Yugoslav Journal of Operations Research
5 (1995), Number 1, 55-64

A CONTRIBUTION TO METHODS OF EXTERIOR PENALTY FUNCTIONS FOR NONLINEAR PROGRAMMING

Aleksandar ZEČEVIĆ

Business School
27. March 149, 11000 Belgrade, Yugoslavia

Abstract: This paper presents a new penalty functions method for nonlinear programming with the change of penalty parameter at each iteration. The efficiency of the method is illustrated by a number of test problems.

Keywords: Exterior penalty functions, nonlinear programming.

1. INTRODUCTION

We consider the nonlinear programming problem of the form:

$$\min \{ f(x) : x \in X \}, \qquad X = \{ x \in E^n : g_i(x) \leq 0, i = 1, \dots, m \}. \tag{C}$$

The basic idea of the method of exterior penalty functions is to replace the constrained problem of the form (C) by a sequence of unconstrained problems of the form (C_k)

$$\min \{ F_k(x) = f(x) + P_k(x) : x \in E^n \}, \qquad k = 1, 2, \dots \tag{C_k}$$

Let $\{ x_k \}$ be the sequence of solutions of (C_k). Under certain conditions, the accumulation points of $\{ x_k \}$ are solutions of (C). The sequence of exterior penalty functions $P_k(x)$ has to satisfy the well known conditions (a) $P_k(x) = 0$, $\forall x \in X$, (b) $P_k(x) > 0$, $\forall x \notin X$, (c) $P_{k+1}(x) > P_k(x)$, $\forall x \notin X$, (d) $P_k(x) \to \infty$, $k \to \infty$, $\forall x \notin X$.

In the sequel we shall use the following choice for $P_k(x)$:

$$P_k(x) = A_k \, p(x), \quad p(x) = \sum_{i=1}^{m} [\max \{0, g_i(x)\}]^2 \tag{1}$$

where A_k is the penalty parameter, $A_k \to \infty$, $k \to \infty$, $A_k < A_{k+1}$, $A_k > 0$.

For each k we need to solve the problem (C_k). To this end we may use any unconstrained minimization method, say method A, for example the Newton's method for unconstrained minimization, the method of conjugate directions, etc.

The problem (C_k) can be solved in practice only approximately which influences directly the convergence of the sequence $\{x_k\}$. In addition, the convergence is greatly influenced by the choice of penalty parameter A_k. A faster increase in A_k, $k \to \infty$, can cause numerical stability problems. On the other hand a slower increase in A_k for $k \to \infty$ slows down the process of minimization and convergence to the feasible region X where the optimal point is located.

In the classical application of methods of penalty functions the change of penalty parameter is made periodically. Namely, for each fixed k a few steps of the method A are applied to the function $F_k(x)$ and k is then replaced by $k+1$. A typical example of such an approach is the method of exterior penalty functions by Polak [6]. The general algorithm of Polak's method is the following:

ALGORITHM 1. (Polak)

Step 0. Choose the initial data:

$$A > 0,\ \alpha \in (0, 1/2),\ \delta > 1,\ x_0 \in E^n,$$

Step 1. Let $A_0 = A,\ i = 0,\ k = 0$

Step 2. Compute $S(x_i, A_k) = -[\nabla f(x_i) + A_k \nabla p(x_i)]$

Step 3. If $\| S(x_i, A_k) \| > 1/A_k$ go to Step 4, otherwise set $A_{k+1} = A_k\, \delta,\ x_{k+1} = x_i$, $k := k+1$ and go to Step 2.

Step 4. Using Polak's algorithm for one–dimensional minimization [6] compute β_i such that $-\beta_i (1-\alpha) \| S(x_i, A_k) \|^2 \le f(x_i + \beta_i S(x_i, A_k)) + A_k p(x_i + \beta_i S(x_i, A_k)) - f(x_i) - A_k p(x_i) \le -\beta_i \alpha \| S(x_i, A_k) \|^2$

Step 5. Set $x_{i+1} = x_i + \beta_i S(x_i, A_k),\ i := i+1$. Go to Step 2.

Polak shows the following convergence result.

LEMMA. (Polak) [6]. Algorithm 1 generates sequences $\{x_k\}$ and $\{A_k\}$ such that $S(x_k, A_k) \to 0$ when $k \to \infty$.

