# The Linearization Method: Principal Concepts and Perspective Directions

B. N. PSHENICHNYI and A. A. SOSNOVSKY
*V. M. Glushkov Institute of Cybernetics, Ukrainian Academy of Sciences, Kiev, Ukraine*

(Received: 5 October 1990; accepted: 13 November 1992)

**Abstract.** The linearization method, for solving the general problem of nonlinear programming and its various modifications, is considered. On the basic ideas of the linearization method, the algorithms for solving the various problems of mathematical programming are constructed for (a) solving systems of equalities and inequalities, (b) multiobjective programming and (c) complementary problem.

**Key words.** The linearization method, systems of equalities and inequalities, multiobjective programming, complementarity problem.

## 1. Introduction

The linearization method as the method of a numerical solution of systems of nonlinear equalities and inequalities was first formulated in [1]. The algorithm constructed in [1] was the generalization of Newton's well-known method for general systems of equalities and inequalities and it employed the basic idea of this method – the replacement of the nonlinear system by its first linear approximation at each step. To this end, a need arose for including additional mechanisms which allows one to choose a single solution from a set of solutions of the linearized system. A smooth penalty for great deviations from the current approximation was chosen as such a mechanism, since under great deviations, the precision of approximation of a nonlinear system by a linear one is somewhat curtailed.

In [2] the linearization method was formulated as applied to solving the general problem of nonlinear programming. To provide global convergence, the nonsmooth penalty function was used here. The algorithm of the paper [2] is the effective working tool in numerical solutions of nonlinear programming problems.

In recent years the linearization method was analyzed from a theoretical point of view; a set of problems being solved by this method was extended, its modifications were suggested, and rules of changing algorithm parameters were improved [3, 4, 5, 6–9]. Wide experience in practical solutions of problems has been gained which shows a high efficiency of the method. These changes and peculiarities, analysis of the state of the linearization method in a concise form, are presented in [14].

It turned out that the basic ideas of the linearization method are useful for solving not only the general problem of nonlinear programming but also a number of problems closely connected to the usual problem of mathematical

*Journal of Global Optimization* **3**: 483–500, 1993.

programming – the system of equalities and inequalities, the multiobjective programming and the complementary problem. It is the aim of this paper to present the computational methods for solving these problems and a comprehensive investigation into the properties of the formulated algorithms. We present little or no proofs since the facts given below either have already been published or may be obtained with the help of a simple change of the known constructions.

## 2. Notation

The subsequent presentation will be carried out in the $n$-dimensional Euclidian space $\mathbb{R}^n$ with the usual inner product $x^T y$ of $x, y \in \mathbb{R}^n$ and norm $\|x\|$ of $x \in \mathbb{R}^n$. Components of vectors are denoted by $x^T = (x_1, x_2, \dots, x_n)$ where the superscript $T$ denotes the opposite. Thus $x$ is a column-vector. Inequalities for vectors will be understood as inequalities for all corresponding elements of vectors. Gradients of smooth functions $f_i(x)$, $g_i(x)$, $i \in \{1, 2, \dots, n\}$ are denoted by $\nabla f_i(x)$, $\nabla g_i(x)$, $i \in \{1, 2, \dots, n\}$ and are the row vectors. The matrix of first derivatives (Jacobian matrix) is denoted by $\nabla f(x)$ and $\nabla g(x)$ for the respective functions.

## 3. Principal Concepts

The main problem for the solution of which the linearization method was formulated is the general nonlinear programming problem. It is required to minimize the function $f_0(x)$, $x \in \mathbb{R}^n$, with constraints $f_i(x) \leq 0$, $i \in \mathbf{I}$, where $\mathbf{I}$ is a finite set of indices, i.e., in a concise form,

$$\mathbf{NP}\colon \min\{f_0(x) \mid f_i(x) \leq 0\,, \quad i = 1, 2, \dots, m\}\,.$$

As in the majority of methods for solving nonlinear problems, each step of the linearization method requires the choice of direction of the shift and the value of this shift. The efficiency of the algorithm depends on a successful solution of the problem of this choice.

To extend Newton's method to a solution of the problem **NP**, it is primarily natural to replace the problem **NP** by the linear approximation

$$\mathbf{NP}_x\colon \min\{f_0(x) + \nabla f_0(x)p \mid f_i(x) + \nabla f_i(x)p \leq 0\,, \quad i = 1, 2, \dots, m\}$$

with a subsequent transfer to a new point $x + p$.

It is clear that such an attempt results in a number of difficulties. The principal difficulty arises in the fact that even though the constraints of the problem $\mathbf{NP}_x$ are consistent, the value of the lower bound equals $-\infty$, and $\mathbf{NP}_x$ has no bounded solution. One of the methods for overcoming this difficulty is found in that constraints of the problem $\mathbf{NP}_x$ should be supplemented with the inequality $\|p\| \leq \varepsilon$. Unfortunately, this leads to the inconsistency of the problem constraints for small $\varepsilon > 0$ and requires the additional procedure of changing $\varepsilon$. This difficulty has not been adequately overcome. Besides, as a rule, the solution $p(x)$ of the problem $\mathbf{NP}_x$, supplemented with the constraint $\|p\| \leq \varepsilon$, has a norm equal to $\varepsilon$ (this constraint is active) that leads to the low rate of convergence because of the need to subdivide the length of the shift along the direction $p(x)$, up to zero.

The second way to modify the problem of $\mathbf{NP}_x$ with the aim that its solution exist and be finite, is to reduce the number of constraints of this problem and take into account the fact that large $p$ poorly approximates the initial problem of **NP**, through introducing an additional quadratic term into the objective function of the problem $\mathbf{NP}_x$. Thus we come to the main auxiliary problem

$$\mathbf{NP}_x^\delta\colon \min\{\nabla f_0(x)p + \tfrac{1}{2}\|p\|^2 \mid f_i(x) + \nabla f_i(x)p \leq 0\,, \quad i \in \mathbf{I}_\delta(x)\}\,,$$

where $\mathbf{I}_\delta(x) = \{i \mid F(x) \geq f_i(x) - \delta\}$, $\delta > 0$, $F(x) = \max\{0, f_1(x), \ldots, f_m(x)\}$.

This problem is associated with the dual problem

$$\mathbf{DNP}_x^\delta\colon \max\left\{-\frac{1}{2}\|\nabla f_0(x) + \sum_{i\in\mathbf{I}_\delta(x)} u^i \nabla f_i(x)\|^2 + \sum_{i\in\mathbf{I}_\delta(x)} u^i f_i(x) \mid u_i \geq 0\,, \quad i \in \mathbf{I}_\delta(x)\right\}.$$

If constraints of the problem $\mathbf{NP}_x^\delta$ are consistent it has a unique solution: $p(x)$. Denote the corresponding solution of the dual problem by $u^i(x)$. (If the solution $\mathbf{DNP}_x^\delta$ is not unique we choose any one.)

