International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

Mathematical Methods for the Analysis of Hierarchical Systems. 1. Problem Formulation, and Stochastic Algorithms for Solving Minimax and Multiobjective Problems

Ereshko, F., Fedorov, V. and Zavriev, S.K.

IIASA Collaborative Paper
May 1984

NOT FOR QUOTATION
WITHOUT PERMISSION
OF THE AUTHOR

MATHEMATICAL METHODS FOR THE ANALYSIS OF HIERARCHICAL SYSTEMS
I. PROBLEM FORMULATION, AND STOCHASTIC ALGORITHMS FOR SOLVING MINIMAX AND MULTIOBJECTIVE PROBLEMS

F.I. Ereshko
V.V. Fedorov
S.K. Zavriev

May 1984
CP-84-19

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
2361 Laxenburg, Austria

PREFACE

This is the first of two papers dealing with mathematical methods that can be used to analyze hierarchical systems.

In this paper, the authors look at the situation that arises when certain decision-making powers are delegated to various elements within a hierarchical structure. It is found that these elements inevitably begin to operate in accordance with their own interests, which are not necessarily those of the system as a whole. Thus we have the problem of how to distribute the decision-making functions between the central body and the other parts of the system in such a way that the efficiency of the control system is maximized with respect to the global criterion.

The authors take a game-theoretical approach to this problem, looking first at two-level hierarchical systems and using Germeyer's games as a model. They derive a number of methods for solving the problem thus formulated, and give some numerical results obtained using two of the resulting algorithms.

ANDRZEJ WIERZBICKI
Chairman
System and Decision Sciences

MATHEMATICAL METHODS FOR THE ANALYSIS OF HIERARCHICAL SYSTEMS I. Problem Formulation, and Stochastic Algorithms for Solving Minimax and Multiobjective Problems

F.I. Ereshko, V.V. Fedorov** and S.K. Zavriev***

*Computing Center of the USSR Academy of Sciences, Moscow, USSR
**Moscow University, Moscow, USSR

1. INTRODUCTION

Hierarchical control systems form one of the most interesting classes of large systems with regard to theoretical and practical applications. Hierarchical control problems were first formulated in connection with the need to distribute the right to process information and the responsibility for making decisions among the various elements of the control system. Problems arise due to the fact that, when different elements of the system have these rights and responsibilities and can exercise them independently, these elements inevitably begin to operate according to their own interests, which generally differ from the global objectives of the system. Thus it is necessary to distribute the decision-making functions between the central body and the separate parts of the system in such a way that the efficiency of the control system is maximized with respect to the global criterion (we shall assume that this criterion coincides with that of the central body). This problem may be divided into two parts [1–3]: the problem of analysis, i.e., the choice of a reasonable control for a fixed hierarchical system, and the problem of synthesis, i.e., the choice of the best structure for the control system.

Game theory seems to provide the best approach to such problems. However, traditional game theory does not consider a number of questions which arise in this particular case, e.g., how to deal with problems caused by the sharing of information between different elements in the hierarchy, priorities in decision making, and lack of knowledge of the objective function by some elements. We shall therefore begin by introducing a class of games in which moves are taken in a fixed order and the process of information transfer is quite similar to that found in some hierarchical systems.

2. HIERARCHICAL TWO-PERSON GAMES

Hierarchical two-person games describe the simplest two-level hierarchical system. This is the most thoroughly investigated hierarchical structure, and is of considerable importance. Let the objective of player 1 (representing the upper level of the hierarchy) be to increase the value of the criterion $F(x,y)$ using decision variable $x \in X$, and the objective of player 2 be to increase the value of the criterion $G(x,y)$ using decision variable $y \in Y$. The principle behind the second player's move is to maximize his gain, given that the outcome depends on his action only.

It is assumed that player 1 has the first move and knows the principle on which the second player will act, as well as being acquainted with F, G, X, Y.

There are various formulations of the games now known as Germeyer's games [1] which depend on the information available to player 1 about the decision of player 2.

Game G_1. Player 1 will not have any information on the choice made by player 2: his strategy is to choose a certain $x^1 \in X$ and report it to player 2.

Then the best guaranteed result of player 1 is

$$v_1 = \sup_{x^1 \in X^1} \inf_{y^1 \in B^1(x^1)} F(x^1, y^1) \quad , \tag{2.1}$$

where

$$B^1(x^1) = \{y^1 \in Y^1 \mid G(x^1, y^1) = \max_{z \in Y^1} G(x^1, z)\} \, , \quad X^1 = X \, , \quad Y^1 = Y \quad .$$

Game G_2. Player 1 will know the choice $y^2 \in Y$ made by player 2: his strategy is to choose the mapping $\tilde{X}_2 = \{\tilde{x}_2 : Y \to X\}$.

The best guaranteed result of player 1 is

$$v_2 = \sup_{\tilde{x}^2 \in \tilde{X}^2} \inf_{y^2 \in B^2(\tilde{x}^2)} F(\tilde{x}^2, y^2) \quad . \tag{2.2}$$

Game G_3. Player 2 formulates his action as a function $y(x)$, i.e., he chooses a mapping $\tilde{y}_3 \in \tilde{Y}^3 = \{\tilde{y}^3 : X^1 \to Y\}$. Player 1 has the first move and since he will know $\tilde{y}_3$ he reports to player 2 the mapping $\tilde{x}^3$ which is an element of the set $\tilde{X}^3 = \{\tilde{x}^3 : \tilde{Y}^3 \to X^1\}$.

