сообщения
объединенного
института
ядерных
исследований
дубна

E11-92-218

V.P.Gergel*, V.A.Grishagin*, V.V.Kukhtin,
E.A.Rogatneva*, R.G.Strongin*, I.N.Vysotskaya*

COMPLICATED PROBLEM SOLUTION TECHNIQUES IN OPTIMAL PARAMETER SEARCHING

*Scientific Research Institute of Applied Mathematics and Cybernetics,
Nizhnij Novgorod State University, Russia

1992

1. INTRODUCTION

Optimal decision problems including approximation (restoration) of dependences by experimental results, optimal design problems, solution of nonlinear equation systems etc. are reduced to the solution of complicated optimization problems of one or several functions usually named as efficiency criteria which define a "quality" of the decision obtained. In this case additional conditions described by corresponding functions-constraints may be imposed on varying parameters (arguments of functions).

A full analysis of an object on the basis of a comparison of all possible variants with an aim of a selection of the best ones is practically inaccessible, because real objects and their mathematical models are extremely complicated (nonlinear, nonmonotonous, multiextremal and so on) and a number of variants is very great. Therefore, effective procedures of economical teleological search of optimal decision are required to solve optimization problems. A main point of the teleological search is the analysis only a small part of variants in order to eliminate many a priori unpromising cases and to concentrate a further search in a subset containing the best variant. One of such teleological search is a local search which is based on an assumption of a definite monotony of optimized criteria mathematically expressed by linearity, a quadratic form, convexity and so on. However, if constraint functions exhibit no monotony features, local-optimal solutions turn out to be inadequate.

In this publication we suggest the procedures of a global search for numerical solution of multidimensional multiextremal multicriteria optimization problems with complicated constraints. The procedures suggested provide teleological search assuming boundedness of object characteristic changes at restricted changes of its parameters (Lipschitz condition).

2. CLASSES OF PROBLEMS TO BE SOLVED

Typical problems of optimal decisions corresponding to complicated multiparametric multicriteria optimization problems with multiextremal criteria and nonconvex constraints are the following:

1) the identification of object model characteristics by experimental data;

2) the approximation of given characteristics of studied or designed object;

3) the selection of parameters which are referred to optimal (including over several criteria) object characteristics under conditions of object normal functioning. These parameters are defined by linear inequalities relating feasible values of parameters.

The identification of object model characteristics by experimental data can be reduced to a minimization problem of the difference between an object model and experimental data. The dependence of this difference on varying parameters is as usual nonmonotonous and multiextremal (oscillating). That is why an accurate identification needs application of global optimization method with nonlinear constraints (if modal parameters should satisfy relations of equality and inequality types). An additional complication arises in the case, when the identification comprises a number of possible modes of the object functioning or a number of condition variants of its functioning (or simultaneously both). An account of the variety mentioned leads to minimax optimization problems, internal maximum of the difference between the response and the experimental data being taken over all possible solutions and conditions of functioning, i.e. an estimate of the difference corresponds to the worst possible case, therefore, this internal maximum is to be calculated by global methods, otherwise the corresponding estimate loses the meaning of a guaranteed result.

The approximation problems arising in designing and modelling objects with the given characteristics are also multiextremal ones, i.e. the residual of the given and model characteristics depends monotonously on the model parameters. Multifunctionality of the object and the variety of its functioning lead as in the case of the identification to minimax optimization problems. Model parameter functional relations accounted as constraints of the optimization problem are also possible.

Optimal decision in designing complicated object includes as usual estimates of a number of criteria of its efficiency (in

different modes and conditions).In relation to some criteria a requirement not to exceed the given tolerances is a sufficient one. Other criteria require to succeed their extremal values. As far as the criteria are as usual conflicting (an improvement of some of them leads to a degradation of others), a selection of the best variant is based on a compromise.

The complexity of concrete optimization problems from the considered class may sufficiently differ (a number of criteria and constraints, a number of varying parameters, computational complexity of functionals entering in criteria and constraints). Each concrete problem from the class under consideration is related to one of the following types:

- multiextremal problem with a single criterion of efficiency, with convex or linear-unimodal [1] constrains and a few number (up to five) of varying parameters (a particular problem);
- multiextremal multicriteria problem with complicated constraints and a number of varying parameters up to ten (a general problem).

Such a division makes possible to fit the complexity of applied algorithms with the complexity of concrete optimization problems.