Let us point out that in Polak's method the change of the penalty parameter A_k is performed when $\| S(x_k, A_k) \| \le 1/A_k$, $k = 0, 1, 2, ...$

In our method instead of the functions $F_k(x)$ we shall use the set of Tihonov's functions of the form:

$$T_k(x) = f(x) + A_k p(x) + \alpha_k \Omega(x), \qquad k = 1, 2, ... \tag{2}$$

where $A_k > 0,\ \lim_{k\to\infty} A_k = \infty,\ \alpha_k > 0,\ \lim_{k\to\infty} \alpha_k = 0$.

The function $\Omega(x)$ is called the stabilizer and it is usually a convex function, for example $\Omega(x) = \| x^2 \| / 2$. We construct a sequence $\{x_k\}$ using the relation:

$$x_{k+1} = x_k + \beta_k S_k, \qquad k = 0, 1, 2, ... \tag{3}$$

where x_0 is the initial point, $x_0 \notin X$, S_k is the direction of the decrease of the function $T_k(x)$ at the point x_k and β_k is the choice of the step-length in the selected direction.

Consider now the case when parameters A_k, α_k and β_k are varied at every iteration. The idea is: if we coordinated in a proper way parameters α_k, A_k and β_k at the iteration k, $k = 0, 1, 2, \ldots$ then it would be possible to get the minimization sequence $\{x_k\}$ which would converge to the solution of the problem (C).

Such a modern approach to the use of the methods of exterior penalty functions was given by F.P. Vasilev [2]. He assumes that the following assumptions are satisfied: X_0 is a convex and closed set in E^n, the functions $f(x)$, $g_1(x)$, $g_2(x)$, ... $g_m(x)$ are convex and differentiable on X_0 and the set X is of the form:

$$X = \{ x : x \in X_0, g_i(x) \le 0, i = 1, \ldots, m \} \tag{4}$$

For the sake of simplicity, in the sequel we shall take $X_0 = E^n$. Vasilev takes as penalty function the function of the form (1) and the stabilizer of the form $\Omega(x) = \| x^2 \| / 2$ and he shows that it is possible to coordinate the parameters A_k, α_k and β_k at the iteration k so that under certain assumptions the sequence $\{x_k\}$ converges to the solution of the problem (C). The author proposes the choice of the parameters α_k, β_k and A_k of the form: $A_k = (k+1)^{1/A}$, $\alpha_k = (k+1)^{-1/\alpha}$, $\beta_k = (k+1)^{-1/\beta}$, $k = 0, 1, 2, \ldots$ where A, α and β are any natural numbers which satisfy the inequalities $2\alpha^{-1} + \beta^{-1} + A^{-1} < 1$, $\alpha^{-1} + 2A^{-1} < \beta^{-1}$, $A < \alpha$(ex. $A = 6, \alpha = 8, \beta = 2$). Such a choice satisfies the conditions of the following convergence theorem:

THEOREM 1. (Vasilev) [2]. Suppose that:

1. The functions $f(x), g_1(x), g_2(x), \ldots g_m(x)$ are convex and differentiable on X_0; the gradients of $f(x)$ and $p(x)$ are such that

   $$\max \{ \| \nabla f(x) \|, \| \nabla p(x) \| \} < L_0 (1 + \| x \|), \quad x \in X_0$$

   where L_0 is a positive constant; the set (4) is non-empty; $f_* = \inf_{x \in X} f(x) > -\infty$ and $X_* = \{ x \in X, f(x) = f_* \} \neq \varnothing$.

2. Lagrange function $L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$, $x \in X_0$,

   $\lambda \in \Lambda_0 = \{ (\lambda_1, \lambda_2, \ldots, \lambda_m) : \lambda_1 \ge 0, \lambda_2 \ge 0, \ldots, \lambda_m \ge 0 \}$ has a saddle point $(x^*, \lambda^*) \in X_0 \times \Lambda_0$ with $L(x^*, \lambda) \le L(x^*, \lambda^*) \le L(x, \lambda^*)$ $\forall x \in X_0$ and $\forall \lambda \in \Lambda_0$

3. The sequences $\{A_k\}$, $\{\alpha_k\}$ and $\{\beta_k\}$ satisfy the assumptions $0 < A_k < A_{k+1}$,

   $$\alpha_k > 0, \beta_k > 0, \lim_{k\to\infty} A_k^{-1} = \lim_{k\to\infty} \alpha_k = \lim_{k\to\infty} \beta_k = \lim_{k\to\infty} A_k^{-1} \alpha_k^{-1} =$$
   $$\lim_{k\to\infty} (A_{k+1} - A_k) \alpha_k^{-2} \beta_k^{-1} = \lim_{k\to\infty} \beta_k A_k^2 \alpha_k^{-1} = \lim_{k\to\infty} | \alpha_{k+1} - \alpha_k | \alpha_k^{-2} \beta_k^{-1} = 0$$

Then the sequence $\{x_k\}$ determined by (3) with $S_k = -\nabla T_k(x_k)$ converges to the solution x^{**} of the problem (C) for which $\Omega(x^{**})$ is minimal.