The best guaranteed result of player 1 in such games is

$$v_3 = \sup_{\tilde{x}^3 \in \tilde{X}^3} \inf_{\tilde{y}^3 \in B^3(\tilde{x}^3)} F(\tilde{x}^3, \tilde{y}^3) \quad .$$

In games G_2 and G_3 the sets of multivalued mappings $B^2(\tilde{x}^2)$ and $B^3(\tilde{x}^3)$ are defined (like $B^1(x^1)$) as the sets of possible answers of the second player, given that the strategy of the first player is fixed.

Increasing the number of iterations we can formulate games G_{2n}, G_{2n+1}, $n \geq 2$.

The sets of players' strategies in game G_{2n} are

$$\tilde{X}^{2n} = \{\tilde{x}^{2n} : \tilde{Y}^{2n} \to \tilde{X}^{2n-2}\} , \quad \tilde{Y}^{2n} = \{\tilde{y}^{2n} : \tilde{X}^{2n-2} \to \tilde{Y}^{2n-2}\}$$

and the best guaranteed result of player 1 is

$$v_{2n} = \sup_{\tilde{x}^{2n} \in \tilde{X}^{2n}} \inf_{\tilde{y}^{2n} \in B^{2n}(\tilde{x}^{2n})} F(\tilde{x}^{2n}, \tilde{y}^{2n}) \quad .$$

In game G_{2n+1} we have

$$\tilde{X}^{2n+1} = \{\tilde{x}^{2n+1} : \tilde{Y}^{2n+1} \to \tilde{X}^{2n-1}\} , \quad \tilde{Y}^{2n+1} = \{\tilde{y}^{2n+1} : \tilde{X}^{2n-1} \to \tilde{Y}^{2n-1}\}$$

$$v_{2n+1} = \sup_{\tilde{x}^{2n+1} \in \tilde{X}^{2n+1}} \inf_{\tilde{y}^{2n+1} \in B^{2n+1}(\tilde{x}^{2n+1})} F(\tilde{x}^{2n+1}, \tilde{y}^{2n+1}) \quad ,$$

where

$$B^k(\tilde{x}^k) = \{\tilde{y}^k \in \tilde{Y}^k \mid G(\tilde{x}^k, \tilde{y}^k) = \max_{z \in \tilde{Y}^k} G(\tilde{x}^k, z)\} \quad .$$

The following relationships hold for $n \geq 2$ [4]:

$$v_{2n} = v_2 , \quad v_{2n+1} = v_3 \quad .$$

Thus from the point of view of player 1 there is no point in having a strategy more complicated than in games G_1, G_2, G_3. In other words, the first three games can be regarded as basic and we shall confine ourselves to a consideration of these games only.

Games G_1, G_2, G_3 have a natural economic interpretation in the framework of the "Center–Producer" system [5].

1. The setting of prices x^1 for the output y of the producer. The natural approach here is game G_1, as in this case prices are chosen without any information about y.

2. Decisions on fixed payments x^2 (subsidies, premiums, assignments and so on). As accounts with the producer are settled on receiving the final product, he may be informed beforehand of the chosen system of fixed payment (i.e., how the amount paid depends on the results of his work). Here we have game G_2 on the set of strategies $\tilde{X}^2$.
3. Allocation of resources x^3 (raw material, equipment, labor and so on). It is obvious that resources must be allocated before the production process begins, and formally the producer has the right to dictate his terms: $\tilde{y}^3 = y(x^3)$. However, since the center has the first move he may report his strategy as the mapping $\tilde{x}^3$: $\tilde{Y}^3 \to X$. This is a typical G_3 formulation, although game G_1 is also possible here. The guaranteed result of player 1 in games G_1, G_2, G_3 satisfies the relationship $v_1 \le v_3 \le v_2$, and thus the allocation of resources to the producer in a game G_3 formulation is more profitable to the center than in G_1.

3. ANALYSIS OF TWO-LEVEL HIERARCHICAL SYSTEMS

Since Germeyer's games may be taken as models of two-level hierarchical systems, the analysis is reduced to the question of finding the solutions of the games formulated in Section 2.

Game G_1. The problem of solving game G_1 is reduced to that of solving a maximin problem with linked variables (see (2.1)).

Assume that the criteria F and G are continuous on compact sets X, Y. Then the inner infimum in (2.1) can be replaced by a minimum. However, in the general case the function

$$f(x) = \min_{y \in B^1(x)} F(x,y)$$

is discontinuous. Consider the simple example $F = y - x^2$, $G = xy$, $X = Y = [-1,1]$. Here $f(x)$ has a discontinuity at point $x = 0$ and the first player has no optimal strategy. This means that ε-optimal strategies x_ε should be found which satisfy the inequality $f(x_\varepsilon) \ge v_1 - \varepsilon$ for given $\varepsilon > 0$. With these assumptions $f(x)$ is lower semicontinuous; in general it is multiextremal.

