WSSIAA 5 (1995) pp. 51–66

DUAL BARRIER-PROJECTION METHODS IN LINEAR PROGRAMMING[1]

Yuri G. Evtushenko and Vitali G. Zhadan

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

Revised 25 March 1995

Abstract

A surjective space transformation technique is used to convert an original dual linear programming problem with equality and inequality constraints into a problem involving only equality constraints. Continuous and discrete versions of the stable gradient projection method are applied to the reduced problem. The numerical methods involve performing inverse transformations. The convergence rate analysis for dual linear programming methods is presented. By choosing a particular exponential space-transformation function we obtain the dual affine scaling algorithm. Variants of methods which have linear local convergence are given.

1 INTRODUCTION

Since 1973, we have developed a family of numerical methods based on space transformation techniques. Using a space transformation, we convert the original problem with equality and inequality constraints to a problem with equality constraints only. This is an old notion commonly used in the optimization literature. Numerous variants of this basic idea exist. In [4] – [12] we used a surjective space transformation and then applied the gradient projection method and Newton's method to solve the reduced nonlinear programming problem. After an inverse transformation to the original space a family of numerical methods for solving optimization problems with equality and inequality constraints was obtained. The proposed algorithms are based on the numerical integration of systems of ordinary differential equations. As a result of a 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. In our algorithms the barrier functions are continuous and equal zero on a boundary. The space transformations are carried out without using conventional barrier or penalty functions and this feature provides a high rate of convergence.

Different numerical methods are obtained by different choices of the space transformations. For example, if we choose an exponential space transformation in the linear programming case, we obtain the Dikin's algorithm [3] from the family of primal barrier-projection methods. This algorithm, however, does not posses local convergence properties and it converges only for starting points inside the feasible set. Furthermore, the discrete version has a less than linear rate of convergence. In [8]-[10] it was shown that if we apply stable versions of the gradient projection algorithm and use the quadratic space transformation, then we obtain local linear convergence. A survey of our results in this field is given in [11] and applications to linear programming are presented in [12].

[1]Research supported by the grant N 94-01-01379 from Russian Scientific fund.

The content of this paper is similar to that of [12], but in contrast to [12], we focus our attention on the solution of the dual linear programming problem, though the methods which we propose permit us to obtain the solution of the primal problem simultaneously with the solution of the dual problem.

For the local convergence analysis we use the Lyapunov linearization principle of determining the stability from the equation of the first approximation about an equilibrium state [2]. Non-local convergence is investigated by using the second (direct) method of Lyapunov.

In Section 2, we describe a family of dual barrier-projection methods. These methods are described by systems of ordinary differential equations. Numerical algorithms are obtained as discretizations of dynamical systems. Sufficient conditions for local convergence of continuous and discrete versions of numerical methods are given. We show that, if the quadratic space transformation is used then we obtain an exponential rate of convergence for continuous methods and a linear rate of convergence for discrete versions.

In Section 3, we use a non-conventional representation of the dual linear programming problem and we propose a different set of algorithms. After some simplification and after choosing a particular exponential space transformation function we obtain the dual affine scaling method proposed by I. Adler, N. Karmarkar, M. Resende and G. Veiga [1]. In Section 4, we investigate non-local convergence properties.

2 BASIC APPROACH AND OUTLINE OF THE METHODS

Consider a linear programming problem in standard form:

$$\text{minimize } c^\top x \text{ subject to } x \in X = \{x \in \mathbb{R}^n : \ b - Ax = 0_m, \ x \geq 0_n\} \tag{2.1}$$

and its dual problem

$$\text{maximize } b^\top u \text{ subject to } u \in U = \{u \in \mathbb{R}^m : \ v = c - A^\top u \geq 0_n\}, \tag{2.2}$$

where $A \in \mathbb{R}^{m\times n}$, $(m < n)$; $c, x, v \in \mathbb{R}^n$; $b, u \in \mathbb{R}^m$ and $\text{rank}(A) = m$; 0_s is the $s-$dimensional null vector.

We define the interior set of U as:

$$U_0 = \{u \in \mathbb{R}^m : \ v = c - A^\top u > 0_n\},$$

and assume that this set is nonempty.

Throughout the paper we assume that the problems have nonempty solution sets X_* and U_*, respectively. We also introduce the following sets:

$$\begin{aligned} V &= \{v \in \mathbb{R}^n : \text{there exists } u \in \mathbb{R}^m \text{ such that } v = c - A^\top u\}, \\ V_U &= \{v \in \mathbb{R}^n : \text{there exists } u \in U \text{ such that } v = c - A^\top u\}. \end{aligned}$$

Here v is the $n-$vector of slack variables. The set V_U is the image of U under the mapping $v(u) = c - A^\top u$. Therefore, $V_U = V \bigcap \mathbb{R}^n_+$, where $\mathbb{R}^n_+$ is the nonnegative orthant of $\mathbb{R}^n$.

We denote the components of a vector by using superscripts and the iterate numbers by using subscripts; $D(z)$ denotes the diagonal matrix whose entries are the components of z. The dimensionality of this matrix is determined by the dimensionality of z.