The use of the above theorem is rather limited. On one hand, the strong assumptions narrow down the class of problems where it can be applied, while, on the other hand, a practical performance of the corresponding method is modest at best. Namely, slow changes of the parameters A_k, β_k and α_k (imposed by assumption 3. of Theorem 1), directly imply a slow rate of convergence of the method to the solution of the problem (C).

2. A NEW METHOD AS THE CONTRIBUTION TO METHODS OF EXTERIOR PENALTY FUNCTIONS

We will propose a new method which has a good practical performance. Three versions of the method will be described. They differ in the choice of the stabilizer $\Omega(x)$ in $T_k(x)$, the choice of the initial values of parameters α_k, β_k and A_k and their change at each iteration and the choice of the direction S_k.

2.1. THE CHOICE OF THE STABILIZER $\Omega(X)$

We recommend the following choice of $\Omega(x)$ in the Tihonov's function (2)

$$\Omega(x) = \frac{||Q\,x||^2}{2} \tag{5}$$

where Q is a positive definite matrix. For example Q can be chosen as the Hessian matrix of $f(x)+P_0(x)$ at the point x_0. The stabilizer $\Omega(x)$ can also be in the form (6) or (7)

$$\Omega(x) = \frac{p^2(x)}{2} \tag{6}$$

$$\Omega(x) = e^{p(x)} \tag{7}$$

2.2. THE CHOICE OF PARAMETERS

The following three possibilities will be considered:

2.2.1. Let:

$$a_0 = \frac{K_1\,||\nabla p(x_0)||}{||\nabla f(x_0)||}, \quad x_0 \notin X,$$

where $K_1 = 10^t$, and $t \in \mathrm{N} \cup \{0\}$ is the smallest number such that $a_0 > 1$, $A_0 = a_0^{1/6}$, $\alpha_0 = a_0^{-1/8}$ and $\beta_0 = a_0^{-1/2}$. At the iteration k, $a_k = a_{k-1} + K$, $\beta_k = a_k^{-1/2}$, $\alpha_k = a_k^{-1/8}$, $A_k = a_k^{1/6}$ $(k = 1, 2, \ldots)$, where $K = 10^s$, $s \in \mathrm{N}$. If $x_0 \in X$ then a_0 is any constant satisfying $a_0 > 1$.

2.2.2. Let:

$$a_0 = \frac{K_2\,||\nabla p(x_0)||}{||\nabla f(x_0)||}, \quad x_0 \notin X,$$

where $K_2 = 10^{-t}$, and $t \in \mathrm{N} \cup \{0\}$ is the smallest number such that $a_0 < 1$. The initial values are $\beta_0 = a_0^{1/4}$, $A_0 = a_0^{-1/5}$, $\alpha_0 = a_0^{1/6}$. At the k-th iteration $a_k = a_{k-1}K$, $\beta_k = a_k^{1/4}$, $A_k = a_k^{-1/5}$, and $\alpha_k = a_k^{1/6}$, $k = 1, 2, \dots$, where $K < 1$ (for example $K = 0.9$). If $x_0 \in X$ then a_0 is any constant satisfying $a_0 < 1$.

2.2.3. Let:

$$a_0 = \frac{K_3 \, ||\nabla p(x_0)||}{||\nabla f(x_0)||}, \quad x_0 \notin X,$$

where $K_3 = 10^{-t}$, and $t \in \mathrm{N} \cup \{0\}$ is the smallest number such that $a_0 < 1$, $\beta_0 = 0.7937\, a_0$, $\alpha_0 = 1.0293\, a_0$. At the iteration k, $a_k = a_{k-1}K$, $A_k = 1/a_k$, $\alpha_k = 1.0293\, a_k$, and $\beta_k = 0.7937\, a_k$ $(k = 1, 2, \dots)$. The constant K is determined by the formula:

$$K = 1 - \frac{1}{K_4 \sqrt[3]{m}}$$

where m is the number of constraints in the problem (C) and $K_4 > 1$ (for example $K_4 = 5$). If $x_0 \in X$ then $a_0 = 1$.

