Comput. Maths. Math. Phys., Vol.17, pp. 73–87
Pergamon Press Ltd. 1978
Printed in Great Britain.

A RELAXATION METHOD FOR NON-LINEAR PROGRAMMING PROBLEM[1]

Yu.G. EVTUSHENKO and V.G. ZHADAN
Moscow

A relaxation method for finding local solution of the general nonlinear programming problem is described. The convergence is proved and the rate of convergence is investigated. The extension to finding of saddle-points is given. The results of numerical computations are presented. The paper develops the approach described in [1, 2]. Other versions of relaxation methods are given in [3, 4].

1 CONTINUOUS VERSION OF THE METHOD

We consider the general nonlinear programming problem

$$\min_{x\in X} F(x), \qquad X = \{x \in \mathbb{E}^n : g(x) = 0_{e1},\ \ h(x) \le 0_{c1},\ \ 0_{n1} \le x\}, \tag{1.1}$$

where $\mathbb{E}^i$ is i-dimensional Euclidean space and 0_{ij} is the $i \times j$ null-matrix, whose elements are all zero. The continuously differentiable functions $F(x)$, $g(x)$, $h(x)$ realize the mappings $F : \mathbb{E}^n \to \mathbb{E}^1$, $g : \mathbb{E}^n \to \mathbb{E}^e$, $h : \mathbb{E}^n \to \mathbb{E}^c$. We introduce the set

$$X_0 = \{x \in \mathbb{E}^n : g(x) = 0_{e1},\ \ h(x) < 0_{c1},\ \ 0_{n1} < x\}.$$

Here X_0 is a relative interior set. Each $x \in X_0$ is called the interior point. If $x \in X \backslash X_0$, then we say that x is a boundary point. It will be convenient to combine constraints of the equality and the inequality types in the single symbol $R = [g, h] = [R^1, \dots, R^m]$. The vector function $R(x)$ thus realizes the mapping $\mathbb{E}^n \to \mathbb{E}^m$, where $m = e + c$. Let $R_x(x)$ be the $n \times m$ matrix in which the (i, j)-th element is $\partial R^j(x)/\partial x^i$. We denote by $D(z)$, $D^{1/2}(z)$ diagonal matrices in which the i-th diagonal elements are, respectively, z^i, $(z^i)^{1/2}$; the dimensionalities of these matrices are determined by the dimensionality of the vector z. The superscript "$\top$" on a vector or matrix denotes transposition.

Denote $\xi(R) = \{\xi(R^1), \dots, \xi(R^m)\}$, $\xi(x) = \{\xi(x^1), \dots, \xi(x^n)\}$, where the function $\xi(z)$ of a scalar argument is defined and continuous for all $z \ge 0$, $z(0) = 0$ and $\xi(z) > 0$ for $z > 0$. Hence $\xi(-h^i(x)) > 0$, $\xi(x) > 0_{n1}$ for any $x \in X_0$. We can put e.g., $\xi(z) = z$. To obtain the numerical solution of problem (1.1), we propose to find the limit (as $t \to \infty$) points of the solution of the Cauchy problem for the system

$$\dot{x} = -D(\xi(x))[F_x(x) + R_x(x)v], \qquad x(0) = x_0 \in X_0. \tag{1.2}$$

Here, the point above a letter denotes differentiation with respect to the independent variable t, while the vector $v \in \mathbb{E}^m$ is found by solving the system of linear equations

$$G(x)v + R_x^\top(x)D(\xi(x))F_x(x) = 0_{m1}, \tag{1.3}$$

[1] *U.S.S.R. Comput. Maths. Math. Phys.*, **17**, 73–87, 1978.

where

$$G(x) = R_x^\top(x) D(\xi(x)) R_x(x) + D(\xi(-R(x))).$$

If there are no constraints apart from $x \geq 0_{n1}$, the method (1.2) has to be written as

$$\dot{x} = -D(\xi(x)) F_x(x). \tag{1.4}$$

We find v from (1.3) and substitute into the right-hand side of system (1.2); then (1.2) can be written as

$$\dot{x} = -M(x) F_x(x), \tag{1.5}$$

where

$$M(x) = D(\xi(x)) \left\{ I_n - R_x [R_x^\top D(\xi(x)) R_x + D(\xi(-R))]^{-1} R_x^\top D(\xi(x)) \right\},$$

and I_i is the identity $i \times i$ matrix. We introduce the index set:

$$\sigma(x) = \{i : R^i(x) = 0, \ \ i = 1, 2, \dots, m\}.$$

Definition 1. *The constraint $R(x) \leq 0_{m1}$ satisfies the constraint qualification (CQ) at the point x if the vector function $R(x)$ is continuously differentiable at x, and all vectors $R_x^i(x)$, where $i \in \sigma(x)$, are linearly independent.*

Lemma 1. *If at every point $x \in X \backslash X_0$ the constraint $R(x) \leq 0_{m1}$ satisfies the CQ, then the matrix $Q(x) = R_x^\top(x) R_x(x) + D(\xi(-R(x)))$ is non-singular, and positive semi-definite for all $x \in X$.*

Proof. We write the matrix $Q(x)$ as a product of a rectangular matrix $B(x)$ of dimensionality $m \times (n+m)$ and the transposed matrix $B^\top(x)$, where the matrix $B(x)$ consists of two block matrices:

$$B(x) = \| R_x^\top(x) \ \ D^{1/2}(\xi(-R(x))) \|, \qquad Q(x) = B(x) B^\top(x).$$

The lemma will be proved if we show that, for any $x \in X$, the rank of $B(x)$ is equal to m, since, if the rank of $B(x)$ is maximal (equal to m), then it will follow that $Q(x)$ is non-singular and positive semi-definite. If there are no equality type constraints, then at every interior point $x \in X_0$ the rank of the matrix $B(x)$ is equal to m since in this case, as the non-zero minor of $B(x)$ we can take the diagonal matrix $D^{1/2}(\xi(-R(x)))$. The lemma is also obvious if $R(x) = 0_{m1}$. Since the vectors $R_x^i(x)$ are linearly independent, the matrix $R_x^\top(x)$ will have a non-zero minor of order m.