In theory the problem may be solved using the penalty function method, which reduces it to an unconstrained optimization problem [1,6,7]. Consider the penalty function

$$J(x,y) = \begin{cases} 0 & \text{if } y \in B^1(x) \\ d & \text{if } y \notin B^1(x) \end{cases} ,$$

where $d > 0$. The reduction of problem (2.1) to a maximin problem with separable variables is based on the following theorem:

Theorem 1 [6,7]. *If x_c yields a solution of* $\max\min(F + cJ)$ *at fixed* c, *then for any sequence $c_k \to \infty$ the points x_{c_k} form an ε-optimal sequence of strategies for the first player.*

A number of methods can be used to solve problems of the form $\max\min\{F(x,y) + cJ\}$, including stochastic programming methods [7,8] and non-smooth optimization methods [9–11]. In addition to the nondifferentiability of the objective function there may be some difficulties connected with the multiple extrema of the problem, which make it necessary to develop appropriate optimization algorithms [12–14].

The use of numerical methods to search for v_1 and the ε-optimal strategy of the first player is complicated by the fact that problem (2.1) is not necessarily stated correctly with respect to the functional, in that any small variations in the second player's strategy $G(x,y)$ (due to errors in computations, for example) can cause variations in the first player's guaranteed result.

In the same way, for $F \equiv y$, $G = f(x)$, $X = Y = [0,1]$ the optimal result of the first player in game G_1 is zero. If the second player's criterion is $G_\varepsilon = G + \varepsilon(y-1)$, where ε may take any small positive value, then the guaranteed result will be equal to 1, since

$$B^\varepsilon(x) = \{y \in Y \mid G_\varepsilon(x,y) = \max_{z \in Y} G_\varepsilon(x,z)\} \equiv \{1\}$$

for any $x \in X$.

To obtain a numerically stable procedure for computing the best guaranteed results, it is necessary to regularize problem (2.1) using the method described in [15].

Game G_2. We shall make use of the following values, sets and functions:

$$L_2 = \max_{y \in Y} G(x^P(y),y) = \max_{y \in Y} \min_{x \in X} G(x,y)$$

$$E_2 = \{y \in Y \mid L_2 = G(x^P(y),y)\}, \quad D_2 = \{(x,y) \mid G(x,y) > L_2\}$$

$$K_2 = \sup_{(x,y)\in D_2} F(x,y) \le F(x_\varepsilon,y_\varepsilon) + \varepsilon$$

$$M_2 = \min_{y\in E_2} \max_{x\in X} F(x,y) \,, \quad F(x^a(y),y) = \max_{x\in X} F(x,y) \quad .$$

Here $x^p(y)$ is a penalizing strategy and $x^a(y)$ is the absolutely optimal strategy of the first player.

Theorem 2 [1]. *Let* $v_2 = \max(K_2,M_2)$. *Then the strategy*

$$\tilde{X}_\varepsilon = \begin{cases} x_\varepsilon & \text{if } y = y_\varepsilon \,,\ K_2 > M_2 \\ x^a(y) & \text{if } y \in E_2 \,,\ K_2 \le M_2 \\ x^p(y) & \text{otherwise} \end{cases}$$

is the ε*-optimal strategy of the first player in game* G_2.

The case $K_2 > M_2$ is particularly interesting: it corresponds to the situation in which the objectives of both players (the levels of the hierarchical system) are in some sense similar.

The theorem formulated above shows that the problem of constructing the optimal strategy in game G_2 is reduced to that of solving a nonlinear programming problem and a maximin problem with separable variables.

Game G_3. Let us define

$$D_3 = \{(x,y) \mid G(x,y) > L_3 = \min_{x\in X} \max_{y\in Y} G(x,y)\} \quad ,$$

$$K_3 = \sup_{(x,y)\in D_3} F(x,y) \le F(x_\varepsilon,y_\varepsilon) + \varepsilon \,, \quad B = \{x \in X \mid \max_{y\in Y} G(x,y) = L_2\}$$

$$B(x) = \{y \in Y \mid G(x,y) = \max_{z\in Y} G(x,z)\}$$

$$M_3 = \sup_{x\in B} \min_{y\in B(x)} F(x,y) \le \min_{y\in B(x_\varepsilon^*)} F(x_\varepsilon^*,y) + \varepsilon$$

Theorem 3 [1,4]. *Let* $v_3 = \max(K_3,M_3)$. *Then the strategy*

$$\tilde{x}_\varepsilon = \begin{cases} x_\varepsilon & \text{if } \tilde{y} = \tilde{y}_\varepsilon \text{ or } K_3 > M_3 \\ x_\varepsilon^* & \text{if } \tilde{y} \ne \tilde{y}_\varepsilon \text{ or } K_3 \le M_3 \end{cases}$$

is the ε*-optimal strategy of the first player in game* G_3.

Here $\tilde{y}_\varepsilon$ is the strategy of the second player, which consists in choosing point $y_\varepsilon \in Y$, and $x_\varepsilon^* \in B$ plays the role of a penalizing strategy. Thus the problem of finding the optimal strategy in game G_3 is reduced to that of solving a mathematical programming problem and a maximin problem with linked variables (value M_3 and strategy $x_\varepsilon^* \in B$).