We now introduce a new n-dimensional space with the coordinates $[w^1, \ldots, w^n]$ and define a differentiable transformation from this space to the original one: $v = \varphi(w)$. This surjective transformation maps $\mathbb{R}^n$ onto $\mathbb{R}^n_+$ or $\text{int}\mathbb{R}^n_+$, i.e. $\mathbb{R}^n_+ = \overline{\varphi(\mathbb{R}^n)}$, where $\bar{B}$ denotes the closure of B. In this case every vector in $\mathbb{R}^n_+$ is the image of at least one vector in $\mathbb{R}^n$ or it is the image of a limit point of a sequence in $\mathbb{R}^n$. In other words for each $v \in \mathbb{R}^n_+$ there exists a $w \in \mathbb{R}^n$ such that $v = \varphi(w)$ or $v = \lim_{i\to\infty} \varphi(w_i)$, where $w_i \in \mathbb{R}^n$.

For the sake of simplicity we use a componentwise space transformation where

$$\varphi(w) = [\varphi^1(w^1), \varphi^2(w^2), \ldots, \varphi^n(w^n)].$$

Let $w^i = \psi^i(v^i)$ denote the inverse transformation of $\varphi^i(w^i)$. This transformation exists at least in a neighborhood of a point $v^i_0 = \varphi^i(w^i_0)$, as long as $\dot{\varphi}^i(w^i_0) \neq 0$.

We introduce a n-vector $\theta(v)$ and a $n \times n$ matrix $G(v)$:

$$\theta(v) = [\theta^1(v^1), \theta^2(v^2), \ldots, \theta^n(v^n)], \qquad G(v) = D(\theta(v)),$$

where $\theta^i(v^i) = (\gamma^i(v^i))^2$, $\gamma^i(v^i) = \dot{\varphi}^i(\psi^i(v^i))$, $1 \leq i \leq n$.

We impose the following conditions on the space transformation $\varphi(v)$:

- $\mathbf{C}_1$. *The functions $\theta^i(v^i)$ are defined and continuous in some neighborhood of $\mathbb{R}^1_+$ and $\theta^i(v^i) = 0$ if and only if $v^i = 0$, where $1 \leq i \leq n$.*

- $\mathbf{C}_2$. *The functions $\theta^i(v^i)$ are continuously differentiable in some neighborhood of $\mathbb{R}^1_+$ and $\dot{\theta}^i(0) > 0$, $1 \leq i \leq n$.*

Different numerical methods with various convergence properties can be obtained from different choices of the space transformation functions. Here we consider only two simple surjective transformations

$$v = \frac{1}{4}D(w)(w), \quad v = e^{-w}, \tag{2.3}$$

where the ith component of the $n-$vector e^{-w} is e^{-w^i}. We shall refer to these transformations as to quadratic and exponential space transformations, respectively.

For the transformations (2.3) we obtain, respectively,

$$\theta(v) = v, \quad G(v) = D(v); \qquad \theta(v) = D(v)v, \quad G(v) = D^2(v).$$

In both cases the Jacobian matrix is singular on the boundary of the set $\mathbb{R}^n_+$. These transformations satisfy $\mathbf{C}_1$. Condition $\mathbf{C}_2$ holds only for the first quadratic transformation (2.3).

By extension of the space and by converting the inequality constraints to equalities, we transform the original dual problem (2.2) into the following equivalent problem

$$\text{maximize } b^\top u \text{ with respect to } u \text{ and } w \text{ subject to } \varphi(w) + A^\top u - c = 0_n. \tag{2.4}$$

The Lagrangian associated with this problem is defined by

$$\tilde{L}(u, w, x) = b^\top u + x^\top[\varphi(w) + A^\top u - c].$$

For solving problem (2.4) we use the stable version of the gradient projection method which is described in [16]. The method is stated as an initial-value problem involving the following system of ordinary differential equations

$$\frac{du}{dt} = \tilde{L}_u(u, w, x(u, w)), \qquad \frac{dw}{dt} = \tilde{L}_w(u, w, x(u, w)). \tag{2.5}$$

The function $x(u,w)$ is chosen to satisfy the following condition:

$$\tilde{L}_{xu}(u,w,x)\dot{u} + \tilde{L}_{xw}(u,w,x)\dot{w} = -\tau\tilde{L}_x(u,w,x). \tag{2.6}$$

Since $\dot{v} = \varphi_w\dot{w}$, (2.5) can be rewritten in terms of u and v as follows:

$$\frac{du}{dt} = b - Ax(u,v), \quad \frac{dv}{dt} = -G(v)x(u,v), \tag{2.7}$$

where $\Phi(v)x(u,v) = A^\top b + \tau(v + A^\top u - c)$ and $\Phi(v) = G(v) + A^\top A$.

We say that an extreme point u of the feasible set U is nondegenerate if the vector $v(u)$ has only m zero components.

Lemma 1. *Let the space transformation $\varphi(w)$ satisfy $\mathbf{C}_1$. Then the matrix $\Phi(v(u))$ is positive definite for all $u \in U_0$.*

Lemma 2. *Let the space transformation $\varphi(w)$ satisfy $\mathbf{C}_1$. Assume that a point $u \in U$ can be represented as*

$$u = \sum_{j=1}^{s} \alpha^j u_j, \quad \alpha^j > 0, \quad \sum_{j=1}^{s} \alpha^j = 1,$$

where u_j, $1 \le j \le s$, are extreme points of U, and at least one point u_j is nondegenerate. Then the matrix $\Phi(v(u))$ is positive definite.