Let k components, $e < k < m$, of the vector function $R(x)$ be zero at $x \in X$. It can be assumed without loss of generality that these components are $R^1(x), \dots, R^k(x)$. Then, the values of the functions $R^{k+1}(x), \dots, R^m(x)$ are strictly less than zero. We isolate in the $k \times n$ matrix

$$V_1(x) = \left\| \begin{array}{c} (R_x^1(x))^\top \\ \dots\dots\dots \\ (R_x^k(x))^\top \end{array} \right\|$$

a $k \times k$ matrix $C(x)$ such that its determinant, which is a minor of the matrix $V_1(x)$ of order k, is non-zero. Such a minor exists, by the constraint qualification. The determinant of the $m \times m$ matrix

$$V_2(x) = \left\| \begin{array}{c|cc} C(x) & \multicolumn{2}{c}{0_{k \times (m-k)}} \\ \hline & \xi^{1/2}(-R^{k+1}(x)) & 0 \\ 0_{(m-k)\times k} & \dots & \dots \\ & 0 & \xi^{1/2}(-R^m(x)) \end{array} \right\|$$

is non-zero. But the determinant of the matrix $V_2(x)$ is at the same time a minor of order m of the matrix $B(x)$. Hence the rank of the matrix $B(x)$ is maximal, i.e., is equal to m. The lemma is proved. $\square$

Let e^i denote the i-th unit vector.

Definition 2. *The constraints $R(x) \le 0_{m1}$ and $0_{n1} \le x$ satisfy the CQ at the point x if the vector function $R(x)$ is continuously differentiable at x and all the vectors $R^j_x(x)$, where $j \in \sigma(x)$, and the vectors e^i, such that $x^i = 0$, are linearly independent.*

By this definition, if the constraints satisfy the CQ at the point x, then the number of components of the vectors $R(x)$ and x, which vanish simultaneously, is not greater than n.

Lemma 2. *If at every point $x \in X \backslash X_0$ the constraints $R(x) \le 0_{m1}$, $0_{n1} \le x$ satisfy the CQ, then $G(x)$ is a non-singular positive semi-definite matrix for all $x \in X$.*

Proof. Let us show that, for any $x \in X$, the rank of the $m \times (n+m)$ matrix

$$\Psi(x) = \| R_x^\top(x) D^{1/2}(\xi(x)) \quad D^{1/2}(\xi(-R(x))) \|$$

is equal to m. It will then follow, from the representation $G(x) = \Psi(x)\Psi^\top(x)$, that the matrix $G(x)$ is non-singular and positive semi-definite. The lemma is obvious if $x = 0_{n1}$ because of in this case we necessarily have $R(x) < 0_{m1}$ and $G(x) = D(\xi(-R))$. Now let $x \ne 0_{n1}$. The pair set $R(x)$, $-x$ can be regarded as a single vector function $(R(x), -x)$, and Lemma 1 can be applied to it. In this case we have an $(m+n) \times (m+2n)$ matrix B_1 which is can be written in the block form:

$$B_1 = \left\| \begin{array}{ccc} R_x^\top & D^{1/2}(\xi(-R)) & 0_{mn} \\ -I_n & 0_{nm} & D^{1/2}(\xi(x)) \end{array} \right\|.$$

Assume that, among the components of the vector x, there are s ones which are equal to zero, $s > 0$. Let the vector x can be splitted in two vectors $x = \{y, z\}$, where $y \ne 0_{k1}$, $y \in \mathbb{E}^k$, $z = 0_{s1}$, $k = n - s$. In the same way, $R_x^\top = \{R_y^\top, R_z^\top\}$, where $R_y^\top$, $R_z^\top$ are $m \times k$ and $m \times s$ matrices, respectively. We write the matrix B_1 as

$$B_1 = \left\| \begin{array}{ccccc} R_y^\top & R_z^\top & D^{1/2}(\xi(-R)) & 0_{mk} & 0_{ms} \\ -I_k & 0_{ks} & 0_{km} & D^{1/2}(\xi(y)) & 0_{ks} \\ 0_{sk} & -I_s & 0_{sm} & 0_{sk} & 0_{ss} \end{array} \right\|.$$

By Lemma 1, the rank of the matrix B_1 is maximal and equal to $m+n$. Hence we can extract from B_1 a square matrix B_2 of order $n+m$, the determinant of which is a non-zero minor of the matrix B_1. Since all the matrices in the lower matrix-row of B_1 are null-matrices, apart from $-I_s$, the matrix B_2 must necessarily contain a column

$$T = \left\| \begin{array}{c} R_z^\top \\ 0_{ks} \\ -I_s \end{array} \right\| \qquad \text{and} \qquad B_2 = \left\| \begin{array}{c|c} B_3 & \multirow{2}{*}{T} \\ \cline{1-1} 0_{sk} & \end{array} \right\|.$$

The $(m+k) \times (m+k)$ matrix B_3 is non-singular. This follows from the Frobenius formula, since $|B_2| = |-I_s|\,|B_3|$. Omitting the column T in the matrix B_1, along with the bottom matrix-row and the right-hand null-matrix-column, we get the matrix

$$B_4 = \left\| \begin{array}{ccc} R_y^\top & D^{1/2}(\xi(-R)) & 0_{mk} \\ -I_k & 0_{km} & D^{1/2}(\xi(y)) \end{array} \right\|$$

of dimensionality $(m+k)\times(m+2k)$, from the elements of the matrix B_4. The matrix B_3 can be formed. The rank of B_4 is, therefore, maximum. If we omit in B_4 any k rows, then the rank of the resulting matrix will also be a maximum. Removing from B_4 the bottom matrix-row and the right-hand null-matrix-column, we obtain the matrix of rank m

$$B_5 = \| R_y^\top \;\; D^{1/2}(\xi(-R)) \|.$$

We multiply B_5 on the right by the non-singular diagonal matrix

$$B_6 = \left\| \begin{matrix} D^{1/2}(\xi(y)) & 0_{km} \\ 0_{mk} & I_m \end{matrix} \right\|$$

and complement the product by the matrix $R_z^\top D^{1/2}(\xi(z))$; both these operations do not change its rank. As a result we obtain the matrix

$$\| R_y^\top D^{1/2}(\xi(y)) \;\; R_z^\top D^{1/2}(\xi(z)) \;\; D^{1/2}(\xi(-R)) \| = \Psi(x).$$

If $x>0$, i.e., $k=n$, $s=0$, then the second and last matrix-columns and the last matrix-row disappear from the matrix B_1. Then, in order to obtain $\Psi(x)$, we have to omit in B_1 the bottom matrix-row, the right-hand null-matrix-column, and the remaining matrix B_5 has to be multiplied by B_6. The rank of the matrix $\Psi(x)$ is equal to m. □

We find from Lemma 2 that, if the CQ is satisfied at every boundary point of X, then the matrix $G(x)$ has an inverse at any point of the set X, and the right-hand sides of the system (1.2) are defined everywhere on X.