4. A COMBINED PENALTY AND STOCHASTIC GRADIENT METHOD (CPSGM)

In the previous section we showed that a necessary step in the analysis of games $G_{1,2,3}$ is the solution of the following minimax problem: Find $x \in X_0$ and u_0, where

$$X_0 = \{x \in A \mid \min_{y \in Y} F(x,y) = u_0\}$$

$$u_0 = \max_{x \in A} \min_{y \in Y} F(x,y) \tag{4.1}$$

$$A = \{x \in X \mid \varphi_i(x) \geq 0 \,,\, i = 1,\dots,m\} \quad .$$

Let us consider certain stochastic algorithms for solving problem (4.1). We may assume without loss of generality that

$$X = O_R(0) = \{x \in E_k \mid \|x\| \leq R\} \,, \quad R > 0$$

and also that functions $F(x,y)$, $\varphi_i(x)$, $i = 1,\dots,m$, are continuous together with their derivatives with respect to x on set $X' \times Y$, $X' = O_{\varepsilon_0}(X)$. In addition, we assume that Y is a compact set from E_1, $A \neq \phi$, $\varepsilon_0 > 0$.

It is clear that

$$A = \{x \in E_k \mid \varphi_i(x) \geq 0 \,,\, i = 1,\dots,m+1\} \quad ,$$

where $\varphi_{m+1}(x) = R - \|x\|$. Now introduce

$$L_q(r,c) = M\, l_q(r,c,y,i) \quad , \tag{4.2}$$

where

$$l_q(r,c,y,i) = u - c_1 |\min(0\,,\, F(x,y) - u)|^q -$$

$$-(c_1/p_i) |\min(0, \varphi_i(x))|^q - c_2 |\min(0, \varphi_{m+1}(x))|^2$$

$$c = (c_1,c_2)\,, \quad r = (x,u)\,, \quad q \geq 2 \quad .$$

Here M represents the mathematical expectation, i is a random number whose values are taken from set $\{1,\ldots,m\}$ with probabilities $p_1,\ldots,p_m$; y is a random number distributed on Y according to measure μ in such a way that any non-empty intersection of y with any open set has positive measure.

It is shown in [1] that problem (4.1) can be reduced to a sequence of problems in which it is required to maximize function (4.2) with $c^n = (c_1^n , c_2^n) \uparrow \infty$ (this is the penalty function method).

The stochastic gradient method [8] can be used to search for the maximum of function L_q at fixed c. If the algorithm allows for penalty parameters c_1, c_2 to increase, then we obtain the following iterative procedure:

$$r^{n+1} = r^n + a_n \xi^n$$

$$\kappa^{n+1} = \kappa^n + b_n (\partial l_q (r^n , c^n , y^n , i^n)/ \partial r - \kappa^n) \tag{4.3}$$

$$n = 1,\ldots,$$

where

$$\xi^n = \begin{cases} \partial l_q (x^n , u^n , c^n , y^n , i^n)/ \partial r , & \text{if } x^n \in O_\varepsilon(X) \\ \partial l_q (x^0 , u^n , c^n , y^n , i^n)/ \partial r , & \text{if } x^n \notin O_\varepsilon(X) \end{cases} ,$$

vector $x^0 \in X$ and value ε, $0 < \varepsilon < \varepsilon_0$, are both chosen arbitrarily; (y^n, i^n) are the values of the random numbers (y,i) during the n-th independent test; $r^1 = (x^1, u^1)$ is the initial approximation; and $\{a_n\}$, $\{b_n\}$, $\{c^n\}$ are control sequences.

Theorem 4 [16]. *Let functions $\varphi_i(x)$, $i = 1,\ldots,m+1$ satisfy the condition*

$$0 \notin \mathrm{co}\{\partial \varphi_i(x)/ \partial x , i \in I_1^{m+1}(x)\} , \tag{4.4}$$

where

$$I_1^{m+1}(x) = \{i \mid \varphi_i(x) \le 0\}$$

for any point $x \in X$, and the control sequences satisfy the following conditions:

$$a_n , b_n , c_1^n , c_2^n > 0 , \quad \sum_{n=1}^{\infty} a_n = \infty , \quad c_1^n , c_2^n \to \infty$$

$$c_1^n / c_2^n \to 0 , \quad b_n \to 0 , \quad a_n \|c^n\|^2 / b_n \to 0 \tag{4.5}$$

$$a_n \|c^n\|^3 \to 0 \ , \quad \|c^{n+1} - c^n\| / a_n \to 0 \ , \quad \sum_{n=1}^{\infty} b_n^2 \|c^n\|^4 < \infty \quad .$$

Then for any initial approximation (r^1, x^1), *sequences* $\{r^n\}$, $\{x^n\}$ *of solutions of algorithm* (4.3) *exist such that, with probability one:*

(1) *A subsequence of the natural series of numbers* $\{n_t\}$ *exists such that*

$$\lim_{t \to \infty} \kappa^{n_t} = 0 \quad .$$

(2) *It follows from* $\lim_{s \to \infty} \kappa^{n_s} = 0$ *that the limit points of sequence* $\{r^{n_s}\}$ *belong to the set of stationary points* [10] *of problem* (4.1).