Proof. To show the nonsingularity of $\Phi(v)$, where $v = v(u)$ it suffices to show that the nullspace of $\Phi(v)$ contains only the point 0_n. Consider the linear system of equations

$$\Phi(v)\bar{x} = G(v)\bar{x} + A^\top A\bar{x} = 0_n. \tag{2.8}$$

By multiplying (2.8) on the left by $\bar{x}^\top$, we obtain

$$\bar{x}^\top G(v)\bar{x} + \bar{x}^\top A^\top A\bar{x} = 0.$$

Both expressions are nonnegative and, therefore,

$$\bar{x}^\top G(v)\bar{x} = 0, \quad \bar{x}^\top A^\top A\bar{x} = 0. \tag{2.9}$$

Let $S_j = \{1 \le i \le n : a_i^\top u_j = c^i\}$, $S = \bigcap_{j=1}^{s} S_j$, where a_i is the ith column of the matrix A. Without loss of generality we assume that $S = \{1, 2, \ldots, k\}$. We select the first k columns and denote them by B. We partition A, $\bar{x}$ and v as

$$A = [B \mid N], \quad \bar{x} = \begin{bmatrix} \bar{x}^B \\ \bar{x}^N \end{bmatrix}, \quad v = \begin{bmatrix} v^B \\ v^N \end{bmatrix}.$$

Since at least one extreme point is nondegenerate, it follows that $k \le m$ and B has full rank. Since $v^B = 0_k$ and $v^N > 0_{n-k}$, we obtain from (2.9) that $\bar{x}^N = 0_{n-k}$. Hence $B\bar{x}^B = 0_m$ and we conclude that $\bar{x} = 0_n$. $\Box$

Corollary 1. *If an extreme point u of a polytope U is nondegenerate, then $\Phi(v(u))$ is positive definite.*

Corollary 2. *If all extreme points of the bounded set U are nondegenerate, then $\Phi(v(u))$ is positive definite for all $u \in U$.*

According to Corollary 2, if $v \in V_U$, then the matrix $\Phi(v)$ is invertible. Because of the continuity it is also invertible in some neighborhood of V_U. For all points from this set we get

$$x(u,v) = \left(G(v) + A^\top A\right)^{-1}\left(A^\top b + \tau(v + A^\top u - c)\right).$$

Substituting this formula into the right-hand side of (2.7) we find that the system takes the explicit form

$$\begin{aligned} du/dt &= b - A\left(G(v) + A^\top A\right)^{-1}\left(A^\top b + \tau(v + A^\top u - c)\right), \\ dv/dt &= -G(v)\left(G(v) + A^\top A\right)^{-1}\left(A^\top b + \tau(v + A^\top u - c)\right). \end{aligned} \tag{2.10}$$

Let $[u(t, z_0), v(t, z_0)]$ denote the solution of the Cauchy problem (2.10) with initial conditions $u(0, z_0) = u_0$, $v(0, z_0) = v_0$, $z_0^\top = [u_0^\top, v_0^\top]$. Let $y(u, v) = c - A^\top u - v$. The condition (2.6) can be written as

$$\frac{dy(u, v)}{dt} = y_u^\top(u, v)\dot{u} + y_v^\top(u, v)\dot{v} = -\tau y.$$

Hence, the system of ordinary differential equations (2.10) has the first integral

$$c - A^\top u(t, z_0) - v(t, z_0) = \left(c - A^\top u_0 - v_0\right) e^{-\tau t}.$$

This implies that $c - A^\top u(t, z_0) - v(t, z_0) \to 0_n$ as $t \to +\infty$. Moreover, along the trajectories of the system (2.10) we get

$$\begin{aligned} b^\top \frac{du}{dt} &= b^\top (b - Ax(u, v)) = \|b - Ax(u, v)\|^2 + x^\top(u, v)A^\top (b - Ax(u, v)) = \\ &= \|b - Ax(u, v)\|^2 + x^\top(u, v)G(v)x(u, v) + \tau x^\top(u, v)\left(c - A^\top u - v\right). \end{aligned}$$

From the second equation of (2.10) it follows that if the transformation $\varphi(v)$ satisfies condition $\mathbf{C}_1$, then each component of the vector $v(t, z_0)$ does not change its sign. Hence, if $v_0 \geq 0$, then $v(t, z_0) \geq 0$ on the entire trajectory. We obtain this important property because of the matrix $G(v)$ in the right-hand side of (2.10), which plays the role of a "barrier", preventing the trajectory $v(t, z_0)$ from passing through the boundary of $\mathbb{R}^n_+$. Hence, we call (2.10) a "dual barrier-projection method".

Note that $y(u(t, z_0), v(t, z_0)) \equiv 0_n$ if $y(u_0, v_0) = 0_n$. We conclude that if $u_0 \in U$, then we can get rid of the equation for v and this way simplify systems (2.7) and (2.10). In this case, (2.7) can be expressed as

$$\frac{du}{dt} = b - Ax(u), \quad \left(G(v(u)) + A^\top A\right) x(u) = A^\top b, \tag{2.11}$$

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

For this system we obtain the following inequality

$$b^\top \frac{du}{dt} = \|b - Ax(u)\|^2 + x^\top(u)G(v(u))x(u) \geq 0.$$

Hence the objective function of the dual problem monotonically increases on a feasible set.

By applying the Euler numerical integration method we obtain the following iterative algorithm

$$\begin{gathered} u_{k+1} = u_k + \alpha_k(b - Ax_k), \quad v_{k+1} = v_k + \alpha_k G(v_k)x_k, \\ \left(G(v_k) + A^\top A\right) x_k = A^\top b + \tau\left(v_k + A^\top u_k - c\right). \end{gathered} \tag{2.12}$$

Similarly for the system (2.11) we have

$$u_{k+1} = u_k + \alpha_k\left(b - Ax_k\right), \quad \left(G(v_k), +A^\top A\right) x_k = A^\top b, \tag{2.13}$$

where $v_k = v(u_k)$. Both variants solve the primal and dual problems simultaneously.