We shall say henceforth that the CQ holds everywhere on X, if it holds at each boundary point of the set X.

Lemma 3. *If the conditions of Lemma 2 hold, then the symmetric matrix $M(x)$ is positive semi-definite for all $x\in X$.*

Proof. We introduce the matrices

$$\begin{aligned} K_1 &= -[R_x^\top D(\xi(x))R_x + D(\xi(-R))]^{-1} R_x^\top D(\xi(x)), \\ K &= \left\| \begin{matrix} D^{1/2}(\xi(x))(I_n + R_x K_1) \\ D^{1/2}(\xi(-R))K_1 \end{matrix} \right\| \end{aligned}$$

of dimensionalities $m\times n$ and $(m+n)\times n$, respectively. The proof of the lemma follows from the representation $M = K^\top K$, which can be shown to hold by direct calculations. □

If x_* is a local solution of problem (1.1) and the constraints $R(x)\le 0_{m1}$, $0_{n1}\le x$ satisfy the CQ, then the vector $v_*\in\mathbb{E}^m$ must exist, such that

$$\begin{array}{ll} D(x_*)(F_x(x_*) + R_x(x_*)v_*) = 0_{n1}, & v_*^i \ge 0, \quad 1+e\le i\le m, \\ D(R(x_*))v_* = 0_{m1}, & F_x(x_*) + R_x(x_*)v_* \ge 0_{n1}. \end{array} \tag{1.6}$$

In the case of convex programming problem (when the constraints $g(x)$ are linear, and the functions $h(x)$ and $F(x)$ are convex), conditions (1.6) are sufficient for a minimum in problem (1.1).

We introduce the set $Z=\{x : F(x)\le F(x_0),\;\; x\in X\}$. Throughout what follows, we shall assume that system (1.2) defines a unique solution $x(x_0,t)$ for every $x_0\in X$.

Lemma 4. *Assume that the functions $F(x)$ and $R(x)$ are continuously differentiable everywhere on the compact set X, the constraints $R(x)\le 0_{m1}$ and $0_{n1}\le x$ satisfy the CQ. Then the* *solutions $x(x_0,t)$ of system* (1.2) *can be prolonged as $t \to \infty$, and the sets X, Z are positive-invariant (i.e., $x_0 \in X$ implies that $x(x_0,t) \in X$ and $x_0 \in Z$ implies $x(x_0,t) \in Z$ for all $t \geq 0$).*

Proof. Computing the first derivatives of the functions $R(x)$ and $F(x)$ along the trajectories of system (1.2), we obtain

$$\dot{R} = D(v)\xi(-R), \qquad \dot{F} = -\|D^{1/2}(\xi(x))H_x(x,v)\|_n^2 - \|D^{1/2}(\xi(-R))v\|_m^2. \tag{1.7}$$

Here $H_x(x,v) = F_x(x) + R_x(x)v$, $\|z\|_i^2 = z^\top z$, and i is the dimensionality of the vector z.

System (1.2) has a solution at least for those t for which $x(x_0,t) \in X$. Let us show that the solution $x(x_0,t)$ does not leave the set X for any $t \geq 0$. Suppose that $R^j(x(x_0,t)) > 0$ for some $t > 0$. Then an instant t_1 exists, such that $R^j(x(x_0,t_1)) = 0$ and $\dot{R}^j(x(x_0,t_1)) > 0$, which contradicts (1.7), since $\|v(t_1)\|_m < \infty$. Similarly, it follows from the system (1.2) that $x(x_0,t) \geq 0_{n1}$. The functions $\xi(x)$, $\xi(-R)$ introduced above thus play the role of "barriers", preventing the trajectory $x(x_0,t)$ from passing through the surface $x = 0_{n1}$, $R(x) = 0_{m1}$. The trajectory $x(x_0,t)$ can approach the boundary points only as $t \to \infty$. If the initial point x_0 is on the boundary, then the entire trajectory of system (1.2) belongs to the boundary. The functions $g^i(x)$ which defining equality-type constraints are integrals of system (1.2). Hence, since the set X is bounded, the solutions of system (1.2) are extendable as $t \to \infty$, and the set X is positive-invariant with respect to (1.2). From this and (1.7) it follows that the set Z is also positive-invariant. □

Denote by $\bar{x}$ the points of X at which the right-hand sides of system (1.2) vanish; they will be referred to as stationary points. The corresponding values $v(\bar{x})$ will be denoted by $\bar{v}$. At the point $\bar{x}$ the following conditions are satisfied:

$$\begin{aligned} &D(\xi(\bar{x}))H_x(\bar{x},\bar{v}) = 0_{n1}, \qquad D(\xi(-R(\bar{x})))\bar{v} = 0_{m1}, \\ &R_x^\top(\bar{x})D(\xi(\bar{x}))R_x(\bar{x})\bar{v} + R_x^\top(\bar{x})D(\xi(\bar{x}))F_x(\bar{x}) = 0_{m1}. \end{aligned} \tag{1.8}$$

Each point x_* being a local solution of problem (1.1) is a stationary point. For, otherwise, taking x_* as an initial point for the system (1.2), we would have that the solution $x(x_*,t) \in X$ and $F(x(x_*,t)) < F(x_*)$ for $t > 0$, since $dF(x_*)/dt < 0$. But this contradicts the conditions for a local minimum of the function $F(x)$.