Parameters A_k, α_k and β_k in Versions 1–3 of the new method are such that the following relation is satisfied:

$$\alpha_k > \frac{1}{A_k} > \beta_k, \quad (k = 0, 1, 2, \dots)$$

2.3. THE CHOICE OF THE DIRECTION

We choose the direction by the following three rules:

$$S_k = -\nabla T_k(x_k), \qquad k = 0, 1, 2, \dots \tag{8}$$

$$S_k = -H_k(x_k)\nabla T_k(x_k), \qquad k = 0, 1, 2, \dots \tag{9}$$

where $H_k(x_k)$ is the inverse of the Hessian matrix at the point x_k, $k = 0, 1, 2, \dots$.

$$S_0 = -\nabla T_0(x_0)$$

$$S_k = m_k \cdot S_{k-1} - \nabla T_k(x_k), \qquad k = 1, 2, \dots$$

where

$$m_k = \begin{cases} 0 & , \quad k = 0, n, 2n, \dots \\ \dfrac{< \nabla T_k(x_k), \nabla T_{k-1}(x_{k-1}) - \nabla T_k(x_k) >}{||\nabla T_{k-1}(x_{k-1})||^2} & , \quad k \neq 0, n, 2n, \dots \end{cases}$$

2.4. THE STOPPING RULE

Let ε_1, ε_2 and ε be constants of the form 10^{-t}, $t \in \mathrm{N}$. The algorithm stops at the k-th iteration if the following conditions are satisfied:

$$||x_{k+1} - x_k|| \le \varepsilon_1, \; |T_{k+1}(x_{k+1}) - T_k(x_k)| \le \varepsilon_2, \; ||S_k|| \le \varepsilon .$$

2.5. THE ALGORITHM OF THE NEW METHOD

The following algorithm describes the structure of the new method.

ALGORITHM 2.

Step 0. Choose the initial data

$x_0 \notin X,\ A_0,\ \alpha_0$ and β_0 ; set $k = 0$

Step 1. Determine

$T_k(x_k),\ \nabla T_k(x_k)$

Step 2. Determine the direction S_k

Step 3. Compute $x_{k+1} = x_k + \beta_k S_k$

Step 4. Compute the parameters $\alpha_{k+1}, \beta_{k+1}, A_{k+1}$

Step 5. Compute $T_{k+1}(x_{k+1}), \nabla T_{k+1}(x_{k+1})$

Step 6. Let $FTOL = | T_{k+1}(x_{k+1}) - T_k(x_k) |$, $XTOL = \| x_{k+1} - x_k \|$, $NORMS = \| S_k \|$

Step 7. If the following conditions are satisfied: $XTOL \le \varepsilon_1$, $FTOL \le \varepsilon_2$ and $NORMS \le \varepsilon$ then go to Step 10, otherwise go to Step 8

Step 8. Set $k := k+1$

Step 9. Go to Step 2

Step 10. Set $\hat{x} = x_{k+1}$ and STOP; $\hat{x}$ is the solution of the problem (C) with the precision ε_1, ε_2 and ε.

Three versions of the general algorithm will be considered. Version 1 is a special case of the algorithm where the stabilizer $\Omega(x)$ is determined by (5) and the choice of parameters is according to 2.2.1. In Version 2 the stabilizer $\Omega(x)$ is determined by (5) or (6) and the choice of parameters is according to 2.2.2. Finally, Version 3 is a special case of the algorithm where the stabilizer $\Omega(x)$ is determined by (5), (6) or (7) and the choice of parameters is according to 2.2.3.

The choice of A_0, α_0 and β_0 as well as the rule for generating the sequences $\{ A_k \}$, $\{ \alpha_k \}$ and $\{ \beta_k \}$ described in 2.2 is less restrictive than in Vasilev's method, which makes it possible to be more flexible in varying parameters and results in a better practical performance of the method. For the time being the proof of the convergence of the sequence $\{ x_k \}$ generated by Algorithm 2 to the solution of problem (C) does not exist. However, if the method is modified in such a way that iterations k', $2k'$, ... are performed according to Polak's method it is easy to derive the appropriate convergence result. If k' is sufficiently large it is clear that this modification does not influence the practical performance of the method.