Now we choose $p(x)$ as a direction of the shift from the point $x$ to construct the next approximation: $x + \alpha p(x)$. The second main concept now involves choosing a size of step $\alpha$ from condition of the descending penalty function $\Phi_N(x) = f_0(x) + \mathbf{N}F(x)$ under a suitable value $\mathbf{N} > 0$. The question is what value $\mathbf{N}$ is suitable. The following result gives the answer to this question: if $\mathbf{N} > \Sigma_{i\in\mathbf{I}_\delta(x)} u^i$, then the successive halving of $\alpha = 1$ up to the first fulfillment of the inequality

$$\Phi_N(x + \alpha p(x)) \leq \Phi_N(x) - \varepsilon\alpha\|p(x)\|^2\,, \quad \varepsilon \in (0, 1) \tag{3.1}$$

will be completed in the finite number of steps.

Thus due to the fact that the auxiliary problem $\mathbf{NP}_x^\delta$ does not depend on $\mathbf{N}$, its solution answers the question of whether the value $\mathbf{N}$ is chosen correctly. The fundamental role of the fact of independence of $\mathbf{NP}_x^\delta$ of $\mathbf{N}$ is emphasized by the following theorem which serves as the basis for constructing the methods with accelerated convergence.

THEOREM 3.1. *For the point $x$ to satisfy the constraints of the problem* **NP** *and establish the minimum necessary conditions of the form*

$$\nabla f_0(x) + \sum_{i=1}^{m} u^i \nabla f_i(x) = 0 ,$$

$$u^i f_i(x) = 0 , \quad u^i \geqslant 0 , \quad i = 1, 2, \ldots , m$$

*it is necessary and sufficient that* $p(x) = 0$. ■

Thus, $p(x)$ not only serves for construction of successive approximations but at the same time is an indicator of a solution.

On the basis of the above we may now formulate the algorithm to solve the problem **NP**. For its operation some parameters $\mathbf{N} > 0$, $\delta > 0$, $\varepsilon \in (0, 1)$ and the initial point $x^0$ should be established. We assume the following:

1. The domain $\Omega_N(x^0) = \{x \mid \Phi_N(x) \leqslant \Phi_N(x^0)\}$ is compact;
2. in this domain, the gradients of all functions entering into the problem satisfy the Lipschitz condition;
3. for $x \in \Omega_N(x^0)$ the problem $\mathbf{NP}_x^\delta$ is solvable and therein may be found a solution for the dual problem $\mathbf{DNP}_x^\delta$ that $\Sigma_{i \in \mathbf{I}_\delta(x)} u^i < \mathbf{N}$.

The steps of the linearization method algorithm, when the point $x^k$ is already constructed, consists of the following:

1. We solve the problem $\mathbf{NP}_x^\delta k$ and find $p(x^k)$.
2. By halving $\alpha = 1$ successfully we find $\alpha^k$ as the first for which the inequality (3.1) is satisfied for $x = x^k$.
3. We put $x^{k+1} = x^k + \alpha^k p(x^k)$.

The properties of the given algorithm are summarized in the following theorem.

THEOREM 3.2. *Under the above assumptions the sequence* $\{x^k\}$ *of points generated by the algorithm has the following properties:*

(a) *the sequence* $\{x^k\}$ *is bounded*;

(b) $F(x^k) \to 0$ *and* $p(x^k) \to 0$;

(c) *there exists the number* $\bar{\alpha} > 0$ *such that* $\alpha^k \geqslant \bar{\alpha}$;

(d) *any limit point* $x^k$ *of the sequence* $\{x^k\}$ *satisfies the constraints of the initial problem* **NP** *and the necessary minimum conditions;*

(e) *the linear programming problem is solved in a finite number of steps, and from some moment on* $\alpha^k = 1$;

(f) *if we put* $u^i(x^k) = 0$, $i \not\in \mathbf{I}_\delta(x^k)$, *then vectors* $u(x^k) = (u^1(x^k), \ldots , u^m(x^k))^T$ *are bounded and their limit point is Lagrangian multipliers for the initial problem* **NP**. ■

Thus, the theorem shows the algorithm convergence in a universally adopted sense. The discussion of the assumptions under which this convergence is proved and the estimation of its rate can be found in [10–13] or in concise form in [14].

## 4. Systems of Equalities and Inequalities

In this section, the linearization method is applied to solving systems of equalities and inequalities. It proves that in this case one can succeed in constructing effective algorithms which have a fast rate of convergence.

Given two finite sets of indices $\mathbf{I}^-$ and $\mathbf{I}^0$ and functions $f_i(x)$, $x \in \mathbb{R}^n$, find the solution of the following systems:

$$f_i(x) \leqslant 0 , \quad i \in \mathbf{I}^- , \quad f_i(x) = 0 , \quad i \in \mathbf{I}^0 . \tag{4.1}$$

Suppose that functions $f_i(x)$ have continuous gradients $\nabla f_i(x)$ and also that the gradients satisfy the Lipschitz condition with constant $\mathbf{L}$: $\|\nabla f_i(x^1) - \nabla f_i(x^2)\| \leqslant \mathbf{L}\|x^1 - x^2\|$.

We use the notation:

$$F(x) = \max\left\{\max_{i \in \mathbf{I}^-} f_i(x),\ \max_{i \in \mathbf{I}^0} | f_i(x)|\right\} ,$$

$$\mathbf{I}_\delta^-(x) = \{i \,|\, i \in \mathbf{I}^- ,\ f_i(x) \geqslant F(x) - \delta\} ,$$

$$\mathbf{I}_\delta^0(x) = \{i \,|\, i \in \mathbf{I}^0 ,\ | f_i(x) | \geqslant F(x) - \delta\} .$$

We choose an initial point $x^0$ and assume that for all $x$ that satisfy the inequality $F(x) \leqslant F(x^0)$, the gradients $\nabla f_i(x)$ are limited in norm by the constant $\mathbf{K}$.