Remarks

1. Condition (4.4) is satisfied if $y(x)$, $i = 1,...m$, are concave and Slater's condition is satisfied.
2. The following are examples of sequences which satisfy conditions (4.5):

$$a_n = n^{-19/20} \ , \ b_n = n^{-3/4} \ , \ c_1^n = n^{1/22} \ , \ c_2^n = n^{1/21} \quad .$$

3. The parameter κ^n is introduced into (4.3) to follow the value of $\partial L_q / \partial r$ and to provide a means of finding the elements of the sequence $\{r^{n_s}\}$ which converges to the set of stationary points. (If F , φ_i are concave with respect to x, then sequence $\{r^n\}$ will converge to the set of solutions of problem (4.1) and there is no need to follow parameter κ^n.)
4. Theorems similar to Theorem 4 but with different restrictions on sequences (4.5) and rather more rigorous restrictions on functions F , φ_i have been proved in [7,17,18].

5. A STOCHASTIC "ERRORS" METHOD FOR FINDING A MAXIMIN

Let us consider problem (4.1), assuming that functions $F, \varphi_i(x)$, $i = 1,...,m$, are concave with respect to x on convex compact set $X \subset E_k$ for any $y \in Y$ (where $Y \in E_l$ is a compact set) and that both functions $F, \varphi_i(x)$ and their partial derivatives with respect to x are continuous on $X \times Y$, $A \neq \phi$.

This problem can be reduced to the following mathematical programming problem [1]: Find $r = (x, u)$ which solves

$$\max_{x,u} U \tag{5.1}$$

subject to

$$(x,u) \in X \times U$$

$$\Phi_q(r) = -\int_Y |\min(0, F(x,y) - u)|^q \mu(dy) - \sum_{i=1}^m |\min(0, \varphi_i(x))|^q \geq 0 \quad ,$$

where $q \geq 1$, U is a line segment which includes

$$[\min_{(x,y)\in X\times Y} F(x,y) ; \max_{(x,y)\in X\times Y} F(x,y)] \quad ,$$

and measure μ satisfies the conditions given on p. 8 in Section 4.

Problem (5.1) is equivalent to the following problem: From the points $r = (x,u)$ for which

$$\max_{x\in X} \Phi_q(r) = 0 \quad , \tag{5.2}$$

find the point with the largest value of u.

Function $\Phi_q(r)$ can be treated as an "error" which characterizes the distance of the point r from the feasible set of problem (5.1). This approach to solving problem (4.1) was suggested for the first time in [19].

Note that, as in (4.2), we have $\Phi_q(r) = M\varphi_q(r,y,i)$, where

$$\varphi_q(r,y,i) = -|\min(0, F(x,y) - u)|^q - (1/p_i)|\min(0, \varphi_i(x))|^q \quad ,$$

and random numbers y, i are as defined in Section 4.

We can now formulate the following iterative algorithm:

$$\begin{aligned} r^{n+1} &= \pi_R(r^n + a_n \xi^n) \\ \varphi^{n+1} &= l_n[\varphi^n + b_n(\varphi_q(r^n, y^n, i^n) - \varphi^n)] \\ n &= 1,2,\ldots, \end{aligned} \tag{5.3}$$

where π_R is the projection operator on $R = X \times V$ and vector ξ^n is defined by the formula

$$\xi^n = \begin{cases} e_1^{(k+1)} & \text{if } \varphi^n \geq -d_n \\ \varphi_q^{\varepsilon_n}(r^n, y^n, i^n) & \text{if } \varphi^n < -d_n \end{cases} \quad . \tag{5.4}$$

Here

$$e_1^{(k+1)} = (0,\ldots,0,1) \in E_{k+1} \quad ;$$

$\varphi_q^{\varepsilon_n}(r^n, y^n, i^n)$ is a conditional ε_n-subgradient of function $\varphi_q(\cdot, y^n, i^n)$ at point r^n from set R, $\varepsilon_n \geq 0$; y^n, i^n are the values of random numbers y and i during the n-th independent test; and (r^1, y^1) is the initial approximation.

Parameter φ^n in algorithm (5.3) follows the value of the error $\Phi_q(r^n)$, $\lim_{n \to \infty} | \varphi^n - \varphi_q(r^n)| = 0$ P-a.s. At the n-th step, if the value of the error is near zero ($\varphi^n \geq -d_n$) the value of u increases in accordance with (5.3), but if $\varphi^n < -d_n$ then the value of u changes in accordance with the stochastic quasi-gradient of the error function.

