Optimization Methods and Software, 1994, v.3, No.1-3, pp. 237–256
Reprints available directly from the publisher
Photocopying permitted by license only

STABLE BARRIER-PROJECTION AND BARRIER-NEWTON METHODS IN NONLINEAR PROGRAMMING[1]

YURI G. EVTUSHENKO and VITALI G. ZHADAN

Computing Center, Russian Academy of Sciences,
40 Vavilov Str., 117967 Moscow GSP-1, Russia

(Revised 28 May 2001)

The present paper is devoted to the application of the space transformation techniques for solving nonlinear programming problems. By using surjective mapping the original constrained optimization problem is transformed to a problem in a new space with only equality constraints. For the numerical solution of the latter problem the stable version of the gradient-projection and Newton's methods are used. After inverse transformation to the original space a family of numerical methods for solving optimization problems with equality and inequality constraints is obtained. The proposed algorithms are based on the numerical integration of the systems of ordinary differential equations. These algorithms do not require feasibility of starting and current points, but they preserve feasibility. As a result of space transformation the vector fields of differential equations are changed and additional terms are introduced which serve as a barrier preventing the trajectories from leaving the feasible set. A proof of convergence is given.

KEY WORDS: constrained minimization, nonlinear programming, space transformation, gradient-projection method, Newton's method, interior point technique, barrier function, Karmarkar's method

1 INTRODUCTION

The purpose of this paper is to call attention to the role of the space transformation techniques in the development of new numerical methods. On the basis of space transformation the original problem with inequality constraints is reduced to a problem with equality constraints. The stable versions of the gradient-projection and Newton's methods are used for solving this reduced problem. The numerical methods are found after performing an inverse transformation. These methods are described by systems of ordinary differential equations. As a result of space transformation the right-hand side of the differential equation is multiplied by some matrices which prevent the trajectories from crossing the boundary of the feasible set. Therefore, these matrices play the role of barrier and we term these methods "barrier-projection" and "barrier-Newton" methods. The space transformation is carried out without using penalty functions and this feature provides a high rate of convergence. The analysis of the method is made on the basis of the stability theory of the solutions of ordinary differential equations. Numerical algorithms are obtained as discretization of dynamical systems. We prove that the barrier-projection method has linear convergence and does not require feasibility. In the linear programming case after some simplifications and after choosing a particular exponential space-transformation function we obtain Dikin's algorithm [3], sometimes called the "variation of Karmarkar's algorithm".

[1]Research was supported by the grant N° 93-012-450 from Russian Scientific fund

The results given here can be found in the papers and books by the authors [4]–[11]. Though some of these papers are translated into English, nevertheless they are essentially unknown in the West. Our approach bears resemblance with some publications which develop the Karmarkar's method (see, for example, [1, 2, 12, 13, 14, 19]). However, there are four main differences:

1. Starting from 1973, we paid attention mainly to the general nonlinear case. Linear programming was considered as a particular case and the polynomiality of the proposed algorithms was not investigated.

2. From 1983 we developed a stable version of the projection method. Therefore, we did not restrict ourselves to the interior point techniques. In our methods the current points are often infeasible, but if the starting points or the current points are feasible, then the subsequent trajectory remains in the feasible set, i.e. the feasibility is preserved.

3. In all proposed methods we did not resort to a penalty-type approach.

4. From 1973 we considered the steepest descent variants of our methods where the trajectory could move along the boundary of the feasible set.

Analyzing the recent publications, we see that many authors are trying to modify and explain Karmarkar's method as a classical method. We think that moving along this direction leads to our results. Here we consider mainly the barrier-projection method in the nonlinear case. Detailed investigations connected with linear programming will be published in English in future. At the present time some of these results are available only in Russian [11].

The family of numerical methods which we propose for the general nonlinear case has been implemented, tested and included in the library of algorithms at the Computing Center of the Russian Academy of Sciences. Computer codes were used for solving various practical nonlinear programming problems. These methods proved to be most efficient in the case where the equality constraints were linear because in this case we could take big stepsizes without violating these constraints.

2 STABLE VERSION OF THE GRADIENT-PROJECTION METHOD

In this section we confine ourselves to the nonlinear programming problem:

$$\text{minimize } f(x) \text{ subject to } x \in X, \tag{1}$$

where X denotes the feasible region which is given by equality constraints:

$$X = \{x \in \mathbb{R}^n : g(x) = 0_m\}. \tag{2}$$

Here $\mathbb{R}^n$ denotes the vector space formed by n-dimensional column vectors with real entries. The functions f and g are continuously differentiable, $f(x)$ maps $\mathbb{R}^n$ onto $\mathbb{R}^1$ and $g(x)$ maps $\mathbb{R}^n$ onto $\mathbb{R}^m$, 0_s is the s-dimensional null vector, 0_{nm} is the $n \times m$ rectangular null matrix. The feasible set X and the set of solutions X_* are supposed to be nonempty. We assume differentiability whenever it is helpful to do so.

Definition 1. *The constraint qualification (CQ) for Problem (1) holds at a point x if the vectors $g^i_x(x)$, $1 \le i \le m$, are linearly independent.*

Definition 2. *We say that x is a regular point for Problem (1), if the CQ holds at x.*

Definition 2 differs from the definition of a regular point of the constraints (see, for example, [15]) because in our case the point x may not belong to the set X.

To obtain the numerical solution of Problem (1) we seek the limit points of the solutions of the system described by the following vector differential equation

$$\frac{dx}{dt} = -L_x(x, u(x)), \tag{3}$$

where $L(x,u) = f(x) + u^\top g(x)$, $L_x(x,u) = f_x(x) + g_x^\top(x)u$, $g_x(x)$ is the $m \times n$ Jacobian matrix of $g(x)$ with respect to x, $a^\top$ denotes the transpose of a.

Let $x(t, x_0)$ denote the solution of the Cauchy problem (3) with initial condition $x_0 = x(0, x_0)$. The function $u(x)$ is chosen to satisfy the following condition

$$\frac{dg}{dt} = -g_x(x)L_x(x, u(x)) = -\tau g(x), \qquad \tau > 0. \tag{4}$$

If the constraint qualification holds on the set

$$Z = \{x \in \mathbb{R}^n : ||g(x)|| \leq ||g(x_0)||\},$$

then the Jacobian matrix $g_x(x)$ is of full rank, the Gram matrix $G(x) = g_x(x)g_x^\top(x)$ is nonsingular and the function $u(x)$ is uniquely determined from condition (4)

$$u(x) = G^{-1}(x)[\tau g(x) - g_x(x) f_x(x)]. \tag{5}$$

By differentiating $f(x(t, x_0))$ with respect to t we obtain

$$\frac{df}{dt} = -||L_x(x, u(x))||^2 + \tau u^\top(x) g(x). \tag{6}$$

If $x_0 \in X$, then the trajectory $x(t, x_0)$ of (3) remains in the feasible set X because $g(x(t, x_0)) \equiv 0_m$ and from (6) it follows that the objective function $f(x(t, x_0))$ monotonically decreases along the trajectory $x(t, x_0) \in X$. If $x_0 \in X$ and/or $\tau = 0$, then the method (3) coincides with the gradient-projection method which has been used by many authors (see, for example, [15, 16]). Nevertheless there is a significant difference between these two approaches. The system of ordinary differential equations (3), where $u(x)$ is given by (5), has the first integral

$$g(x(t, x_0)) = g(x_0)e^{-\tau t}. \tag{7}$$

This means that if $\tau > 0$, then method (3) has a remarkable property: all its trajectories approach the feasible set as t tends to infinity and the feasible set X is an asymptotically stable set of the system (see [5, 10, 11, 18]).