3. COMPUTATIONAL RESULTS

The new method has been tested on the following standard test problems taken from the literature.

PROBLEM 1. [3]

$$\min f(x) = -x_1 \cdot x_2$$

$$x_1 + x_2^2 - 1 \le 0, \quad -x_1 - x_2 \le 0$$

The solution: $f(x^*) = -0.3849$, $x^* = (0.666, 0.577)$, $x_o = (-0.1, -0.1)$, $x_o \notin X$.

PROBLEM 2. [7]

$$\min f(x) = [\,-(x_3 - 1)^{\sin x_1} - (x_4 - x_2)^2\,]$$

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 - 3 \le 0, \qquad x_1 + x_2 + x_3 + x_1 x_2 + x_3 x_4 + x_2 x_3 - 2 \le 0,$$

$$x_1^2 - x_1 x_2 + x_2^2 + x_4^2 x_3 - x_1 x_2^2 - 6 \le 0,$$

$$0 \le x_1 \le 2, \qquad -1 \le x_2 \le 1, \qquad 1.05 \le x_3 \le 2, \qquad 0 \le x_4 \le 1.$$

The solution: $f(x^*) = -4.795$, $x^* = (0, -0.974, 1.05, 0.974)$, $x_o = (-0.1, -1, 0.2, 1)$, $x_o \notin X$.

PROBLEM 3. [1]

$$\min f(x) = 0.0204\, x_1 x_4 (x_1 + x_2 + x_3) + 0.0187\, x_2 x_3 (x_1 + 1.57\, x_2 + x_4) + 0.0607\, x_1 x_5^2 (x_1 + x_2 + x_3) + 0.0437\, x_2 x_3 x_6^2 (x_1 + 1.57\, x_2 + x_4)$$

$$2.07 - 0.001\, x_1 x_2 x_3 x_4 x_5 x_6 \le 0$$

$$0.00062\, x_1 x_4 x_5^2 (x_1 + x_2 + x_3) + 0.00058\, x_2 x_3 x_6^2 (x_1 + 1.5\, x_2 + x_6) - 1 \le 0$$

$$x_1 \ge 0, \quad x_2 \ge 0, \quad x_3 \ge 0, \quad x_4 \ge 0, \quad x_5 \ge 0, \quad x_6 \ge 0.$$

The solution: $f(x^*) = 134.9$, $x^* = (5.53, 4.66, 10.43, 12.08, 0.752, 0.878)$, $x_o = (5.59, 4.3, 12.02, 11.2, 0.8, 1.1)$, $x_o \notin X$.

PROBLEM 4. [5]

$$\min f(x) = -x_1 x_2 x_3$$

$$x_1 - 42 \le 0, \qquad x_2 - 42 \le 0, \qquad x_3 - 42 \le 0,$$

$$0 \le x_1 + 2x_2 + 2x_3 \le 72$$

The solution: $f(x^*) = -3456.00$, $x^* = (24.000, 12.000, 12.000)$, $x_o = (25, 15, 15)$, $x_o \notin X$.

We test and compare the method of Polak, the method of Vasilev with the choice of parameters $A = 6$, $\alpha = 8$, $\beta = 2$, $A_k = (k+1)^{1/A}$, $\alpha_k = (k+1)^{-1/\alpha}$, $\beta_k = (k+1)^{-1/\beta}$ and Versions 1–3 of the new method. To this end Polak's and Vasilev's method are modified to include a stopping rule analogous to the one defined in 2.4. The performance of the method is measured by the number of function and partial derivatives evaluations (ZB) needed to achieve the precision defined by the parameters ε_1, ε_2 and ε. The method stops if ZB exceeds 600. The tables contain the information about the approximate solution $\hat{x}$, the value $f(\hat{x})$ and the relative error

$|f(\hat{x})-f(x^*)|/f(x^*)$. The tests are done in FORTRAN on PC 286 using single precision arithmetic.

Problem 1.

Method	Vasilev	Polak	New Method		
			Version 1	Version 2	Version 3
$f(\hat{x})$	-0.4622	-0.3904	-0.436	-0.3837	-0.3849
$\hat{x}_1$	0.700	0.665	0.670	0.665	0.666
$\hat{x}_2$	0.646	0.587	0.651	0.578	0.577
$\left\lvert\frac{f(\hat{x})-f(x^*)}{f(x^*)}\right\rvert$	0.20083	0.01429	0.1328	0.0012	$0.5 \cdot 10^{-3}$
ZB	600	600	600	516	507

Remark: We choose the direction by (8), $\varepsilon_1=10^{-3}$, $\varepsilon_2=10^{-3}$, $\varepsilon=10^{-3}$.