3. DESCRIPTION OF APPROACH TO OPTIMAL DECISION PROBLEMS

3.1. Optimization problem formal statement

3.1.1. Object model

A mathematical model of the optimization object is characterized by a set of parameters

$$y = (y_1, \ldots, y_N) \tag{1}$$

and by the vector function of characteristics

$$W(y) = (W_1(y), \ldots, W_n(y)) \quad , \tag{2}$$

defined in such a way, that a decrease of characteristic values corresponds to a better decision.

It is assumed that:

1) vector coordinates $y = (y_1, \dots, y_N)$ may vary in the given limits defined by a start and an end vectors

$$a = (a_1, \dots, a_N), \quad b = (b_1, \dots, b_N), \tag{3}$$

which set hyperinterval

$$D = \{ y \in R^N : a_i \le y_i \le b_i, \ 1 \le i \le N \} \tag{4}$$

of vector y possible values;

2) each characteristic $w_i(y)$, $1 \le i \le n$, is Lipschitz function with corresponding Lipschitz constant, which may not be given.

3.1.2. Functional constraints

In relation to a part of coordinate functions w_j from (2), the numbers of which are defined by set

$$G = \{ j_1, \dots, j_m \} \subset \{ 1, \dots, n \}, \tag{5}$$

a condition is stated on decreasing their values up to some given tolerances q_i, i.e. functional constraints are imposed on possible values of set y from (1)

$$W_{j_i}(y) \le q_i, \quad j_i \in G. \tag{6}$$

Sets $y \in D$ satisfying conditions (6) are called feasible. Symbols

$$g_i(y) = W_{j_i}(y) - q_i, \quad 1 \le i \le m, \tag{7}$$

make possible to transform inequality (6) in equivalent form

$$g_i(y) \le 0, \quad 1 \le i \le m. \tag{8}$$

Set

$$G = \{ y \in D : g_i(y) \le 0, \ 1 \le i \le m \} \tag{9}$$

is called a set of feasible variants or a feasible area. When $m = 0$, i.e. at the absence of constraints, $Q = D$.

In the considered class of general multiextremal problems there are concrete problems, for which one can point out a priori some subareas of area D from (4) which do not contain optimal decisions. The following system of area D subsets can be used for an approximate description of such subareas which are excluded from consideration in the decision-making process.

Hyperinterval D is divided by N hyperplanes parallel to coordinate hyperplanes and passing through median points of hyperinterval D edges, orthogonal to these planes, on 2^N hyperintervals. Let us callthese hyperintervals as hyperintervals of the first division. Each hyperinterval of the first division is divided by N different planes passing through median points of orthogonal edges. As a result, each hyperinterval of the first division is divided on 2^N hyperintervals of the second division. By analogy one can form hyperintervals of any μ-th division.

Each hyperinterval is identified by N-tuple

$$< \mu; \nu_1, \dots, \nu_N > \tag{10}$$

and is designated via $D(\mu; \nu_1, \dots, \nu_N)$, where μ is a number of the division containing the given hyperinterval, ν_i, $1 \le \nu_i \le 2^\mu$, $1 \le i \le N$, is a number of subinterval in line segment $[a_i, b_i]$.

A list of hyperintervals of the type mentioned

$$S = \{ < \mu(l); \nu_1(l), \dots \nu_N(l) > : 1 \le l \le L \}, \tag{11}$$

approximating the subarea of area D , which a priori does not contain optimal dicisions, is a tool of fixation of additional constraints. In this way one can define "reduced" area

$$\mathcal{D}_S = \mathcal{D} \setminus \bigcup_{\ell=1}^{L} \mathcal{D}(\mu(\ell); \nu_1(\ell), \dots, \nu_N(\ell)), \tag{12}$$

in which the optimization problem is to be solved. Let us name list S from (11) as a list of forbidden subintervals.

3.1.3. Vector criterion of efficiency

A part of coordinate functions w_i from (2), which numbers are defined by set

$$F = \{ i_1, \dots, i_k \} \subset \{ 1, \dots, n \}, \tag{13}$$

is considered as vector criterion of efficiency

$$f(y) = (f_1(y), \dots, f_k(y)), \tag{14}$$

where

$$f_j(y) = w_{i_j}(y), \quad j \in F. \tag{15}$$

If k > 1, the model and the relating optimization problem are called as multicriteria and the case k = 1 corresponds to scalar (single-criterion) problem of optimal decision.