Therefore, we call method (3) "the stable version of the gradient-projection method". On the contrary the well-known gradient-projection method is neutrally stable with respect to the feasibility condition. It means that if $g(x_0) = c$, then $g(x(t, x_0)) \equiv c$ for all $t \geq 0$ and we have to introduce an additional correction procedure in order to satisfy feasibility. This procedure increases the computation time. The gradient-projection method and the method described above can be considered as particular cases of the interior point techniques. There is another motivation in the literature for the proposed method. Let us introduce the augmented Lagrangian

$$M = L(x,u) - \frac{1}{2\tau}||L_x(x,u)||^2, \qquad \tau > 0.$$

If we maximize M with respect to u, then we obtain exactly the same formula as (5) (see [20]).

If x_* is a solution of Problem (1), and there exists a corresponding multiplier u_*, then the pair $[x_*, u_*]$ is a Kuhn–Tucker point, i.e. at this point from the first-order necessary conditions for a minimum it holds that

$$L_x(x_*, u_*) = f_x(x_*) + g_x^\top(x_*)u_* = 0_n, \qquad g(x_*) = 0_m. \tag{8}$$

We say that a point x_*, is an equilibrium point of system (3), if the right-hand side evaluated at x_* is a null vector. The right-hand side of system (3) defines a vector field which vanishes at equilibrium points. We prove now that at every regular point this field is nonvanishing except at points x_* such that $[x_*, u(x_*)]$ forms a Kuhn–Tucker pair.

Lemma 2.1. *Let x_* be a regular point for Problem (1). Then x_* is an equilibrium point of system (3) if and only if the pair $[x_*, u_*]$, where $u_* = u(x_*)$, satisfies (8).*

Proof. Let W be a $m \times n$ rectangular matrix whose rank is m. We introduce the pseudoinverse matrix $W^+ = W^\top(WW^\top)^{-1}$ and the orthogonal projector $\pi(W) = I_n - W^+W$, where I_n is the $n \times n$ identity matrix.

We substitute (5) into the right-hand side of (3) and after some transformation we can rewrite system (3) in the following projective form

$$\frac{dx}{dt} = -\pi[g_x(x)]f_x(x) - \tau[g_x(x)]^+g(x), \tag{9}$$

where $\pi[g_x(x)] = I_n - [g_x(x)]^+g_x(x)$. This operator projects any vector onto the null-space of the matrix $g_x(x)$:

$$g_x^\perp(x) = \ker g_x(x) = \{z \in \mathbb{R}^n : g_x(x)z = 0_m\}.$$

If $g(x) = 0_m$, then $\pi(g_x(x))f_x(x) = L_x(x, u(x))$. Therefore, if conditions (8) holds, then the right-hand side of (9) is the null vector and the statement "if" is proved. Now we prove "only if".

The first vector $\pi(g_x(x))f_x(x)$ at the right-hand side of (9) belongs to the null-space of the matrix $g_x(x)$. The second vector $[g_x(x)]^+g(x)$ belongs the row subspace of the matrix $g_x(x)$. Therefore, these vectors are orthogonal. If x_* is an equilibrium point, then both the mentioned vectors evaluated at this point are the null vectors and $g_x^\top(x_*)z = 0_n$, where $z = G^{-1}(x_*)g(x_*)$. Taking into account the constraint qualification, we come to the conclusion that $g_x^\top(x_*)$ has maximal rank, $G(x_*)$ is nonsingular and, therefore, $g(x_*) = 0_m$. Hence we obtain both conditions (8). $\square$

Integrating (3) using Euler method, we obtain the following iterative process:

$$x_{k+1} = x_k - h_k L_x(x_k, u_k), \qquad u_k = u(x_k), \tag{10}$$

where $h_k > 0$ is a stepsize, function $u(x)$ is defined by (5).

Each equilibrium point x_* of system (9) is a fixed point of iteration (10), i.e. $x_k = x_*$ implies $x_{k+1} = x_*$ and if iterates (10) converge to a regular point x_*, then the pair $[x_*, u(x_*)]$ satisfies the Kuhn–Tucker conditions (8).

Theorem 2.1. *Let $[x_*, u_*]$ be a Kuhn–Tucker point of Problem (1), (2), let the CQ hold at x_* and assume that for any nonzero vector $z \in g_x^\perp(x_*)$ the inequality $z^\top L_{xx}(x_*, u_*)z > 0$ holds. Then system (3), (5) with $\tau > 0$ is asymptotically stable at the isolated local solution point x_* and the pair $[x(t, x_0), u(x(t, x_0))]$ tends to $[x_*, u_*]$, if the starting point x_0 is close enough to the point x_*. There exists a number h_* such that for any fixed $0 < h_k < h_*$ the iterations defined by (10) converge locally with a linear rate to a Kuhn–Tucker point $[x_*, u_*]$.*

The statement of this theorem follows from Theorem 3.2 which will be proved in the next section.

We briefly discuss the global behavior of the method. Assume that the CQ holds on the compact set Z, where Problem (1) has a unique minimal point x_*. The standard existence and uniqueness theorem for the Cauchy problem guarantees that the solution of (9) exists locally. From (7) it follows that the solution $x(t, x_0)$ is bounded, remains in the set Z and can he prolonged for $0 \leq t < \infty$. Therefore, the unique limit point of $x(t, x_0)$ is x_* and the set Z belongs to the domain of attraction of the solution point x_*. We note that method (3) does not require feasibility of the starting and current points, but it preserves feasibility.

3 SPACE TRANSFORMATION

We introduce the additional inequality constraint $x \in P$, where P is assumed to have a nonempty interior. Consider the following problem:

$$\text{minimize } f(x) \text{ subject to } x \in X = \{x \in \mathbb{R}^n : g(x) = 0_m, \ \ x \in P\}. \tag{11}$$

We introduce a new n-dimensional space with the coordinates $[y^1, \ldots, y^n]$ and make a differentiable transformation from this space to the original one: $x = \xi(y)$. This surjective transformation maps $\mathbb{R}^n$ onto P or int P, i.e. $P = \overline{\xi(\mathbb{R}^n)}$, where $\bar{B}$ is the closure of B.

Consider the transformed minimization problem

$$\text{minimize } \tilde{f}(y) = f(\xi(y)) \text{ subject to } y \in Y, \tag{12}$$

where $Y = \{y \in \mathbb{R}^n : \tilde{g}(y) = g(\xi(y)) = 0_m\}$.

