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

Yu.G. EVTUSHENKO, V. G. ZHADAN
Computing Center
40 Vavilov Str.
117967 Moscow GSP-1
Russia

Abstract: A space transformation technique is used for the reduction of constrained minimization problems to minimization problems without inequality constraints. The continuous and discrete versions of stable barrier-projection method and Newton's method are applied for solving such reduced LP and NLP problems. The space transformation modifies these methods and introduces additional matrices which play the role of a multiplicative barrier, preventing the trajectories from crossing the boundary of the feasible set. The proposed algorithms are based on the numerical integration of systems of ordinary differential equations. These algorithms do not require feasibility of starting and current points, but they preserve feasibility. Some results about convergence rate analysis for continuous and discrete versions of the methods are presented. We describe primal barrier-projection methods, primal barrier-Newton methods and primal-dual barrier-Newton methods. For LP we develop dual barrier-projection and barrier-Newton methods.

1. Introduction

The purpose of this paper is to indicate the advantages and capabilities of the space transformation techniques for solving linear programming (LP) and nonlinear programming (NLP) problems. Here we give a survey of the most important results and numerical methods which were obtained and published mainly in Russian. We describe the primal barrier-projection and barrier-Newton methods (sections 2, 3), primal-dual barrier-Newton method (section 5). For linear programming we developed dual barrier-projection and barrier-Newton methods (section 4).

The analysis of the methods is made on the basis of the qualitative theory of dynamical systems [17, 18] and the stability theory of the solutions of ordinary differential equations [3, 8]. Numerical algorithms are obtained as discretization of dynamical systems. For local convergence analysis we use the Lyapunov linearization principle of determining the stability from the equation of the first approximation about an equilibrium state which is valid for continuous and discrete systems. Nonlocal convergence is investigated by using the second (direct) method of Lyapunov. This theory is quite general, and, although already published in 1892, it remains one of the most powerful technique for detecting stability and convergence [3, 8, 10].

Starting from 1973, we have developed a family of numerical methods based on the space transformation techniques [7, 9, 11, 13, 14]. Such an approach enabled us to reduce the original

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

NLP problem to a problem without inequality 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 and stated as initial-value problems.

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 term "barrier" has often been misunderstood. We do not utilize the barriers which are described, for example, in the book [19] written by A. Fiacco and G. McCormick. Numerical methods which are based on penalty functions are inherently unstable, since one has to increase the penalty parameter without bound in order to obtain convergence. In contrast, the unconventional multiplicative barriers which we use do not tend to infinity when a current point approaches the boundary of a feasible set. In our algorithms the barrier functions are continuous and equal to zero on a boundary. These barriers imply the feasibility of the trajectories. Therefore, we need not introduce any penalty coefficients.

In the second section we consider a family of barrier-projection methods. In the linear programming case after simplifications and choosing an exponential space-transformation function we obtain Dikin's algorithm [6], sometimes called the "variation on Karmarkar's algorithm". The analysis of this algorithm was given in numerous papers (see, for example, [2, 4, 21, 23, 24, 34, 35]). However, there are four main differences with our approaches:

1. We considered LP and NLP problems.

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.

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

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

The primal-dual barrier-Newton method is obtained after applying Newton's method to solving the nonlinear system of equations that is derived from the generalized Kuhn-Tucker stationary condition. We present some results about convergence rate analysis for continuous and discrete versions and show that superlinear and quadratic convergence can be attained. In a particular case our algorithm bears a resemblance to the algorithm of [38], which was developed for LP. But there are also significant differences:

1. Our algorithms can start the computation from the infeasible region, although they preserve feasibility.

2. Our algorithms enable us to take different stepsizes in the primal space and in the dual space, which is proved to be important for computation.

3. We do not use any penalty or usual barrier functions.

4. There exists a set of initial points such that the method solves a linear programming problem in a finite number of iterates.

The numerical algorithms which we propose for linear and nonlinear programming problems have 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 problems.

2. Space Transformation Technique and Primal Barrier-Projection Methods

Consider the following NLP problem:

$$\text{minimize } f(x) \text{ subject to } x \in X = \{x \in \mathbb{R}^n : g(x) = 0_m, \quad x \in P\}, \tag{1}$$

where the functions f and g are twice continuously differentiable, $f(x)$ maps $\mathbb{R}^n$ onto $\mathbb{R}^1$ and $g(x)$ maps $\mathbb{R}^n$ onto $\mathbb{R}^m$, P is a convex set with nonempty interior, 0_s is the s-dimensional null vector, 0_{sk} is the $s \times k$ rectangular null matrix.

We assume differentiability whenever it is helpful to do so. Subscripts will be used to distinguish values of quantities at a particular iteration and superscripts will indicate components of vectors.

Primal and dual linear programming problems are stated in the standard form:

$$\text{minimize } c^T x \text{ subject to } x \in X = \{x \in \mathbb{R}^n : g(x) = b - Ax = 0_m, \quad x \geq 0_n\}, \tag{2}$$

$$\text{maximize } b^T u \text{ subject to } u \in U = \{u \in \mathbb{R}^m : v = L_x(x,u) = c - A^T u \geq 0_n\}, \tag{3}$$

where c is a n-vector, b is a m-vector, A is a full rank $m \times n$ matrix, and $v \in \mathbb{R}^n$ is the vector of dual slack variables.

Define

$$\mathbb{R}^n_{++} = \{x \in \mathbb{R}^n : x > 0_n\}, \quad \text{int}U = \left\{u \in \mathbb{R}^m : v = c - A^T u > 0_n\right\},$$

$$\text{ri}X = \left\{x \in \mathbb{R}^n_{++} : Ax = b\right\},$$

We assume that ri$X$ and int$U$ are nonempty and that primal and dual nondegeneracy holds. In this case both problems have unique solutions x_* and u_*, respectively.

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

$$\text{minimize } \tilde{f}(y) = f(\xi(y)) \text{ subject to } y \in Y, \tag{4}$$

where $Y = \{y \in \mathbb{R}^n : \tilde{g}(y) = g(\xi(y)) = 0_m\}$. The Lagrangian associated with Problem (4) is defined by $\tilde{L}(y, \tilde{u}) = \tilde{f}(y) + \tilde{u}^T \tilde{g}(y)$.

We can use the stable version of the gradient projection method for solving Problem (4). The method is stated as an initial-value problem involving the following system of ordinary differential equations

$$\frac{dy}{dt} = -\tilde{L}_y(y, \tilde{u}(y)), \quad \tilde{L}_y(y, \tilde{u}) = \tilde{f}_y(y) + \tilde{g}_y^T(y)\tilde{u}. \tag{5}$$

The function $\tilde{u}(y)$ is chosen to satisfy the following condition [33]

$$\frac{d\tilde{g}}{dt} = \tilde{g}_y \frac{dy}{dt} = -\tau \tilde{g}(y), \quad \tau > 0. \tag{6}$$

From this condition we obtain the system of linear algebraic equations

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

where $\tilde{f}_y = \tilde{J}^T f_x$, $\tilde{g}_y = g_x\tilde{J}$, $\tilde{J} = dx/dy$ and $\tilde{J}$ is the Jacobian matrix of the transformation $x = \xi(y)$ with respect to y. If $\tilde{J}$ is a 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 (5), we have

$$\frac{dx}{dt} = \frac{d\xi}{dy}\frac{dy}{dt} = J(x)\frac{dy}{dt} = -G(x)L_x(x, u(x)), \quad x(0, x_0) = x_0 \in P, \tag{7}$$

$$\Gamma(x)u(x) + g_x(x)G(x)f_x(x) = \tau g(x), \tag{8}$$

where we have introduced the two Gram matrices:

$$\Gamma(x) = g_x(x)G(x)g_x^T(x), \quad G(x) = J(x)J^T(x).$$

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 $x \in P$:

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

the conjugate cone to the cone F:

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

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

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

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

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 1 *Let the space transformation $\xi(y)$ satisfy C_1, and let the CQ for Problem (1) hold at a point $x \in P$. Then the Gram matrix $\Gamma(x)$ is invertible and positive definite.*

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

If at a point x the conditions of Lemma 1 hold, then we can find from (8) the function $u(x)$, substitute it into the right-hand side of (7) and write (7) in the following projective form

$$\frac{dx}{dt} = -J(x)\left\{\pi[g_x(x)J(x)]J^T(x)f_x(x) + \tau[g_x(x)J(x)]^+g(x)\right\}. \tag{9}$$

Let $x(t, x_0)$ denote the solution of the Cauchy Problem (7) with initial condition $x_0 = x(0, x_0)$. For the following, we assume that the initial-value problem under consideration is always uniquely solvable. By differentiating $f(x(t, x_0))$ with respect to t we obtain

$$\frac{df}{dt} = -\|J^T(x)L_x(x, u(x))\|^2 + \tau u^T(x)g(x). \tag{10}$$

Hence 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 $\|g(x(t, x_0))\|$ is sufficiently small.