3.1.4. Optimal decision

In minimization problem of efficiency vector criterion (14) on feasible set Q denoted as

$$f(y) \to \min, \quad y \in Q, \tag{16}$$

partial criteria are as usual conflicting. A substitution of vector y' by vector y'' may lead at the same time to the de-

crease of partial criterion f_i and to the increase of other partial criterion f_j :

$$f_i(y'') < f_i(y'), \quad f_j(y'') > f_j(y'). \tag{17}$$

In this connection a notion of effective non-improvable decision (see [2]) is used when characterizing various solutions of problem (16).

Vector f(y*) is called an effective one (non-improvable, optimal according to Pareto), if for any $y \in Q$ from conditions

$$f_i(y) \le f_i(y^*), \quad 1 \le i \le k, \tag{18}$$

it is followed that f(y) = f (y*). I.e. set Q has no decisions y exceeding y* by some partial criteria and conceding y* at the same time by any partial criterion. Points $y^* \in Q$, which correspond to effective vectors, are also will be named as effective points.

If all criteria are positive in area Q , i.e.

$$f_i(y) > 0, \quad 1 \le i \le k, \quad y \in Q, \tag{19}$$

point $y^* \in Q$ is effective in that and only that case when there exists weight vector

$$\lambda = (\lambda_1, \dots, \lambda_k), \tag{20}$$

$$\sum_{i=1}^{k} \lambda_i > 0; \quad \lambda_i \ge 0, \quad 1 \le i \le k, \tag{21}$$

that

$$y^* = y^*_\lambda = \arg\min_{y \in Q} \left(\max_{1 \le i \le k} \lambda_i f_i(y) + \alpha \sum_{i=1}^{k} f_i(y) \right), \tag{22}$$

where positive coefficient α is sufficiently small. Coordinates λ_i of vector λ can be considered as coefficients of partial criteria importance.

We shall call effective point $y^* = y^*_\lambda$, the solution of scalar optimization problem (22) at given vector λ , as the partial solution of multicriteria problem (16) corresponding to weight vector λ . We shall call a set of all effective points (Pareto set) as a complete solution of multicriteria problem (16).

For $k = 1$, vector criterion $f(y)$ is transformed into the scalar one : $f(y) = f_1(y)$ and problem (16) becomes a nonlinear programming problem. In this case, according to (21), $\lambda_1 > 0$, and a set of effective points consists of points of function $f_1(y)$ global minimum on set Q .

3.2. Estimate methods of optimal parameters

The estimate method of multicriteria or scalar global-optimal decision is based on the successive reduction of the initial multidimensional multicriteria problem with nonlinear non-convex constraints to a set of scalar one-dimensional problems. A possibility of such a reduction is secured by decomposition schemes of the vector problem to scalar ones [2] , by special modifications of constraint accounting schemes and by schemes of dimension reduction. In this case different schemes of dimension reduction are used for general and partial problems as well as different algorithms for the solution of reduced (one-dimensional) sub - problems. [3]

All algorithms used have been theoretically investigated and they provide a convergence to a global solution of corresponding multiextremal problem with constraints.

For the solution of scalar unimodal problems or the search of local-optimal decisions in scalar multiextremal problems it is assumed to apply the well known and sufficiently effective method of golden section [4] in modification of [5] and Helder-Meed method of deformable polyhedron[6].

3.2.1. Reduction of multicriteria problem to scalar optimization problem

A search problem of vector criterion effective point (2) corresponding to some value of vector λ from (20) is reduced for the case (19) to the scalar solution of problem

$$F(y) \to min, \quad y \in Q, \tag{23}$$

where

$$F(y) = \max_{1 \le i \le k} \lambda_i f_i(y) + \alpha \sum_{i=1}^{k} f_i(y), \tag{24}$$

i.e. the problem of global minimization of convolution (24) of vector criterion (2) on set Q from (9) (α - sufficiently small parameter). If (16) is a one-criterion problem, then $\lambda_1 = 1$, $\alpha = 0$ in (24), i.e. $F(y) = f_1(y)$.

In the case of one-criterion statement of the general problem (16), problem (23) coincides with that of (16) and their solutions coincide. If the problem statement is a multicriteria one, then the complete solution of (16) is attained by the solution of basic problem set with different λ .