BASIC ASSUMPTION. There are numbers $\delta > 0$ and $\mathbf{C} > 0$ such that for all $x$ for which $F(x) > 0$, $F(x) \leqslant F(x^0)$ the following system is solvable for $p$:

$$\begin{aligned} &\nabla f_i(x)p + f_i(x) \leqslant 0 , \quad i \in \mathbf{I}_\delta^- , \\ &\nabla f_i(x)p + f_i(x) = 0 , \quad i \in \mathbf{I}_\delta^0 . \end{aligned} \tag{4.2}$$

Let $p(x)$ be the solution of (4.2) that has the minimum norm. Then for $x$ such that $F(x) > 0$:

$$\|p(x)\| \leqslant \mathbf{C}F(x) . \tag{4.3}$$

The inequality (4.3) characterizes, to a certain extent, the regular solvability of system (4.2). In particular, if the system (4.2) is transformed into a system of $n$ equations in $n$ unknowns, condition (4.3) is equivalent to the assumption that the matrix of the corresponding systems is nonsingular. As will be shown further on, (4.3) holds if the gradients $\nabla f_i(x)$, $i \in \mathbf{I}_\delta^-(x) \cup i \in \mathbf{I}_\delta^0(x)$ are linearly independent for all $x$, $F(x) > 0$.

We turn now to the construction of the algorithm. The successive approximations are constructed by the formula

$$x^{k+1} = x^k + \alpha^k p(x^k) , \quad p^k = p(x^k) \tag{4.4}$$

where parameter $\alpha^k$ is chosen by sequentially halving unity until the following inequality is satisfied:

$$F(x^k + \alpha^k p(x^k)) \leq (1 - \varepsilon\alpha^k)F(x^k) \tag{4.5}$$

where $\varepsilon$ is any number, chosen from the beginning, $0 < \varepsilon < 1$. Clearly, formula (4.4) is applicable if $F(x) > 0$. Otherwise, the process stops and $x^k$ is the solution of problem (4.1).

CONVERGENCE OF THE ALGORITHM. The implementation of the algorithm proposed is characterized by the following theorem.

THEOREM 4.1. *Let all the assumptions of the preceding subsection be fulfilled. Then sequence* $\{x^k\}$, $k = 0, 2, \ldots,$ *generated by the algorithm according to formula* (4.4) *converges to* $x^*$, *the solution of system* (4.1), *and at the same time*

(a) *for a sufficiently great* $k$, $\alpha^k = 1$;

(b) *for a sufficiently great* $k$,

$$F(x^{k+1}) \leq \mathbf{LC}^2 F^2(x^k)\,;$$

(c) *for any* $q$, $0 < q < 1$ *there is a number* $k(q)$ *such that*

$$\|x^* - x^k\| \leq \frac{q^{2^{k-k(q)}}}{\mathbf{LC}(1 - q)} \tag{4.6}$$

*for all* $k \geq k(q)$. ■

REMARK 4.1. Let us solve a system of $n$ equations $f_i(x) = 0$, $i = 1, 2, \ldots, n$, where $x \in \mathbb{R}^n$. Then $f_i(x) \geq F(x) - \delta$, $i = 1, \ldots, n$, $F(x) = \max_{1 \leq i \leq n} |f_i(x)|$ for any $\delta$, provided $x$ is sufficiently close to the solution $x^*$. Therefore, $\mathbf{I}_\delta^0(x) = \{1, 2, \ldots, n\}$ and system (4.2) takes the form

$$\nabla f_i(x)p + f_i(x) = 0\,, \quad i = 1, 2, \ldots, n\,. \tag{4.7}$$

Therefore the method proposed coincides with Newton's method in which iterations are performed by the formula $x^{k+1} = x^k + p(x^k)$, where $p(x^k)$ is the solution of system (4.7). The condition for the convergence of Newton's method is the nonsingularity at point $x^*$ of matrix $\nabla f(x^*)$, where $\nabla f(x^*)$ is an $n \times n$ matrix whose rows are $\nabla f_i(x^*)$. In this case, $p(x) = -(\nabla f(x))^{-1} f(x)$, where $f(x)$ is a column-vector whose components are $f_i(x)$. But it follows from the last formula that $\|p(x)\| \leq \|(\nabla f(x))^{-1}\|\,\|f(x)\| \leq \mathbf{C}_0\|(\nabla f(x))^{-1}\|F(x)$ where $\mathbf{C}_0$ is a constant. It can be seen from this inequality that (4.3) holds in a certain neighborhood of point $x^*$.

Thus it follows from the theorem proved that the usual Newton's method is locally convergent in solving a system of $n$ equations with $n$ unknowns.

REMARK 4.2. If only one equation $f(x) = 0$ in $n$ unknowns is to be solved, then system (4.2) takes the form

$$\nabla f(x)p + f(x) = 0\,, \tag{4.8}$$

and it is required to find the solution of this equation with a minimum norm, i.e. to find the minimum of $\|p(x)\|^2$ with constraints (4.8). Using the rule of Lagrange multipliers, we have in this case

$$p(x) = -\frac{f(x)}{\|\nabla f(x)\|^2}\nabla f(x),\ \text{hence}\ \|p(x)\| = \frac{1}{\|\nabla f(x)\|}\,|\nabla f(x)|\,.$$

Clearly, formula (4.3) will be satisfied if $\|\nabla f(x)\| \geq \gamma$ for all $x$.

SUFFICIENT CONDITIONS OF CONVERGENCE. The main condition (4.3) which guarantees the convergence of the algorithm is not easy to check. This subsection describes conditions that can be checked more effectively. In particular, for the convex case if there is an interior point in the domain defined by expressions (4.1), the conditions guarantee the convergence of the algorithm.

Let the system contain only inequality constraints, i.e.

$$f_i(x) \leq 0\,, \quad i \in \mathbf{I}^-\,. \tag{4.9}$$

Then the subsidiary system (4.2) takes the form

$$\nabla f_i(x)p + f_i(x) \leq 0\,, \quad i \in \mathbf{I}_\delta^-\,. \tag{4.10}$$

Clearly, this system can be solved with $F(x) > 0$ if the system

$$\nabla f_i(x)p + F(x) \leq 0\,, \quad i \in \mathbf{I}_\delta^- \tag{4.11}$$

is solvable.

LEMMA 4.1. *If $F(x) > 0$, then system* (4.11) *has a solution if and only if*

$$\mathbf{L}_\delta(x) = \min_{\lambda_i \geq 0}\Big\| \sum_{i \in \mathbf{I}_\delta(x)} \lambda_i \nabla f_i(x)\Big\| > 0$$

*where the minimum is taken over all $\lambda_i \geq 0$ such that $\Sigma_{i\in \mathbf{I}_\delta^-(x)} \lambda_i = 1$. Then the solution $p^*(x)$ of system* (4.11) *with a minimum norm satisfies the equality $\|p^*(x)\| = 1/[\mathbf{L}_\delta(x)]F(x)$.* ■

THEOREM 4.2. *Let all the assumptions of the subsection hold, except the basic* *one. Moreover, let* $\mathbf{L}_\delta(x) \geq \gamma > 0$ *for all* $x$ *such that* $0 < F(x) \leq F(x^k)$. *Then the conditions of the basic assumption are fulfilled too and all the results of Theorem* 4.1 *hold for problem* (4.9). ∎

Note that the condition $\mathbf{L}_\delta(x) \geq \gamma > 0$ is natural enough, for it requires linear independence of vectors $\nabla f_i(x)$, $i \in \mathbf{I}_\delta^-(x)$.