The system of ordinary differential equations (7), where $u(x)$ is given by (8), has the first integral

$$g(x(t, x_0)) = g(x_0)e^{-\tau t}. \tag{11}$$

This means that if $\tau > 0$, then method (7) has a remarkable property: all its trajectories approach the manifold $g(x) = 0_m$ as t tends to infinity and this manifold is an asymptotically stable attractor of the system (see [8, 15, 33]). If $x_0 \in X$, then the trajectory $x(t, x_0)$ of (7) remains in the feasible set X because $g(x(t, x_0)) \equiv 0_m$ and condition C_1 implies $x(t, x_0) \in P$. Therefore, if $x_0 \in X$ or $\tau = 0$, then the trajectories of (9) coincide with the trajectories of the following system

$$\frac{dx}{dt} = -J(x)\pi[g_x(x)J(x)]J^T(x)f_x(x). \tag{12}$$

If the condition $x \in P$ is absent in Problem (1), then the space transformation is trivial: $x = y$ and, taking $J(x) = I_n$ in (12), we obtain the well-known gradient projection method [29].

Definition 2 *The pair $[x_*, u_*]$, where $x_* \in P$, is a Kuhn-Tucker pair of Problem (1), if*

$$J^T(x_*)L_x(x_*, u_*) = 0_n, \quad g(x_*) = 0_m. \tag{13}$$

Lemma 2 *Let conditions of Lemma 1 be satisfied at a point $x_* \in P$. Then x_* is an equilibrium point of system (7) if and only if the pair $[x_*, u_*]$, where $u_* = u(x_*)$, satisfies (13).*

Definition 3 *The strict complementary condition (SCC) holds at a point $[x_*, u_*] \in P \times \mathbb{R}^m$, if*

$$L_x(x_*, u_*) \in \ ri\ F^*(x_* \mid P), \tag{14}$$

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 (1).

Theorem 1 *Assume that f and g are twice-differentiable functions and the space transformation $\xi(y)$ satisfies C_1. Sufficient conditions that a point $x_* \in P$ be an isolated local minimum of Problem (1) are that there exists a Lagrange multiplier vector u_* satisfying the Kuhn-Tucker conditions (13), that the SCC holds at $[x_*, u_*]$ and that $z^TJ^T(x_*)L_{xx}(x_*, u_*)J(x_*)z > 0$ 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 (1), 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 $\xi(y)$

$$x^i = \xi^i(y^i), \quad 1 \le i \le n. \tag{15}$$

For such a transformation the corresponding Jacobian matrix is diagonal and

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

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

$$J(x) = D(\dot{\xi}(y))\mid_{y=\delta(x)}, \quad G(x) = J^2(x) = D(\theta(x))$$

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

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

C$_2$. *The vector function $\theta(x)$ is defined at each point $x \in \mathbb{R}^n_+$, and $\theta^i(x^i) = 0$ if and only if $x^i = 0$, where $1 \le i \le n$.*

C$_3$. *The vector function $\theta(x)$ is differentiable in some neighborhood of $\mathbb{R}^n_+$ and $\dot{\theta}^i(0) > 0$, $1 \le\ i \le\ n$.*

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

$$\begin{array}{rclrclrcl} x^i & = & \xi^i(y^i) = \frac{1}{4}(y^i)^2, & J(x) & = & D^{1/2}(x), & G(x) & = & D(x), \\ x^i & = & \xi^i(y^i) = e^{y^i}, & J(x) & = & D(x), & G(x) & = & D^2(x). \end{array}$$

In these two cases the Jacobian matrix is singular on the boundary of the set $P = \mathbb{R}^n_+$. These transformations satisfy C_1, C_2. Condition C_3 holds only for transformation (). 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 (7) interesting particular cases are derived. In order to distill the essence of multiplicative barriers we consider the simplest case where $X = P = \mathbb{R}^n_+$. If we use the space transformation functions () and (), then we obtain

$$\frac{dx}{dt} = -D^\alpha(x) f_x(x), \tag{16}$$

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

Let $x^i(t, x_0) = 0$. Then we have $\theta^i(x^i) = 0$ and $dx^i(t, x_0)/dt = 0$ in (16). From the last equality it follows that the trajectory $x(t, x_0)$ of system (16) cannot cross the boundary $x^i = 0$. Thus the Gram matrix $G(x) = D^\alpha(x)$ plays the role of a "barrier", preventing the trajectory $x(t, x_0)$ from passing through the boundary of P. Therefore, we call (7) a "primal barrier-projection method".

In our first publication in this field [9] we used the quadratic space transformation () and considered the resource allocation problem where $X = \{x \in \mathbb{R}^n_+ : \quad \sum_{i=1}^n x^i = 1\}, x_0 \in X$. In this case $m = 1$ and system (8) was solved explicitly, while system (9) had the form

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

where $e \in \mathbb{R}^n$ is the vector of ones, $d(x) = x^T f_x(x)$. More general cases were considered in subsequent papers [7, 11, 13].

If $P = \mathbb{R}^n_+$, then we can simplify definitions 1, 3 and the statement of Theorem 3. Let e^i denote the n-th order unit vector whose i-th component is equal to one and let

$$K(x) = \{\bar{x} \in \mathbb{R}^n : g_x(x) D(\theta(x)) \bar{x} = 0_m\}.$$

Definition 4 *The constraint qualification (CQ) for Problem (1), where $P = \mathbb{R}^n_+$, holds at a point x, if all vectors $g^i(x)$, $1 \leq i \leq m$, and all e^j, such that $x^j = 0$, are linearly independent. We say that x is a regular point for Problem (1), if the CQ holds at x. The strict complementary condition (SCC) holds at a point $[x, u]$, if $L_{x^i}(x, u) > 0$ for all i such that $x^i = 0$.*

Theorem 2 *Assume that f and g are twice-differentiable functions and the space transformation $\xi(y)$ satisfies C_2. Sufficient conditions for a point $x_* \in P = R[n]_+$ to be an isolated local minimum of Problem (1) are that there exists a Lagrange multiplier vector u_* such that $[x_*, u_*]$ satisfies the Kuhn-Tucker conditions (13), that the SCC holds and that $z^T D(\theta(x_*)) L_{xx}(x_*, u_*) D(\theta(x_*)) z > 0$ for all $z \in K(x_*)$, satisfying $D(\theta(x_*)) z \neq 0_n$.*

Applying the Euler method for solving system (9), we obtain

$$x_{k+1} = x_k - h_k J(x_k) \left\{ \pi[g_x(x_k) J(x_k)] J^T(x_k) f_x(x_k) + \tau [g_x(x_k) J(x_k)]^+ g(x_k) \right\}, \qquad (17)$$

where a stepsize $h_k > 0$.

Theorem 3 *Let $[x_*, u_*]$ be a Kuhn-Tucker pair of Problem (1), where the CQ and the second-order sufficiency conditions of Theorem 2 hold. Let the space transformation $\xi(y)$ satisfy conditions C_2, C_3 and let $\tau > 0$. Then x_* is an asymptotically stable equilibrium state of system (9); there exists a number h_* such that for any fixed $0 < h_k < h_*$ the sequence $\{x_k\}$, generated by (17), converges locally with a linear rate to x_* while the corresponding sequence $\{u_k\}$ converges to u_*.*

The proof of this theorem is given in [13, 16]. It is based on Lyapunov linearization principle. The conditions for asymptotic stability are expressed by mean of the characteristic multipliers of the matrix of the equations of the first approximation about the equilibrium point x_*. The asymptotic stability of a point x_* implies the local convergence of trajectory $x(t, x_0)$ to point x_*. The corresponding statement about the convergence of discrete variants follows from Theorem 2.3.7 (see [8]).

Applying method (7) for solving Problem (2), we obtain the following continuous and discrete versions

$$\frac{dx}{dt} = -G(x)[c - A^T u(x)], \quad x(0, x_0) = x_0 > 0_n, \qquad (18)$$

$$x_{k+1} = x_k - h_k G(x_k)[c - A^T u(x_k)], \quad x_0 > 0_n, \qquad (19)$$

where the function $u(x)$ is found from linear equation (8) which can be rewritten as follows

$$AG(x)A^T u(x) - AG(x)c = \tau(b - Ax). \qquad (20)$$

By differentiating the objective function with respect to t, we obtain

$$c^T \frac{dx}{dt} = -\|J(x)(c - A^T u(x))\|^2 + \tau u^T(x)(b - Ax).$$

The system of ordinary differential equations (18) has the first integral

$$Ax(t,x_0) = b + (Ax_0 - b)e^{-\tau t}. \tag{21}$$

From (21) it follows that, if $Ax_0 = b$, then $Ax(t,x_0) \equiv b$ for all $t \geq 0$ and the trajectory $x(t,x_0)$ of (18) remains in the feasible set X, the objective function monotonically decreases along the trajectories.