3.2.2. Constraint account

The constraint account is realized by penalty function method[7]. According to this method problem (23) is substituted by a solution of problem

$$\Phi(y) \to min, \quad y \in D, \tag{25}$$

where

$$\Phi(y) = F(y) + c \cdot G(y). \tag{26}$$

Penalty constant $C > 0$, and penalty function $G(y)$ is such, that

$$G(y)=0,\ y\in Q,\qquad G(y)>0,\ y\notin Q. \tag{27}$$

In solving the general problem function $G(y)$ has the following form

$$G(y)=\sum_{j=1}^{m}\gamma_j Q_j(y), \tag{28}$$

$$Q_j(y)=|g_j(y)|+g_j(y),\quad 1\le j\le m, \tag{29}$$

and weight coefficient vector $\gamma=(\gamma_1,\ldots,\gamma_m)$ belongs to simplex Γ :

$$\Gamma=\{\gamma\in R^m:\gamma_j\ge 0,\ 1\le j\le m,\ \sum_{j=1}^{m}\gamma_j=1\}. \tag{30}$$

Weight coefficients γ_j, $1\le j\le m$, should "flatten" an influence of different constraints $g_j(y)$ on the penalty function.

In solving a partial problem, function $G(y)$ has the form:

$$G(y)=\max\{0;\ g_1(y),\ldots,g_m(y)\}. \tag{31}$$

3.2.3. Dimension reduction

The solution of multidimensional (a number of varying parameters $N>1$) problems, in which it is required to find a global-optimal solution, is carried out by reduction of initial multfdimensional problem (23) to a set of more simle one-dimensional subproblems.

Different schemes are used for reduction of general and partial problems.

The general problem is reduced to one-dimensional one with a help of a curve called as Peano curve(or development) [3]. It has been proved the existance of curves $y(x)$ given by continuous coordinate functions $y_i(x)$, $x\in[0,1]$, $1\le i\le N$, which make single-valued mapping line segment $[0,1]$ onto the hyperinterval from (4) , i.e.

$$\mathcal{D}=\{y\in R^N:\ a_i\le y_i\le b_i,\ 1\le i\le N\}=\{y(x),\ 0\le x\le 1\}. \tag{32}$$

Such curves make possible to reduce the multidimensional minimization problem of continuous function $\varphi(y)$ in area D to the one-dimensional minimization problem of continuous function $\varphi(y(x))$ in line segment $[0,1]$, so far as, according to (32)

$$\min\{\varphi(y):\ y\in\mathcal{D}\}=\min\{\varphi(y(x)):\ x\in[0,1]\}. \tag{33}$$

In solving the partial problem, a multistep optimization procedure is used as a reduction scheme ("nested" optimization scheme), which, by example of solving problem (23) without constraints, i.e. when $Q=D$, consists in the following.

The following relation is true for continuous function $F(y)$

$$\min_{y\in\mathcal{D}}\mathcal{F}(y)=\min_{a_1\le y_1\le b_1}\ \min_{a_2\le y_2\le b_2}\ \ldots\ \min_{a_N\le y_N\le b_N}\mathcal{F}(y). \tag{34}$$

Let us introduce functions

$$\mathcal{F}_N(y)=\mathcal{F}(y),\qquad \mathcal{F}_i(u_i)=\min_{a_{i+1}\le y_{i+1}\le b_{i+1}}\mathcal{F}_{i+1}(u_{i+1}), \tag{35}$$

where vector $U_i=(y_1,\ldots,y_i)\in R^i$, $1\le i\le N$. Then, the solution of problem (23) is reduced according to (34) to the solution of one-dimentional problem

$$\mathcal{F}_1(y_1)\to\min,\qquad y_1\in[a_1,b_1], \tag{36}$$

in which every calculation of function $F_1(y_1)$ at given y_1 is the solution of one-dimensional problem

$$F_2(y_1, y_2) \to \min, \quad y_2 \in [a_2, b_2]. \tag{37}$$

In its turn, each calculation of function $F_2(y_1, y_2)$ at fixed y_1 and y_2 assumes the solution of one-dimensional problem

$$F_3(y_1, y_2, y_3) \to \min, \quad y_3 \in [a_3, b_3] \tag{38}$$

and so on. Thereby, the solution of multidimensional problem (23) is reduced to the solution of a set of "nested" one-dimensional subproblems.