Theorem 1. *Let the conditions of Lemma* 4 *be satisfied, and let all the stationary points of the set X be isolated. Then for any non-stationary initial points $x_0 \in X_0$ the solution $x(x_0,t)$ of system* (1.2) *converges to an admissible stationary point, at which the necessary (and also the sufficient, in the case of a convex programming problem) conditions for a minimum* (1.6) *are satisfied.*

Proof. Let x_0 be an arbitrary point in X_0, and let $x(x_0,t)$ be a solution of the Cauchy problem (1.2). Since the set X is compact, the set of ω-limit points Ω for the solution $x(x_0,t)$ is non-empty. Let us show that ω belongs to the set of admissible stationary points. Since $F(x)$ is bounded below on X, and $F(x(x_0,t))$ is a monotonically decreasing function, then, by [5, 6], all the points of ω lie on the same level surface of the function F. Let $\tilde{x} \in \omega$. We draw a trajectory $x(\tilde{x},t)$ through $\tilde{x}$. Any point of it will also belong to ω, so that $\dot{F}(x(\tilde{x},t)) \equiv 0$ and hence $\dot{F}(\tilde{x}) = 0$. But it is clear from (1.7) that this is only possible when $\tilde{x}$ is a stationary point for system (1.2). Hence, since all the stationary points of X are isolated, we find that ω consists of the single admissible stationary point $\tilde{x}$, to which $x(x_0,t)$ converges as $t \to \infty$.

For every $x = x(x_0,t)$ we can define from (1.3) $\varphi^j(t) = v^j(x(x_0,t))$, $j = 1,2,\ldots,m$, and evaluate $\psi^i(t) = H_{x^i}(x(x_0,t), v(x(x_0,t)))$, $i = 1,2,\ldots,n$. These functions are continuous, so that the existence of the limit

$$x_1 = \lim_{t\to\infty} x(x_0, t)$$

implies the existence of the limits

$$\bar{\varphi}^j = \lim_{t\to\infty} \varphi^j(t), \qquad \bar{\psi}^i = \lim_{t\to\infty} \psi^i(t).$$

Let us show that, for $j = e+1, \dots, m$, $i = 1, \dots, n$ they are non-negative. At the limit point x_1 conditions (1.8) hold, and $\bar{\psi}^i = 0$, if $x_1^i > 0$. Now consider the case when $x_1^i = 0$. We find from the system (1.2) that

$$x^i(x_0, t) = x_0^i \exp(-\Phi^i(t)), \qquad \Phi_i(t) = \int_0^t \frac{\xi(x^i(x_0,\tau))}{x^i(x_0,\tau)} \psi^i(\tau) d\tau.$$

Assume that $\bar{\psi}^i < 0$; then a $\bar{t} > 0$ can be found such that $\psi^i(t) < 0$ for all $t > \bar{t}$, and hence $\Phi^i(t) < \Phi^i(\bar{t})$ for the same t. But this contradicts the condition $x_1^i = 0$. Thus, $\bar{\psi}^i \geq 0$, $i = 1, 2, \dots, n$. We can show in the same way, with the aid of (1.7) and (1.8), that $\bar{\varphi}^j \geq 0$, $j = e+1, \dots, m$. In view of this and (1.8), it follows that the necessary conditions (1.6) hold at x_1. The theorem is proved. $\square$

It can easily be seen that the conditions of this theorem can be relaxed by requiring that they are all satisfied on the set Z, rather than on X.

In the case of convex programming problems the requirement that the solution of problem (1.1) be unique does not seem to be essential. We proved this fact in one particular case. The following result takes place.

Theorem 2. *Let $F(x)$ be a convex, continuously differentiable function, let $g(x)$, $h(x)$ be linear, and let the constraints $R(x) \leq 0_{m1}$, $0_{n1} \leq x$ satisfy the CQ everywhere on the compact set X. Then the solutions of system* (1.2) *converge to the solutions set of problem* (1.1) *for any $x_0 \in X_0$.*

The method (1.2) allows considerable arbitrariness in the choice of the function $\xi(z)$. Each of the constraints can be taken into account for with the aid of a special function. In cases when there is no constraint $x \geq 0_{n1}$ in problem (1.1), the matrix $D(\xi(x))$ is replaced in (1.2) – (1.8) by the identity matrix. If instead of the condition $x \geq 0_{n1}$ the constraint $x \geq a$ is imposed, then we write $D(\xi(x-a))$ instead of $D(\xi(x))$ in the expression mentioned. If the constraints have the form $a^i \leq x^i \leq b^i$ or $c^j \leq h^j(x) \leq d^j$, then we introduce two barrier functions $\xi_1(x)$, $\xi_2(h(x))$, in which the i-th and j-th components are e.g., $\xi_1^i(x) = (x^i - a^i)(b^i - x^i)$, $\xi_2^j(h(x)) = (h^j(x) - c^j)(d^j - h^j(x))$. Systems (1.2) and (1.3) have the form

$$\dot{x} = -D(\xi_1(x))(F_x + R_x v), \qquad \left[R_x^\top D(\xi_1(x)) R_x + D(\xi_2(R(x)))\right] v + R_x^\top D(\xi_1(x)) F_x = 0_{m1}.$$

Hence constraints of this type do not increase the order of the linear system (1.3).

If there are no constraints of inequality type in problem (1.1) ($m = e$, $R(x) = g(x)$, $R_x(x) = g_x(x)$, $X_0 = \emptyset$), then (1.2) is the same as the method described in [7, Chapter 2, Section 3]. In this case the matrix M has the form

$$M(x) = I_n - g_x(x)(g_x^\top(x) g(x))^{-1} g_x^\top(x).$$

It is obvious that $M(x) g_x(x) = 0_{n1}$, $MM = M$, and hence $M(x)$ projects any vector $z \in \mathbb{E}^n$ onto the tangent manifold to the set X at the point x, i.e., onto the orthogonal complement of the subspace generated by the vectors $g_x^1(x), \dots, g_x^e(x)$. The stationarity condition $M(x)F_x(x) =$

$= 0_{n1}$ implies that the projection of the vector $F_x(x)$ onto the tangent manifold is equal to zero (the necessary condition for an extremum). The vector $\dot{x}$ lies in the tangent manifold, with the result that the $g^i(x)$ are integrals of system (1.2).