Theorem 4 *Let x_*, u_* be unique solutions of Problems (2) and (3), respectively. Let the space transformation $\xi(y)$ satisfy the conditions C_2, C_3 and let $\tau > 0$. Then x_* is asymptotically stable equilibrium state of system (18). There exists $h_* > 0$ such that for any fixed $0 < h_k < h_*$ the sequence $\{x_k\}$, generated by (19), converges locally with a linear rate to x_* while the corresponding sequence $\{u_k\}$ converges to u_*.*

If we use the quadratic and exponential space transformations (), (), then from (18), (20) we obtain, respectively,

$$\begin{array}{rclcrcl} \frac{dx}{dt} & = & D(x)(A^Tu(x) - c), & \quad & AD(x)A^Tu(x) & = & AD(x)c + \tau(b - Ax), \\ \frac{dx}{dt} & = & D^2(x)(A^Tu(x) - c), & \quad & AD^2(x)A^Tu(x) & = & AD^2(x)c + \tau(b - Ax). \end{array}$$

If $\tau > 0$, then method () converges locally to x_*, i.e. for any $\epsilon > 0$ there is $\delta > 0$ such that $\|x_0 - x_*\| < \delta$ implies

$$\lim_{t\to\infty} \|x(t,x_0) - x_*\| = 0. \tag{22}$$

Method () converges locally on $\mathbb{R}^n_{++}$, i.e., if $\|x_0 - x_*\| < \delta$ and $x_0 \in \mathbb{R}^n_{++}$, then (22) holds. Theorem 4 cannot be used in this case because condition C_3 is not satisfied for exponential space transformation.

Denote $\mu = \max_{1\leq i\leq n} v^i_*$. In [14], using Lyapunov's linearization principle, we proved that the discrete version of () converges linearly if the stepsize h_k is fixed and is such that $h_k < h_* = 2\min[1/\tau, 1/\mu]$.

If we set $\tau = 0$, then (), () yield

$$\begin{array}{rclcrcl} \frac{dx}{dt} & = & D(x)(A^Tu(x) - c), & \quad & AD(x)A^Tu(x) & = & AD(x)c, \\ \frac{dx}{dt} & = & D^2(x)(A^Tu(x) - c), & \quad & AD^2(x)A^Tu(x) & = & AD^2(x)c. \end{array}$$

Method () converges locally on the set $Ax = b$, i.e., if x_0 belongs to this set and $\|x_0 - x_*\| < \delta$, then (22) takes place. Method () was proposed in 1978 (see [11, 8]). Nonlocal convergence analysis of this method was made in [14] on the basis of the second Lyapunov method. System () was also considered in [21].

The discrete and continuous versions of method () were investigated in numerous papers (see, for example, [2, 4, 6, 21, 23, 34, 35]). In [2] the discrete version was called "a variation on Karmarkar's algorithm". We should remark that method () does not possess the local convergence property. Here the convergence takes place only if x_0 belongs to the relative interior of X. This result was proved by G.Smirnov on the basis of the vector Lyapunov function. He investigated method () and proved that for the vector of nonbasic components x_N the following estimate holds: $\|x_N(t,x_0)\| \approx O(t^{-1})$ as $t \to \infty$. If we use method (), then $\|x_N(t,x_0)\| \approx O(e^{-\lambda_* t})$, where $\lambda_* > 0$. Hence the trajectories of system () with the quadratic transformation converge locally faster than the trajectories of system () with the exponential transformation. Therefore, in our papers and codes we used mainly the quadratic space transformation ().

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

$$x^i = [a^i + b^i + (b^i - a^i)\sin y^i]/2, \quad G(x) = D(x-a)D(b-x)$$

and method (7) is written as follows

$$\frac{dx}{dt} = -D(x-a)D(b-x)L_x(x, u(x)), \quad x(0, x_0) = x_0 \in P,$$

$$g_x(x)D(x-a)D(b-x)\left[f_x(x) + g_x^T u(x)\right] = \tau g(x).$$

The statements of Theorems 3 and 4 are generalized for this case.

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

$$\text{minimize } f(x) \text{ subject to } x \in X = \{x \in \mathbb{R}^n : g(x) = 0_m, \quad h(x) \leq 0_c\}, \tag{23}$$

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

In Problem (23) 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 $p \in \mathbb{R}^c$, define $q = m + c$, combine primal, dual variables and all constraints:

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

Then the original Problem (23) is transformed into the equivalent Problem

$$\text{minimize } f(x) \text{ subject to } z \in Z = \{z \in \mathbb{R}^{n+c} : \Phi(z) = 0_q, \quad p \in \mathbb{R}^c_+\}. \tag{24}$$

This problem is similar to (1). 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:

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

Combining variables and constraints for the reduced problem, let us define

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

Problems (23) and (24) can be formulated as follows:

$$\text{minimize } f(x) \text{ subject to } \hat{z} \in \hat{Z} = \{\hat{z} \in \mathbb{R}^{n+c} : \hat{\Phi}(\hat{z}) = 0_q\}. \tag{25}$$

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

$$\frac{dz}{dt} = -\tilde{G}(p)L_z(z, w(z)). \tag{26}$$

Here

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

$$\Phi_z(z)\tilde{G}(p)L_z(z,w(z))=\tau\Phi(z), \tag{27}$$

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

System (26) can be rewritten in the more detailed form

$$\frac{dx}{dt}=-L_x(x,w(z)), \quad \frac{dp}{dt}=-G(p)v, \tag{28}$$

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

$$\Gamma(z)w(z)+\Phi_x(x)f_x(x)=\tau\Phi(z), \quad \Gamma(z)=\Phi_x(x)\Phi_x^T(x)+\left[\begin{array}{cc} 0_{mm} & 0_{mc} \\ 0_{cm} & G(p) \end{array}\right].$$

Condition (6) can be written as

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

Therefore system (28) has two first integrals:

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

$$h(x(t,z_0))+p(t,z_0)=(h(x_0)+p_0)e^{-\tau t}, \quad z_0^T=[x_0^T,p_0^T]. \tag{29}$$

Similarly to (10) we obtain

$$\frac{df}{dt}=-\|L_x\|^2-\|J^T(p)v\|^2+\tau[u^Tg+v^T(h+p)]. \tag{30}$$

Let us introduce the index set $\sigma(p)=\{i\in[1:c]:p^i=0\}$. The set P in Problem (24) has the form $P=\{[x,p]\in\mathbb{R}^{n+c}: \quad p\geq 0_c\}$. Therefore, definition 1 of CQ for this problem can be reformulated as follows.

Definition 5 *The constraint qualification (CQ) for Problem (24) 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 C_1:

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

$$\ker J^T(p)=\text{ lin } F^*(p\mid\mathbb{R}^c_+)=\{b\in\mathbb{R}^c: \text{ if } i\in\sigma(p), \text{ then } b^i=0\}. \tag{31}$$

Lemma 3 *Let the mapping $\varphi(y)$ satisfy C_4 and let the CQ for Problem (24) hold at a point $z\in P$. Then the Gram matrix $\Gamma(z)$ is invertible and positive definite.*

We can use all results given in this section for Problem (24) and for the corresponding numerical method (28). It follows from CQ for Problem (24) that at a point z the vectors $\Phi_z^i(z)$, $1\leq i\leq q$, are linearly independent. Consequently z is a regular point for] Problem (24). The pair $[z_*,w_*]$ is a Kuhn-Tucker pair of Problem (24), if

$$L_x(z_*,w_*)=0_n, \quad \Phi(z_*)=0_q, \quad D(p_*)v_*=0_c. \tag{32}$$

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 *Let conditions of Lemma 3 be satisfied at a point* $z_* \in \mathbb{R}^n \times \mathbb{R}^c_+$. *Then* z_* *is an equilibrium point of system (28) if and only if the pair* $[z_*, w_*]$ *satisfies (32).*

It is convenient to use for Problem (24) a component-wise mapping $p^i = \varphi^i(y^i)$, $1 \leq i \leq c$. Therefore, $\varphi_y = D(\dot{\varphi})$, where $\dot{\varphi}(y)$ is a c-dimensional vector. We denote $\gamma^i(p^i) = \dot{\varphi}^i(\psi^i(p^i))$, $\theta(p) = [\theta^1(p^1), \ldots, \theta^c(p^c)]$, $\theta^i(p^i) = [\gamma^i(p^i)]^2$, $1 \leq i \leq c$. For this mapping the matrix $G(p)$ has the diagonal form $G(p) = D(\theta(p))$. Introduce the following conditions:

$\mathbf{C}_5$. *The vector function* $\theta(p)$ *is defined at each point* $p \in \mathbb{R}^c_+$, *and* $\theta^i(p^i) = 0$ *if and only if* $p^i = 0$, *where* $1 \leq i \leq c$.