Multistep procedure (34) can also be expand to the problems with constraints. Such a development of the multistep procedure is described in detail in [1].

3.2.4. Solution of reduced problems

One-dimensional problems of types (33), (36)-(38) are solved using an information-statistical algorithm of a global search from [3]: $\psi(x) \to \min, \quad x \in [a, b]$.

The algorithm executes successively iterations in points of the most probable location of the global minimum, the points being estimated by corresponding interpretation of trial results in the frames of some stochastic model.

The first trial is carried out in arbitrary point $x_1 \in (a,b)$. The point of any subsequent $(k+1)$-th trial is defined by expression

$$x^{k+1} = \frac{x_t + x_{t-1}}{2} - \begin{cases} (z_t - z_{t-1})/2r\mu, & 1 < t \leq k, \\ 0, & t = 1,\ t = k+1, \end{cases} \tag{39}$$

$$z_i = \psi(x_i),$$

number t being determined from condition

$$R(t) = \max\{R(i) : 1 \leq i \leq k+1\}, \tag{40}$$

where for $1 \leq i \leq k$

$$\Delta_i = x_i - x_{i-1}, \tag{41}$$

$$R(i) = \Delta_i + (z_i - z_{i-1})^2/\mu^2\Delta_i - 2(z_i + z_{i-1})/r\mu, \tag{42}$$

$$R(1) = 2\Delta_1 - 4z_1/r\mu, \tag{43}$$

$$R(k+1) = 2\Delta_{k+1} - 4z_k/r\mu, \tag{44}$$

$$\mu = \max\{|z_i - z_{i-1}|/\Delta_i : 1 < i \leq k\}. \tag{45}$$

In case when (45) has zero value, as well as at $k = 1$, it is assumed that $\mu = 1$. Number r from (39), (42)-(44) is the method parameter, the following inequality is to be valid : $r > 1$.

Convergence conditions

If this algorithm is applied for minimization of Lipschitz (with constant K) function ψ, then for any limit point x' of sequence $\{x^k\}$ generated by this algorithm, the following is true [3] :

1) point x' is if only local-optimal point of function ψ, if this function has a finite number of local minima ;

2) if there is other point x'' of sequence $\{x^k\}$, then $\psi(x'') = \psi(x')$, i.e. a simultaneous convergence to different

function values is impossible and, therefore, the method generates a non-uniform net while minimizing functions distinct from a constant;

3) $z^k = \Psi(y^k) \geqslant \Psi(y')$, $k = 1,2, \ldots$,i.e. the algorithm cannot generate convergence to points, where the function value exceeds the result of some realized trial ;

4) if at some step

$$r\mathcal{M} > 2k,$$

then y' is a point of function Ψ global minimum and, moreover, the set of all limit points of sequence $\{y^k\}$ coincides with the set of points of function Ψ global minimum.

3.3. Illustra tions

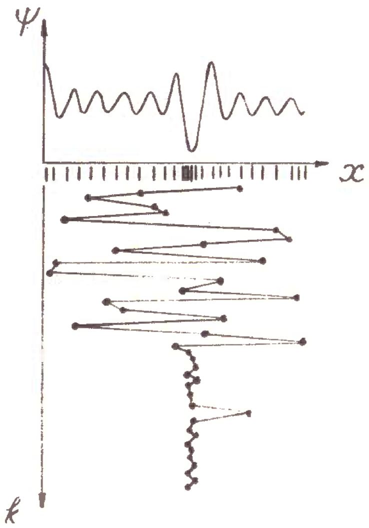

Fig.1

The diagram of Fig.1 illustrates the solving process of a concrete example with a help of the method described (when $r = 1,4$). Vertical dashes under abscissa axis give points of 42 trials, which form an essentially non-uniform net crowded in the neighbourhood of the global minimum . Segments of the lower broken line connect successively (from up to down) points (marked by dark circles) corresponding to pairs (y^k, k), (y^{k+1}, $k+1$), where y^k is coordinate , k - number of the iteration.

Fig. 2

Fig.2 illustrates results of minimizing trial function $\varphi(y_1,y_2)$ (construction procedure of such functions is described in [8]) with a help of the algorithm described. The figure shows: search area (quadrate), lines of constant function level (function levels are noted), points of 131 trials (dark circles).

As it is followed from the figure, points of successive trials generated by the algorithm form essentially non-uniform net crowding in the neighbourhood of global optimum.