In method (1.2), the matrix $M(x)$ projects the vector $F_x(x)$ at the boundary points X onto the tangent manifold to the set

$$Y(x) = \{z \in \mathbb{E}^n : R^i(z) = 0,\ \ i \in \sigma(x),\ \ z^j = 0,\ \text{ if } x^j = 0\},$$

so that there is no possibility of the trajectory $x(x_0, t)$ crossing the boundary of the set X. At the points x where certain $R^i(x) < 0$, the projection $M(x)F_x(x)$ onto $R^i_x(x)$ is, by (1.7), equal to $\xi(-R^i(x))v^i$. For numerical calculations it is usual to take $\xi(-R^i(x)) = -R^i(x)$. Hence the rate of motion of the trajectories in the direction of the boundary $R^i(x) = 0$ tends to zero as the boundary is approached. Different "barrier" functions $\xi(-R)$ lead to different types of this rate. The introduction of the functions $\xi(x)$, $\xi(-R)$ considerably simplifies numerical computations as compared with the gradient projection method [3], by automatically changing the direction of the vector $\dot{x}(x_0, t)$ near the boundary.

Far away from the hypersurface $h^i(x) = 0$, when $h^i(x) \ll 0$, one need not fear that the trajectory $x(x_0, t)$ intersects on a small interval $(t, t+\delta)$, and the function $h^i(x)$ and its derivative can be omitted in the expression for $M(x)$, while they are introduced and the constraint is made active only when $-\varepsilon < h^i(x(x_0, t)) < 0$, where $\varepsilon > 0$ is chosen, depending on the integration step of the system (1.2). This device enables the order of system (1.3) to be reduced. Moreover, the introduction of the barrier functions $\xi(-R)$ can lead to the fact that $h^i(x(x_0, t)) \equiv 0$ for $h^i(x_0) = 0$, or to the fact that $|h^i(x(x_0, t))|$ becomes extremely small, in spite of the fact that $\dot{h}^i < 0$. To avoid this drawback, we omit the functions h^i, h^i_x in the expressions for M, and calculate on the basis of the new system the derivative $\dot{h}^i = (h^i_x)^\top \dot{x}$. If it turns out that the derivative is negative, we continue the movement along the trajectory of this system. In other words, when eliminating the "barrier", we check whether this can be done without violating the constraints. This procedure proves especially simple in the case when $m = 0$ and there is only the constraint $x \geq 0_{n1}$ in system (1.1); system (1.4) is replaced by the following:

$$\dot{x}^i = \begin{cases} -F_{x^i}, & \text{if } x^i > \varepsilon > 0 \text{ or } 0 \leq x^i \leq \varepsilon \text{ and } F_{x^i} \leq 0, \\ -x^i F_{x^i}/\varepsilon, & \text{if } 0 \leq x^i \leq \varepsilon \text{ and } F_{x^i} > 0. \end{cases}$$

Here, $i = 1, 2, \ldots, n$. If F_x is continuous, then the right-hand sides of the system are also continuous, and no sliding modes are present in the system.

In some particular cases, (1.2) transforms into the methods described in [1, 2].

Consider the linear programming problem defined in the standard form: to find

$$\min_{x \in X} c^\top x, \qquad X = \{x : Ax = b,\ \ x \geq 0_{n1}\}, \tag{1.9}$$

where $x, c \in \mathbb{E}^n$, $b \in \mathbb{E}^m$, and A is an $m \times n$ matrix. The dual problem to (1.9) consists in finding

$$\max_{p \in P} b^\top p, \qquad P = \{p \in \mathbb{E}^m : A^\top p \leq c\}.$$

Setting $\xi(z) = z$ we find that the method (1.2) for solving the primal problem leads to the system

$$\dot{x} = D(x)(A^\top p - c), \ \text{ where } \ AD(x)A^\top p = AD(x)c, \quad x(0) = x_0. \tag{1.10}$$

In this case, $c^\top \dot{x} = -\|D^{1/2}(x)(c - A^\top p\|^2_n \leq 0$, if $x_0 > 0_{n1}$, $Ax_0 = b$. Similarly, for the dual problem

$$\begin{aligned} &\dot{p} = b - Ax, \ \text{ where } \ [A^\top A - D(A^\top p - c)]x = A^\top b, \quad p(0) = p_0, \\ &b^\top \dot{p} = \|b - Ax\|^2_m + x^\top D(c - A^\top p) \geq 0, \end{aligned} \tag{1.11}$$

if $A^\top p_0 < c$. The relaxation method for solving the primal problem (1.9) is effective, when $n \gg m$. Similarly, method (1.11) is more suitable if $n \ll m$. Methods (1.10) and (1.11) need to be only slightly modified in the case of quadratic programming problems.

For the algorithm to operate, it is important that the initial point $x_0 \in X_0$. If this point is not known, it can be found by using the algorithm in the following way. Let

$$\begin{array}{rclrcl} W^1(x) & = & \{i : h^i(x) < 0, \quad 1 \le i \le c\}, & W^2(x) & = & \{i : h^i(x) \ge 0, \quad 1 \le i \le c\}, \\ W^3(x) & = & \{i : |g^i(x)| \le \varepsilon, \quad 1 \le i \le e\}, & W^4(x) & = & \{i : |g^i(x)| \ge \varepsilon, \quad 1 \le i \le e\}. \end{array}$$

Here, ε is the admissible accuracy of satisfying the equality type constraints. We solve the minimization problem for

$$f(x) = \sum_{i \in W^2(x)} h^i(x) + \sum_{j \in W^4(x)} [g^i(x)]^2$$

under the constraints $x \in W^5$, where

$$W^5 = \{x : x \ge 0_{n1}, \ \ h^i(x) \le 0, \ \ i \in W^1(x), \ \ g^j(x) = \text{const}, \ \ j \in W^3(x)\}.$$

If, during the computations, it is found that, for certain values of x, the superscripts i do not all belong to the set $W^2(x)$, then we regard the functions $h^i(x)$ as constraints of the inequality type, and exclude them from the expressions for $f(x)$. We proceed in a similar way with constraints of the equation type, including those for which $|g^j(x)| \ge \varepsilon$, in the expression for $f(x)$, and excluding those for which this condition is infringed. The process is continued until we obtain an admissible point.