$\mathbf{C}_6$. *The vector function* $\theta(p)$ *is differentiable in some neighborhood of* $\mathbb{R}^c_+$ *and* $\dot{\theta}^i(0) > 0$, $1 \leq i \leq c$.

Let us consider the tangent cone

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

Theorem 5 *Let* $[z_*, w_*]$ *be a Kuhn-Tucker pair of Problem (24), 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* $\bar{x}^T L_{xx}(x_*, w_*)\bar{x} > 0$ *holds. Then system (28) with* $\tau > 0$ *is asymptotically stable at the equilibrium point* z_*. *The discrete version of the method converges locally with a linear rate, if the stepsize is fixed and sufficiently small.*

Consider the simplified version of method (28). Suppose that along the trajectories of system (28) 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 (28) the additional vector p and integrate the system which does not employ this vector:

$$\frac{dx}{dt} = -L_x(x, w(x)), \tag{33}$$

where

$$\Gamma(x)w(x) + \Phi_x(x)f_x(x) = \tau \left[\begin{array}{c} g(x) \\ 0_c \end{array} \right], \tag{34}$$

$$\Gamma(x) = \Phi_x(x)\Phi_x^T(x) + \left[\begin{array}{cc} 0_{mm} & 0_{mc} \\ 0_{cm} & G(-h(x)) \end{array} \right].$$

Along the trajectories of (33) we have

$$\frac{dg}{dt} = -\tau g(x), \quad \frac{dh}{dt} = -G(-h(x))v(x), \tag{35}$$

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

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 (35) since $\theta^i(0) = 0$. Hence $x(t, x_0) \in X$ for all $t \geq 0$. Thus the Gram matrix $G(-h(x))$ 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 (33) belongs to the boundary.

Method (33) is closely related to method (7). Let us consider Problem (1), assuming that $P = \mathbb{R}^n_+$. We have two alternatives: we can use methods (7) or (33). 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 (8) of m equations or system (34) 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 (34), we can reduce the computational time by using the Frobenius formula for an inverse matrix [8]. After some transformations we find that formulas (33), (34) can be written as (7), (8), respectively, if in the last formulas we take

$$G(x) = D(\mu(x)), \quad \mu^i(x^i) = \theta^i(x^i)/[1+\theta^i(x^i)], \quad 1 \le i \le n.$$

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

3. Primal Barrier-Newton Method

Let us consider Problem (1) 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 (15).

Equation (8) can be rewritten as

$$g_x(x)D(\theta(x))L_x(x,u(x)) = \tau g(x). \tag{36}$$

Therefore if the space transformation $\xi(y)$ satisfies 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 (1). In the previous section 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

$$\Lambda(x)\frac{dx}{dt} = -D(\alpha)D(\theta(x))L_x(x,u(x)), \quad x(t,x_0) = x_0, \tag{37}$$

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:

$$\Lambda(x) = D(\dot\theta)D(L_x) + D(\theta)L_{xx} + D(\theta)g_x^T\frac{du}{dx}. \tag{38}$$

Here all matrices and vectors are evaluated at a point x and the function $u(x)$ is defined from (36). By differentiating equality (36) with respect to x, we obtain

$$g_x\left[D(\dot\theta)D(L_x) + D(\theta)L_{xx} + D(\theta)g_x^T\frac{du}{dx}\right] + E = \tau g_x, \tag{39}$$

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 (1). From (36) and (39) we find that

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

$$\frac{du}{dx} = \Gamma^{-1}[\tau g_x - g_x D(\dot{\theta}) D(L_x) - g_x D(\theta) L_{xx} - E]. \tag{40}$$

Using the notations $T(x) = g_x^T(x)\Gamma^{-1}(x)$, $\Omega(x) = T(x) g_x(x)$, we obtain after substitution of (40) into (38)

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

Lemma 5 *Let conditions of Theorem 3 be satisfied at the point $[x_*, u_*]$. Then $\Lambda(x_*)$ is a nonsingular matrix.*

Theorem 6 *Let conditions of Theorem 3 be satisfied at the point $[x_*, u_*]$. Then for any scaling vector $\alpha > 0_n$ and any $\tau > 0$ the solution point x_* is an asymptotically stable equilibrium point of system (37) and the discrete version*

$$x_{k+1} = x_k - h_k \Lambda^{-1}(x_k) D(\alpha) D(\theta(x_k)) L_x(x_k, u_k), \quad u_k = u(x_k) \tag{41}$$

locally converges with at least linear rate to the point x_* if the stepsize h_k is fixed and $h_k < 2/\max_{1 \le i \le n} \alpha_i$. 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_*.

System (37) 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 $u_0 = u(x_0)$, $D(e^{-\alpha t})$ is a diagonal matrix whose i-th diagonal element is $e^{-\alpha^i t}$. Taking into account (36), 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 (1) is such that $g_x D(\theta) T E = 0_{mn}$. This condition is satisfied, for example, for a linear programming problem. It is easy to prove that in this case

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

Therefore, by differentiating $g(x)$ along the solutions of (37), 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 (36) 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 (41) was considered in [37] in the case where $P = \mathbb{R}^n$ and $D(\theta(x)) \equiv I_n$.

We apply the primal barrier-Newton method (37) for solving the linear programming Problem (2). In this case we have

$$\Lambda(x) = [I_n - H(x)]D(\dot{\theta}(x))D(c - A^T u(x)) + \tau H(x), \tag{42}$$

$$H(x) = D(\theta(x))A^T(AD(\theta(x))A^T)^{-1}A. \tag{43}$$

As in the previous section the vector-function $u(x)$ is found from the linear equation (20).

Introduce a Lebesque level set in $\mathbb{R}^n$

$$\Omega = \{x \in \mathbb{R}^n_+ : \ \|Ax - b\| \leq \|Ax_0 - b\|, \ 0_n \leq D(\theta(x))(c - A^T u(x)) \leq D(\theta(x_0))v_0\},$$

where x_0 is an initial point in (37), $v_0 = c - A^T u_0$, $u_0 = u(x_0)$.

Theorem 7 *Suppose that the set Ω is compact and contains a unique stationary point x_*. Assume that the space transformation $\xi(y)$ satisfies C_2 and is such that the matrix $\Lambda(x)$ is nonsingular everywhere on Ω. If the starting point x_0 is such that $x_0 > 0_n$, $v_0 > 0_n$, then*

$$\lim_{t\to\infty} x(t, x_0) = x_*, \quad \lim_{t\to\infty} u(x(t, x_0)) = u_*, \tag{44}$$

where x_*, u_* are the solutions of Problem (2) and (3), respectively.

Integrating (37) using the Euler method, we obtain the following iterative process:

$$x_{k+1} = x_k - h\Lambda^{-1}(x_k)D(\theta(x_k))(c - A^T u(x_k)), \tag{45}$$

where $h > 0$ is a stepsize.

Each equilibrium point x_* of system (37) is a fixed point of iterations (45), i.e. $x_k = x_*$ implies $x_{k+1} = x_*$, and if iterates (45) converge to a regular point x_*, then the pair $[x_*, u(x_*)]$ satisfies the Kuhn-Tucker conditions.

If the conditions of Theorem 2 hold and the space transformation function satisfies conditions C_2, C_3, then the matrix $\Lambda(x_*)$ is nonsingular. Therefore, if the stepsize h is fixed and $0 < h < 2$, then the discrete version (45) locally converges to the point x_* with at least linear rate. If matrix $\Lambda(x)$ satisfies the Lipschitz condition in a neighborhood of x_* and $h = 1$, then the sequence $\{x_k\}$ converges quadratically to x_*.

4. Dual Barrier-Projection and Barrier-Newton Methods

For the sake of simplicity we consider in this section only the dual linear programming problem. If we apply method (28) to (3), then we obtain the following dual barrier-projection method

$$\frac{du}{dt} = b - Ax(u, v), \quad \frac{dv}{dt} = -G(v)x(u, v), \tag{46}$$

$$(G(v) + A^T A)x(u, v) = A^T b + \tau(v + A^T u - c),$$

where $u(t, z_0) = u_0$, $v(t, z_0) = v_0$, $z_0^T = [u_0^T, v_0^T]$. A very similar calculation as the one in (29), (30) yields the following formulas

$$v(t, z_0) + A^T u(t, z_0) - c = (v_0 + A^T u_0 - c)e^{-\tau t},$$

$$b^T \frac{du}{dt} = \|b - Ax\|^2 + x^T G(v)x + \tau x^T(c - A^T u - v). \tag{47}$$

For the component-wise space transformation $\varphi(y)$ method (46) can be written in the form

$$\frac{du}{dt} = b - Ax(u,v), \quad \frac{dv}{dt} = -D(\theta(v))x(u,v), \tag{48}$$

$$(D(\theta(v)) + A^T A)x(u,v) = A^T b + \tau(v + A^T u - c).$$

If condition C_5 holds, then from (48) we conclude that each component of the vector $v(t, u_0)$ does not change sign. Hence, if $v_0 = c - A^T u_0 > 0_n$, then on the entire trajectory $v(t, z_0) > 0_n$.