Theorem 5 [20]. *Assume that a constant $k > 0$ exists such that $\|\xi\| < k$, $n = 1,\ldots$ for any $\gamma \in \Gamma = \{(y^1, i^1, \ldots, y^n, i^n \ldots)\}$, that sequences $\{a_n\}$, $\{b_n\}$ and $\{d_n\}$, $\{l_n\}$, $\{\varepsilon_n\}$ exist such that*

$$a_n = c_1 n^{-t_1}, \quad b_n = c_2 n^{-t_2}, \quad d_n = cn^{-t}$$

$$t_1, t_2, t, c_1, c_2, c = \text{const} > 0$$

$$\sum_{n=1}^{\infty} a_n d_n = \infty, \quad a_n / b_n \to 0, \quad \sum_{n=1}^{\infty} b_n^2 < \infty, \quad \varepsilon_n / d_n \to 0$$

and that one of the following conditions is satisfied:

$$e_n = 1, \quad n = 1,\ldots, \quad 2t < \min\,(t_2 - 1/2\,,\, 1 - t_2\,,\, 2(t_1 - t_2)) \quad .$$

$$e_n = n^t / (n+1)^t, \quad n = 1,\ldots, \quad 2t < \min\,(t_2 - 1/2\,,\, 2(1 - t_2)\,,\, (2t_1 - t_2)) \quad .$$

Then for any initial approximation (r^1, φ^1), the sequence r^n defined by algorithm (5.3) converges to the set of solutions of problem (4.1) with probability one.

Remark. The following are examples of sequences which satisfy the conditions of the theorem:

$$a_n = n^{-8/10}, \quad b_n = n^{-7/10}, \quad d_n = n^{-1/11}, \quad \varepsilon_n = n^{-1/10} \quad .$$

6. ε-SUBGRADIENT DESCENT ALGORITHM FOR APPROXIMATION OF THE PARETO SET

Consider the following parametric programming problem: Find $x(\alpha) \in X_0(\alpha)$, where

$$X_0(\alpha) = \{x \in X \mid f(x,\alpha) = \max_{x' \in X} f(x',\alpha)\} \tag{6.1}$$

for all $\alpha \in A$.

Function $F(x,\alpha)$ is assumed to be continuous on convex compact set $X \in E_k$ for any $\alpha \in A$, where $A \in E_s$ is a bounded set. We say that point x^* is a solution of problem (6.1) at $\alpha = \alpha^*$ with accuracy (δ,Δ) if $\rho^2(x^*, X_\delta(\alpha^*)) < \Delta$, where ρ is a metric and

$$X_\delta(\alpha) = \{x \in X \mid f(x,\alpha) \geq \max_{x' \in X} f(x',\alpha) - \delta\}\ , \quad \alpha \in A \quad .$$

Assume that values δ_0 , Δ_0 , $\alpha_0 > 0$ are given. Let us construct an algorithm for finding (δ_0 , Δ_0), the approximate solutions of problem (6.1) at all d-nets $Ad = \{d_1,\ldots,d_N\}$ on A such that $d \leq d_0$ and

$$\|\alpha^i - \alpha^{i-1}\| < 2d\ , \quad i = 1,\ldots,N \quad . \tag{6.2}$$

Here $\|a\| = \max_i (a_i)$, $a \in E_S$.

We shall assume that $f(x,\alpha)$ is concave with respect to x on X for any $\alpha \in A$, diam $X \leq D$ and

$$|f(x,\alpha') - f(x,\alpha'')| \leq L\|\alpha' - \alpha''\| \quad ,$$

where L = const > 0. Let the solution of problem (6.1) be known with accuracy (δ_0 , Δ_0) at values of parameter $\alpha = \alpha_1$ from d-net Ad.

We shall determine the solution of problem (6.1) at the nodes of net Ad using the formula

$$x^{x+1} = \pi_X(\kappa^n + a\xi_\varepsilon^n)\ , \quad n = 1,\ldots,N-1 \quad , \tag{6.3}$$

where π_X is the projection operator on X; ξ_ε^n is the conditional ε-subgradient of concave function $f(\cdot\ , \alpha^n)$ at point x^n on set X , $\varepsilon > 0$; and a is a step-size multiplier.

Theorem 6 [20]. *If parameters a, d, ε of algorithm* (6.2)–(6.3) *satisfy the following conditions:*

$$a + 2a\varepsilon + a^2K^2 + \frac{8LD^2}{\delta_0} d + \frac{16L^2D^2}{\delta_0^2} d^2 < \Delta_0$$

$$2a\varepsilon + a^2K^2 + \frac{8LD^2}{\delta_0} d + \frac{16L^2D^2}{\delta_0^2} d^2 < \frac{2\alpha\delta_0}{D}\sqrt{\alpha} \qquad (6.4)$$

$$0 < d \le d_0 , \quad a > 0 , \quad \varepsilon \ge 0 , \quad \alpha > 0 \quad ,$$

where K is a constant, so that $\|\xi_\varepsilon^n\| < K$, $n = 1,\dots,N$ in (6.3), *and $\rho^2(x^1, X_{\delta_0}(\alpha^1)) < \Delta_0$, then all subsequent x^n, $n = 2,\dots,$ will satisfy condition $\rho^2(x^n, X_{\delta_0}(\alpha^n)) < \Delta_0$.*

Thus, using algorithm (6.2)–(6.3) we can obtain the solution of problem (6.1) with precision (δ_0, Δ_0) on d-net Ad, $0 < d \le d_0$. For any fixed δ_0, Δ_0 it is always possible to find values of a, α and ε which are sufficiently small that inequalities (6.4) are satisfied.

We shall now show how algorithm (6.2)–(6.3) may be applied to vector optimization problems.