THEOREM 4.3. *Let functions* $f_i(x)$ *in problem* (4.9) *be convex and continuously differentiable. Besides, let the domain defined by the inequality* $F(x) \leq F(x^0)$ *be compact, the gradients* $\nabla f_i(x)$ *in this domain satisfy Lipschitz condition and there is a point* $x^*$ *such that* $F(x^*) = \gamma < 0$. *Then with* $\delta < -\gamma$ *all the conditions of Theorem* 4.1 *are fulfilled.* ∎

The next problem, as shown below, is strongly related to the problem of solving systems of equalities and inequalities.

## 5. Nonlinear Complementarity Problem

In the traditional setting the nonlinear complementarity problem (NCP) can be defined as follows.

Let a function $f(x)$: $\mathbb{R}^n \to \mathbb{R}^n$ be given. Find a point $x \in \mathbb{R}^n$, such that

$$x \geq 0\,, \quad f(x) \geq 0\,, \quad x^T f(x) = 0\,. \tag{5.1}$$

When considering only a linear function $f(x) = \mathbf{M}x + q$, where the matrix $\mathbf{M} \in \mathbb{R}^n \times \mathbb{R}^n$, and the vectors $x$, $q \in \mathbb{R}^n$, we have the linear complementarity problem (LCP) of finding an $x \in \mathbb{R}^n$ such that for a given $\mathbf{M}$ and $q$

$$x \geq 0\,, \quad \mathbf{M}x + q \geq 0\,, \quad x^T(\mathbf{M}x + q) = 0\,. \tag{5.2}$$

The complementarity problem (CP) appears in various fields such as mathematical programming, game theory, economic equilibrium, etc.

### THE RELATIONSHIP OF THE CP WITH THE LINEAR PROBLEM OF MOMENTS

The idea of the NCP (5.1) solution lies in not solving the problem itself, but the equivalent system of nonlinear inequalities:

$$f(x) \geq 0\,, \quad x \geq 0\,, \quad x^T f(x) \leq 0\,. \tag{5.3}$$

The linearization algorithm [10] is used for this purpose. The auxiliary problem of the latter has the form:

$$
\begin{aligned}
&\min_p\{\tfrac{1}{2}\|p\|^2\}, \text{ with constraints}\\
&\nabla f_i(x)p + f_i(x) \geqslant 0\,, \quad i \in \mathbf{I}\,,\\
&\qquad p_i + x_i \geqslant 0\,, \quad i \in \mathbf{I}\,,\\
&\nabla l(x)p + l(x) \leqslant 0\,,
\end{aligned}
\tag{5.4}
$$

where $\mathbf{I} = \{1, 2, \ldots, n\}$.

Here, by $l(x)$, the expression $x^T f(x)$ is denoted. It should be noted that $\nabla l(x) = x^T \nabla f(x) + f(x)^T$.

The solution to this problem, in a sense, is equivalent to the solution of a linear problem of moments [13]. In [15] this fact has been considered in detail. Some generalizations of the results obtained in [15] can be found in [16] and [17].

Let us write the Lagrange expression for the constraints of problem (5.4): $L(p, u, v, u_0) = \mathbf{v}^T(p + x) + \mathbf{u}^T(\nabla f(x)p + f(x)) - \mathbf{u}_0(\nabla l(x)p + l(x)) = (\mathbf{v}^T + \mathbf{u}^T\nabla f(x) - \mathbf{u}_0\nabla l(x))p + \mathbf{v}_x^T + \mathbf{u}^T f(x) - \mathbf{u}_0 l(x)$.

Here $\mathbf{u} \geqslant 0$, $\mathbf{v} \geqslant 0$, $\mathbf{u}_0 \geqslant 0$ are Lagrange multipliers for the respective constraints. From Theorem 5.10 [13] it follows that if there exists a constant $\mathbf{C} > 0$ such that the inequality

$$
\mathbf{C}\|\mathbf{v}^T + \mathbf{u}^T\nabla f(x) - \mathbf{u}_0\nabla l(x)\| \geqslant -\mathbf{v}^T x - \mathbf{u}^T f(x) + \mathbf{u}_0 l(x) \tag{5.5}
$$

is satisfied for any $\mathbf{u} \geqslant 0$, $\mathbf{v} \geqslant 0$, $\mathbf{u}_0 \geqslant 0$ then the system of linear inequalities of the problem (5.4) is solvable and $\|p\| \leqslant \mathbf{C}$.

Thus, taking account of the above-said, it is easy to formulate the following proposition.

THEOREM 5.1. *If the domain $\Omega_\varepsilon = \{x \geqslant 0\colon f(x) \geqslant -\varepsilon,\ x^T f(x) \leqslant \varepsilon\}$, where $\varepsilon > 0$, is a compact set and there is a constant $\mathbf{C} > 0$ such that inequality (5.5) is satisfied at all points of the domain $\Omega_\varepsilon$ for any vectors $\mathbf{u} \geqslant 0$, $\mathbf{v} \geqslant 0$ and scalar value $\mathbf{u}_0 \geqslant 0$, then the auxiliary problem (5.4) has a solution bounded in norm by the constant $\mathbf{C}$ everywhere in $\Omega_\varepsilon$.* ∎

In our opinion this theorem is essential because NCP can now be seen from a relatively new viewpoint. Namely, it reveals the relation of the latter to important mathematical problems, i.e., the linear problem of moments.

And there, the solvability condition in form (5.5) for the auxiliary problem (5.4), formulated by Theorem 5.1, is actually a sufficient condition for existence of the NCP solution, and as shown later, it is strongly related to other well-known similar conditions.

BASIC ASSUMPTIONS. Henceforth, besides assumptions of Theorem 5.1, we shall make the following assumptions:

(1) All functions $f_i(x)$, $i \in \mathbf{I}$ are continuously differentiable;

(2) The gradients $\nabla f_i(x)$, $i \in \mathbf{I}$ and the vector-function $f(x)$ satisfy the Lipschitz condition on the compact set $\Omega_\varepsilon$ with the constant $\mathbf{L}$.

(3) The solution $x^*$ of NCP satisfies the relations

$$\begin{cases} x_i^* = 0, & f_i(x^*) > 0, \quad i \in \bar{\mathbf{I}} \subset \mathbf{I}, \\ x_i^* > 0, & f_i(x^*) = 0, \quad i \in \bar{\bar{\mathbf{I}}} \equiv \mathbf{I} \backslash \bar{\mathbf{I}}. \end{cases}$$

(4) The gradients $\nabla f_i(x)$, $i \in \bar{\bar{\mathbf{I}}}$ at the point $x^*$ are linearly independent.