Since we now have only equality constraints, we can use the numerical method described in the previous section for solving (12). The Lagrangian associated with Problem (12) is defined by $\tilde{L}(y, u) = \tilde{f}(y) + u^\top \tilde{g}(y)$. Conditions (3), (5) are written as follows

$$\frac{dy}{dt} = -\tilde{L}_y(y, u(y)), \qquad \tilde{L}_y(y, u) = \tilde{f}_y(y) + \tilde{g}_y^\top(y) u, \tag{13}$$

$$\tilde{g}_y(y)\tilde{g}_y^\top(y)u(y) + \tilde{g}_y(y)\tilde{f}_y(y) = \tau \tilde{g}(y), \qquad y_0 \in \mathbb{R}^n,$$

where $\tilde{f}_y = \tilde{J}^\top f_x$, $\tilde{g}_y = g_x \tilde{J}$, $\tilde{J} = dx/dy$ and $\tilde{J}$ is the Jacobian matrix of the transformation $x = \xi(y)$ with respect to y. If $\tilde{J}$ is nonsingular matrix, then there exists an inverse transformation $y = \delta(x)$, so it is possible to return from the y-space to the x-space and we obtain in this way a matrix $J(x) = \tilde{J}(\delta(x))$ which is now a function of x. By differentiating $\xi(y)$ with respect to y and taking into account (13), we have

$$\frac{dx}{dt} = \frac{d\xi}{dy}\frac{dy}{dt} = J(x)\frac{dy}{dt} = -G(x)L_x(x, u(x)), \qquad x(0, x_0) = x_0 \in P, \tag{14}$$

$$\Gamma(x)u(x) + g_x(x)G(x)f_x(x) = \tau g(x), \tag{15}$$

where we have introduced the two Gram matrices:

$$\Gamma(x) = g_x(x)G(x)g_x^\top(x), \qquad G(x) = J(x)J^\top(x).$$

If the condition $x \in P$ is absent in Problem (11), then the space transformation is trivial: $x = y$ and, taking $G(x) = I_n$ in (14), (15), we obtain formulas (3), (4).

We define the following sets: the nullspace of the matrix $g_x(x)J(x)$ at x:

$$K(x) = \{z \in \mathbb{R}^n : g_x(x)J(x)z = 0_m\},$$

the cone of feasible directions at $x \in P$:

$$F(x\,|P) = \{z \in \mathbb{R}^n : \exists \lambda(z) > 0 \text{ such that } x + \lambda z \in P, \;\; 0 < \lambda \leq \lambda(z)\},$$

the conjugate cone to the cone F:

$$F^*(x\,|P) = \{z \in \mathbb{R}^n : z^\top y \geq 0 \quad \forall y \in F(x\,|P)\},$$

the linear hull of the cone F^* at $x \in P$:

$$S(x) = \operatorname{lin} F^*(x\,|P) = \{z \in \mathbb{R}^n : z = \sum_{i=1}^{s} \lambda^i z_i, \;\; \lambda^i \in \mathbb{R}; \;\; z_i \in F^*(x\,|P), \;\; 1 \leq i \leq s, \;\; s = 1, 2, \ldots\}.$$

Definition 3. *The constraint qualification (CQ) for Problem (11) holds at a point $x \in P$ if all vectors $g^i_x(x)$, $1 \leq i \leq m$, and any nonzero vector $p \in S(x)$ are linearly independent. We say that x is a regular point for Problem (11) if the CQ holds at x.*

We impose the following condition on the transformation $\xi(y)$:
$\mathbf{C}_1$. *At each point $x \in P$ the Jacobian $J(x)$ is defined and* $\ker J^\top(x) = S(x)$.
From this condition it follows that the Jacobian $J(x)$ is nonsingular in the interior of P, it is singular only on the boundary of P.

Lemma 3.1. *Let the space transformation $\xi(y)$ satisfy $\mathbf{C}_1$, and let the CQ for Problem (11) hold at a point $x \in P$. Then the Gram matrix $\Gamma(x)$ is invertible and positive definite.*

Proof. We prove first that the rank of $B(x) = J^\top(x) g^\top_x(x)$ is equal to m. If $x \in \operatorname{int} P$, then it is obvious because $J(x)$ is nonsingular and $g_x(x)$ has maximal rank according to the constraint qualification.

Let x belong to the boundary of P. If rank of $B(x)$ is less than m, then there exists a nonzero vector $z \in \mathbb{R}^m$ such that $B(x)z = J^\top(x) g^\top_x(x) z = 0_n$. According to condition $\mathbf{C}_1$ the nonzero vector $p = g^\top_x(x) z$ belongs to $S(x)$, therefore, vectors g^i_x, $1 \leq i \leq m$, and p are linearly dependent. This contradicts to CQ. We obtain that $B(x)$ has full rank and from the representation $\Gamma(x) = B^\top(x) B(x)$ it follows that $\Gamma(x)$ is a positive definite matrix. $\square$

If at a point x the conditions of Lemma 3.1 hold, then we can find from (15) the function $u(x)$, substitute it into the right-hand side of (14) and write (14) in the following projective form

$$\frac{dx}{dt} = -J(x) \left\{ \pi[g_x(x)J(x)] J^\top(x) f_x(x) + \tau [g_x(x) J(x)]^+ g(x) \right\}. \tag{16}$$

Similarly to (6) we have

$$\frac{df}{dt} = -||J^\top(x) L_x(x, u(x))||^2 + \tau u^\top(x) g(x). \tag{17}$$

As before the objective function $f(x(t, x_0))$ monotonically decreases on the feasible set X and if the trajectory is close to X, i.e. if $\|g(x)\|$ is sufficiently small.

Definition 4. *The pair $[x_*, u_*]$, where $x_* \in P$, is a Kuhn–Tucker pair of Problem (11), if*

$$J^\top(x_*) L_x(x_*, u_*) = 0_n, \qquad g(x_*) = 0_m. \tag{18}$$

Lemma 3.2. *Let conditions of Lemma 3.1 be satisfied at a point $x_* \in P$. Then x_* is an equilibrium point of system (14) if and only if the pair $[x_*, u_*]$, where $u_* = u(x_*)$, satisfies (18).*

Proof. Let x_* be an equilibrium point of system (14). The matrices $G(x_*)$ and $J^\top(x_*)$ have the same nullspace. Therefore, from the inclusion $L_x(x_*, u_*) \in \ker G(x_*)$ we conclude that $L_x(x_*, u_*) \in \ker J^\top(x_*)$. Hence the first condition in (18) is true. The validity of the second condition follows from equality (15). $\square$

Definition 5. *The strict complementary condition (SCC) holds at a point $[x_*, u_*] \in \in P \times \mathbb{R}^m$, if*

$$L_x(x_*, u_*) \in \operatorname{ri} F^*(x_*|P), \tag{19}$$

where ri A is a relative interior of the set A.

The space transformation described above can be used to derive the following second-order sufficient conditions for a point to be an isolated local minimum in Problem (11).

Theorem 3.1. *Assume that f and g are twice-differentiable functions and the space transformation $\xi(y)$ satisfies $\mathbf{C}_1$. Sufficient conditions that a point $x_* \in P$ be an isolated local minimum of Problem (11) are that there exists a Lagrange multiplier vector u_* satisfying the Kuhn–Tucker conditions (18), that the SCC holds at $[x_*, u_*]$ and that $z^\top J^\top(x_*) L_{xx}(x_*, u_*) J(x_*) z > 0$ for every $z \in K(x_*)$ such that $J(x_*)z \neq 0_n$.*

We denote by $D(z)$ the diagonal matrix containing the components of a vector z. The dimensionality of this matrix is determined by the dimensionality of z.

For the sake of simplicity we consider now the particular case of Problem (11), where the set P is the positive orthant, i.e. $P = \mathbb{R}^n_+$. It is convenient for this set P to use a component-wise differentiable space transformation $\xi(y)$

$$x^i = \xi^i(y^i), \qquad 1 \leq i \leq n. \tag{20}$$

For such a transformation the corresponding Jacobian matrix is diagonal and

$$\tilde{J}(y) = D(\dot{\xi}(y)), \qquad \dot{\xi}(y) = [\dot{\xi}^1(y^1), \dot{\xi}^2(y^2), \ldots, \dot{\xi}^n(y^n)]^\top.$$

Let $\delta(y)$ be the inverse transformation. Denote

$$J(x) = D(\dot{\xi}(y))|_{y=\delta(x)}, \qquad G(x) = J^2(x) = D(\theta(x))$$

with the vector $\theta(x) = [(\dot{\xi}^1(y^1))^2, (\dot{\xi}^2(y^2))^2, \ldots, (\dot{\xi}^n(y^n))^2]|_{y=\delta(x)}$.