Newton's method can be used to solve problem (1.1). Then, for the case $\xi(z) = z$ we obtain the following system of $n + m$ equations:

$$\begin{array}{rcl} \dot{x} & = & -U^{-1}D(x)(H_x + R_x\dot{v}), \qquad U = D(H_x) + D(x)H_{xx}, \\ \dot{v} & = & -[R_x^\top U^{-1}D(x)R_x]^{-1}[R_x^\top U^{-1}D(x)H_x - D(R)v]. \end{array}$$

This system is written in such a way that

$$D(x(t))H_x(x(t), v(t)) = D(x_0)H_x(x_0, v(x_0))e^{-t}, \qquad \dot{R} = -D(v)R.$$

The system is more complicated and we shall not dwell on it here.

2 ESTIMATION OF THE RATE OF CONVERGENCE

Method (1.2) can be used for finding a saddle-point. In this case, however, we cannot use the proof of convergence described in Section 1, since it was based on the monotonic property of the function $F(x(x_0, t))$. This property is absent in saddle-point problems. In short, different arguments are needed to justify the method. We give below an alternative proof of the convergence of method (1.2), which extends to the case of finding saddle-points and enables the rate of convergence to be estimated. We shall confine ourselves to a summary and merely sketch out the underlying ideas of the proofs of the theorems. For simplicity, we assume that the constraints $x \gg 0_{n1}$ are absent in problem (1.1) and that $\xi(z) = z$. When referring to problem (1.1) and to the expressions of Section 1, this is the case we shall have in mind. As mentioned above, we have to put $D(\xi(x)) = I_n$, $D(\xi(-h)) = -D(h)$ everywhere in the expressions of Section 1.

For points $x \in X$, $v \in \mathbb{E}^m$ the following condition **S** can be stated: given any non-zero vector $z \in \mathbb{E}^{n+c}$ such that

$$N(x)z = 0_{m1}, \tag{2.1}$$

we have the inequality

$$z^\top H_{zz}(x,v)z > 0.$$

Here,

$$\begin{aligned} N(x) &= \left\| \begin{array}{cc} g_x^\top(x) & 0_{mc} \\ h_x^\top(x) & D([-2h(x)]^{1/2}) \end{array} \right\|, \\ H_{zz}(x,v) &= \left\| \begin{array}{cc} H_{xx}(x,v) & 0_{nc} \\ 0_{cn} & D(\tilde{v}) \end{array} \right\| \end{aligned}$$

are $m \times (n+c)$ and $(n+c)\times(n+c)$ matrices, respectively, while

$$H_{xx}(x,v) = F_{xx}(x) + \sum_{i=1}^{m} v^i R^i_{xx}(x)$$

is an $n \times n$ matrix, and $\tilde{v} = [v^{e+1}, v^{e+2}, \ldots, v^m]$.

Let us state the sufficient conditions for a minimum in problem (1.1).

Theorem 3. *Let in order for the admissible point x_* to be a local, isolated minimum of problem (1.1), where $F(x)$ and $R(x)$ are be twice continuously differentiable functions, it is sufficient that a vector $v_* \in \mathbb{E}^m$ exist, such that the stationary conditions*

$$H_x(x_*, v_*) = F_x(x_*) + R_x(x_*)v_* = 0_{n1}, \qquad D(x_*)R(x_*) = 0_{m1}$$

hold at the point x_*, and the condition **S** *holds at the point (x_*, v_*).*

Notice that it follows from the assumptions of the theorem that $\tilde{v}_* \geq 0_{m-e,1}$ and that the strict complementary condition holds at the point x_*.

We introduce the vector $y \in \mathbb{E}^c$ by the relation

$$2h^i(x) + (y^i)^2 = 0, \qquad i = 1, 2, \ldots, c. \tag{2.2}$$

We replace problem (1.1) by the following: to find $\min F(z)$ with respect to $z = [x, y]$, $x \in \mathbb{E}^n$, $y \in \mathbb{E}^c$, in the presence of the equality-type constraints $g(x) = 0_{e1}$ and (2.2). This approach has been widely used in other papers, see e.g., [8, 9]. The remaining arguments are similar to those in the proof of Theorem 4 in [10].

Definition 3. *A solution $x(t)$ of system (1.2) is said to be conditionally asymptotically stable in the neighborhood of the point x_* if for any given $\varepsilon > 0$ there exists $\delta > 0$ such that every solution $x(x_0, t)$ satisfying the condition $\|x_0 - x_*\|_n < \delta$ is such that*

$$\|x(x_0,t) - x_*\|_n < \varepsilon, \qquad \lim_{t\to\infty} x(x_0,t) = x_* \qquad \text{for all } t \geq 0.$$

Theorem 4. *Let the sufficient conditions for a minimum, given in Theorem 3, hold at the point x_*, and let the constraints satisfy the CQ on X. Then the solution $x(x_0, t)$ of system (1.2) is conditionally asymptotically stable in the neighborhood of the point x_*. If, in addition,*

$$z^\top H_{zz}(x_*, v_*)z \geq \gamma \|z\|^2_{n+c} > 0 \quad \text{for all } z, N(x_*)z = 0_{m1}, \tag{2.3}$$

such that $N(x_*)z = 0_{m1}$, then we have the following exponential estimation:

$$\begin{aligned} 0 &\leq -\dot{F}(x(x_0,t)) = \|H_x(x(x_0,t), v(x(t)))\|^2_n + \\ &+ \|D^{1/2}(-R(x(x_0,t)))v(x(t))\|^2_m \leq [\|H_x(x_0, v(x_0))\|^2_n + \\ &+ \|D^{1/2}(-R(x_0))v(x_0)\|^2_m]e^{-\gamma t}. \end{aligned} \tag{2.4}$$

Proof. We form the non-negative function

$$\varphi(x, v) = \frac{1}{2} H_z^\top(x, v) H_z(x, v).$$

Here

$$H_z^\top(x, v) = [H_x^\top(x, v), \tilde{v}^\top(x(t, x_0)) D^{1/2}(-h(x))]$$

is a $1 \times (n + c)$ matrix row.

The function $v(x(t))$, defined by (1.3), is differentiable. It can, therefore, be differentiated along the trajectories of system (1.2). Recalling that $\dot{R} = -R_x^\top H_x = -D(v)R$, we obtain $\dot{\varphi} \leq -H_z^\top H_{zz} H_z$.