Theorem 1. *Let x_* and u_* be unique nondegenerate solutions of Problems (2.1) and (2.2), respectively, and let $v_* = c - A^\top u_*$. Assume that the space transformation $\varphi(w)$ satisfies conditions $\mathbf{C}_1$, $\mathbf{C}_2$ and $\tau > 0$. Then the following statements are true:*

1. *The pair $[u_*, v_*]$ is an asymptotically stable equilibrium state of system (2.10).*

2. *The solutions $u(t, z_0), v(t, z_0)$ of system (2.10) converge locally 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 (2.1).*

3. *The point u_* is an asymptotically stable equilibrium state of system (2.11).*

4. *The solutions $u(t, u_0)$ of system (2.11) converge locally to the optimal solution u_* of the dual problem (2.2). The corresponding function $x(u(t, u_0))$ converges to the optimal solution x_* of the primal problem (2.1).*

5. *There exists an $\alpha_* > 0$ such that for any fixed $0 < \alpha_k < \alpha_*$ the sequence $\{u_k, v_k\}$ generated by (2.12) converges locally with a linear rate to $[u_*, v_*]$ while the corresponding sequence $\{x_k\}$ converges to x_*.*

6. *There exists an $\alpha_* > 0$ such that for any fixed $0 < \alpha_k < \alpha_*$ the sequence $\{u_k\}$ generated by (2.13) converges locally with a linear rate to u_* while the corresponding sequence $\{x_k\}$ converges to x_*.*

Proof. Let $\delta z^\top = [\delta u^\top, \delta v^\top]$, $\delta u = u(t, z_0) - u_*$ and $\delta v = v(t, z_0) - v_*$. We linearize system (2.10) in the neighborhood of the point $z_*^\top = [u_*^\top, v_*^\top]$. Then we obtain the first approximation of (2.10) about point z_*:

$$\delta \dot{z} = -Q \delta z,$$

where

$$Q = \left[\begin{array}{cc} \tau A \Phi^{-1} A^\top & A\Phi^{-1}\left(\tau I_n - D(\dot{\theta}(v_*))D(x_*)\right) \\ \tau G(v_*)\Phi^{-1} A^\top & \left(I_n - G(v_*)\Phi^{-1}\right) D(\dot{\theta}(v_*))D(x_*) + \tau G(v_*)\Phi^{-1} \end{array} \right],$$

$\Phi = G(v_*) + A^\top A$ and I_n is the $n \times n$ identity matrix.

Suppose that the first m columns of A are linearly independent and denote the $m \times m$ matrix determined by these columns as B. Assume that B is the optimal basis. With respect to this partition we can write

$$x_* = \left[\begin{array}{c} x_*^B \\ x_*^N \end{array} \right], \quad v_* = \left[\begin{array}{c} v_*^B \\ v_*^N \end{array} \right], \quad A = [B \mid N], \tag{2.14}$$

$$G(v_*) = \left[\begin{array}{cc} 0_{mn} & 0_{md} \\ 0_{dm} & G_N \end{array} \right], \quad \Phi = \left[\begin{array}{cc} B^\top B & B^\top N \\ N^\top B & G_N + N^\top N \end{array} \right],$$

where $x_*^B > 0_m$, $v_*^B = 0_m$, $v_*^N > 0_d$, $d = n - m$ and $G_N = D(\theta(v_*^N))$ is the $d \times d$ matrix. Using the Frobenious formula we can find Φ^{-1} and obtain

$$\Phi^{-1} A^\top = \left[\begin{array}{c} B^{-1} \\ 0_{dm} \end{array} \right], \quad Q = \left[\begin{array}{cc} \tau I_m & Q_2 \\ 0_{nm} & Q_1 \end{array} \right], Q_1 = \left[\begin{array}{cc} D(\dot{\theta}(v_*^B))D(x_*^B) & 0_{md} \\ Q_3 & \tau I_d \end{array} \right],$$

where the matrices Q_2 and Q_3 are not essential.

It is obvious that the characteristic equation

$$\det\left(Q - \lambda I_{n+m}\right) = 0$$

has following roots: $\lambda_i = \dot{\theta}^i(0) x_*^i$, $\lambda_j = \tau$, $1 \le i \le m$, $m+1 \le j \le n+n$. Since the transformation $\varphi(w)$ satisfies $\mathbf{C}_2$, and since x_* is a nondegenerate optimal solution of the problem (2.1), all these roots are positive and the smallest root is

$$\lambda_* = \min\left[\tau, \min_{1 \le i \le m} \dot{\theta}^i(0) x_*^i\right] > 0.$$

Hence, according to Lyapunov's linearization principle, the equilibrium point z_* is asymptotically stable and the following estimate holds

$$\limsup_{t\to\infty} \frac{\ln \|z(t, z_0) - z_*\|}{t} < -\lambda_*.$$

Denote

$$\alpha_* = 2/\lambda^*, \quad \lambda^* = \max\left[\tau, \max_{1\le i\le m} \dot{\theta}^i(0)x_*^i\right].$$

If the stepsize $\alpha_k < \alpha_*$, then by Theorem 2.3.7 from [5] the linear convergence of the discrete versions (2.12) follows from the proof given above.