We impose on a space transformation $\xi(y)$ the following conditions:

$\mathbf{C}_2$. *$\theta^i(x^i) = 0$ if and only if $x^i = 0$, where $1 \leq i \leq n$.*

$\mathbf{C}_3$. *The space transformation $\xi(y)$ satisfies condition $\mathbf{C}_2$ and $\dot{\theta}^i(0) > 0$, $1 \leq i \leq n$.*

Different numerical methods are obtained by different choices of the space transformations. As a rule we perform the following quadratic and exponential transformations

$$x^i = \xi^i(y^i) = \frac{1}{4}(y^i)^2, \quad J(x) = D^{1/2}(x), \quad G(x) = D(x), \tag{21}$$

$$x^i = \xi^i(y^i) = e^{y^i}, \qquad J(x) = D(x), \qquad G(x) = D^2(x). \tag{22}$$

In these two cases the Jacobian matrix is singular on the boundary of the set P. These transformations satisfy $\mathbf{C}_1$, $\mathbf{C}_2$. Condition $\mathbf{C}_3$ holds only for transformation (21). If the SCC holds at $[x_*, u_*]$, then for these transformations we obtain that $L_{x^i}(x_*, u_*) > 0$ for all i such that $x_*^i = 0$.

From method (16) interesting particular cases are derived. Let $X = P = \mathbb{R}^n_+$. If we use the neutral stable version of (3) with the space transformation functions (21), (22), then we obtain

$$\frac{dx}{dt} = -D^\alpha(x) f_x(x), \tag{23}$$

where $x_0 > 0_n$, $\alpha = 1$ for transformation (21) and $\alpha = 2$ for (22).

Let $x^i(t, x_0) = 0$. Then we have $\theta^i(x^i) = 0$ and $dx^i(t, x_0)/dt = 0$. From the last equality it follows that the trajectory $x(t, x_0)$ of system (23) cannot cross the boundary $x^i = 0$. Thus a transformation function plays the role of a "barrier", preventing the trajectory $x(t, x_0)$ from passing through the boundary of P. Therefore, we call (16) a "barrier-projection method".

In our first publication in this field [6] we used the quadratic space transformation (21) and considered the resource allocation problem where $X = \{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x^i = 1\}$. In this case $m = 1$ and system (15) was solved explicitly, while system (16) had the form

$$\frac{dx}{dt} = -D(x)[f_x(x) - c(x)e], \qquad x_0 \in X,$$

where $e \in \mathbb{R}^n$ is the vector of ones, $c(x) = x^\top f_x(x)$. More general cases were considered in subsequent papers [4, 5, 7].

Applying the Euler method for solving system (16), we obtain

$$x_{k+1} = x_k - h_k J(x_k)\{\pi[g_x(x_k)J(x_k)]J^\top(x_k) f_x(x_k) + \tau[g_x(x_k)J(x_k)]^+ g(x_k)\}. \tag{24}$$

Theorem 3.2. *Let $[x_*, u_*]$ be a Kuhn–Tucker pair of Problem (11), where the CQ and the second-order sufficiency conditions of Theorem 3.1 hold. Let the space transformation $\xi(y)$ satisfy condition $\mathbf{C}_3$. Then the system (16) with $\tau > 0$ is asymptotically stable at the isolated local solution point x_*. There exists a number h_* such that for any fixed $0 < h_k < h_*$ the sequence $\{x_k\}$, generated by (24), converges locally with a linear rate to x_* while the corresponding sequence u_k converges to u_*.*

Proof. Denote $\delta x(t) = x(t, x_0) - x_*$ and linearize system (16) in the neighborhood of the point x_*. Then we obtain the equation of the first approximation of (16) about the equilibrium point x_*:

$$\delta\dot{x} = -Q\delta x, \tag{25}$$

where $Q = \tilde{M}[GL_{xx} + D(\dot{\theta})D(L_x)] + \tau\tilde{P}$, $\tilde{M} = I_n - \tilde{P}$, $\tilde{P} = Gg_x^\top(g_x G g_x^\top)^{-1} g_x$. Here all functions are evaluated at the points $x = x_*$, $u = u_* = u(x_*)$.

The stability of system (25) is determined by the properties of the roots of the characteristic equation

$$\det(Q - \lambda I_n) = 0. \tag{26}$$

For proof we split the vector x_* in two vectors $x_*^\top = [x_B^\top, x_N^\top]$, $x_B \in \mathbb{R}^s$, $x_N \in \mathbb{R}^d$, $d = n - s$. All components of the vector x_N are equal to zero and all components of x_B are interior, i.e. $x_B > 0$. In a similar way we represent vectors $\theta^\top(x_*) = [\theta_B^\top, \theta_N^\top]$, $L_x^\top = [(L_x^B)^\top, (L_x^N)^\top]$ and matrices:

$$L_{xx} = \begin{bmatrix} L_{xx}^B & L_{xx}^{BN} \\ L_{xx}^{NB} & L_{xx}^N \end{bmatrix}, \qquad \tilde{P} = \begin{bmatrix} \tilde{P}_B & \tilde{P}_{BN} \\ \tilde{P}_{NB} & \tilde{P}_N \end{bmatrix}, \qquad J = \begin{bmatrix} J_B & 0_{sd} \\ 0_{ds} & J_N \end{bmatrix},$$

$$G_B = D(\theta_B) = J_B J_B.$$

From $\mathbf{C}_2$ and the strict complimentary condition (19) it follows that $\theta_N = 0_d$, $L_x^B = 0_s$ and $\tilde{P}_{BN}$, $\tilde{P}_{NB}$, $\tilde{P}_{NN}$ are null matrices. Hence the matrix Q can be decomposed into the following blocks

$$Q = \left[\begin{array}{cc} Q_1 & Q_3 \\ 0_{ds} & Q_2 \end{array} \right],$$

where the matrix Q_3 is not essential and

$$Q_1 = J_B \bar{M} J_B L_{xx}^B + \tau J_B \bar{P} J_B^{-1}, \qquad Q_2 = D(\dot{\theta}_N) D(L_x^N),$$

$$\bar{M} = I_s - \bar{P}, \qquad \bar{P} = J_B g_{x_B}^\top (g_{x_B} G_B g_{x_B}^\top)^{-1} g_{x_B} J_B.$$

The characteristic equation (26) is equivalent to the two equations:

$$|Q_1 - \lambda I_s| = 0, \qquad |Q_2 - \lambda I_d| = 0.$$

The solutions of the second equation are found explicitly: $\lambda_i = \dot{\theta}^i(x_*^i) L_{x^i}(x_*, u_*)$, $s+1 \leq i \leq n$.

From $\mathbf{C}_3$ and the strict complementarity condition we obtain:

$$\hat{\lambda} = \min_{s+1 \leq i \leq n} \lambda_i > 0. \tag{27}$$

The matrix Q_1 is similar to $S_1 = J_B^{-1} Q_1 J_B$, therefore, they have the same eigenvalues and we can consider the following characteristic equation

$$|S_1 - \lambda I_s| = |\bar{M} \bar{L}_{xx}^B + \tau \bar{P} - \lambda I_s| = 0,$$

where $\bar{L}_{xx}^B = J_B L_{xx}^B J_B$.