Let vector criterion

$$W(x) = (w_1(x),\dots,w_m(x))$$

be defined and positive on $X \subset E_k$, and

$$W = \{w \in E_m \mid w = w(x) , x \in X\} \quad .$$

Let $\Pi(w)$ be the set of efficient vectors from W (Pareto-optimal vectors), where

$$\Pi(X) = \{x \in X \mid w(x) \in \Pi(w)\} \quad .$$

We shall use the following notation:

$$\Lambda^m = \{\lambda = (\lambda_1,\dots,\lambda_m) \mid \sum_{i=1}^{m} \lambda_i = 1 , \lambda_i \ge 0 , i = 1,\dots,m\}$$

$$X^\gamma(\lambda) = \{x \in X \mid F^\gamma(x,\lambda) = \max_{x' \in X} F^\gamma(x', \lambda)\} \quad ,$$

where

$$F^{\gamma}(\lambda) = \min_{1 \le i \le m} \lambda_i w_i(x) + \gamma \sum_{i=1}^{m} w_i(x)$$

$$w^{\gamma}(\lambda) = \{w \in W \mid w = w(x)\,,\, x \in X^{\gamma}(\lambda)\}\,,\quad \lambda \in \Lambda^m\,,\quad \gamma > 0\quad .$$

It is shown in [21] that for $\forall \varepsilon > 0\ \exists \gamma, d_0$, the set

$$\bigcup_{\lambda \in \Lambda_d^m} \{w(\lambda)\}\quad ,$$

where $w(\lambda)$ is an arbitrary point from $w^{\gamma}(\lambda)$ and Λ_d^m is an arbitrary d-net on Λ^m, $0 < d \le d_0$, is an ε-net on $\Pi(w)$. Thus to find an ε-net on $\Pi(w)$ it is sufficient to solve the following parametric programming problem: Find

$$x(\lambda) \in X^{\gamma}(\lambda) \tag{6.5}$$

for all $\lambda \in \Lambda_d^m$, $0 < d \le d_0$.

If functions $w_i(x)$, $i = 1,\dots,m$ are concave and continuous on convex compact set X, then an approximate solution of problem (6.5) can be found using algorithm (6.3).

7. NUMERICAL RESULTS

The proposed algorithms were implemented and then tested on some simple problems in order to investigate their practical efficiency.

7.1. CPSGM

Algorithm (4.3) (with certain modifications) has been used to solve (4.1) with functions

$$F_1(x,y) = \cos\,(0.25(x_1 + x_2 + x_3) + y_1 - 0.5) +$$

$$+ \cos\,(0.25(x_1 + 2x_2 + x_3) + y_2 - 0.5) + \cos\,(0.5(x_1 + x_2) + y_3 - 0.5)$$

$$F_2(x,y) = 1 + (x_1 - 0.5)(y_1 - 0.5) + (x_2 - 0.5)(y_2 - 0.5) +$$

$$+ (x_3 - 0.5)(y_3 - 0.5) + (x_4 - 0.5)(y_4 - 0.5)$$

defined on the product of unit cubes. The following control sequences were used: $\{b_n\} = n^{-3/4}$, $\{a_n c^n\} = n^{-11/20}$, $q = 1$, $\varepsilon = 0$ for F_1, and $q = 2$, $a_n = n^{-0.95}$, $b_n = n^{-0.72}$, $c_n = n^{0.2}$ for F_2.

The results are presented in Table 1.

Table 1. The results obtained with the CPSGM algorithm.

<table>
<tr><th rowspan="2">Function</th><th rowspan="2">Initial approximation</th><th rowspan="2">Number of iterations</th><th rowspan="2">Precise solution</th><th colspan="2">Approximate solution</th></tr>
<tr><th>With precise gradient</th><th>With approx. gradient</th></tr>
<tr><td rowspan="4">F_1</td><td rowspan="4">x^0=(0.000, 0.000,1.00)
u^0=1.000</td><td>400</td><td rowspan="4">x^*=(0.000, 0.000,0.000)
u^*=2.634</td><td>x=(0.009,0.017, 0.009);u=2.634</td><td>x=(0.000,0.000 0.018);u=2.606</td></tr>
<tr><td>800</td><td>x=(0.012,0.018, 0.012);u=2.626</td><td>x=(0.000,0.000, 0.097);u=1.549</td></tr>
<tr><td>2400</td><td>x=(0.000,0.000, 0.000);u=2.619</td><td>x=(0.000,0.000, 0.016);u=2.613</td></tr>
<tr><td>3200</td><td>x=(0.000,0.000, 0.000);u=2.634</td><td>x=(0.000,0.000, 0.000);u=2.624</td></tr>
<tr><td rowspan="4">F_2</td><td rowspan="4">x^0=(0.6,0.6, 0.6,0.6)
u^0=2</td><td>500</td><td rowspan="4">x^*=(0.5000, 0.5000, 0.5000, 0.5000)
u^*=1</td><td colspan="2">x=(0.4704,0.4668, u=1.069</td></tr>
<tr><td>1500</td><td colspan="2">x=(0.5118,0.4974,0.5014,0.5003); u=1.059</td></tr>
<tr><td>9000</td><td colspan="2">x=(0.5005,0.4988,0.5006,0.4975); u=1.041</td></tr>
<tr><td>17000</td><td colspan="2">x=(0.5002,0.5012,0.4989,0.5030); u=1.036</td></tr>
</table>

It can be observed that a good approximation to solution x and the first approximation to u are obtained reasonably quickly. However, further refinement of the solution takes place very slowly.