If $u_0 \in U$, then the solutions of (2.11) coincide with corresponding solutions of (2.10), if in (2.10) we take $v_0 = v(u_0)$. Therefore, the local exponential convergence of (2.10) implies the local exponential convergence of (2.11). In the similar way the linear convergence of (2.12) implies the linear convergence of (2.13). □

We proposed the dual method (2.11) in 1977. It was described in [7] where we also gave the following primal method

$$\frac{dx}{dt} = G(x)\left[c - A^\top u(x)\right], \tag{2.15}$$

where $AG(x)A^\top u(x) = AG(x)c$, $G(x) = D(x)$. Both methods are similar and both solve primal and dual problems simultaneously. The method (2.15) is very popular now. It was reinvented recently in [13], [15] and analyzed in the book [14].

3 OTHER VARIANTS OF DUAL METHODS

As before we assume that A has full rank, therefore, the nullspace of A has dimension $d = n-m$. Let P be a full rank $d\times n$ matrix such that $AP^\top = 0_{md}$. Therefore, the columns of $P^\top$ are linearly independent and form a basis for the nullspace of A. We partition A as $A = [B, N]$, where the square matrix B is nonsingular. We can now write the matrix P as

$$P = \left[-N(B^\top)^{-1} \mid I_d\right].$$

The definitions of the sets V and V_U can be rewritten as follows:

$$V = \{v \in \mathbb{R}^n : P(v-c) = 0_d\}, \quad V_U = \left\{v \in \mathbb{R}^n_+ : P(v-c) = 0_d\right\}.$$

Let $\bar{x} \in \mathbb{R}^n$ be an arbitrary vector which satisfies the constraint $A\bar{x} = b$. Then

$$\max_{u\in U} b^\top u = \max_{u\in U} \bar{x}^\top A^\top u = \max_{v\in V_U} \bar{x}^\top (c-v) = \bar{x}^\top c - \min_{v\in V_U} \bar{x}^\top v.$$

Hence the solution of the dual problem (2.2) can be substituted by the following equivalent minimization problem:

$$\min_{v\in V_U} \bar{x}^\top v.$$

Applying the stable barrier-projection method [12] to this problem, we obtain

$$\frac{dv}{dt} = -G(v)\left(\bar{x} - P^\top x(v)\right), \tag{3.1}$$

$$PG(v)P^\top x(v) = PG(v)\bar{x} + \tau P(c-v). \tag{3.2}$$

If a point v is such that the matrix $PG(v)P^\top$ is invertible, then we can solve the linear equation (3.2) and obtain

$$x(v) = \left(PG(v)P^\top\right)^{-1} \left(PG(v)\bar{x} + \tau P(c-v)\right).$$

Let $H(v) = G^{1/2}(v)$ and introduce the pseudoinverse matrix $(PH)^+ = (PH)^\top(PGP^\top)^{-1}$ and the projection matrix $(PH)^\sharp = (PH)^+PH$. The system (3.1), (3.2) can be rewritten in the following projective form:

$$\frac{dv}{dt} = H\left[\tau(PH)^+P(c-v) - \left(I_n - (PH)^\sharp\right)H\bar{x}\right]. \tag{3.3}$$

The first vector in the square brackets belongs to the null space of AH^{-1} and the second vector belongs to the row space of this matrix. Furthermore,

$$P\frac{dv}{dt} = \tau P(c-v), \quad P(c - v(t,v_0)) = P(c-v_0)e^{-\tau t}.$$

Hence, the trajectories $v(t,v_0)$ approach the set V as $t \to \infty$.

If $v_0 \in V_U$ and $v_0 > 0$, then the entire trajectory does not leave the feasible set V_U, the objective function $\bar{x}^\top v(t,v_0)$ is a monotonically decreasing function of t and (3.3) can be rewritten as follows:

$$\frac{dv}{dt} = -G(v)\left(I_n - P^\top\left(PG(v)P^\top\right)^{-1}PG(v)\right)\bar{x}, \quad v_0 \in \mathrm{ri}V_U. \tag{3.4}$$

Theorem 2. *Suppose that the conditions of Theorem 1 hold. Then:*

1. *The point v_* is an asymptotically stable equilibrium point of system (3.1).*
2. *The solutions $v(t,v_0)$ of (3.3) converge locally to v_* with an exponential rate of convergence.*
3. *There exists an $\alpha_* > 0$ such that for any fixed $0 < \alpha_k < \alpha_*$ the discrete version*

$$v_{k+1} = v_k - \alpha_k G(v_k)\left(\bar{x} - P^\top x_k\right), \quad x_k = x(v_k) \tag{3.5}$$

converges locally with a linear rate to v_* while the corresponding sequence $\{x_k\}$ converges to x_*.

The proof is very similar to that of Theorem 1.

Since for system (3.4) $P\dot{v} = 0_d$, it follows that the vector $\dot{v}$ belongs to null-space of P which coincides with the row space of A. Therefore, there exists a vector $\lambda \in \mathbb{R}^m$ such that

$$\dot{v} = A^\top\lambda. \tag{3.6}$$

If $v > 0_n$, then after left multiplying both sides of (3.6) with $AG^{-1}(v)$ and in view of (3.4) we obtain

$$\lambda = -\left(AG^{-1}(v)A^\top\right)^{-1}A\bar{x} = -\left(AG^{-1}(v)A^\top\right)^{-1}b.$$

Hence, on the set $\mathrm{ri}V_U$ the method (3.4) takes the form

$$\frac{dv}{dt} = -A^\top\left(AG^{-1}(v)A^\top\right)^{-1}b, \quad v_0 \in \mathrm{ri}V_U.$$

In u-space this method can be written as

$$\frac{du}{dt} = \left(AG^{-1}(v(u))A^{\top}\right)^{-1} b, \quad u_0 \in U_0.$$

If we use the quadratic and exponential space transformations (2.3), we obtain

$$\frac{du}{dt} = \left(AD^{-1}(v(u))A^{\top}\right)^{-1} b, \quad u_0 \in U_0, \tag{3.7}$$

and

$$\frac{du}{dt} = \left(AD^{-2}(v(u))A^{\top}\right)^{-1} b, \quad u_0 \in U_0, \tag{3.8}$$

respectively. The system (3.8) coincides with the continuous version of the dual affine scaling method proposed by I. Adler, N. Karmarkar, M. Resende and G. Veiga in 1989 (see [1]).