Problem 2.

Method	Polak	New Method		
		Version 1	Version 2	Version 3
$f(\hat{x})$	-6.942	-5.551	-4.7986	-4.9747
$\hat{x}_1$	0.0881	0.001	0	0
$\hat{x}_2$	-1.2188	-1.166	-0.975	-0.974
$\hat{x}_3$	0.7232	0.668	1.049	1.05
$\hat{x}_4$	1.2188	1.166	0.974	0.974
$\left\lvert\frac{f(\hat{x})-f(x^*)}{f(x^*)}\right\rvert$	0.395	0.1097	0.0354	10^{-4}
ZB	600	329	300	465

Remark: We choose the direction by (10), $\varepsilon_1=10^{-4}$, $\varepsilon_2=10^{-4}$, $\varepsilon=10^{-3}$.

Vasilev's method gives satisfactory results only in the case of Problem 1. In Problems 2–4 the method diverges and the results are omitted. In Problems 1–3 the best results are obtained when Version 3 of the new method is used, while in Problem 4 Version 2 gives the best results. It should be noted that different choice of parameters in Vasilev's method might give satisfactory results, but experiments in this direction were not performed.

Problem 3.

Method	Polak	New Method		
		Version 1	Version 2	Version 3
$f(\hat{x})$	117.012	127.700	135.8974	134.372
$\hat{x}_1$	5.425	5.53	5.673	5.46
$\hat{x}_2$	4.122	4.24	4.315	4.375
$\hat{x}_3$	11.740	11.96	12.067	12.002
$\hat{x}_4$	11.124	11.16	11.071	11.041
$\hat{x}_5$	0.672	0.642	0.671	0.6934
$\hat{x}_6$	0.845	0.929	0.936	0.9277
$\left\lvert\frac{f(\hat{x})-f(x^*)}{f(x^*)}\right\rvert$	0.1326	0.0534	$0.74 \cdot 10^{-2}$	$0.39 \cdot 10^{-2}$
ZB	600	490	428	147

Remark: We choose the direction by (9), $\varepsilon_1 = 10^{-2}$, $\varepsilon_2 = 10^{-2}$, $\varepsilon = 10^{-2}$.

Problem 4.

Method	Polak	New Method		
		Version 1	Version 2	Version 3
$f(\hat{x})$	−4521.9901	−3319.782	−3455.9999	−3458.0054
$\hat{x}_1$	26.03	23.911	24.004	23.748
$\hat{x}_2$	13.238	11.783	11.999	12.067
$\hat{x}_3$	13.123	11.783	11.999	12.067
$\left\lvert\frac{f(\hat{x})-f(x^*)}{f(x^*)}\right\rvert$	0.308	0.0394	$2.9 \cdot 10^{-8}$	$0.58 \cdot 10^{-3}$
ZB	600	600	450	456

Remark: We choose the direction by (10), $\varepsilon_1 = 10^{-4}$, $\varepsilon_2 = 10^{-4}$, $\varepsilon = 10^{-4}$.

REFERENCES

[1] Biggs,B.M.C., "A numerical comparation between two approaches to the nonlinear programming problem", Technical Report No.77, The Hatfield Polytechnic, Great Britain, 1976.

[2] Васильев,Ф.П., *Численые методы решения оптимизации*, Наука, Москва, 1986.

[3] Fiacco,A.V., and McCormick,G.P., *Nonlinear Programming Sequential Unconstrained Minimization Techniques*, John Wiley and Sons, Toronto, 1986.

[4] Gonzaga,C.C., "An algorithm for solving linear programming problem in O(n^{3}L) operations", in: N. Megiddo (ed.), *Progress in Mathematical Programming, Interior Point and Related Methods*, Springer, New York, 1988, 1–28.

[5] Kuester,J.K., and Mize,J.H., *Optimization Techniques with Fortran*, Mc Graw–Hill, New York, 1980.

[6] Polak,E., *Computational Methods in Optimization*, Academic Press, New York, 1971.

[7] Скоков,В.А., "Стандарная программа минимизации функций многих переменных при наличии ограничений общего вида", МГУ – Вычеслительний центр выпуск 23, Москва, 1971.