When the gradient of the efficiency function is computed using the difference scheme, the rate of convergence of the algorithm is the same as when the precise gradient is used.

7.2. Errors Method

The errors method (with parameters $\varepsilon_n = 0$, $a_n = n^{-0.8}$, $b_n = n^{-0.7}$, $d_n = 0.01n^{-0.1}$) was used to find the maximin of functions

$$F_i(x,y) = \sum_{j=1}^{i+1} \cos\,(x_i + y_i - 0.5)\ ;\ i = 1,2$$

defined on unit cubes.

The results of the computations are presented in Table 2.

Table 2. The results obtained using the errors method.

Function	Initial approximation	No. of iterations	Precise solution	Approximate solution
F_1	x^0=(0.000, 0.000); u^0=0.000	200	x^*=(0.000, 0.000); u^*=1.756	x=(0.0277,0.0289); u=1.809
		600		x=(0.000,0.0028); u=1.900
		5400		x=(0.0133,0.0089); u=1.860
		10600		x=(0.0037,0.0066); u=1.815
F_2	x^0=(0.000, 0.000,1.000); u^0 = 0.000	200	x^*=(0.000, 0.000,0.000); u^*=2.634	x=(0.0458,0.0000,0.0133); u=2.721
		600		
		42000		x=(0.0196,0.0066,0.0172); u=2.831
		84000		x=(0.0046,0.0076,0.0119); u=2.791

REFERENCES

1. Y.B. Germeyer. *Games with Non-Antagonistic Interests*. Nauka, Moscow, 1976.
2. Y.B. Germeyer and N.N. Moiseev. "On some problems in the theory of hierarchical control systems", in *Problems in Applied Mathematics and Mechanics*. Nauka, Moscow, 1971.
3. V.N. Burkov. Models and methods of functioning of hierarchical systems. *Automatika i Telemechanica*, 11, pp. 106–131, 1977.
4. N.N. Kukushkin. The role of mutual information in two-person games with non-antagonistic interests. *Zhurnal Vychislitelnoj Matematiki i Matematicheskoj Fiziki*, 12(4), pp. 1029–1034, 1972.
5. I.A. Vortel and F.I. Ereshko. *Mathematics of Conflict and Cooperation*, Znanie, Moscow, 1973.
6. V.A. Gorelik. Approximate determination of maximins with constraints linking the variables. *Zhurnal Vychislitelnoj Matematiki i Matematicheskoj Fiziki*, 12(2), pp. 510–517, 1972.
7. V.V. Fedorov. *Computational Maximin Methods*. Nauka, Moscow, 1979.
8. Y.M. Ermoliev. *Methods of Stochastic Programming*. Nauka, Moscow, 1976.
9. N.Z. Shor. New directions in the development of methods for nonsmooth optimization. *Kibernetika*, 6, pp. 87–91, 1977.
10. V.F. Demyanov and V.H. Malozemov. *Introduction to Minimax*. Nauka, Moscow, 1972.
11. E.A. Nurminski. *Numerical Methods for Solving Deterministic and Stochastic Minimax Problems*. Naukova Dumka, Kiev, 1979.
12. R.G. Strongin. *Numerical Methods in Multiextremal Problems*. Nauka, Moscow, 1978.
13. F.I. Ereshko and A.S. Zlobin. An algorithm for centralized allocation of resources among active subsystems. *Economics and Mathematical Methods*, 4, pp. 703–713, 1977.
14. A.G. Sukharev. *Optimal Search for Optima*. Moscow University Press, 1975.
15. D.A. Molodtzov and V.V. Fedorov. Approximation of two-person games with information transfer. *Zhurnal Vychislitelnoj Matematiki i Matematicheskoj Fiziki*, 13(6), pp. 1469–1484, 1973.
16. S.K. Zavriev. On searching for stationary points in a maximin problem with constraints. *Vestnik MGU, Series Computational Mathematics and Cybernetics*, 2, pp. 48–57, 1980.
17. S.K. Zavriev. A combined penalty and stochastic gradient method for the determination of maximin. *Zhurnal Vychislitelnoj Matematiki i Matematicheskoj Fiziki*, 19(2), pp. 329–342, 1979.
18. N.M. Novikova. A stochastic quasigradient method for the determination of maximin. *Zhurnal Vychislitelnoj Matematiki i Matematicheskoj Fiziki*, 17(1), pp. 91–99, 1977.
19. Y.B. Germeyer and I.A. Krylov. Determination of maximin by a "discrepancies" method. *USSR Computational Mathematics and Mathematical Physics*, 12(4), pp. 871–881, 1972.

20. S.K. Zavriev. *Stochastic Gradient Algorithms for Solving Minimax Problems*. Ph.D. Thesis, Moscow University, 1981.

21. P.S. Krasnoschekov et al. Decomposition in problems of technical design. *Izvestia of the USSR Academy of Sciences, Series Technical Cybernetics*, 2, pp. 7–17, 1979.