According to Theorem 2, the solution of (3.1) converges locally with an exponential rate to equilibrium point $v_* = v(u_*)$. Therefore, the solutions of (3.7) also converge to the point u_* in the set U_0.

The discrete version of (3.7) consists of the iteration

$$u_{k+1} = u_k + \alpha_k \left(AD^{-1}(v_k)A^{\top}\right)^{-1} b, \quad u_0 \in U_0, \tag{3.9}$$

where $v_k = v(u_k)$. Taking into account Theorem 2.3.7 from [5] we conclude that the exponential rate of convergence of (3.7) insures local linear convergence of the discrete variant (3.9) if the step length α_k is sufficient small.

4 NON-LOCAL CONVERGENCE ANALYSIS

In this section we consider the global convergence of the dual barrier-projection method (3.9) on the set U. Suppose that the problem (2.2) is such that

$$A\beta = 0_m, \tag{4.1}$$

where β is a vector of ones in $\mathbb{R}^n$.

We assume that the dual problem (2.2) has a unique solution u_*. Let $v_* = v(u_*)$ and $J_*^N = \{1 \le i \le n : v_*^i > 0\}$. Then

$$0 < \sum_{i \in J_*^N} v_*^i = \beta^{\top} v_* = \beta^{\top} c = C.$$

Here we denoted $C = \sum_{i}^{n} c^i$ and showed that $C > 0$.

Condition (4.1) implies that along all trajectories of the system (3.7) the following property holds:

$$\sum_{j=1}^{n} v^j(u(t, u_0)) = \text{const}. \tag{4.2}$$

Introduce the Lyapunov function

$$F(u) = \sum_{i \in J_*^N} v_*^i \left(\ln v_*^i - \ln v^i(u)\right). \tag{4.3}$$

This function is well-defined and continuously differentiable everywhere on the set

$$U_1 = \left\{u \in U : v^i(u) > 0_n, \quad i \in J_*^N\right\}.$$

Moreover $F(u_*) = 0$ and $F(u) > 0$ for all $u \in U_1$ such that $u \neq u_*$. This follows from the well-known inequality

$$F(u) = -C \sum_{i \in J_*^N} \frac{v_*^i}{C} \ln \frac{v^i(u)}{v_*^i} = -C \ln \prod_{i \in J_*^N} \left(\frac{v^i(u)}{v_*^i} \right)^{v_*^i/C} > -C \ln \sum_{i \in J_*^N} \frac{v^i(u)}{C} = 0.$$

The derivative of the Lyapunov function (4.3) along the solutions of (3.7) is

$$\frac{dF(u)}{dt} = F_u^\top \dot{u} = v_*^\top D^{-1}(v(u)) A^\top \left(A D^{-1}(v(u)) A^\top \right)^{-1} b.$$

Let

$$p(u) = \left(A D^{-1}(v(u)) A^\top \right)^{-1} b, \quad x(u) = D^{-1}(v(u)) A^\top p(u).$$

These functions satisfy the following conditions:

$$Ax(u) = b, \quad x^\top(u) v(u) = \beta^\top A^\top p(u) = 0.$$

Hence

$$\frac{dF(u)}{dt} = v_*^\top x(u) = x^\top(u) \left(c - A^\top u_* \right) = b^\top u - b^\top u_* \leq 0, \tag{4.4}$$

where equality holds only if $u = u_*$.

For an arbitrary $u_0 \in U_0$ define a Lebesgue level set $Q = \{u \in U_1 : F(u) \leq F(u_0)\}$. In view of (4.2) the set V_U is compact. Hence, U and Q are also compact. The set Q does not contain any vertex from U other than u_*. The inequality (4.4) implies that $u(t, u_0) \in Q$ for all $t \geq 0$.

Let

$$K = \inf_{u \in Q} \frac{\langle b, u_* - u \rangle}{F(u)}. \tag{4.5}$$

Here, $\langle \cdot, \cdot \rangle$ stands for the standard scalar product in $\mathbb{R}^n$. Using (4.4) and (4.5), we obtain

$$F(u(t, u_0)) \leq F(u_0) e^{-Kt}$$

for all $t \geq 0$.

Lemma 3. *Suppose that the dual problem (2.2) has a unique nondegenerate solution u_*. Then the following estimate holds:*

$$K \geq \bar{K}(u_0) = \frac{1 - e^{-F(u_0)/C}}{F(u_0)} \min_{1 \leq j \leq m} s_j > 0, \tag{4.6}$$

where $s_j = b^\top (u_* - u_j)$ and u_j is a vertex of U adjacent to u_*.

Proof. We introduce the variable $z = u - u_*$ and write $F(u)$ and K as

$$F(u_* + z) = \tilde{F}(z) = - \sum_{i \in J_*^N} v_*^i \ln \left(1 - \frac{\langle a_i, z \rangle}{v_*^i} \right), \quad K = - \sup_{z \in Q_1} \frac{\langle b, z \rangle}{\tilde{F}(z)},$$

where $Q_1 = \{z \in Z : \tilde{F}(z) \leq F(u_o)\}$, $Z = \{z \in \mathbb{R}^m : A^\top z \leq v_*\}$ and a_i is the ith column of A.