The matrices $\bar{M}$ and $\bar{P}$ are projection matrices for the tangent subspace

$$\bar{K}(x_*) = \{z \in \mathbb{R}^s : g_{x_B}(x_*) J_B(x_*) z = 0_m\}$$

and its orthogonal complement, respectively. Furthermore, we have

$$\bar{P}\bar{M} = 0_{ss}, \qquad \bar{P}\bar{P} = \bar{P}, \qquad \bar{M}\bar{M} = \bar{M}. \tag{28}$$

Let λ_i and z_i be an eigenvalue and a corresponding nonzero eigenvector of the matrix S_1. Then

$$(\bar{M}\bar{L}_{xx} + \tau \bar{P}) z_i = \lambda_i z_i, \qquad z_i \in \mathbb{R}^s. \tag{29}$$

If $\bar{P} z_i \neq 0_s$, then premultiplying (29) by the matrix $\bar{P}$ and taking into account (28), we obtain $\lambda_i = \tau$. If $\bar{P} z_i = 0_s$, then $z_i \in \bar{K}(x_*)$ and multiplying (29) by $z_i^\top$ we have

$$\lambda_i = \frac{z_i^\top J_B(x_*) L_{xx}^B(x_*, u_*) J_B(x_*) z_i}{||z_i|||^2}, \qquad 1 \leq i \leq s, \quad \tilde{\lambda} = \min_{1 \leq i \leq s} \lambda_i > 0.$$

These results imply that all roots of the characteristic equation for the matrix Q are real and the smallest root $\lambda_* = \min\,[\tilde{\lambda}, \hat{\lambda}]$ is positive. Hence, according to Lyapunov's linearization principle [5], the equilibrium point x_* is asymptotically stable and the following estimation holds:

$$\lim_{t \to \infty} \sup \frac{||x(t, x_0) - x_*||}{t} \leq -\lambda_*.$$

Denote $h_* = 2/\lambda^*$, where $\lambda^* = \max_{1 \leq i \leq n} \lambda_i$. If the stepsize $h_k < h_*$, then, by [5, Theorem 2.3.7], the linear convergence of the discrete version (24) follows from the proof given above. $\square$

Theorem 3.1 cannot be used for the space transformation (22) because this transformation does not satisfy condition $\mathbf{C}_3$. This case was considered by G. Smirnov on the basis of the vector Lyapunov function [17].

Suppose that (11) is a linear programming problem, i.e.

$$f(x) = c^\top x, \qquad g(x) = b - Ax, \qquad P = \mathbb{R}^n_+,$$

where A is a $m \times n$ rectangular matrix. If we use the exponential space transformation (22) and put $\tau = 0$, then from (16) we obtain

$$\frac{dx}{dt} = D^2(x)(A^\top u(x) - c), \qquad AD^2(x)A^\top u(x) = AD^2(x)c. \tag{30}$$

The discrete and continuous versions of this method were investigated in various papers (see, for example, [1, 2, 3, 12, 13, 19]). In [1] the discrete version was called "a variation on Karmarkar's algorithm". We should remark that method (30) does not possess the local convergence property. The convergence takes place only if x_0 belongs to the relative interior of X, i.e. $g(x_0) = 0_m$, $x_0 > 0_n$. Detailed analysis of the method in the case of a linear programming problem is given in [11].

There is another interesting case, where P is an n-dimensional box, i.e. $P = \{x \in \mathbb{R}^n : a \leq \leq x \leq b\}$. Then we can use the following transformation

$$x^i = \frac{[a^i + b^i + (b^i - a^i)\sin y^i]}{2}, \qquad G(x) = D(x-a)D(b-x).$$

The statement of the theorem can be generalized for this set.

4 PROBLEMS WITH GENERAL INEQUALITY CONSTRAINTS

The preceding results and algorithms admit straightforward extensions for problems involving general inequality constraints by using space dilation. Consider Problem

$$\text{minimize}\, f(x) \text{ subject to } x \in X = \{x \in \mathbb{R}^n : g(x) = 0_m,\;\; h(x) \leq 0_c\}, \tag{31}$$

where $h(x)$ maps $\mathbb{R}^n$ into $\mathbb{R}^c$.

In Problem (31) we do not have nonnegativity constraints on the separated variables. Nevertheless our approach can be used in this case by extension of the space and by converting the inequality constraints to equalities. We introduce an additional variable $p \in \mathbb{R}^c$, denote $q = m + c$, combine primal, dual variables and all constraints:

$$z = \begin{bmatrix} x \\ p \end{bmatrix} \in \mathbb{R}^{n+c}, \qquad w = \begin{bmatrix} u \\ v \end{bmatrix} \in \mathbb{R}^q, \qquad \Phi(z) = \begin{bmatrix} g(x) \\ h(x) + p \end{bmatrix}, \qquad \Phi : \mathbb{R}^{n+c} \to \mathbb{R}^q.$$

Then the original Problem (31) is transformed into the equivalent Problem

$$\text{minimize}\, f(x) \text{ subject to } z \in Z = \{z \in \mathbb{R}^{n+c} : \Phi(z) = 0_q,\;\; p \in \mathbb{R}^c_+\}. \tag{32}$$

This Problem is similar to (11). In order to take into account the constraint $p \geq 0_c$ we introduce a surjective differentiable mapping $\varphi : \mathbb{R}^c \to \mathbb{R}^c_+$ and make the space transformation $p = \varphi(y)$, where $y \in \mathbb{R}^c$, $\overline{\varphi(\mathbb{R}^c)} = \mathbb{R}^c_+$. Let φ_y be the square $c \times c$ Jacobian matrix of the mapping $\varphi(y)$ with respect to y. We assume that it is possible to define the inverse transformation $y = \psi(p)$ and hence we obtain the $c \times c$ Jacobian and Gram matrices:

$$J(p) = \varphi_y(y)|_{y=\psi(p)}, \qquad G(p) = J(p)J^\top(p).$$

Combining variables and constraints for the reduced Problem, denote

$$\hat{z} = \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^{n+c}, \qquad \hat{\Phi}(\hat{z}) = \begin{bmatrix} g(x) \\ h(x) + \varphi(y) \end{bmatrix}, \quad \hat{\Phi}_{\hat{z}} = \begin{bmatrix} g(x) & 0_{mc} \\ h(x) & \varphi_y \end{bmatrix}.$$

Problems (31) and (32) can be formulated as follows:

$$\text{minimize } f(x) \text{ subject to } \hat{z} \in \hat{Z} = \{\hat{z} \in \mathbb{R}^{n+c} : \hat{\Phi}(\hat{z}) = 0_q\}. \tag{33}$$

In the last Problem we have only equality constraints, therefore, we can use classical optimality conditions and the numerical method described in the second section. After inverse transformation to the space of x and p we obtain

$$\frac{dz}{dt} = -\tilde{G}(p)L_z(z, w(z)). \tag{34}$$

Here $L(z, w) = f(x) + w^\top \Phi(z),\ L_z(z, w) = f_z(z) + \Phi_z^\top(z)w,$

$$\Phi_z(z)\tilde{G}(p)L_z(z, w(z)) = \tau\Phi(z), \tag{35}$$

$$\tilde{G}(p) = \begin{bmatrix} I_n & 0_{nc} \\ 0_{cn} & G(p) \end{bmatrix}, \qquad \Phi_z = \left[\Phi_x \left| \begin{matrix} 0_{mc} \\ I_c \end{matrix} \right. \right], \Phi_x = \begin{bmatrix} g_x \\ h_x \end{bmatrix}, \qquad f_z = \begin{bmatrix} f_x \\ 0_c \end{bmatrix}.$$