Thus, the general conditions (1) and (2) of linearization algorithm [10], are supplemented by two more sufficient new assumptions. However, these hypotheses are not too onerous. In fact, the assumptions (3) and (4) quite logically require only a certain solution-regularity and, as we shall see later allow for a transition in a certain neighborhood of the point $x^*$ to a procedure which ensures quadratic convergence of the below defined algorithm.

## THE FORMULATION OF THE ALGORITHM

Now to formulate the computation procedure, offering a solution to a system of nonlinear inequalities (5.3) and, consequently, of NCP (5.1) itself.

Let the numbers $0 < \gamma < 1$, $0 < \xi < 1$ be chosen. We describe the general step of an algorithm. Let $x^k$ be given.

1. Solve the auxiliary problem (5.4) with $x = x^k$ and determine $p^k = p(x^k)$.
2. Divide a set of indices $\mathbf{I}$ into two groups $\bar{\mathbf{I}}$ and $\bar{\bar{\mathbf{I}}}$ in the following way:

$$\bar{\mathbf{I}} = \{i \in \mathbf{I}\colon \gamma f_i(x^k) \geq x_i^k\},$$
$$\bar{\bar{\mathbf{I}}} = \{i \in \mathbf{I}\colon \gamma x_i^k \geq f_i(x^k)\}.$$

Verify the satisfaction of the condition $\bar{\mathbf{I}} \cup \bar{\bar{\mathbf{I}}} = \mathbf{I}$. If it is not fulfilled, go to step 4. If the latter is true, then calculate $y^k$ by solving the following system of linear equations:

$$\begin{cases} y_i^k + x_i^k = 0, & i \in \bar{\mathbf{I}}, \\ \nabla f_i(x^k) y^k + f_i(x^k) = 0, & i \in \bar{\bar{\mathbf{I}}}. \end{cases} \tag{5.6}$$

If the system has no solution, then go to step 4.

3. If the system (5.6) has a solution $y^k$, then we set $\bar{x} = x^k + y^k$. Verify the validity of a system of inequalities

$$\begin{cases} \gamma f_i(\bar{x}) \geq \bar{x}_i, & i \in \bar{\mathbf{I}}, \\ \gamma \bar{x}_i \geq f_i(\bar{x}), & i \in \bar{\bar{\mathbf{I}}}. \end{cases}$$

If the latter is satisfied, then we set $x^{k+1} = \bar{x}$ and go back to step 1, else we go to step 4.

4. At first set $\alpha = 1$. Divide $\alpha$ by $2^k$ $k = 1, 2, \ldots$ *until the inequality* $Y(x^{k+1}) \leq$

$(1-\alpha\xi)\mathbf{Y}(x^k)$ is satisfied, where $\mathbf{Y}(x)=\max\{0, -f_1(x), \ldots, -f_n(x), l(x)\}$, and $x^{k+1}=x^k+\alpha_k p^k$.

Go back to step 1.

THEOREM 5.2. *Let conditions of Theorem* 5.1 *and the main assumptions* (1) *and* (2) *be satisfied. Then the suggested algorithm generates a sequence* $\{x^k\}$ *for which* $\mathbf{Y}(x^k)\to 0$. *If in addition, the conditions* (3) *and* (4) *are satisfied, and the point* $x^*$ *is the unique solution of* NCP, *then the sequence* $\{x^k\}$ *converges quadratically to* $x^*$. ■

REMARK 5.1. As can be seen from an algorithm construction, transition from $x^k$ to $x^{k+1}$ is stipulated by steps 3 or 4. The latter corresponds to the general linearization algorithm [10] and guarantees only the existence of such a constant $\mathbf{C}>0$, that $\mathbf{Y}(x^k)\leq \mathbf{C}/k$. Thus convergence is liable to be slow by virtue of the first statement of Theorem 5.2.

REMARK 5.2. To prove the second statement of Theorem 5.2, it is necessary to evaluate the expression $\|x^{k+1}-x^*\|$, where $x^{k+1}$ is chosen in step 3 of the algorithm. It is obvious that transition from $x^k$ to $x^{k+1}$ in step 3 results in transition to a quadratic rate of convergence. It is just at this step that local properties of the complementarity problem are taken into account, and a faster convergence rate is possible only in a local neighborhood of solution $x^*$. Namely, in the region where a system of the type (5.6) makes sense. Note that the latter is correctly defined in a certain neighborhood of solution $x^*$ by virtue of assumption (3) and continuity of the function $f(x)$. Assumption (4) guarantees solvability of the latter in the neighborhood of the point $x^*$.

REMARK 5.3. The criterion of proximity to a solution in the form of determination of a set of indices $\bar{\mathbf{I}}$ and $\bar{\bar{\mathbf{I}}}$ permits us to solve in a reasonable way the problem of combining a global convergence property for the general NCP algorithm with a quadratically convergent local Newton's procedure.

EXISTENCE THEORY. It is not difficult to see that the auxiliary problem (5.4) representing the quadratic programming problem, is an essential link in formulating the NCP algorithm.

Note that the condition of existence for the solution to the auxiliary problem (5.4) formulated by the Theorem 5.1 is actually the condition of existence of the NCP solution. In fact, if the inequality (5.5) of Theorem 5.1 holds, then the auxiliary problem (5.4) is solvable, and this fact, in turn, brings about the existence of a solution to problem (5.3). The proof of the latter fact is given by Theorem 5.2 by establishing the convergence of the algorithm to such a solution.

In this sense it is possible to talk about a constructive existence theory for the NCP. As mentioned above, condition (5.5) results from the relation of the last problem to the linear problem of moments. The next theorem investigates the question of how intrinsic this relationship is.

Now let us recall that the first general theorems of the NCP solution existence used various forms of function monotonicity $f(x)$. Thus, in [18] Karamardian proved the existence and uniqueness of a solution under the condition that mapping $f(x)$ is continuously and strongly monotonic, i.e. there exists such $\mu > 0$, that for any $x, y \geqslant 0$ the inequality $(f(x) - f(y))^T(x - y) \geqslant \mu \|x - y\|^2$ is valid.

In the special case when the mapping $f(x): \mathbb{R}^n \rightarrow \mathbb{R}^n$ is continuously differentiable, the existence condition for the solution of the auxiliary problem (5.4) stated by Theorem 5.1 follows from the condition of strong monotonicity of $f(x)$.