The function $\tilde{F}(z)$ is convex on the set Q_1. Furthermore, $\tilde{F}(0) = 0$, $\tilde{F}(z) > 0$ and $b^\top z < 0$ for all $z \in Z$, $z \neq 0_m$. Thus, for any point $\bar{z} \in S = \{z \in Q_1 : \tilde{F}(z) = F(u_0)\}$ and any $0 \leq \alpha \leq 1$ the inequality $\tilde{F}(\alpha \bar{z}) \leq \alpha \tilde{F}(\bar{z})$ holds. Hence,

$$\frac{\langle b, \alpha \bar{z} \rangle}{\tilde{F}(\alpha \bar{z})} \leq \frac{\langle b, \bar{z} \rangle}{\tilde{F}(\bar{z})}, \qquad K = - \frac{1}{F(u_0)} \max_{z \in S} \langle b, z \rangle. \tag{4.7}$$

The point $z = 0$ is a vertex of the polytope Z. Let z_j be another adjacent vertex of this polytope and let β_j be a solution of the following equality:

$$\sum_{i \in J_*^N} v_*^i \ln(1 - \beta_j q_{ij}) + F(u_0) = 0, \tag{4.8}$$

where $q_{ij} = a_i^\top z_j / v_*^i$. Since $\tilde{F}(z_j) = +\infty$, we obtain that $0 < \beta_j < 1$ and

$$\max_{z \in S} \langle b, z \rangle = \max_{1 \le j \le m} \beta_j \langle b, z_j \rangle = \max_{1 \le j \le m} -\beta_j s_j < 0. \tag{4.9}$$

The inequality $A^\top z \le v_*$ implies that $q_{ij} \le 1$ for all $i \in J_*^N$. Moreover, for at least one i we have $q_{ij} = 1$. Therefore,

$$\ln(1 - \beta_j q_{ij}) \ge \ln(1 - \beta_j).$$

Hence any β_j which satisfies (4.8) is such that $\beta_j \ge \bar{\beta}$, where $\bar{\beta}$ is a solution of the following equation:

$$\ln(1 - \bar{\beta}) \sum_{i \in J_*^N} v_*^i + F(u_0) = 0.$$

We conclude that $\bar{\beta} = 1 - e^{-F(u_0)/C}$. Taking into account (4.7) and (4.9) we obtain the estimate (4.6). $\square$

Let $\mu(u) = \max_{1 \le i \le n} x^i(u)$. We note that for any $u \in U_0$ the inequality $\mu(u) > 0$ holds. By contradiction, assume that $\mu(u) \le 0$. Then $x(u) \le 0$, and also $x^i(u) < 0$ at least for one i. For any $\alpha > 0$ we have $\alpha x(u) \le 0_n < \beta$. Multiplying this inequality by $D(v(u))$ we obtain $\alpha A^\top (AD^{-1}(v)A^\top)^{-1} b \le v(u)$ or, equivalently,

$$A^\top \left(u + \alpha(AD^{-1}(v)A^\top)^{-1} b\right) \le c. \tag{4.10}$$

Thus we must have that $u + \alpha(AD^{-1}(v)A^\top)^{-1}b \in U$ for any $\alpha > 0$. This contradicts the compactness of the set U. From (4.10) it follows that the value $1/\mu(u)$ is the upper bound for α such that $u + \alpha x(u) \in U$.

Theorem 3. *Let a stepsize α_k in (3.9) be chosen such that*

$$0 < \alpha_k = \gamma / \mu(u_k), \quad 0 < \gamma < 1. \tag{4.11}$$

Then for any $u_0 \in U_0$ there exists $\gamma(u_0)$ such that $0 < \gamma(u_0) < 1$ and for all $0 < \gamma \le \gamma(u_0)$, $k \ge 0$ the following estimate holds:

$$F(u_{k+1}) \le F(u_k) \left(1 - \frac{\alpha_k K}{2}\right), \tag{4.12}$$

where K is defined from (4.5).

Proof. Denote

$$W(u, \alpha) = \alpha^{-1} \sum_{i \in J_*^N} v_*^i \ln(1 - \alpha x(u)).$$

It follows from (3.9) that

$$F(u_{k+1}) = F(u_k) - \alpha_k W(u_k, \alpha_k). \tag{4.13}$$

Using the Taylor series expansion, we obtain

$$W(u, \alpha) = -v_*^\top x(u) - \frac{\alpha}{2} \sum_{i \in J_*^N} \frac{v_*^i (x^i(u))^2}{(1 - \alpha \zeta^i(u) x^i(u))^2},$$

where $0 \leq \zeta^i(u) \leq 1$, $i \in J_*^N$. The last equation and (4.4) imply that for any $\alpha \leq \gamma/\mu(u)$ we have

$$W(u,\alpha) \geq b^\top(u_* - u) - \frac{\gamma}{2(1-\gamma)^2\mu(u)} \sum_{i \in J_*^N} v_*^i (x^i(u))^2. \tag{4.14}$$