System (34) can be rewritten in the more detailed form

$$\frac{dx}{dt} = -L_x(x, w(z)), \qquad \frac{dp}{dt} = -G(p)v, \tag{36}$$

where the function $w(z)$ is found from the following linear system of q equations

$$\Gamma(z)w(z) + \Phi_x(x)f_x(x) = \tau\Phi(z), \qquad \Gamma(z) = \Phi_x(x)\Phi_x^\top(x) + \begin{bmatrix} 0_{mm} & 0_{mc} \\ 0_{cm} & G(p) \end{bmatrix}.$$

Condition (4) can be written as

$$\frac{dg(x)}{dt} = -\tau g(x), \qquad \frac{d(h(x)+p)}{dt} = -\tau(h(x)+p).$$

Therefore, system (36) has two first integrals:

$$g(x(t, z_0)) = g(x_0)e^{-\tau t},$$

$$h(x(t, z_0)) + p(t, z_0) = (h(x_0) + p_0)e^{-\tau t}, \qquad z_0^\top = [x_0^\top, p_0^\top].$$

Similarly to (6) we obtain

$$\frac{df}{dt} = -||L_x||^2 - ||J^\top(p)v||^2 + \tau[u^\top g + v^\top(h + p)].$$

Introduce the index set $\sigma(p) = \{i \in [1 : c] : p^i = 0\}$. The set P in Problem (32) has the form $P = \{[x, p] \in \mathbb{R}^{n+c} : p \geq 0\}$. Therefore, Definition 3 of CQ for this problem can be reformulated as follows.

Definition 6. *The constraint qualification (CQ) for Problem (32) holds at a point $z \in P$, if all vectors $g_x^i(x)$, $1 \leq i \leq m$, and vectors $h_x^i(x)$, $i \in \sigma(p)$, are linearly independent.*

Impose on the mapping $\varphi(y)$ a new condition which is similar to $\mathbf{C}_1$:

$\mathbf{C}_4$. *At each point $p \in \mathbb{R}^c_+$ the Jacobian $J(p)$ is defined and*

$$\ker J^\top(p) = \operatorname{lin} F^*(p|\mathbb{R}^c_+) = \{b \in \mathbb{R}^c : \text{ if } i \in \sigma(p), \quad \text{then } b^i = 0\}. \tag{37}$$

Lemma 4.1. *Let the mapping $\varphi(y)$ satisfy $\mathbf{C}_4$ and let the CQ for Problem (32) hold at a point $z \in P$. Then the Gram matrix $\Gamma(z)$ is invertible and positive definite.*

Proof. If we denote

$$H(z) = \begin{bmatrix} g_x(x) & 0_{mc} \\ h_x(x) & J(p) \end{bmatrix},$$

then $\Gamma(z) = H(z)H^\top(z)$. Hence the matrix $\Gamma(z)$ is positive definite and invertible at any point z, where the matrix $H(z)$ has a maximal rank.

Suppose that $\sigma(p) = \{1, 2, \dots, s\},\ 0 \le s \le c$. Then, according to $\mathbf{C}_4$, the matrix $J^\top(p)$ has the form

$$J^\top(p) = [0_{cs} \ \ W],$$

where W is a $c \times (c - s)$ matrix of full rank. So the following representation holds:

$$H^\top(z) = \begin{bmatrix} g_x^\top & h_x^1, \dots, h_x^s & h_x^{s+1}, \dots, h_x^c \\ 0_{cm} & 0_{cs} & W \end{bmatrix}.$$

If matrix $H^\top(z)$ is not of full rank, then there exists a nonzero vector $a \in \mathbb{R}^q$ such that $H^\top(z)a = 0_{n+c}$. Taking into account that the matrix W has a maximal rank, we obtain $a^i = 0$, $n + s + 1 \le i \le n + c$, while the other components are not all zero. Consequently the linear combination of vectors $g_x^i,\ 1 \le i \le m$, and $h_x^j,\ j \in \sigma(p)$, is equal to zero. This contradicts to CQ for Problem (32). $\square$

We can use all results given in the second section for Problem (32) and for the corresponding numerical method (36). It follows from CQ for Problem (32) that at a point z the vectors $\Phi_z^i(z)$, $1 \le i \le q$, are linearly independent. Consequently z is a regular point for Problem (32). The pair $[z_*, w_*]$ is a Kuhn–Tucker pair of Problem (32), if

$$L_x(z_*, w_*) = 0_n, \qquad \Phi(z_*) = 0_q, \qquad D(p_*)v_* = 0_c. \tag{38}$$

If $v_*^i > 0$ for i such that $p_*^i = 0$, then the strict complementarity condition (SCC) is fulfilled at the point z_*.

Lemma 4.2. *Let conditions of Lemma 4.1 be satisfied at a point $z_* \in \mathbb{R}^n \times \mathbb{R}^c_+$. Then z_* is an equilibrium point of system (36) if and only if the pair $[z_*, w_*]$ satisfies (38).*

Proof. Let z_* be the stationary point of (36). We need only to prove the validity of the last two conditions (38). It follows from (36) that $G(p_*)v_* = 0_c$. The nullspace of the matrix $G(p_*) = J(p_*)J^\top(p_*)$ coincides with the nullspace of the matrix $J^\top(p_*)$ and consequently $v_* \in \ker J^\top(p_*)$. Taking into account (37), we obtain $D(p_*)v_* = 0_c$. The equality $\Phi(z_*) = 0_q$ follows from (35). $\square$

It is convenient to use for Problem (32) a component-wise mapping $p^i = \varphi^i(y^i),\ 1 \le i \le c$. Therefore, $\varphi_y = D(\dot\varphi)$, where $\dot\varphi(y)$ is a c-dimensional vector. We denote $\gamma^i(p^i) = \dot\varphi^i(\psi^i(p^i))$, $\theta(p) = [\theta^1(p^1), \dots, \theta^c(p^c)],\ \theta^i(p^i) = [\gamma^i(p^i)]^2,\ 1 \le i \le c$. For this mapping the matrix $G(p)$ has the diagonal form $G(p) = D(\theta(p))$. Introduce the following conditions:

$\mathbf{C}_5$: $\gamma^i(p^i) = 0$ *if and only if* $p^i = 0$;

C$_6$: *$\theta^i(p^i)$ is a differentiable function and $\dot{\theta}^i(0) > 0$.*

Let us consider the tangent cone

$$\tilde{K}(x) = \{\bar{x} \in \mathbb{R}^n : g_x(x)\bar{x} = 0_m; \;\; \langle h^i_x, \bar{x} \rangle = 0, \;\; i \in \sigma(h(x))\}.$$

Theorem 4.1. *Let $[z_*, w_*]$ be a Kuhn–Tucker pair of Problem (32), where the CQ and the SCC hold. Let the space transformation $\varphi(y)$ satisfy conditions* **C$_5$, C$_6$** *and assume that for any nonzero vector $\bar{x} \in \tilde{K}(x_*)$ the inequality $\bar{x}^\top L_{xx}(x_*, w_*)\bar{x} > 0$ holds. Then system (36) with $\tau > 0$ is asymptotically stable at the isolated local solution point z_*. The discrete version of the method converges locally with a linear rate, if the stepsize is constant and sufficiently small.*

The proof is nearly identical to the proof of the Theorem 3.2.

