in [1:m]} (F^s(x_j) - F^s(x_i)), \quad x_i \in A_k. \tag{26}$$

The inequality $h^i_{jk} > 0$ holds if there exists at least one point $x_c \in A_k$ such that $F(x_c) < F(x_j)$, otherwise $h^i_{jk} = 0$ and the radius of the covering ball is minimal, i.e. it is equal to $\rho = \varepsilon/L$.

As in the third section, we can take into account constraint restrictions. Suppose that problem (21) is replaced by the following problem

$$\text{global} \min_{x \in P \cap X} F(x). \tag{27}$$

Let X_* denote the set of global solutions of this problem. This set is defined by (22) with additional requirements: $x \in P \cap X$, $w \in P \cap X$.

We extend the admissible set by introducing the set $Z_\varepsilon = P \cap X_\varepsilon$, where X_ε is given by (15). The definition of an ε-optimal solution carries over to the case of problem (27) with the following change: instead of the condition $A \subset X$ it is required that $A \subset Z_\varepsilon$. The rule for determining A_k is also changed. Suppose that N_k is a set of k points belonging to P.

RULE 2 FOR CONSTRUCTING A_k. Assume that N_k, N_{k+1}, and A_k are known (A_k may be empty). At the point $x_{k+1} \in N_{k+1}$ it is checked whether $x_{k+1} \in Z_\varepsilon$. If not, then x_{k+1} is not included in A_k, and A_k is then denoted by A_{k+1}; otherwise, the same arguments as in the construction of A_k are carried out with a check of the three cases which were described above.

Denote

$$\bar{B}(x_j, F(x_j)) = \{x \in \mathbb{R}^n : \|x - x_j\| \le \bar{r}_{jk}\},$$
$$\bar{r}^i_{jk} = (1/L) \max[\varepsilon + h^i_{jk}, \Psi(x_j) - \delta],$$

where $0 < \delta < \varepsilon$ and i, h^i_{jk} are given by (26), the feasible set X is defined by (13).

Theorem 7. *Suppose that the set of global solutions X_* of the multicriterial problem* (21) *is nonempty. Assume that the vector-valued function F and the function Ψ satisfy a Lipschitz condition on P. Let the set N_k of points in P be such that*

$$P \subseteq \bigcup_{j=1}^{k} \bar{B}(x_j, F(x_i)). \tag{28}$$

Then the set A_k constructed by the second rule forms an ε-optimal solution of the multicriterial problem (27)*.*

The proof is very similar to that of Theorem 5 and, therefore, is omitted. Any radius $\bar{r}_{jk}$ cannot be less that the quantity $(\varepsilon - \delta)/L > 0$. Therefore, finite sets N_k and A_k satisfying the conditions of Theorem 7 exist in the case where the set P is compact. Theorem 7 is very interesting: it provides a simple rule for finding global ε-solution of a multicriterial problem. A set N_k satisfying condition (28) can be constructed by using diverse variants of the non-uniform covering method which were developed for finding a global extremum of a function of several variables. As before, local methods for improving the current ε-optimal solution N_k can be used. Some variants of local iterative methods for solving multicriterial problem are described in Zhadan (1988).

For an approximate solution of the problem, it suffices to implement a covering of an admissible set P by a non-uniform net. This is an essential advantage of such an approach in comparison with the well-known scalarization schemes. We mention for example the reference point approximation, the method of successive concessions, the method of inequality and other traditional methods which require, for their realization, a multiple global minimization of auxiliary functions on the feasible set X (see, for example Jahn, Merkel (1992)).

The set A_k obtained by computer calculations is transmitted to a user (a designer solving the multicriterial problem). The designer may wish to examine some or all the alternatives, make his tradeoffs analysis, make his judgment, or develop his preference before making rational decision. As the final solution, the user chooses a concrete point from the set A_k, starting from the specifics of the problem, or from some additional considerations not reflected in the statement of the problem (21).

The main result of this section is that, for the constructive solution of multicriterial optimization problems, it is possible to use the non-uniform covering technique developed in research on global extrema of functions of several variables. The approach presented here is developed in Evtushenko (1985), Evtushenko and Potapov (1984,1987), Potapov (1984).