We introduce the function

$$r(u) = \mu(u) \frac{\langle b, u_* - u \rangle}{\sum\limits_{i \in J_*^N} v_*^i (x^i(u))^2}$$

and prove that

$$\bar{r} = \inf_{u \in Q} r(u) > 0. \tag{4.15}$$

Consider a minimizing sequence $\{u_s\}$ such that all $u_s \in Q$, $\lim\limits_{s \to \infty} u_s = \bar{u}$ and $\bar{r} = \lim\limits_{s \to \infty} r(u_s)$. If $\bar{u} \neq u_*$, then $\bar{r} > 0$. We prove that if $\bar{u} = u_*$, then $\bar{r} > 0$. Suppose that the partition (2.14) holds, where $v_*^B = 0_m$, $v_*^N > 0_d$. The same partition will be used for vector $v(u)$ and for matrix A. Denote $\Gamma^B(u) = BD^{-1}(v^B(u))B^\top$, $\Gamma^N(u) = ND^{-1}(v^N(u))N^\top$. Since the matrix B is nonsingular we have

$$\Gamma(u) = AD^{-1}(v(u))A^\top = \Gamma^B(u) + \Gamma^N(u) = \Gamma^B(u)\left[I + (\Gamma^B(u))^{-1}\Gamma^N(u)\right],$$

$$\Gamma^{-1}(u) = (\Gamma^B(u))^{-1} + \Phi(u),$$

where $\|\Phi(u)\| = \mathrm{o}(\|u - u_*\|)$. Hence we obtain

$$x^B(u) = D^{-1}(v^B)B^\top(\Gamma^B(u))^{-1}b + D^{-1}(v^B)B^\top\Phi(u)b = x_*^B + \phi_1(u),$$

$$x^N(u) = D^{-1}(v^N)N^\top\Gamma^{-1}(u)b = \phi_2(u),$$

$$\mu(u) = \max_{1 \leq i \leq n} x^i(u) = \max_{1 \leq i \leq m} x_*^i(u) + \phi_3(u),$$

where $\|\phi_i(u)\| = \mathrm{O}(\|u - u_*\|)$, $i = 1, 2, 3$.

Assume the contrary, i.e. $\bar{r} = 0$, then $r(u_s) < 1$ for all s sufficiently large. It follows from $x^N(u) = \mathrm{O}(\|u - u_*\|)$ that

$$\sum_{i \in J_*^N} v_*^i (x^i(u))^2 = \mathrm{o}(\|u - u_*\|).$$

This means that the inequality $r(u_s) < 1$ does not hold for all s sufficiently large. Therefore, $\bar{r} > 0$. From (4.15) it follows that there exists sufficiently small $0 < \gamma(u_0) < 1$ such that for all $0 < \gamma < \gamma(u_0)$ and $u \in Q$ we have

$$\frac{\gamma}{(1-\gamma)^2\mu(u)} \sum_{i \in J_*^N} v_*^i (x^i(u))^2 \leq b^\top(u_* - u).$$

Hence, for such γ, u and $\alpha \leq \gamma/\mu(u)$ we obtain from (4.14) that $W(u,\alpha) \geq b^\top(u_* - u)/2$. From this inequality and in view of the inequality $b^\top(u_* - u) \leq KF(u)$ and (4.13) we conclude that (4.12) holds for any $u_k \in Q$. $\square$

Let

$$S(u_0) = \max_{u \in Q} \max_{1 \leq i \leq n} x^i(u).$$

If the stepsize α_k is such that $\alpha_k = \gamma/\mu(u_k)$, then $\alpha_k \geq \bar{\alpha}(u_0) = \gamma/S(u_0)$ for all $k \geq 0$. Hence we have

$$V(u_{k+1}) \leq V(u_k)\left[1 - \frac{\alpha K}{2}\right], \tag{4.16}$$

where $0 < \alpha \leq \bar{\alpha}(u_0)$.

Let ϵ be the tolerance for the Lyapunov function. Then it follows from (4.6) and (4.16) that the total number of iterations performed by algorithm (3.9), (4.11) is no greater than

$$\frac{2}{\bar{\alpha}(u_0)\bar{K}(u_0)} \ln\left(\frac{V(u_0)}{\epsilon}\right).$$

5 ACKNOLEGEMENTS

The authors would like to thank K.L. Teo and V. Rehbock for very useful corrections and comments. The work described in this paper is a part of research project sponsored by Russian Scientific fund.

Bibliography

[1] I. Adler, N. Karmarkar, M.G.C. Resende and G. Veiga, An Implementation of Karmarkar's Algorithm for Linear Programming, *Math. Programming*, **44**, 297–335 (1989).

[2] 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).

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

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

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

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

[7] Yu.G. Evtushenko and V.G. Zhadan, A relaxation method for solving problems of nonlinear programming, *U.S.S.R. Comput. Maths. Math. Phys.*, **17**(4), 73–87 (1978).

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

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

[10] Yu.G. Evtushenko and V.G. Zhadan, Stable Barrier-Projection and Barrier-Newton Methods in Nonlinear Programming, *Optimization Methods and Software* **3**(1-3), 237–256 (1994).

[11] Yu.G. Evtushenko and V.G. Zhadan, Stable Barrier-Projection and Barrier-Newton Methods for Linear and Nonlinear Programming, In: *Algorithms for Continuous Optimization*, (Edited by E.Spedicato), NATO ASI Series, 255-285, Kluwer Academic Publishers (1994).

[12] Yu.G. Evtushenko and V.G. Zhadan, Stable Barrier-Projection and Barrier-Newton Methods in Linear Programming, *Computational Optimization and Applications* **3**(4), 289–303 (1994).

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

[14] U. Helmke and J.B. Moore, *Optimization and Dynamical Systems*, Springer-Verlag (1994).

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

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