Consider the simplified version of method (36). Suppose that along the trajectories of system (36) the following condition holds

$$h(x(t, z_0)) + p(t, z_0) \equiv 0_c.$$

From this equality we can define p as a function of h. We exclude from system (36) the additional vector p and integrate the system which does not employ this vector:

$$\frac{dx}{dt} = -L_x(x, w(x)), \tag{39}$$

where

$$\Gamma(x)w(x) + \Phi_x(x)f_x(x) = \tau \begin{bmatrix} g(x) \\ 0_c \end{bmatrix}, \tag{40}$$

$$\Gamma(x) = \Phi_x(x)\Phi_x^\top(x) + \begin{bmatrix} 0_{mm} & 0_{mc} \\ 0_{cm} & G(-h(x)) \end{bmatrix}.$$

Along the trajectories of (39) we have

$$\frac{dg}{dt} = -\tau g(x), \qquad \frac{dh}{dt} = -G(-h(x))v(x), \tag{41}$$

$$\frac{df}{dt} = -\|L_x(x, w(x))\|^2 - \|J^\top(-h(x))v(x)\|^2 + \tau u^\top(x)g(x).$$

Let us show that the solution $x(t, x_0)$ does not leave the set X for any $t > 0$, if $x_0 \in X$. Suppose this is not true and let $h^i(x(t, x_0)) > 0$ for some $t > 0$. Then there is a time $0 < t_1 < t$ such that $h^i(x(t_1, x_0)) = 0$ and $\dot{h}^i(x(t_1, x_0)) > 0$. This contradicts (41) since $\theta^i(0) = 0$. Hence $x(t, x_0) \in X$ for all $t \geq 0$. Thus the Jacobian $J(-h(x))$ plays the role of a "barrier" preventing $x(t, x_0)$ from intersecting the hypersurface $h^i(x) = 0$. The trajectory $x(t, x_0)$ 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 (39) belongs to the boundary.

Method (39) is closely related to method (14). Let us consider Problem (11), assuming that $P = \mathbb{R}^n_+$. We have two alternatives: we can use methods (14) or (39). The main computational work required in any numerical integration method is to evaluate the right-hand sides of the systems for various values of x. This could be done by solving the linear system (15) of m equations or system (40) of $m + n$ equations, respectively. One might suspect that the introduction of slack variables p increases the computational work considerably. However, by taking advantage of the simple structure of equation (40), we can reduce the computational time by using the Frobenious formula for an inverse matrix [5]. After some transformations we find that formulas (39), (40) can be written as (14), (15), respectively, if in the last formulas we take

$$G(x) = D(\mu(x)), \qquad \mu^i(x^i) = \frac{\theta^i(x^i)}{1+\theta^i(x^i)}, \qquad 1 \le i \le n.$$

Therefore, the performances of both seemingly unrelated methods are very similar.

5 BARRIER-NEWTON METHOD

Let us consider Problem (11) supposing that all functions $f(x)$, $g^i(x)$, $1 \le i \le m$, are at least twice continuously differentiable. Assume also for simplicity that $P = \mathbb{R}^n_+$ and the transformation $\xi(y)$ has the component-wise form (20).

Equation (15) can be rewritten as

$$g_x(x) D(\theta(x)) L_x(x, u(x)) = \tau g(x). \tag{42}$$

Therefore, if the space transformation $\xi(y)$ satisfies $\mathbf{C}_2$ and x is a regular point such that

$$D(\theta(x)) L_x(x, u(x)) = 0_n,$$

then $[x, u(x)]$ is a Kuhn–Tucker point of Problem (11). In Section 3 we used the gradient method for finding a solution x of this equation. Now we will apply Newton's method for this purpose. The continuous version of Newton's method leads to the initial value problem for the following system of ordinary differential equations

$$\Lambda(x) \frac{dx}{dt} = -D(\alpha) D(\theta(x)) L_x(x, u(x)), \qquad x(t, x_0) = x_0, \tag{43}$$

where $\alpha \in \mathbb{R}^n$ is a scaling vector, $\Lambda(x)$ is the Jacobian matrix of the mapping $D(\theta(x))L_x(x, u(x))$ with respect to x:

$$\Lambda(x) = D(\dot{\theta}) D(L_x) + D(\theta) L_{xx} + D(\theta) g_x^\top \frac{du}{dx}. \tag{44}$$

Here all matrices and vectors are evaluated at a point x and the function $u(x)$ is defined from (42). By differentiating equality (42) with respect to x, we obtain

$$g_x \left[D(\dot{\theta}) D(L_x) + D(\theta) L_{xx} + D(\theta) g_x^\top \frac{du}{dx} \right] + E = \tau g_x, \tag{45}$$

where $E(x)$ denotes the $m \times n$ matrix with the elements

$$e_{ij}(x) = \sum_{k=1}^{n} \theta^k(x) \frac{\partial^2 g^i(x)}{\partial x^k \partial x^j} \frac{\partial L(x, u(x))}{\partial x^k}.$$

Let us assume that x is a regular point of Problem (11). From (42) and (45) we find that

$$u = \Gamma^{-1}[\tau g - g_x D(\theta) f_x], \qquad \Gamma = g_x D(\theta) g_x^\top,$$

$$\frac{du}{dx} = \Gamma^{-1}[\tau g_x - g_x D(\dot{\theta}) D(L_x) - g_x D(\theta) L_{xx} - E]. \tag{46}$$

Using the notations $T(x) = g_x^\top(x) \Gamma^{-1}(x)$, $\Omega(x) = T(x) g_x(x)$, we obtain after substitution of (46) into (44)

$$\Lambda = [I_n - D(\theta)\Omega][D(\theta) L_{xx} + D(\dot{\theta}) D(L_x)] + D(\theta) T [\tau g_x - E].$$

Lemma 5.1. *Let conditions of Theorem 3.2 be satisfied at the point* $[x_*, u_*]$. *Then* $\Lambda(x_*)$ *is a nonsingular matrix.*

Proof. Taking into account that $E(x_*) = 0_{mn}$ at the Kuhn–Tucker point $[x_*, u_*]$, we obtain

$$\Lambda(x_*) = \tau D(\theta(x_*))\Omega(x_*) + [I_n - D(\theta(x_*))\Omega(x_*)][D(\theta(x_*))L_{xx}(x_*, u_*) + D(\dot{\theta}(x_*))D(L_x(x_*, u_*))].$$

Hence the matrix $\Lambda(x_*)$ coincides with the matrix Q which was introduced in (25). It was shown in the proof of Theorem 3.2 that all eigenvalues of this matrix are real and strictly positive. Therefore, $\Lambda(x_*)$ is nonsingular. $\square$

Theorem 5.1. *Let conditions of Theorem 3.2 be satisfied at the point* $[x_*, u_*]$. *Then for any scaled vector* $\alpha > 0_n$ *and any* $\tau > 0$ *the solution point* x_* *is an asymptotically stable equilibrium point of system (43) and the discrete version*

$$x_{k+1} = x_k - h_k \Lambda^{-1}(x_k) D(\alpha) D(\theta(x_k)) L_x(x_k, u_k), \qquad u_k = u(x_k) \tag{47}$$

locally converges with at least linear rate to the point x_* *if the stepsize* h_k *is fixed and* $h_k < \, < \dfrac{2}{\max\limits_{1 \leq i \leq n} \alpha_i}$. *If the matrix* $\Lambda(x)$ *satisfies a Lipschitz condition in a neighborhood of* x_*, *if* $h_k = 1$ *and* $\alpha = e$, *then the sequence* $\{x_k\}$ *converges quadratically to* x_*.