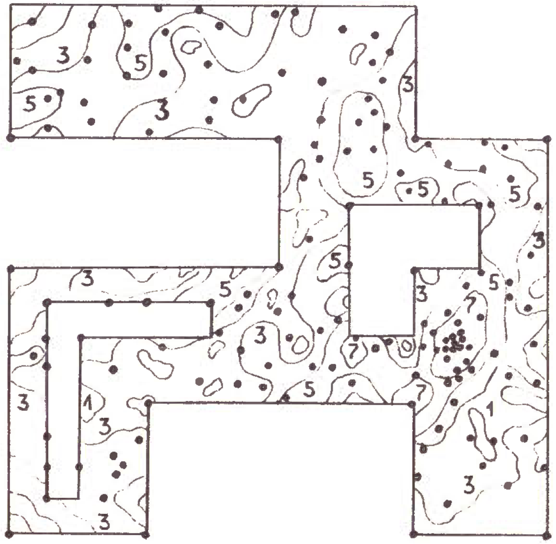

Fig. 3

Fig.3 illustrates results of the algorithm application for maximizing function $\varphi(y_1, y_2)$, belonging to the same class as the function in Fig.2, in not simply connected area D_s (non-hatched part of the quadrate). The figure shows function constant level lines and points of 150 trials (dark circles).

3.4. Computing aspects

The following computing aspects are accounted in computer implementation of supposed methods. It is seen from the decisive rule of the algorithm that values (estimates of differences Δ_i from (41), characteristics R(i) from (42)- (44) and so on) used at every search step change very seldom (or do not change at all), that makes possible to avoide their recalculation by memorizing once calculated values. In solving multidimentional problems the number of iterations may be sufficiently large that impedes storing and collating points x_i corresponding to iterations of the problem and related values z_i, Δ_i, R(i) and so on .

This information may be presented in the form of a matrix, which columns, for example, are :

$$(x_i, \Delta_i, z_i, R(i))^T, \qquad 1 \le i \le k+1. \tag{46}$$

When storing this matrix is divided onto units located in direct access memory segments. Hereby, an insertion of a new column corresponding to point x^{k+1} requires sorting only in segment limits. Characteristic values corresponding to unit columns make possible to forcast its "pumping" in immediate-access memory at the current step of the optimum search. When maximum feasible dimension is succeeded (as a result of column insertion), the unit is divided onto two new ones [9] .

Described methods have been practically approved in solving many applied problems, for example computer-aided design problems of aircraft, digital data transfer systems, power semiconductor devices, new technology products, bandpass filter circuits, in solving nonlinear algebraic equation systems [10] .

REFERENCES

1. Grishagin V.A., Strongin R.G. Multiextremal function optimization at monotonous unimodal constraints // Izv.AN USSR, Technicheskaya Kibernetika, 1984, No.4 (In Russian).

2. Germeier Yu.B. Introduction into the theory of operations research.M.: Nauka,1971 (in Russian).

3. Strongin R.G. Numerical methods in multiextremal problems (Information-statistical algorithms).-M.: Nauka,1978 (in Russian).

4. Wild D.J. Methods of extremum search. M.:Nauka,1967(in Russian).

5. Rogatneva E.A. Unimodal function minimum(maximum) calculation i bounded and unbounded areas using golden section method (Fortran subprogram) // Algorithms and programs. Information bulletin of GFAP USSR. M.: VNTICenter,1979,No.5(31).(in Russian)

6. Himmelblau D. Applied nonlinear programming. M.:Mir,1975

7. Fedorov V.V. Numerical methods for maxminumum problems. -M.: Nauka,1979 (in Russian).

8. Grishagin V.A. Operational characteristics of some global search algorithms / Random search problems.-Vyp.7.-Riga: Zinatne,1978,pp.198-206. (in Russian)

9. Strongin R.G., Gergel' V.P. Database for solving complicated optimization problems in computer-aided design system of electr technical devices / Proceedings of All-Union scientific-technical seminar on integration of packages and databases of compute aided design systems of electrotechnical devices. Tallinn: NII TEZ,1982,pp.125-130. (in Russian).

10. Vysotskaya I.N.,Strongin R.G. Method of nonlinear equation solution using a priori probability root estimates. Journal of calculus mathematics and mathematical physics, 1983,V.23, No.1,pp.3-12. (in Russian).

Received by Publishing Department
on May 25, 1992.