Another approach to the solution of problem (27) can be obtained if we consider the function $\Psi(x)$ as an additional component of an outcome vector, which also should be minimized. We introduce extended multicriterial problem

$$\text{global} \min_{x \in P} \bar{F}(x), \tag{29}$$

where

$$\bar{F}^\top(x) = [F^1(x), \ldots, F^{m+1}(x)], \quad F^{m+1}(x) = \Psi(x).$$

For this problem the set of optimal solutions P_* is defined by the condition

$$P_* = \{x \in P : \text{if it exists } w \in P \text{ such that } \bar{F}(w) \leq \bar{F}(x), \text{ then } \bar{F}(w) = \bar{F}(x)\}.$$

For the solution of problem (27) we can use the method which we used for solving problem (21), where $X = P$. Now the sets N_k and A_k consist of points from the set P. If A_k is an ε-optimal solution of problem (29), then for each point $x_* \in P_*$ there exists a point $x \in A_k$ such that

$$F(x) - \varepsilon e \leq F(x_*), \qquad \Psi(x) - \varepsilon \leq \Psi(x_*).$$

The solutions set X_* of problem (27) belongs to the solutions set P_* of problem (29) because the latter includes the points x_* such that $\Psi(x_*) > 0$. Now the set of ε-optimal solutions can be found in similar way as for problem (21). A third variant of solution of problem (27) can be constructed by analogy with the approach described in Evtushenko (1985).

The multicriterial approach can be used for solving the nonlinear programming problem. In this case we introduce a bicriterial minimization problem. Let the vector-valued function $F(x)$ consist of two components: $F^\top(x) = [f(x), \Psi(x)]$. Instead of the original problem (12) with feasible set X defined by (13) we define problem

$$\text{global} \min_{w \in P} F(x).$$

The Pareto set for this problem coincides with the sensitivity function of problem (12). Using the approach described in this section, we obtain an ε-approximation of the sensitivity function.

5. Solution of a system of nonlinear equalities and inequalities

In this section we confine ourselves to the problem of finding a feasible point which belongs to the set X given by (19). To solve this problem approximately it suffices to find at least one point from the set

$$X_\varepsilon = \{x \in P : \Psi(x) \le \varepsilon\},$$

where $\Psi(x)$ is defined by (20).

We assume that further computations for sharpening the solution will involve local methods. When X_* is empty, the algorithm should guarantee that the assertion concerning the absence of approximate solutions be true. The problem of finding a point $x_* \in X_*$ is equivalent to the minimization of $\Psi(x)$ on P. Define

$$\Psi_* = \min_{x \in P} \Psi(x). \tag{30}$$

If X_* is nonempty, then $\Psi_* = 0$; otherwise, $\Psi_* > 0$. Suppose that the mappings $h(x)$, $g(x)$ satisfy a Lipschitz condition on P with constant L.

Now for finding an approximate solution of problem (30) we can use Theorem 5. For the sequence of points $\{x_k\}$ from P we use (5) to determine the record point x_r, we record the value R_k and define a covering ball B_{sk} with radius r_{sk}:

$$\begin{aligned} B_{sk} &= \{x \in \mathbb{R}^n : \|x - x_s\| \le r_{sk}\}, \\ r_{sk} &= [\Psi(x_s) - \min[R_k - \varepsilon, \delta]]/L, \qquad 0 < \delta < \varepsilon. \end{aligned}$$

The stopping rule of covering procedures consists only in verification of inequality $R_i \le \varepsilon$, $i = 1, 2, \ldots, k$. If for some s it is true, then $x_s \in X_\varepsilon$ and an approximate solution x_s is found, otherwise we have to cover all set P and find a minimal value of $\Psi(x)$ on P with accuracy ε.

Theorem 8. *Suppose that $\Psi(x)$ satisfies a Lipschitz condition and the set N_k of the points from P is such that $P \subseteq \bigcup_{i=1}^{k} B_{ik}$. Then*

1) if X_* is nonempty, then $x_r \in X_\varepsilon$;

2) if $R_k > \varepsilon$, then X_* is empty.

The proof follows from the above observations and is therefore omitted.

If we know that X_* is nonempty, then we can cover P by balls H_j which we defined in the second section and set $\Theta = f_* = 0$, $\delta_2 = \varepsilon$, $0 < \delta_1 < \varepsilon$, then $r_j = [\Psi(x_j) - \delta_1]/L$.

6. Solution of minimax problems

Let $f(x, y)$ be a continuous function of $x \in X \subset \mathbb{R}^n$, $y \in Y \subset \mathbb{R}^m$. We consider the minimax problem

$$f_* = \min_{x \in X} \max_{y \in Y} f(x, y). \tag{31}$$

Here we have internal maximization and external minimization problems. We can rewrite Problem (31) in the following equivalent way

$$f_* = \min_{x \in X} \varphi(x), \tag{32}$$

where $\varphi(x) = f(x, y)$, $y \in W(x)$, $W(x) = \arg\max_{y \in Y} f(x, y)$, i.e. $W(x)$ is the set of all solutions of the internal problem.