If $v_0 = c - A^T u_0$, then $v(t, z_0) = c - A^T u(t, z_0)$ for all $t \geq 0$ and the system (46) can be simplified

$$\frac{du}{dt} = b - Ax(u), \quad (D(\theta(v)) + A^T A)x(u) = A^T b, \tag{49}$$

where $u(0, u_0) = u_0 \in U$.

Applying the Euler method for numerical integration of systems (48) and (49), we obtain, respectively,

$$u_{k+1} = u_k + h_k(b - Ax_k), \quad v_{k+1} = v_k + h_k D(\theta(v_k))x_k, \tag{50}$$
$$(D(\theta(v_k)) + A^T A)x_k = A^T b + \tau(v_k + A^T u_k - c),$$

$$u_{k+1} = u_k + h_k(b - Ax_k), \quad (D(\theta(v_k)) + A^T A)x_k = A^T b. \tag{51}$$

These variants of dual barrier-projection method solve simultaneously both the dual and primal linear programming problems. The same property has the primal barrier-projection method (19).

Theorem 8 *Let x_*, u_* be unique solutions of nondegenerate Problems (2), (3), respectively, and $v_* = c - A^T u_*$. Let the space transformation $\varphi(y)$ satisfy conditions C_5, C_6 and $\tau > 0$. Then the following statements are true:*

1. *The pair $[u_*, v_*]$ is an asymptotically stable equilibrium state of system (48).*

2. *The solutions $u(t, z_0)$, $v(t, z_0)$ of system (48) locally converge to the pair $[u_*, v_*]$. The corresponding function $x(u(t, z_0),\ v(t, z_0))$ converges to the optimal solution x_* of the primal problem (2).*

3. *There exists $h_* > 0$ such that for any fixed $0 < h_k < h_*$ the sequence $\{u_k, v_k\}$ generated by (50) converges locally with a linear rate to $[u_*, v_*]$ while the corresponding sequence $\{x_k\}$ converges to x_*.*

4. *The point u_* is an asymptotically stable equilibrium state of system (49).*

5. *The solutions $u(t, u_0)$ of system (49) locally converge to the optimal solution u_* of dual problem (3). The corresponding function $x(u(t, u_0))$ converges to the optimal solution x_* of primal problem (2).*

6. *There exists $h_* > 0$ such that for any fixed $0 < h_k < h_*$ the sequence $\{u_k\}$ generated by (51) converges locally with a linear rate to u_* while the corresponding sequence $\{x_k\}$ converges to x_*.*

If we use the space transformation $\varphi(y)$ of the forms () and (), then we can rewrite formulas (49) as follows

$$\frac{du}{dt} = b - Ax(u), \quad (D^\alpha(v) + A^T A)x(u) = A^T b. \tag{52}$$

Here $\alpha = 1$ for the quadratic transformation () and $\alpha = 2$ for the exponential transformation ().

From (47) and (52) we obtain that

$$\frac{dv}{dt} = -D^\alpha(v)x, \quad b^T \frac{du}{dt} = \|Ax - b\|^2 + \|D^{\alpha/2}(v)x\|^2 \geq 0.$$

Hence the objective function $b^T u(t, u_0)$ of the dual problem is a monotonically increasing function of t, all components of vector $v(t) = c - A^T u(t, u_0)$ do not change sign.

Denote $\mu = \max\limits_{1 \leq i \leq n} x_*^i$. If $\alpha = 1$, then for algorithm (50) we have $h_k < h_* = 2\min[1/\tau, 1/\mu]$; for algorithm (51) we obtain $h_k < h_* = 2/\mu$. If $\alpha = 1$, then both algorithms converge linearly.

In order to define the right-hand side of the differential equation (49) we have to solve a system of n linear equations. If $u \in \mathrm{int} U$, then the square positive definite matrix $D^\alpha(v) + A^T A$ is nonsingular, and we can use the Sherman-Morrison-Woodbury formula

$$(D^\alpha(v) + A^T A)^{-1} = D^{-\alpha}(v)\left[I_n - A^T\left(I_m + AD^{-\alpha}(v)A^T\right)^{-1} AD^{-\alpha}(v)\right],$$

where the square matrix $I_m + AD^{-\alpha}(v)A^T$ is also nonsingular. Using this formula, we can transform (49) to equivalent the system

$$\frac{du}{dt} = z(u), \quad (I_m + AD^{-\alpha}(v)A^T)z(u) = b. \tag{53}$$

Here we solve a system of m linear equations. Method (49) was proposed in 1978 (see [11], [8]). If $\alpha = 2$, then method (53) is similar to the dual affine scaling algorithm which was investigated in [1].

We can apply Newton's method for solving the system

$$b - Ax(u) = 0_m,$$

where the function $x(u)$ is defined from the condition

$$(D(\theta(v)) + A^T A)x(u) = A^T b. \tag{54}$$

The continuous version of Newton's method leads to the following system of ordinary differential equations

$$\Lambda(u)\frac{du}{dt} = Ax(u) - b, \quad \Lambda(u) = -A\frac{dx}{du}. \tag{55}$$

By differentiating the relation (54) with respect to u, we obtain

$$-D(\dot\theta(v))D(x)A^T + \left(D(\theta(v)) + A^T A\right)\frac{dx}{du} = 0_{nm}.$$

Therefore,

$$\Lambda(u) = -A\left(D(\theta(v)) + A^T A\right)^{-1} D(\dot\theta(v))D(x(u))A^T. \tag{56}$$

Lemma 6 *Let all conditions of Theorem 8 be fulfilled. Then $\Lambda(u_*)$ is a nonsingular matrix.*

Theorem 9 *Let all conditions of Theorem 8 hold. Then u_* is an asymptotically stable equilibrium point of system (55). If the matrix $\Lambda(u)$ satisfies a Lipschitz condition in a neighborhood of u_*, then the discrete version*

$$u_{k+1} = u_k + h_k \Lambda^{-1}(u_k)(Ax_k - b), \quad x_k = x(u_k), \tag{57}$$

locally converges with at least linear rate to the point u_* if the stepsize h_k is fixed and $0 < h_k < 2$. If the matrix $\Lambda(u)$ satisfies a Lipschitz condition in a neighborhood of u_* and $h_k = 1$, then the sequence $\{u_k\}$ converges quadratically to u_*.

If we use the quadratic space transformation (), then formula (56) is simplified and we can rewrite method (55) as follows

$$\frac{du}{dt} = \left[A\left(D(v) + A^T A\right)^{-1} D(x(u))A^T\right]^{-1} (b - Ax(u)).$$

The discrete version of this method is similar to (57).

5. Primal-Dual Barrier-Newton Methods

For solving Problem (1) it is possible to solve the nonlinear system (13). For simplicity we consider here the case where P is a n-dimensional positive orthant, i.e. $P = \mathbb{R}^n_+$ and the component-wise space transformation (15) is used.

Introduce the additional mapping

$$\phi(v) = \left[\phi^1(v^1), \dots, \phi^n(v^n)\right]$$

and impose on this mapping the following conditions:

C$_7$. *The vector function $\phi(v)$ is defined at each point $v \in \mathbb{R}^n_+$, and $\phi^i(v^i) = 0$ if and only if $v^i = 0$, where $1 \leq i \leq n$.*

C$_8$. *The vector function $\phi(v)$ is differentiable in some neighborhood of $\mathbb{R}^n_+$ and $\dot{\phi}^i(0) > 0$, $1 \leq i \leq n$.*

If the functions $\theta(x)$, $\phi(v)$ satisfy conditions C_2 and C_7, respectively, then the necessary conditions (13) can be rewritten in the form

$$D(\theta(x))\phi(L_x(x,u)) = 0_n, \quad g(x) = 0_m, \quad x \in \mathbb{R}^n_+. \tag{58}$$

For solving this system we use the continuous version of Newton's method. The computation process is described by the system of ordinary differential equations

$$W(x,u)\begin{pmatrix} \dot{x} \\ \dot{u} \end{pmatrix} = -\begin{pmatrix} \alpha D(\theta(x))\phi(L_x(x,u)) \\ \tau g(x) \end{pmatrix}, \tag{59}$$

where $\alpha > 0$, $\tau > 0$, W is a square matrix of order $n+m$

$$W(x,u) = \begin{pmatrix} M & D(\theta(x))D(\dot{\phi})g_x^T \\ g_x & 0_{mm} \end{pmatrix}, \quad M = D(\dot{\theta})D(\phi) + D(\theta)D(\dot{\phi})L_{xx}. \tag{60}$$

By following the trajectories satisfying (59), we can theoretically obtain a solution of the system of nonlinear equations (58). In practice, we build the iterative procedures using a discretization of dynamical systems.