THEOREM 5.3. *Let the domain $\Omega_\varepsilon$ be a compact set as before. Suppose now that mapping $f(x)$ is continuously differentiable and strongly monotonic. Then there exists a constant $\mathbf{C} > 0$, such that relation* (5.5) *is satisfied for any $\mathbf{u} \geqslant 0$, $\mathbf{v} \geqslant 0$, $\mathbf{u}_0 \geqslant 0$, at all points of the domain $\Omega_\varepsilon$.* ■

At present there are a great number of publications on various generalizations of theorems of NCP (5.1) solution existence, including the condition of coerciveness of operator $f(x)$ [19], the condition of existence in terms of "exceptional sequences" [20], and others. Theorems 5.1 and 5.3 show that the new condition of an existence of a solution to the complementarity problem formulated in the present work is natural in the context of already-known and similar conditions in the general theory of the problem.

Further confirmation of the above-claim can be supplied by the results concerned with the linear complementarity problem (5.2).

Note that Theorem 5.1 substantiates the LCP solution algorithm. Convergence of the latter can be proved with less strict constraints. Actually, in Theorem 5.2 it is not necessary to require fulfillment of the basic conditions (1) and (2), since they are held automatically for linear functions. Note that strong monotonicity of $f(x)$ implies positive definiteness of the Jacobian matrix $\nabla f(x)$ which coincides with matrix $\mathbf{M}$ in the case of the linear complementarity problem (5.2). In this case one may talk about strict copositiveness of matrix $\mathbf{M}$. Then the following statement is valid.

COROLLARY 5.1. *Let $\Omega_\varepsilon$ be a compact set as before. Let also $f(x) = \mathbf{M}x + q$. Besides, let matrix $\mathbf{M}$ be strictly copositive (i.e. $y^T\mathbf{M}y > 0$ holds for any $y \geqslant 0$, $y \neq 0$). Then there exists a constant $\mathbf{C} > 0$ such that the relation*

$$\begin{aligned} C\|v^T + \mathbf{u}^T\mathbf{M} - \mathbf{u}_0(x^T(\mathbf{M}^T + \mathbf{M}) + q^T)\| &\geqslant -\mathbf{v}^Tx - \mathbf{u}^T(\mathbf{M}x + q) \\ &+ \mathbf{u}_0(x^T\mathbf{M}x + x^Tq) \end{aligned} \tag{5.7}$$

holds at all points of the domain $\Omega_\varepsilon$. ■

To prove this corollary it suffices to note that in linear cases the concepts of strict copositiveness and strong copositiveness of matrix $\mathbf{M}$ coincide.

The following statement reveals a connection of the condition of the LCP solution of the type (5.7) with the similar well-known condition, e.g. the mapping $f(x): \mathbb{R}^n \to \mathbb{R}^n$ being coercive.

COROLLARY 5.2. *Let the requirements of the previous corollary be satisfied. Moreover, let the operator $f(x)$ be coercive (i.e. for any $x \in \mathbb{R}^n$ $x^T f(x) \geq \gamma(\|x\|)\|x\|$ holds, where $\gamma(t) \to +\infty$ and $t \to +\infty$, $t > 0$). Then there exists a constant $\mathbf{C} > 0$ such that* (5.7) *is valid at all points of the domain $\Omega_\varepsilon$.* ■

The proof of the given statement is based on the fact that determinations of coercitiveness of operator $f(x)$ and of strict copositiveness of the matrix $\mathbf{M}$ in the linear case coincide.

Thus, it was shown, that a condition of type (5.7) is natural. Namely, it proved to be closely related to earlier similar known conditions for the existence of a solution for the LCP.

Therefore it seems reasonable to suppose that the proposed algorithm for a wide class of complementarity problems would be reasonably efficient. Several examples of numerical verification of this conclusion can be found in [16].

## 6. Multiobjective Optimization Problem

A vector optimization problem arises at the first stage of the general procedure of decision-making, when from the whole set of admissible alternatives some set is singled out, each element of which meets the requirements of an efficient solution.

Let the functions $f_i(x)$, $i = 1, 2, \ldots, m$, be specified. They form the vector criterion $f(x) = (f_1(x), \ldots, f_m(x))$ of some vector optimization problem

$$\begin{aligned}&\min f(x), \text{ under constraints}\\&\quad x \in \mathbb{X}_g = \{x \in \mathbb{R}^n \mid g_j(x) \leq 0\,, \quad j = 1, 2, \ldots, l\}\,. \end{aligned} \tag{6.1}$$

When considering only linear functions $f_i(x) = c_i^T x$, $i = 1, 2, \ldots, m$ and $g_j(x) = a_j^T x$, $j = 1, 2, \ldots, l$ we may already say about a linear vector optimization problem

$$\begin{aligned}&\min Cx \text{ under constraints}\\&x \in \mathbb{X} = \{x \in \mathbb{R}^n \mid Ax \leq b\}\,,\end{aligned} \tag{6.2}$$

where $C$ is the $(n \times m)$-matrix of objective function coefficients, $A$ is the $(n \times l)$-matrix of constraint coefficients and $b$ is the $l$-vector.

The majority of the known strategies of the vector optimization problem solution lie in characterizing effective solutions in terms of optimal solutions of some corresponding scalar optimization problem. The difference in methods of reducing the vector criterion $f(x)$ to one objective function defines the conventional classification of approaches to a solution of the vector optimization problem.

This section investigates the properties of methods based on the idea of linearization in the aspect of multi-objective optimization. These properties are characterized in terms of weak-efficient, efficient (Pareto-optimal) and proper-efficient solutions (see [16], [21]–[23]). The obtained results are true for a linear case [24].

Let us introduce some definitions [21]. The solution **x** is called weak-efficient (efficient) if there is no such $x \neq x^*$ that $f(x) < (\leq) f(x^*)$ is satisfied.

Admissible solution $x^*$ is called proper-efficient if it is efficient and if there exists such a positive number **M** that for any $i = 1, 2, \ldots, m$ and $x \in \mathbb{X}_g$ for which the inequality $f_i(x^*) > f_i(x)$ is satisfied and some $\nu \in \{1, \ldots, m\}$ such that $f_\nu(x^*) < f_\nu(x)$ the inequality $[(f_i(x^*) - f_i(x))/(f_\nu(x) - f_\nu(x^*)))] \leq \mathbf{M}$ is satisfied.

BASIC ASSUMPTIONS. Let us consider some modification of the linearization algorithm [10] based on necessary conditions of the efficiency of the problem (6.1). Relate the auxiliary problem to every point $x \in \mathbb{X}_g$:

$$\min_{p,\xi} \{\xi + \tfrac{1}{2}\|p\|^2\}, \text{ under constraints}$$
$$\nabla f_i(x)p \leq \xi\,, \quad i \in I\,,$$
$$\nabla g_j(x)p + g_j(x) \leq 0\,, \quad j \in J\,, \tag{6.3}$$

where $I = \{1, 2, \ldots, m\}$, $J = \{1, 2, \ldots, l\}$.