Denote $z^\top = [x^\top, y^\top] \in \Omega = X \times Y$ and $\bar{f}(z) = f(x,y)$. By Theorem 1.5.2 from Evtushenko (1985), if $\bar{f}(z)$ satisfies a Lipschitz condition on $\Omega = X \times Y$ with constant L, then the function $\varphi(x)$, defined by (32), also satisfies a Lipschitz condition with the same constant L. This property opens broad possibilities to use the method of finding global extrema of multiextremal function for the solution of minimax problems. The same method can be used sequentially for solving internal as well as external problems. The subprograms of local search usually are chosen differently, since the functions $f(x,y)$ are often differentiable and their local maximization is carried out using properties of smoothness of f in y. The function φ is only directionally differentiable, and it has to be locally minimized by other methods.

Comparing the problem (31) with the problem of finding the global extremum of f in z on Ω, we can conclude that (31) has an important advantage. Indeed, let the current record value of $\varphi(x_r)$ be known. If at some other point $x_s \in X$ we have to find the value of $\varphi(x_s)$, the process of maximization of f in y can be stopped as soon as at least one point $y_1 \in Y$ has been found such that $f(x_s, y_1) \geq \varphi(x_r)$ since in this case a priori $\varphi(x_s) \geq \varphi(x_r)$ and the knowledge of the exact value $\varphi(x_s)$ will not improve the current record value $\varphi(x_r)$. This property makes it possible in a number of cases to terminate the process of solving the internal problem.

Theoretically, this approach makes it possible to solve sequential minimax problems and opens the door to solving the discrete approximation of differential games. For example, the simplest Isaacs dolichobrachistochrone game problem was solved in Evtushenko (1985, see pages 463, 464). Full details of the first layerwise variant of the covering algorithm and codes in ALGOL-60 are given in Evtushenko (1972).

7. Computational results

In this section we present some computational experiments using the non-uniform covering techniques. The branch and bound algorithm was applied to a number of standard test functions. This algorithm requires an upper bound L_* for the Lipschitz constant L. At the very beginning we run the algorithm with some constants L_i which are much smaller than L_*. Incorrect (diminished) Lipschitz constants are used in order to find good record points and use them in subsequent computations with bigger Lipschitz constants. If we take $L_2 > L_1 \geq L_*$ and run the algorithm, taking $L = L_1$ and $L = L_2$, then the difference between record values must be less than ε. This condition is necessary but not sufficient for the Lipschitz constant L_1 to be greater or equal to the true value.

Everywhere for the local search we use the Brent modification of Powell method (Brent (1973)). The accuracy ε is fixed, we set $\varepsilon = 0.1$. In Tables 1 – 3 we give the Lipschitz constants L_i, the record values R_k, the record points x_r and k — the number of function evaluations.

The following simple examples illustrate the space covering approach.

Example 1. Griewank function (Torn (1989))

$$f(x) = (x^1)^2/200 + (x^2)^2/200 - \cos(x^1\sqrt{2})\cos(x^2\sqrt{2}) + 1,$$
$$x \in \mathbb{R}^2, \quad -2 \leq x^i \leq 3, \quad x_* = [0,0], \quad f(x_*) = 0, \quad x_0 = [3,3], \quad f(x_0) = 0.8851.$$

Table 1. Results for Griewank function

L	R_k	x_k^1	x_k^2	k
0.1	0.1349	-2.0000	-2.0000	100
0.2	0.0491	2.2104	2.2104	96
0.4	0.0491	2.2104	2.2104	26
0.8	0.0000	0.0000	0.0000	117
1.6	0.0000	0.0000	0.0000	362

Example 2. Goldstein-Price function (Goldstein (1971))

$$\begin{aligned} f(x) &= (1+(x^1+x^2+1))^2(19-14x^1+3(x^1)^2+3(x^2)^2-14x^2+6x^1x^2))\cdot \\ &\cdot\ (30+(2x^1-3x^2)^2(18-32x^1+12(x^1)^2+48x^2-36x^1x^2+27(x^2)^2), \\ &\quad x\in\mathbb{R}^2, \quad -2\le x^1\le 3, \quad -3\le x^2\le 2, \\ &\quad x_*=[0,-1], \quad f(x_*)=3, \quad x_0=[0,1], \quad f(x_0)=28611. \end{aligned}$$

Table 2. Results for Goldstein-Price function

L	R_k	x_k^1	x_k^2	k
100	3.0000	0.0000	-0.9999	114
180	3.0000	0.0000	-0.9999	4963

Example 3. Hartman function (Torn (1989))

$$f(x) = -\sum_{j=1}^{4} c^j e^{-[a_1^j(x^1-b_1^j)^2+a_2^j(x^2-b_2^j)^2+a_3(x^3-b_3^j)^2]},$$

where

$$\begin{aligned} &a_1=[3,0.1,3,0.1],\ a_2=[10,10,10,10],\ a_3=[30,35,30,35], \\ &b_1=[0.3689,0.4699,0.1091,0.03815], \quad b_2=[0.117,0.4387,0.8732,0.5743], \\ &b_3=[0.2673,0.747,0.5547,0.8828], \quad c=[1,1.2,3,3.2], \\ &x\in\mathbb{R}^3, \quad -2\le x^i\le 2, \quad i=1,2,3, \\ &x_*=[0.1146,0.5556,0.8526],\ f(x_*)=-3.8628,\ x_0=[2,2,-2],\ f(x_0)=0. \end{aligned}$$