Lemma 7 *Let $[x_*, u_*]$ be a Kuhn-Tucker pair, where the conditions of Theorem 2 are satisfied. Assume that x_* is a regular point for Problem (1) and the functions $\theta(x)$, $\phi(v)$ satisfy conditions C_2, C_3 and C_7, C_8, respectively. Then the matrix $W(x_*, u_*)$ is nonsingular.*

We combine vectors x, u and define

$$z = \begin{pmatrix} x \\ u \end{pmatrix}, \; z_0 = \begin{pmatrix} x_0 \\ u_0 \end{pmatrix}, \; z_* = \begin{pmatrix} x_* \\ u_* \end{pmatrix}, \; R(z) = \begin{pmatrix} D(\theta(x))\phi(L_x) \\ g(x) \end{pmatrix}.$$

Let $x(t, z_0)$, $u(t, z_0)$ denote the solutions of the Cauchy Problem (59) with initial conditions $x_0 = x(0, z_0)$, $u_0 = u(0, z_0)$. Using this notation, we rewrite the system of equations (59) as

$$W(z)\frac{dz}{dt} = -D(\gamma)R(z), \quad z(0, z_0) = z_0, \tag{61}$$

where γ has the first n components equal to α and all other components equal to τ. We denote $\gamma_* = \min[\alpha, \tau]$.

Theorem 10 *Suppose that the conditions of Lemma 7 hold. Then for any $\alpha > 0$, $\tau > 0$ the pair $z_*^T = [x_*^T, u_*^T]$ is an asymptotically stable equilibrium point of system (61). If the stepsize h_k is fixed and $0 < h_k < 2/\gamma_*$, then the discrete version*

$$z_{k+1} = z_k - h_k W^{-1}(z_k)D(\gamma)R(z_k) \tag{62}$$

locally converges to the point z_* with at least linear rate. If $W(z)$ satisfies a Lipschitz condition in a neighborhood of z_* and $h_k = \alpha = \tau = 1$, then the sequence $\{z_k\}$ converges quadratically to z_*.

According to Theorem 10 the convergence of methods (61), (62) is guaranteed only for a very restricted set of starting points. Consider the global behavior of method (61).

Lemma 8 *Let x be a regular point, and let the pair $[x, u]$ be such that $L_x^i(x, u) \neq 0, x^i \neq 0$, for all $1 \leq i \leq n$, and $M(x, u)$ is nonsingular. Then $W(x, u)$ is nonsingular.*

Define the nonnegative Lyapunov function

$$F(x, u) = \|D(\theta(x))\phi(L_x(x, u))\| + \|g(x)\|$$

and introduce two sets:

$$\Omega_0 = \{[x, u] : F(x, u) \leq F(x_0, u_0), \quad x \geq 0_n, \quad L_x(x, u) \geq 0_n\},$$

$$\tilde{\Omega}_0 = \{[x, u] \in \Omega_0 : \quad x > 0_n, \quad L_x(x, u) > 0_n\}.$$

Theorem 11 *Suppose that the set Ω_0 is bounded and contains the unique Kuhn-Tucker pair $[x_*, u_*]$. Suppose also that for any pair $[x, u] \in \tilde{\Omega}_0$ the conditions of Lemma 8 are satisfied. Then all trajectories of (61), starting from a pair $[x_0, u_0] \in \tilde{\Omega}_0$, converge to $[x_*, u_*]$.*

Method (59) is fully applicable to linear programming problems (2) and (3). Formula (60) is simplified in this case:

$$W = \begin{pmatrix} M & -D(\theta)D(\dot{\phi})A^T \\ -A & 0_{mm} \end{pmatrix},$$

where $M = D(\dot{\theta}(x))D(\phi(v))$. If M and the matrix $\Gamma = AM^{-1}D(\theta)D(\dot{\phi})A^T$ are nonsingular, then we can use the Frobenius formula for the inverse matrix:

$$W^{-1} = \begin{pmatrix} M^{-1}[I_n - D(\theta)D(\dot{\phi})A^T\Gamma^{-1}AM^{-1}] & -M^{-1}D(\theta)D(\dot{\phi})A^T\Gamma^{-1} \\ -\Gamma^{-1}AM^{-1} & -\Gamma^{-1} \end{pmatrix}. \tag{63}$$

The following theorem will guarantee the local convergence of algorithm (61), (62) in a neighborhood of the solution.

Theorem 12 *Let $[x_*, u_*]$ be a nondegenerate optimal pair for the linear programs (2) and (3). Assume that the functions $\theta(x)$ and $\phi(v)$ satisfy conditions C_2, C_3 and C_7, C_8, respectively. Then:*

1. *$W(x_*, u_*)$ is nonsingular;*
2. *$[x_*, u_*]$ is an asymptotically stable equilibrium pair of system (61), and trajectories $[x(t, z_0), u(t, z_0)]$ of (61) converge to $[x_*, u_*]$ on $\tilde{\Omega}_0$;*
3. *the discrete version (62) locally converges with at least a linear rate to the pair $[x_*, u_*]$;*
4. *if $W(z)$ satisfies the Lipschitz condition in a neighborhood of z_*, $h_k = \alpha = \tau = 1$, then the sequence $\{x_k, u_k\}$ converges quadratically to $[x_*, u_*]$.*

The most important feature of the proposed method is that, in contrast to numerous path-following methods, there exists a set of initial points such that the method solves Problems (2) and (3) in a finite number of iterates. It would be extremely interesting to determine the family of sets of starting points which ensures that the solution is reached after a prescribed number of iterations. We will show three such sets. For the sake of simplicity we consider here the case where the stepsizes are fixed and

$$h_k = \alpha = \tau = 1, \quad \theta(x) = x, \quad \phi(v) = v. \tag{64}$$

In this case algorithm (16) is rewritten as follows:

$$D(v_k)x_{k+1} - D(x_k)A^T u_{k+1} = -D(x_k)A^T u_k, \quad Ax_{k+1} = b. \tag{65}$$

Let T_k be a set of pairs $[x_0, u_0]$ such that, if $[x_0, u_0] \in T_k$, then algorithm (62) solves Problems (2) and (3) in k iterations.

Introduce the sets of indexes

$$\sigma(x) = \left\{1 \le i \le n : x^i = 0\right\}, \quad \sigma(v) = \left\{1 \le i \le n : v^i = 0\right\}.$$

If all components of x and v are non-zero, then we write $\sigma(x) = \sigma(v) = \emptyset$. We define the following three sets of pairs

$$\Omega_1 = \{[x, u] : x = x_*,\ \sigma(x) \cap \sigma(v) = \emptyset\}, \quad \Omega_2 = \{[x, u] : u = u_*,\ \sigma(x) \cap \sigma(v) = \emptyset\},$$

$$\Omega_3 = \{[x, u] : \sigma(x) = \sigma(x_*), \quad \sigma(x) \cap \sigma(v) = \emptyset\}.$$

Theorem 13 *Assume that all conditions of Theorem 12 hold. Let the transformation functions be linear. If $[x_0, u_0] \in \Omega_1$ or $[x_0, u_0] \in \Omega_2$, then method (62) yields an optimal solution in a single step. If $[x_0, u_0] \in \Omega_3$, then method (62) yields an optimal solution in at most two steps.*

Now we consider a particular case of the method in which the trajectories $x(t, z_0)$ and $v(t, z_0)$ belong to the positive orthant $\mathbb{R}^n_+$. We define

$$V = \left\{ v \in \mathbb{R}^n : \text{exists } u \in \mathbb{R}^m \text{such that } v = c - A^T u \right\}. \tag{66}$$

We introduce the vectors $q \in \mathbb{R}^n$, $p \in \mathbb{R}^n$, whose i-th components are $\theta^i/\dot{\theta}^i$, $\phi^i/\dot{\phi}^i$, respectively, and define $H = D^{1/2}(q)D^{-1/2}(p)$.

We say that a point $z \in \mathbb{R}^{n+m}$, $z^T = [x^T, u^T]$ is interior, if $x \in \mathbb{R}^n_{++}$ and $u \in \text{int}U$. If moreover $x \in \text{ri}X$, then the point z is strictly interior.