Introduce the following assumptions. Let there exist such $N > 0$ that

(a) for some $i \in I$ the set $\Omega_N = \{x \in \mathbb{R}^n \mid f_i(x) + N \cdot G(x) \leq C_i\}$ is limited, where $C_i = f_i(x_0) + N \cdot G(x_0)$, $G(x) = \max\{0, g_1(x), \ldots, g_1(x)\}$;

(b) gradients $\nabla f_i(x)$, $i \in I$ and $\nabla g_j(x)$, $j \in J$ in $\Omega_N$ satisfy Lipschitz condition with the constant $L$;

(c) there exist such Lagrange multipliers of the problem (6.3) $v_j(x)$, $j \in J$, that $\Sigma_{j \in J} v_j(x) \leq N$ and the latter is solvable with respect to $p \in \mathbb{R}^n$ for any $x \in \Omega_N$.

It is easy to see that the problem (6.3) is equivalent to the following convex programming problem [10]:

$$\min_p \{\tfrac{1}{2}\|p\|^2 + \max_{i \in I} \{\nabla f_i(x)p\}\}, \text{ under constraints}$$
$$\nabla g_j(x)p + \nabla g_j(x) \leq 0\,, \quad j \in J\,.$$

Now we write necessary and sufficient conditions relating the minimum point to Lagrange multipliers of the problem (6.3). Note that the Lagrange function has the form:

$$L(p, \xi, u, v) = \xi + \tfrac{1}{2}\|p\|^2 + \sum_{i\in I} u_i(x)[\nabla f_i(x)p - \xi]$$

$$+ \sum_{j\in J} v_j(x)[\nabla g_j(x)p + g_j(x)] = \xi\left[1 - \sum_{i\in I} u_i(x)\right] + \tfrac{1}{2}\|p\|^2$$

$$+ \sum_{i\in I} u_i(x)\nabla f_i(x)p + \sum_{j\in J} v_j[\nabla g_j(x)p + g_j(x)]\,.$$

There exist such $u_i(x) \geq 0$, $i \in I$, $v_j(x) \geq 0$, $j \in J$, that

$$u_i(x)[\nabla f_i(x)p - \xi] = 0\,, \quad i \in I\,, \tag{6.4}$$

$$v_j(x)[\nabla g_j(x)p + g_j(x)] = 0\,, \quad j \in J\,, \tag{6.5}$$

$$\sum_{i\in I} u_i(x) = 1\,, \tag{6.6}$$

$$p(x) + \sum u_i(x)\nabla f_i(x) + \sum v_j(x)\nabla g_j(x) = 0\,. \tag{6.7}$$

## ALGORITHM FORMULATION AND PRINCIPAL RESULTS

Now we formulate the computational procedure for solving the vector optimization problem. Let $x^0$ be the initial approximation and $\varepsilon$, $0 < \varepsilon < 1$ be chosen. Let the point $x^k$ be already obtained. Then (1) we solve the auxiliary problem (6.3) for $x = x^k$ and find $p^k = p(x^k)$; (2) we find the first value of $\alpha = 1, 2, \ldots$, for which the inequality

$$\max_{i\in I}\,[f_i(x + \alpha p) - f_i(x)] + NG(x + \alpha p) \leq NG(x) - \alpha\varepsilon\|p\|^2\,, \tag{6.8}$$

will be satisfied for $\alpha = (1/2)^s$. If such $s$ is found then assume $\alpha_k = 2^{-s}$, $x^{k+1} = x^k + \alpha_k p^k$.

From the assumption about the continuity of vector function $f(x)$ components on the non-empty compact $\Omega_N$ the existence of all kinds of effective points follows [21]. Now we formulate the first convergence theorem.

THEOREM 6.1. *Let the assumptions* (a)–(c) *of this section be satisfied. In addition, let the regularity condition be satisfied at any limit point $x^*$ of the sequence $\{x^k\}$ of points generated by the proposed algorithm: there is such point $p \in \mathbb{R}^n$ that for any $j \in J(x^*) = \{j \in J \mid g_j(x^*) = 0\}$ the inequality $\nabla g(x^*)p < 0$ is satisfied. Then the point $x^*$ satisfies the necessary conditions of weak efficiency and* $\|p^k\| \to 0$.

The proof of the given Theorem is based on reducing the extreme necessary conditions of problem (6.3) in solution to the equation

$$\sum_{i\in I} u_i^* \nabla f_i(x^*) + \sum_{j\in J(x^*)} v_j^* \nabla g_j(x^*) = 0\,. \tag{6.9}$$

The latter, together with the regularity condition and the relation (6.6), correspond to the necessary condition of weak-efficiency of the point $x^*$ formulated by the Da Cunha–Polak–Geoffrion theorem presented in [21]. It is shown there that under some assumptions concerning the convex functions $f(x)$ and $g(x)$, the expression (6.9) corresponds to sufficient optimal conditions. Therefore the following statements are true.

COROLLARY 6.1. *Let the vector-function $f(x)$ be pseudoconvex and functions $g_j(x)$ for any $j \in J(x^*)$ be quasi-convex. Then necessary and sufficient conditions of weak-efficiency are satisfied at the point $x^*$ in the conditions of Theorem* 6.1.

COROLLARY 6.2. *If, in addition, the strict quasi-convexity of the vector-function $f(x)$ and the convexity of $\mathbb{X}_g$ are assumed then the point $x^*$ may be stated to satisfy necessary and sufficient conditions of efficiency (Pareto-optimum).*

According to the second part of the Da Cunha–Polak–Geoffrion theorem, the expression (6.9) is a necessary condition of the proper efficiency of solution $x^*$ if the condition $u_i > 0$ for any $i \in I$ and $\Sigma\, u_i = 1$ is satisfied in it. It is not difficult to see that in the general case the suggested algorithm does not provide Lagrange positive multipliers. However, the latter may be guaranteed by having imposed some conditions of generalized regularity (CGR). Namely, let $\nabla f_i(x)$, $i \neq \nu$, $i \in I$ and $\nabla g_j(x)$, $j \in J(x^*)$ be linearly independent at the point $x^* \in \mathbb{X}_g$ for any $\nu \in \mathrm{I}$.

REMARK 6.1. It is clear that the given condition is a certain generalization of the usual condition of regularity of a single-objective mathematical programming problem in form of the condition of a linear independence of the gradients of active constraints.

The assumption made makes it possible to strengthen the result of Theorem 6.1.