We use the **S** property, and put $z = [H_x(x, v), D^{1/2}(-h(x))\tilde{v}] \in \mathbb{E}^{n+c}$, the condition (2.1) will then hold for any $t \geq 0$. The quadratic form in the expression for $\dot{\varphi}$ is strictly negative, if the elements of the matrices H_{xx} and $D(\tilde{v})$ are evaluated at the point (x_*, v_*). But, in view of the continuity of H_{xx} and $D(\tilde{v})$, this property holds if (x, v) lies in a neighborhood of the point (x_*, v_*). Hence $\dot{\varphi} < 0$ and $\dot{\varphi} = 0$ only at a stationary point (x_*, v_*), x_* being a locally isolated solution of problem (1.1). On the basis of (2.3) we get $\varphi(t) \leq \varphi(0)e^{-\gamma t}$. Then, (2.4) follows from (1.7). The theorem is proved. □

3 DISCRETE VERSION OF THE METHOD

Integrating (1.5) by Euler's scheme, we get

$$x_{s+1} = x_s - \alpha_s M(x_s) F_x(x_s), \tag{3.1}$$

where $0 < \alpha_s$ are the integration steps.

If the conditions of Lemma 2 hold then the maximum value of the norm of the matrix M can be defined on X by the relation

$$\lambda = \max_{x \in X} \max_{y \in \mathbb{E}^n} \frac{y^\top M(x) y}{||y||^2} < \infty.$$

Put

$$\mu = \max_i \max_{x \in X} \frac{\partial H(x, v)}{\partial x^i}, \qquad \nu = \max_j \max_{x \in X} v^j(x);$$

here, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$, $v^j(x)$ is given by (1.3). In what follows we shall write $v_s^j = v^j(x_s)$.

The function $F_x(x)$ satisfies a Lipschitz condition on X with the constant L, if,

$$\|F_x(x_1) - F_x(x_2)\|_n \leq L\|x_1 - x_2\|_n. \tag{3.2}$$

for any $x_1 \in X$, $x_2 \in X$.

We shall assume for simplicity that $\xi(z) = z$, and α_s are constants.

Theorem 5. *If the conditions of Theorem* 1 *and condition* (3.2) *are satisfied and the function* $R(x)$ *is linear, then for* $0 < \alpha_s < \min(1/\mu, 1/\nu, 2/\lambda L)$ *and for any non-stationary initial points* $x_0 \in X_0$ *the sequence* $\{x_s\}$ *converges to an admissible stationary point, at which the necessary conditions (and also sufficient conditions in the case of convex programming problem) for a minimum* (1.6) *are satisfied; here,* $x_s \in X_0$, $F(x_{s+1}) \leq F(x_s)$ *for* $s = 0, 1, \ldots$.

If, moreover, $\gamma\|z\|^2_{n+c} \le z^\top H_{zz}(x, v(x))z \le \Gamma\|z\|^2_{n+c}$ *for all* $x \in X$, $z \in \mathbb{E}^{n+c}$, *then given any initial point* $x_0 \in X_0$, *we have*

$$\begin{aligned} &\|H_x(x_s, v_s)\|^2_n + \|D^{1/2}(-R(x_s))v_s\|^2_m \le \\ \le\ & [\|H_x(x_0, v_0)\|^2_n + \|D^{1/2}(-R(x_0))v_0\|^2_m][1 - \alpha_s\gamma + \alpha_s^2\Gamma^2]^s, \end{aligned} \tag{3.3}$$

and the sequence $x_s \to x_*$, *where* x_* *is the unique solution of problem* (1.1).

Proof. Since the functions $R(x)$ are linear, we have

$$R^i(x_{s+1}) = R^i(x_s)[1 - \alpha_s v^i_s], \qquad i = 1, 2, \dots, m. \tag{3.4}$$

Noting that $R^i(x_0) \le 0$, $|v^i_s| < \nu$, we find that $R^i(x_{s+1}) \le R^i(x_0)(1 - \alpha_s\nu)^s$, if $\alpha_s < 1/\nu$. Here, for equation-type constraints $R^i(x)$, $i = 1, 2, \dots, e$, it follows from $R^i(x_s) = 0$ that $R^i(x_{s+1}) = 0$ for any α_s. It can be shown in the same way that, for $\alpha_s < 1/\mu$, the components of the vector x_s have fixed sign for all $s \ge 0$. Hence the set X is positive-invariant with respect to (3.1), as in the continuous case.

From the Newton–Leibniz formula we have

$$F(x_{s+1}) \le F(x_s) - \alpha_s F_x(x_s)M(x_s)F_x(x_s) + \frac{L\alpha_s^2}{2}\|M(x_s)F_x(x_s)\|^2_n. \tag{3.5}$$

The symmetric matrix $M(x_s)$ is non-negative definite, so that M will have a arithmetic square root. Denote it by $M^{1/2}$; then, $M = M^{1/2}M^{1/2}$. Introducing the vector $y = M^{1/2}F_x$, we can transform inequality (3.5) to the form

$$F(x_{s+1}) - F(x_s) \le \alpha_s\|y\|^2_n\left[-1 + \frac{\alpha_s L}{2}\frac{y^\top M(x_s)y}{\|y\|^2_n}\right] \le \alpha_s\|y\|^2_n\left[-1 + \frac{\alpha_s L\lambda}{2}\right].$$

Thus for $\alpha_s < 2/\lambda L$ the sequence $F(x_s)$ monotonically decreases. Since $F(x)$ is bounded from below on X, it follows that the limit of $F(x_s)$ exists. Hence

$$\begin{aligned} &\lim_{s\to\infty}[F(x_{s+1}) - F(x_s)] = 0, \qquad 0 \le F_x^\top(x_s)M(x_s)F_x(x_s) \le \frac{2[F(x_s) - F(x_{s+1})]}{\alpha_s[2 - \alpha_s\lambda L]}, \\ &\lim_{s\to\infty} F_x^\top(x_s)M(x_s)F_x(x_s) = 0. \end{aligned}$$

Using the representation (1.7), we find that

$$\lim_{s\to\infty}\|D(x_s)(F_x(x_s) + R_x(x_s)v_s)\|_n = \lim_{s\to\infty}\|D^{-1/2}(R(x_s))v_s\|_m = 0,$$

i.e., at every limit point of the sequence x_s, the stationarity conditions (1.8) must be satisfied. Since the stationary points are isolated, the following limits exist:

$$\bar{x} = \lim_{s\to\infty} x_s, \qquad \bar{R}^j = \lim_{s\to\infty} R^j(x), \qquad \bar{v}^j = \lim_{s\to\infty} v^j_s, \qquad \bar{H}_{x^i} = \lim_{s\to\infty} H_{x^i}(x_s, v_s).$$

In accordance with (3.4), for $\bar{R}^j = 0$ the infinite product

$$\prod_{s=0}^{\infty}[1 - \alpha_s v^j_s]$$

must be zero. For this it is necessary [11] that

$$\sum_{s=0}^{\infty}\ln[1 - \alpha_s v^j_s] = -\infty,$$

but this is possible only if $\bar{v}^j \geq 0$. If $\bar{R}^j < 0$, then we necessarily have $\bar{v}^j = 0$. It can be shown in just the same way that $\bar{H}_{x^i} \geq 0$.