If z is an interior point and conditions C_2, C_3, C_7, C_8 are satisfied then H and AH^2A^T are nonsingular and we can use formula (63) for the inverse matrix. In this case method (61) can be written as

$$\frac{dx}{dt} = H\left(\tau(AH)^+(b - Ax) - \alpha\pi(AH)Hp\right), \tag{67}$$

$$\frac{du}{dt} = \left(AH^2A^T\right)^{-1}(\alpha Aq + \tau(b - Ax)).$$

Introduce a new matrix $\Lambda \in \mathbb{R}^{d \times n}$. The columns of Λ^T forms a basis for the null space of A, i.e. $A\Lambda^T = 0_{md}$. Now definition (66) can be written equivalently as $V = \{v \in \mathbb{R}^n : \Lambda(v - c) = 0_d\}$ and from (67) we obtain

$$\frac{dx}{dt} = H\left[\tau(AH)^+(b - Ax) - \alpha\pi(AH)Hp\right], \tag{68}$$

$$\frac{dv}{dt} = -A^T\frac{du}{dt} = H^{-1}(AH)^+\left[\tau(Ax - b) - \alpha Aq\right]. \tag{69}$$

Let $x(t, z_0)$, $u(t, z_0)$, $v(t, z_0)$ denote the solutions of systems (67) and (68), (69) with initial conditions $x(0, z_0) = x_0 > 0_n$, $u(0, z_0) = u_0$, $v(0, z_0) = c - A^Tu_0 = v_0 > 0_n$. In system (68), (69) we used derivatives $\dot{x}$ and $\dot{v}$ instead of $\dot{x}$ and $\dot{u}$ as we did before. We can do this if we are sure that the entire trajectory $v(t, z_0)$ belongs to V, i.e. for all $t \geq 0$ $v(t, z_0) - c$ belongs to the range space of A^T. Differentiating the d-dimensional vector Λv along the solutions of (68) and (69), we obtain $\Lambda\dot{v} = 0_d$. Therefore, if $v_0 \in V$, then the trajectory $v(t, z_0)$ belongs to V for all $t \geq 0$ and we can use system (68), (69) for numerical calculations. The trajectories $x(t, z_0)$, $v(t, z_0)$ of (68), (69) start from any positive point and go to the optimum as $t \to \infty$.

For the sake of simplicity we consider the case where $\phi(v) = v$ and $\theta(x)$ is a homogeneous function of x, i.e. $\theta(\beta x) = \beta^\lambda\theta(x)$, $\lambda = \alpha/\tau$. Then according to Euler's formula we have $x = \lambda q(x)$, $\lambda H^2 = D(x)D^{-1}(v)$. Introduce new vectors $\mu \in \mathbb{R}^m$, $\eta \in \mathbb{R}^n$ and a scalar function Φ:

$$\mu = \left(AD(x)D^{-1}(v)A^T\right)^{-1}b, \quad \eta = D^{-1}(v)A^T\mu, \quad \Phi(x, v) = x^Tv. \tag{70}$$

Let e be the vector of ones in $\mathbb{R}^n$. Now systems (67) and (68), (69) can be written as

$$\frac{dx}{dt} = \tau D(x)[\eta - e], \quad \frac{du}{dt} = \alpha\mu, \quad \frac{dv}{dt} = -\alpha A^T\mu. \tag{71}$$

The condition $\dot{\theta}(0) > 0$ from C_3 is not necessarily fulfilled for homogeneous functions $\theta(x)$, nevertheless the following theorem is valid.

Theorem 14 *Let $[x_*, u_*]$ be a nondegenerate optimal pair for linear programs (2) and (3). Suppose that $D(\theta(x)) = D^\lambda(x)$, $\phi(v) = v$, $\alpha > 0$, $\tau > 0$. Assume that Ω_0 is bounded. Let the starting point z_0 be interior, then the trajectories of (71) are such that:*

1. *the matrix $AD(x(t,z_0))D^{-1}(v(t,z_0))A^T$ is nondegenerate for all $t \geq 0$;*
2. *$z(t,z_0) \in \Omega_0$ and $v(t,z_0) \in V$ for all $t > 0$;*
3. *the objective function $b^T u(t,z_0)$ of the dual problem increases monotonically;*
4. *the pair $[x(t,z_0), u(t,x_0)]$ is bounded and converges to $[x_*, u_*]$ as $t \to \infty$;*
5. *all components of vectors $D^\lambda(x(t,z_0))v(t,z_0)$, $Ax(t,z_0)$ change monotonically and*

$$D^\lambda(x(t,z_0))v(t,z_0) = e^{-\alpha t} D^\lambda(x_0)v_0,$$

$$Ax(t,z_0) - b = e^{-\tau t}(Ax_0 - b).$$

From (71) we have

$$\langle \dot{x}, \dot{v} \rangle = \alpha\tau\mu^T(Ax - b). \tag{72}$$

If $Ax_0 = b$, $x_0 > 0_n$, then $Ax(t,z_0) \equiv b$, and from (72) it follows that the vectors $\dot{x}$ and $\dot{v}$ are orthogonal. The vector $\dot{x}$ belongs to the nullspace of AH and $\dot{v}$ belongs to its orthogonal complement, $x(t,z_0) \in \mathrm{ri}X$ and $v(t,z_0) \in V$ for all $t \geq 0$. Hence the point $z(t,z_0)$ is strictly interior.

If we set in (71) $\alpha = 1$, $\tau = 0$ and $x = e$, then we have

$$\frac{dv}{dt} = -A^T(AD^{-1}(v)A^T)^{-1}b, \quad \frac{du}{dt} = (AD^{-1}(v)A^T)^{-1}b.$$

This method was introduced in [14]. It is similar to (53) and to the affine scaling algorithm which was proposed in [1].

Algorithm (71) has one important disadvantage connected with the necessity to know a starting point such that $v_0 \in \mathrm{int}V$. It is possible to get rid of this restriction if we use the barrier-Newton method in the extended space of variables x, u, v. The simplest version of the method is described by following system of $2n+m$ differential equations

$$\frac{dx}{dt} = \tau G(A^T\zeta - c), \quad \frac{dv}{dt} = \alpha(c - v - A^T\zeta), \quad \frac{du}{dt} = \alpha(\zeta - u), \tag{73}$$

where $\zeta = (AGA^T)^{-1}[b - Ax + AGc]$, $G = D(x)D^{-1}(v)$. The essential difference in the requirements on the initial conditions between systems (71) and (73) is that in the latter case we impose only the simplest restrictions: $x_0 > 0_n$, $v_0 > 0_n$. If $v_0 = c - A^T u_0$, then this system is reduced to (71). More detailed consideration of this approach is given in [14].

If we set in (73) $\alpha = 0$, $\tau = 1$, $Ax = b$ and $v = e$, then we obtain ().

By applying the Euler numerical integration method to system (71) we obtain the simplest discrete version of the method:

$$x_{k+1} = D(x_k)(e + \tau_k(\eta_k - e)), \quad v_{k+1} = D(v_k)(e - \alpha_k\eta_k), \quad u_{k+1} = u_k + \alpha_k\mu_k, \tag{74}$$

where $x_0 > 0_n$, $v_0 > 0_n$, and μ_k and η_k are defined by (70).

As before, if $v_0 \in V$, then $v_k \in V$ for all k and the last formula in (74) can be omitted. The objective function $b^T u_k$ also increases monotonically. From (74) we obtain

$$Ax_{k+1} - b = (1 - \tau_k)(Ax_k - b), \tag{75}$$

$$\Phi_{k+1} = (1 - \tau_k)\Phi_k + (\tau_k - \alpha_k)\mu_k^T Ax_k + \alpha_k \tau_k \mu_k^T (Ax_k - b). \tag{76}$$

The iterates produced by algorithm (74) are well-defined if vectors x_k, v_k are strictly positive for all k. In order to ensure the positiveness of x_{k+1} and v_{k+1} we have to choose the step lengths α_k and τ_k such that

$$e \geq \alpha_k \eta_k, \quad e \geq \tau_k(e - \eta_k).$$

It is now straightforward to verify that non-negativity conditions hold if α_k and τ_k satisfy

$$0 < \alpha_k \leq \alpha_k^* = 1/[\eta_k^*]_+, \quad 0 < \tau_k \leq \tau_k^* = 1/[1 - \eta_*^k]_+,$$

where $[\alpha]_+ = \max[0, \alpha]$, η_k^* and η_*^k are maximal and minimal components of the vector η_k, respectively. We will adopt the convention that, if $\eta_*^k \geq 1$, then $\tau_k^* = +\infty$, and, if $\eta_k^* \leq 0$, then $\alpha_k^* = +\infty$.

If we set $\alpha = \tau = 1$ and substitute $b = Ax$ in (74), then formulas (74) coincide with the primal-dual interior point algorithms proposed in [38], if in the latter we ignore the barrier (perturbation) term. In this algorithm the starting point z_0 must be strictly interior. Algorithm (74) does not require feasibility of starting and current points, but according to (75) it preserves feasibility. Another important advantage of algorithm (74) is that it permits us to take different step lengths in the primal space and in the dual space. This property is very useful, especially at the beginning of the computation when the starting point is far from the solution and only one maximal the stepsize, either α_k^* or τ_k^*, is very large. This advantage disappears as $k \to \infty$ because maximal step lengths tend to one.

We specify three classes of procedures for determining the step lengths:

1. step lengths are fixed and small enough, hence the discrete process (74) is close to a continuous one (71);

2. stepsizes are close to one and therefore the discrete process has properties of Newton's method;

3. stepsizes are chosen from steepest descent conditions or from another auxiliary optimization problem.