THEOREM 6.2. *Let the major assumptions* (a)–(c) *of this section be fulfilled. And let the* CGR *be satisfied at the limit point $x^*$ of the sequence $\{x^k\}$ of points generated by the algorithm. Then the point $x^*$ satisfies the necessary conditions of proper efficiency.*

The work in [21] shows that if Lagrange multipliers $u_i^*$, $i \in I$ are positive and the equality in (6.9) is fulfilled then the assumption about the pseudo-convex function $f(x)$ is already insufficient for proper-efficiency of solution as it is the case in Corollary 6.2.

COROLLARY 6.3. *Let the conditions of the previous theorem be satisfied. Then if the vector-function $f(x)$ is convex and functions $g_j(x)$ are quasi-convex for any $j \in J(x^*)$ then the point $x^*$ satisfies the necessary and sufficient conditions of proper-efficiency.*

So far we have considered the problem of vector optimization in the general form, i.e., when the objective and constraint functions have a nonlinear character. It is in this case that we should expect the greatest effect from the application of the suggested algorithm. However, its use in the linear case is quite reasonable [10]. The linear problem of vector optimization possesses its specific peculiarities that allow the conditions of algorithm application be substantially simplified and therefore the finite results be strengthened.

First of all we note that not all assumptions (a)–(c) of this section remain necessary. Thus, the condition (b) is satisfied automatically. The convexity of objective and constraints functions allows on to speak of both necessary and sufficient conditions of efficiency of the limiting point $x^*$ of the algorithm.

If we take into account the fact that in the linear case the sets of Pareto-optimal and proper efficient solutions coincide [21] then the given characteristic of the point $x^*$ is complete. Here one can take from the above-mentioned two conditions, a weaker one.

THEOREM 6.3. *Let assumption* (a) *and* (c) *of this section be fulfilled. In addition, let the regularity condition be satisfied at any limit point $x^*$ of the sequence $\{x^k\}$ of points generated by the suggested algorithm: there is such point $p \in \mathbb{R}^n$ that for any $j \in J(x^\star) = \{j \in J \mid a_j x^\star = b\}$ the inequality $\alpha_j p < 0$ is fulfilled. Then the point $x^*$ satisfies necessary and sufficient conditions of efficiency. The problem is solved in a finite number of steps.*

The first part of this theorem is a simple corollary of Theorem 6.1 proved for the vector optimization problem in the general form. The proof of the fact that the linear vector optimization problem is solved in a finite number of steps is in principle based on the ideas of the proof of the alike fact for the general linearization method in case of the linear programming problem and shown in [24].

## References

1. Pshenichny, B. N. (1970), Newton's Method for Solving Systems of Equalities and Inequalities, *Mat. zametky* **8**(5), 635–640 (in Russian).
2. Pshenichny, B. N. (1970), Algorithm for the General Problem of Mathematical Programming, *Kibernetika* **5**, 120–125 (in Russian).
3. Pshenichny, B. N. and L. A. Sobolenko (1980), Acceleration of the Linearization Method

Convergence for the Problem of Conditional Minimization, *USSR Computational Mathematics and Mathematical Physics* **20**(3), 605–614 (English Translation from JVM & MF).
4. Pshenichny, B. N. and K. I. Nemchenko (1983), Method of Linearization without Analytic Evaluation of Derivatives, *Cybernetics* **19**(2), 206–212 (English Translation from Kibernetika).
5. Pshenichny, B. N. and E. E. Kyryk (1984), Application of the Linearization Method to the Large-Scale Problems, *Kibernetika* **3**, 75–79 (in Russian).
6. Panin, V. M. (1980), Linearization Method for Discrete Minimax Problem, *Kibernetika* **3**, 86–91 (in Russian).
7. Panin, V. M. (1981), A Linearization Technique for a Continuous Minimax Problem, *Kibernetika* **2**, 75–78 (in Russian).
8. Panin, V. M. (1984), Estimates of the Rates of Convergence of the Two Linearization Methods, *Dokl. Acad. of Sci. of the Ukr., serie A* **5**, 77–79 (in Russian).
9. Panin, V. M. (1984), Finite Penalty Methods with Linear Approximation of the Constraints. I, II, *Kibernetika* **2**, 44–55; **4**, 73–81 (in Russian).
10. Pshenichny, B. N. (1983), *The Linearization Method*, Moscow, Nauka (in Russian).
11. Pshenichny, B. N. (1980), *Convex Analysis and Extremal Problems*, Moscow, Nauka (in Russian).
12. Pshenichny, B. N. and Yu. M. Danilin (1978), *Numerical Methods in Extremal Problems*, Moscow, Mir Publishers (in English).
13. Pshenichny, B. N. (1971), *Necessary Conditions for an Extremum*, New York, Marcel Dekker.
14. Pshenichnyj, B. N. (1987), The Linearization Method, *Optimization* **18**(2), 179–196.
15. Pshenichny, B. N. and A. A. Sosnovsky (1989), The Complementarity Problem, *Cybernetics* **24**(4), 437–443 (English Translation from Kibernetika).
16. Sosnovsky, A. A. (1989), Development of the Numerical Methods for Solving the Multi-Objective Optimization and Complementarity Problems, Abstract of Thesis, V. M. Glushkov Inst. of Cybern., Kiev (in Russian).
17. Pshenichnyi, B. N. and A. A. Sosnovsky (1992), Nonlinear Complementarity Problem, *Optimization* **25**, 355–361.
18. Karamardian, S. (1969), The Nonlinear Complementarity Problems with Applications. I, II, *J. Optimiz. Theory and Appl.* **4**(2), 87–98; **3**, 167–181.
19. More J. J. (1974), Coercivity Conditions in Nonlinear Complementarity Problems, *SIAM Review* **16**, 1–15.
20. Smith, T. E. (1984), A Solution Condition for Complementarity Problems: with an Application to Spatial Price Equilibrium, *Appl. Math. and Comput.* **15**(1), 61–69.
21. Podinovskij, V. V. and V. D. Nogin (1982), *Pareto-Optimal Solutions of Multi-Objective Problems*, Moscow, Nauka (in Russian).
22. Pshenichny, B. N. and A. A. Sosnovsky (1987), The Linearization Method for Solving a Multicriterion Optimization Problem, *Kibernetika* **6**, 102–107 (in Russian).
23. Pshenichny, B. N. and A. A. Sosnovsky (1991), A New Algorithm for Vector Optimization Problem, *Lect. Notes in Econ. and Math. Systems* **351**, 83–87.
24. Sosnovsky, A. A. (1988), On One Algorithm of Solving Linear Vector Optimization Problem, *Sb. nauchn. tr.*, V. M. Glushkov Inst. of Cybern., Kiev, 42–47 (in Russian).