In short, at a limit point of the sequence $\{x_s\}$ the Kuhn–Tucker conditions (1.6) are satisfied.

The convergence rate estimation (3.3) is obtained in the same way as (2.4). In the present case, the first difference of the function φ has to be evaluated along the solution of system (3.1). The theorem is proved. □

Method (3.1) is specially effective if the equation-type constraints in (1.1) are linear, since we can then take comparatively large steps α_s and construct different modifications of the method. A steeped descent version of the method has been used successfully in this case. The algorithm is not significantly more complicated if the inequality type constraints are nonlinear.

Method (3.1) can be used for sufficiently small α_s, if the functions $g^i(x)$ are nonlinear then a check of the conditions $|g^i(x_s)| < \varepsilon$ has to be provided, and if they are infringed, the current value of x_s has to be refined.

On the basis of the methods described, three standard programs for solving problem (1.1) were developed at the computational Centre of the USSR Academy of Sciences. In the first program, for the case when the $g(x)$ are nonlinear, system (1.2) was integrated by Euler's scheme with a fixed scheme with a fixed step; in the second, a method with variable step α_s is realized; the step is split up if the condition for the process to be of the relaxation type is infringed, or if the point x_s leaves the set of admissibility. In the third program, unimportant constraints of the inequality type are omitted remote from the boundary. Let us quote some results of numerical computations by the first program. Let $n = 3$, $e = c = 1$, $F(x) = [x^1 + x^2 + x^3]^2 + 4[x^1 - x^2]^2$, $g(x) = x^1 + x^2 + x^3 - 1$, $h(x) = 3 - 4x^3 - 6x^2 + [x^1]^3$, $x \geq 0$. As x_0 we took the vector $(0.1, 0.7, 0.2)$. The system was integrated until the reduction of the function at each step was not greater than 10^{-5}. With $\alpha_s = 0.1$, we performed 83 steps, and obtained $F = 1.8310951$, $x^1 = 0.2937327$, $x^2 = 0.1001667$, $\mathrm{x}^3 = 0.6061005$. With $\alpha_s = 0.5$, the problem was solved after 13 steps; here, $F = 1.831030$, $x^1 = 0.2937386$, $x^2 = 0.1001510$, $x^3 = 0.6061104$.

4 FINDING SADDLE POINTS

All the above methods can be extended in an obvious way to the problem of finding a saddle point of functions in the case of disconnected sets. Assume that we are seeking

$$\min_{x \in X} \max_{y \in Y} F(x, y), \qquad X = \{x \in \mathbb{E}^n : h(x) \leq 0_{p1}\}, \quad Y = \{y \in \mathbb{E}^m : f(y) \leq 0_{q1}\}.$$

We introduce the vectors $v \in \mathbb{E}^p$, $w \in \mathbb{E}^q$ and put

$$\Phi(x, y, v, w) = F(x, y) + \sum_{i=1}^{p} v^i h^i(x) + \sum_{i=1}^{q} w^i f^i(y).$$

In this case, the method (1.2) reduces to the systems $\dot{x} = -\Phi_x$, $\dot{y} = \Phi_y$, where

$$[h_x^\top h_x + D(-h)]v + h_x^\top F_x = 0, \qquad [f_y^\top f_y + D(-f)]w + f_y^\top F_y = 0.$$

The convergence conditions may be stated and proved in the same way as in Theorems 4, 5.

REFERENCES

1. Yu.G. Evtushenko, and V.G. Zhadan, (1973). Numerical methods for solving some operations research problems, *Zh. vȳchisl. Mat. mat. Fiz.*, **13**, No.3, 583–598.
2. Yu.G. Evtushenko, (1974). Two numerical methods for solving problems of nonlinear programming, *Dokl. Akad. Nauk SSSR*, **215**, No.1, 38–40.
3. J.B. Rosen, The gradient projection method for nonlinear programming, *J. Soc. Industr. and Appl. Math.*, Part I, **8**, No.1, 181–217, 1960; Part II, **9**, No.4, 514–532, 1961.
4. V.G. Karmanov, (1975). *Mathematical programming* (Matematicheskoe programmirovanie), Moscow: Nauka, Fizmatgiz.
5. E.A. Barbashin, (1967). *Introduction to the theory of stability* (Vvedenie v teoriyu ustoichivosti), Moscow: Nauka.
6. Yu.G. Evtushenko, and V.G. Zhadan, (1975). Application of the method of Lyapunov functions for studying the convergence of numerical methods, *Zh. vȳchisl. Mat. mat. Fiz.*, **15**, No.1, 101–112.
7. K.J. Arrow et al., (1958). *Studies in linear and nonlinear programming*, Stanford Univ. Press.
8. J.G. Taylor, (1973). A squared-variable transformation approach to nonlinear programming optimality conditions, *Naval. Res. Logist. Quart*, **20**, No.1, 25–29.
9. Hamdy Taha and Guy Curry, (1971). Classical derivation of the necessary and sufficient conditions for optimal linear programs, *Operat. Res.*, **19**, No.4, 1045–1050.
10. A. Fiacco, and G. Maccormick, (1958). *Nonlinear programming*, Wiley.
11. G.M. Fikhtengol'ts, (1959). *Course of differential and integral calculus* (Kurs diffcrentsial'nogo i integral'nogo ischisleniya), II, Moscow: Fizmatgiz.