The convergence and polynomiality of one algorithm from the first class are investigated by G.Smirnov in [32]. Consider the second class. Define the step lengths by the formulas

$$\alpha_k = (1 - \varrho_k)\alpha_k^*, \quad \tau_k = (1 - \varrho_k)\tau_k^*. \tag{77}$$

Here we multiply the maximal stepsizes τ_k^*, α_k^* in primal and dual space by a safety factor $1 - \varrho_k$, $0 < \varrho_k < 1$. The choice $\varrho_k = 0$ corresponds to the case which allows steps to the boundary of the positive orthant and a loss of strict feasibility.

We consider the simplest step length rules, in which the sequence $\{\varrho_k\}$ is generated by:

$$0 < \varrho_k < 1, \quad \lim_{k \to \infty} \varrho_k = 0, \tag{78}$$

$$\varrho_k = \min[\varsigma \Phi_k, 1 - \delta], \quad 0 < \varsigma, \quad 0 < \delta < 1, \tag{79}$$

$$\varrho_k = \varsigma \Phi_k / [1 + \varsigma \Phi_k], \quad 0 < \varsigma. \tag{80}$$

All these choices guarantee that if $x_k > 0_n$ and $v_k > 0_n$, then the new primal and dual slack variables obtained from formulas (74) will remain strictly positive.

Theorem 15 *Let the conditions of Theorem 12 be satisfied. If the step lengths are determined by any rule (77) – (80), and the starting points are such that* $x_0 > 0_n$*,* $v_0 > 0_n$*, then the sequence* $\{x_k, u_k\}$ *generated by algorithm (74) converges locally and component-wise to the solution* $[x_*, u_*]$ *with at least superlinear convergence rate and*

$$\overline{\lim_{k\to\infty}} \frac{|x^i_{k+1} - x^i_*|}{|x^i_k - x^i_*|} = 0, \quad \overline{\lim_{k\to\infty}} \frac{|u^j_{k+1} - u^j_*|}{|u^j_k - u^j_*|} = 0, \quad 1 \le i \le n, \quad 1 \le j \le m,$$

$$\lim_{k\to\infty} \alpha^*_k = 1, \quad \lim_{k\to\infty} \tau^*_k = 1, \quad \lim_{k\to\infty} \mu_k = 0_m, \quad \lim_{k\to\infty} \eta^*_k = 1, \quad \lim_{k\to\infty} \eta^k_* = 0.$$

The components of η_k *converge to one or to zero. If rules (79) or (80) are used, then* $\{x_k, u_k\}$ *converge to* $[x_*, u_*]$ *quadratically.*

The proof of this theorem is given in [14].

Consider the third approach to stepsize choice. We may view Φ_{k+1} and $\|Ax_{k+1} - b\|$ as functions of stepsizes α_k and τ_k. It is desirable to choose α_k, τ_k in order to minimize Φ_{k+1} and $\|Ax_{k+1} - b\|$. For solving this bicriterial minimization problem we can apply well-known methods which are developed in multiobjective analysis. Combining both objective functions in a unique criterion with the help of Chebyshev's convolution function, we obtain the following auxiliary minimization Problem

$$\min_{0 \le \alpha_k \le \omega\alpha^*_k} \ \min_{0 \le \tau_k \le \omega\tau^*_k} F(x_{k+1}, v_{k+1}),$$

where $F(x, v) = \max\{\|Ax - b\|, \Phi(x, v)\}$, $0 < \omega < 1$.

This variant of the method is proved to be the most efficient.

6. Conclusion

In this paper we have shown that a great variety of numerical algorithms can be constructed on the basis of space transformation techniques. Generalization of this approach, computational aspects, optimal choice of a stepsize and applications to practical problems are beyond of the scope of the present paper. We aim to publish all these results in English in a near future. We hope that our approach adds new general insight to Karmarkar's algorithm which is so popular in the West.

7. Acknowledgements

The authors would like to express their sincere thanks to C.G. Broyden, M.C. Bartholomew-Biggs, E. Spedicato and Z. Huang for very valuable comments and suggestions. We have also received much helpful input from many other people, too numerous to list here, which we nevertheless gratefully acknowledge.

References

[1] I. Adler, N. Karmarkar, M. Resende, and G. Veiga, An implementation of Karmarkar's algorithm for linear programming, *Mathematical programming*, **44**, 297-335 (1989).

[2] E. Barnes, A variation on Karmarkar's algorithm for solving linear programming problems, *Mathematical Programming* **36**, 174-182 (1986).

[3] A. Bacciotti, Local stability of nonlinear control systems, *Series on Advances in Mathematics for Applied Sciences*, **8**, World Scientific Publishing Co. Ptc. Ltd., Singapore (1992).

[4] D. Bayer, J. Lagarias, The nonlinear geometry of linear programming. Affine and projective scaling trajectories, *Trans. Amer. Math. Soc.*, **314**, 499-526 (1989).

[5] D. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York (1982).

[6] I. Dikin, Iterative solution of problems of linear and quadratic programming, *Sov. Math. Dokl.*, **8**, 674-675 (1967).

[7] Yu. Evtushenko, Two numerical methods of solving nonlinear programming problems, *Sov. Math. Dokl.*, **15**, 420-423 (1974).

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

[9] Yu. Evtushenko and V. Zhadan, Numerical methods for solving some operations research problems, *U.S.S.R. Comput. Maths. Math. Phys.*, **13**, 56-77 (1973).

[10] Yu. Evtushenko and V. Zhadan, Application of the method of Lyapunov functions to the study of the convergence of numerical methods, *U.S.S.R. Comput. Maths. Math. Phys.*, **15**, 96-108 (1975).

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

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

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

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

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

[16] Yu. Evtushenko and V. Zhadan, Stable barrier-projection and barrier-Newton methods in nonlinear programming, to appear in *Optimization Methods and Software*, (1993).

[17] L. Faybusovich, Dynamical systems which solve optimization problems with linear constraints, *IMA Journal of Mathematical Control and Information*, **8**, 135-149 (1991).

[18] L. Faybusovich, Hamiltonian structure of dynamical systems which solve linear programming problems, *Physica D* **53**, 217-232 (1991).

[19] A. Fiacco and G. McCormic, *Nonlinear programming: Sequential unconstrained minimization techniques*, John Wiley & Sons, New York (1968).

[20] C. Gonzaga, C. Path following methods for linear programming, *SIAM Review*, **34**, 167-224 (1992).

[21] S. Herzel, M. Recchioni, and F. Zirilli, A quadratically convergent method for linear programming, *Linear Algebra and its Applications*, **152**, 255-289 (1991).

[22] F. Jarre, G. Sonnevend, and J. Stoer, *An implementation of the method of analytic centers*, Lect. Notes Control and Information Sci., **111**, Springer-Verlag, New York (1988).

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

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

[25] M. Kojima, S. Mizuno, and A. Yoshise, *A primal-dual interior point method for linear programming*, in: N. Megiddo (ed.), Progress in Mathematical Programming – Interior Point and Related Methods, Springer-Verlag, Berlin, Chap. 2 (1989).

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

[27] N. Megiddo, *Pathways to the optimal set in linear programming*, in: N. Megiddo (ed.), Progress in Mathematical Programming – Interior Point and Related Methods, Springer-Verlag, Berlin, Chap. 8 (1989).

[28] S. Mehrotra, On the implementation of a (primal-dual) interior point method, *Technical Report 90-03*, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois (1990).

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

[30] K. McShane, C. Monma,and D. Shanno, An implementation of primal-dual interior point method for linear programming, *ORSA J. Comput.*, **1**, 70-89 (1989).

[31] G. Smirnov, Convergence of barrier-projection methods of optimization via vector Lyapunov functions, to appear in: *Optimization Methods and Software.*

[32] G. Smirnov, Complexity of the barrier-Newton method in linear programming, to appear in *Optimization Methods and Software.*

[33] K. Tanabe, A geometric method in nonlinear programming, *Journal of Optimization Theory and Applications*, **30**, 181- 210 (1980).

[34] R. Vanderbei, M. Meketon, and B. Freedman, A modification of Karmarkar's linear programming algorithm, *Algorithmica*, **1**, 395-407 (1986).

[35] Zi-luan Wei, An interior point method for linear programming, *Journal of Computing Mathematics*, Oct., 342-350 (1987).

[36] Y. Ye, R. Tapia, and Y. Zhang, (1991) A superlinearly convergent $O(\sqrt{n}L)$-iteration algorithm for linear programming, *Technical Report TR91-22*, Rice University, Houston, Texas (1991).

[37] V. Zhadan, On two classes of methods for solving nonlinear programming problems, *Sov. Math. Dokl*, **22**, 388-392 (1980).

[38] Y. Zhang, R. Tapia, and J. Dennis, On the superlinear and quadratic convergence of primal-dual interior point linear programming algorithms, *Technical Report TR90-6*, Rice University, Houston, Texas (1990).