Proof. The right-hand side of system (43) is a differentiable mapping at $x = x_*$. Therefore, we have

$$D(\theta(x))L_x(x, u(x)) = D(\theta(x_*))L_x(x_*, u(x_*)) + \Lambda(x_*)\delta x + \Phi(\delta x),$$

where $\delta x = x - x_*$, $||\Phi(\delta x)|| = O(||\delta x||^2)$.

Linearizing system (43) at the point x_*, we obtain

$$\delta \dot{x} = -\hat{Q}(x_*)\delta x, \qquad \hat{Q}(x_*) = \Lambda^{-1} D(\alpha)\Lambda(x_*).$$

The stability of this system is determined by the properties of the roots of the characteristic equation

$$\det(\hat{Q}(x_*) - \lambda I_n) = 0.$$

Matrix $\hat{Q}(x_*)$ is similar to matrix $D(\alpha)$; therefore, they have the same eigenvalues $\lambda_i = \, = \alpha^i > 0$, $1 \leq i \leq n$. According to Lyapunov linearization principle the equilibrium point x_* is asymptotically stable. The proof of the other statements of this theorem is nearly identical to the proof of Theorem 3.2 and to the proof of convergence of Newton's method. $\square$

System (43) has the first integral

$$D(\theta(x(t, x_0)))L_x(x(t, x_0), u(x(t, x_0))) = D(e^{-\alpha t})D(\theta(x_0))L_x(x_0, u_0),$$

where $u_0 = u(x_0)$, $D(e^{-\alpha_i t})$ is a diagonal matrix whose i-th diagonal element is $e^{-\alpha^i t}$. Taking into account (42), we obtain

$$g(x(t, x_0)) = \tau^{-1} g_x(x(t, x_0)) D(e^{-\alpha t}) D(\theta(x_0)) L_x(x_0, u_0).$$

Hence, if the trajectory $x(t, x_0)$ remains in a bounded set, where the CQ holds, then $\|g(x(t, x_0))\| \to 0$ as $t \to +\infty$.

Suppose that Problem (11) is such that $g_x D(\theta) TE = 0_{mn}$. This condition is satisfied, for example, for a linear programming problem. It is easy to prove that in this case

$$g_x(x)\Lambda(x) = \tau g_x(x), \qquad g_x(x) = \tau g_x(x)\Lambda^{-1}(x).$$

Therefore, by differentiating $g(x)$ along the solutions of (43), we have

$$\frac{dg}{dt} = -g_x(x)\Lambda^{-1}(x)D(\alpha)D(\theta(x))L_x(x,u(x)) = -\frac{1}{\tau}g_x(x)D(\alpha)D(\theta(x))L_x(x,u(x)).$$

If $\alpha = e \in \mathbb{R}^n$, then using relation (42) we obtain finally:

$$g(x(t,x_0)) = g(x_0)e^{-t}.$$

We come to the conclusion that the feasible manifold $g(x) = 0_m$ is asymptotically stable. The trajectory initiating at a point $x_0 \in X$ does not leave the feasible set. The method (47) was considered in [20] in the case where $P = \mathbb{R}^n$ and $D(\theta(x)) \equiv I_n$.

6 CONCLUSION

In this paper we have shown that various numerical algorithms can be constructed on the basis of space transformation. We hope that our approach adds new general insight to Karmarkar's algorithm which is so popular in the West and up to now is used only for the linear programming problem. Generalization of this approach, computational aspects, choice of a stepsize and application to linear programming are beyond the scope of the present paper. We aim to publish all these results in English in a near future.

REFERENCES

1. E. Barnes, (1986). A variation on Karmarkar's algorithm for solving linear programming problems, *Math. Programming*, **36**, 174–182.

2. D.A. Bayer, J.C. Lagarias, (1989). The nonlinear geometry of linear programming. Affine and projective scaling trajectories, *Trans. Amer. Math. Soc.*, **314**(2), 499–526.

3. I. Dikin, (1967). Iterative solution of problems of linear and quadratic programming, *Sov. Math. Dokl.*, **8**, 674–675.

4. Yu. Evtushenko, (1974). Two numerical methods of solving nonlinear programming problems, *Sov. Math. Dokl.*, **15**(2), 420–423.

5. Yu. Evtushenko, (1985). Numerical Optimization Techniques, *Optimization Software*, Inc. Publications Division, New York.

6. Yu. Evtushenko, V. Zhadan, (1973). Numerical methods for solving some operations research problems, *U.S.S.R. Comput. Maths. Math. Phys.*, **13**(3), 56–77.

7. Yu. Evtushenko, V. Zhadan, (1978). A relaxation method for solving problems of non-linear programming, *U.S.S.R. Comput. Maths. Math. Phys.*, **17**(4), 73–87.

8. Yu. Evtushenko, V. Zhadan, (1989). New approaches in optimization techniques, in: *Lecture Notes in Control and Information Sciences*, 143, System Modelling and Optimization, Proc. of the 14th IFIP Conference, ed. H.-J. Sebastian and K. Tammer, Leipzig: Springer-Verlag, 23–37.

9. Yu. Evtushenko, V. Zhadan, (1991). Barrier-projective and barrier-Newton numerical methods in optimization (the nonlinear programming case), *Reports an Comput. Math.*, (in Russian), Computing Center of the USSR Academy of Sciences.

10. Yu. Evtushenko, V. Zhadan, (1992). The space transformation techniques in mathematical programming, in: *Lecture Notes in Control and Information Sciences*, 180, System modelling and Optimization, Proc. of the 15th IFIP Conference, P. Kall (ed.), Zurich: Springer-Verlag, 292–300.

11. Yu. Evtushenko, V. Zhadan, (1992). Barrier-projective and barrier-Newton numerical methods in optimization (the linear programming case). *Reports an Comput. Math.*, (in Russian), Computing Center of the USSR Academy of Sciences.

12. S. Herzel, M.C. Recchioni, F. Zirilli, (1991). A quadratically convergent method for linear programming, *Linear Algebra and its Applications*, **152**, 255–289.

13. M. Kallio, (1986). *On gradient projection for linear programming*, Working paper 94, Yale School of Organization and Management.

14. N. Karmarkar, (1984). *A new polynomial-time algorithm for linear programming*, Combinatorica, **4**, 373–395.

15. D. Luenberger, (1973). *Introduction to linear and nonlinear programming*, Addison-Wesley Publishing Company, Reading.

16. J. Rosen, (1960). The gradient projection method for nonlinear programming, part 1, linear constraints, *SIAM J. Applied Math.*, **8**, 181–217.

17. G. Smirnov, Convergence of barrier-projection methods of optimization via vector Lyapunov functions, *in this Journal.*

18. K. Tanabe, (1980). A geometric method in nonlinear programming, *Journal of Optimization Theory and Applications*, **30**(2), 181–210.

19. Wei Zi-luan, (1987). An interior point method for linear programming, *Journal of Computing Mathematics*, 342–350.

20. V. Zhadan, (1980). On two classes of methods of solving nonlinear programming problems, *Sov. Math. Dokl.*, **22**(2), 388–392.