Table 3. Results for Hartman function

L	R_k	x_k^1	x_k^2	x_k^3	k
0.1	-3.8628	0.1146	0.5556	0.8526	221
2	-3.8628	0.1146	0.5556	0.8526	6
4	-3.8628	0.1146	0.5556	0.85	1615

We also solved Branin and so-called Camel problems. Starting from various initial points, we found the global solution using the local method from the very beginning of computation. Therefore, these examples were not very interesting for illustration of global covering technique.

8. Conclusion

The non-uniform covering technique has given rise to numerous theoretical results and effective computational procedures for solving various global optimization problems. Recent developments indicate that these results can be generalized and extended significantly for parallel computations.

References

[1] Betro B. (1991), *Bayesian Methods in Global Optimization*, J. of Global Optimization 1, 1–14.

[2] Brent R. (1973), *Algorithms for minimization without derivatives*, Prentice-Hall, Englewood Cliffs.

[3] Evtushenko Yu. (1971), *Numerical methods for finding global extreme, (case of non-uniform mesh)*, U.S.S.R. Comput. Math. and Math. Phys. 11, 1390–1403 (Zh. Vȳchisl. Mat. i Mat. Fiz., 11, 1390–1403).

[4] Evtushenko Yu. (1972), *A numerical method for finding best guaranteed estimates*, U.S.S.R. Comput. Math. and Math. Phys. 12, 109–128 (Zh. Vȳchisl. Mat. i Mat. Fiz., 12, 109–128).

[5] Evtushenko Yu. (1974), *Methods for finding the global extremums in operations research*, Proceedings of Comp. Center of U.S.S.R. Academy of Sciences 4, Moscow, 39–68 (in Russian).

[6] Evtushenko Yu. (1985), *Numerical optimization techniques.* Optimization Software Inc., Publication Division, New-York.

[7] Evtushenko Yu. and Potapov M. (1984), *Global search and methods for solution of operational control problems*, VNIPOU, 128–152 (in Russian).

[8] Evtushenko Yu. and Potapov M. (1985), *Nondifferenliable Optimization: Motivation and Applications* (Laxenburg, 1984), Lecture Notes in Economics and Math. Systems 255, 97–102, Springer-Verlag, Berlin.

[9] Evtushenko Yu. and Potapov M. (1987), *Methods of numerical solutions of multicriterion problems*, Soviet Math. Dokl. 34, 420–423.

[10] Evtushenko Yu. and Potapov M. (1987), *Numerical solution of multicriterion problem*, Cybernetics and Computational Techniques 3, 209–218 (in Russian).

[11] Evtushenko Yu., Potapov M., Korotkikh V. (1992), *Numerical methods for global optimization*, in: Floudas, C.A. and Pardalos, P.M. (eds.), Recent Advances in Global Optimization, Princeton University Press.

[12] Evtushenko Yu. and Ratkin V. (1987), *Bisection method for global optimization*, Izvestija Akademii Nauk AN USSR, Tehnicheskaya Kibernetika, 119–128 (in Russian).

[13] Goldstein A., Price J. (1971), *On descent from local minima*, Mathematics of Computation 25, 569–574.

[14] Horst R. and Tuy H. (1990), *Global Optimization*, Deterministic Approaches, Springer-Verlag, Berlin.

[15] Jahn I. and Merkel A. (1992), *Reference Point Approximation Method for the Solution of Bicriterial Nonlinear Optimization Problems*, J. of Optimization Theory and Application 73, 87–104.

[16] Potapov M. (1984), *Non-uniform covering methods and their use in solving global optimization problems in a dialogue mode*, author's summary of candidate's dissertation, Moscow (in Russian).

[17] Ratschek H. and Rokne J. (1988), *New computer methods for global optimization*, Ellis Horwood Limited, Chichester.

[18] Torn A. and Zilinskas A. (1989), *Global Optimization*, Lecture Notes in Computer Science 350, Springer-Verlag, Berlin.

[19] Volkov E. (1974), *Approximate and exact solutions of systems of nonlinear equation*. Proceedings of Mathematical Institute of USSR Academy of Sciences 131, Moscow (in Russian).

[20] Zhadan V. (1988), *An augmented Lagrange function method for multicriterion optimization problems*, U.S.S.R. Comput. Math. and Math. Phys. 28, 1–11 (Zh. Vȳchisl. Mat. i Mat. Fiz., 28, 1603–1618).