Computational Optimization and Applications, 3, No.4, 289–303 (1994)

Stable Barrier-Projection and Barrier-Newton Methods in Linear Programming[1]

YURI G. EVTUSHENKO AND VITALI G. ZHADAN
Computing Center, Russian Academy of Sciences, 40 Vavilov str., 117967 Moscow GSP-1, Russia

Revised April 5, 2001

Dedicated to Professor George B. Dantzig on the occasion of his eightieth birthday.

Abstract. The present paper is devoted to the application of the space transformation techniques for solving linear programming problems. By using a surjective mapping the original constrained optimization problem is transformed to a problem in a new space with only equality constraints. For the numerical solution of the latter problem the stable version of the gradient-projection and Newton's methods are used. After an inverse transformation to the original space a family of numerical methods for solving optimization problems with equality and inequality constraints is obtained. The proposed algorithms are based on the numerical integration of the systems of ordinary differential equations. These algorithms do not require feasibility of the starting and current points, but they preserve feasibility. 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. A proof of a convergence is given.
Keywords: Linear programming, space transformation, gradient-projection method, Newton's method, interior point technique, barrier function, Karmarkar's method.

1. Introduction

Starting from 1973, we developed a family of numerical methods for solving a nonlinear programming (NLP) problem [5] – [12]. On the basis of a space transformation the original NLP problem with inequality constraints was reduced to a problem with equality constraints. The stable version of the gradient-projection method and Newton's method were used for solving this reduced problem. The numerical methods were found after performing an inverse transformation. These methods were described by systems of ordinary differential equations. As a result of the space transformation we obtained differential equations which prevented the trajectories from crossing the boundary of the feasible set. Therefore, we termed these methods "barrier-projection" and "barrier-Newton" methods. The space transformation was carried out without using conventional barriers or penalty functions and this feature provided a high rate of convergence. The analysis of the method was made on the basis of the stability theory of the solutions of ordinary differential equations. Numerical algorithms were obtained as discretization of dynamical systems. We proved that the barrier-projection method had linear convergence and did not require feasibility of initial vectors. We showed that under standard assumptions the barrier-Newton method converged quadratically.

The purpose of this paper is to apply our results to linear programming (LP) problem. After some simplifications and after choosing a particular exponential space-transformation function we obtain Dikin's algorithm [4] from the barrier-projection method sometimes called the "variation of Karmarkar's algorithm". However, there are some differences between our

[1]Research was supported by the grant N93-012-450 from Russian Scientific Fund.

approach:

1. We use mainly quadratic space transformation and owing to it we get faster local convergence.

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

3. We use multiplicative barrier functions and do not resort to a penalty-type algorithms.

4. In [11] we considered the steepest descent variants of our methods where the trajectory could move along the boundary of the feasible set.

Here we briefly describe our approach. Computational aspects, steepest descent are beyond the scope of the present paper. More detailed analysis is given in [10] and [11].

2. Basic approach and outline of the methods

Consider the following nonlinear programming problem:

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

Here R^n denotes the vector space formed by n-dimensional column vectors with real entries. The set P is assumed to have a nonempty interior. The functions $f(x)$ and $g(x)$ are continuously differentiable, $f(x)$ maps R^n onto R^1 and $g(x)$ maps R^n onto R^m, 0_m is the m-dimensional null vector, 0_{nm} is the $n \times m$ rectangular null matrix. The feasible set X and the set of solutions X_* are supposed to be nonempty. We assume differentiability whenever it is helpful to do so. Subscripts will be used to distinguish values of quantities at a particular iteration and superscripts will indicate components of vectors.

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

$$\text{minimize } \tilde{f}(y) = f(\xi(y)) \text{ subject to } y \in Y, \tag{2}$$

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

The Lagrangians associated with Problem (1) and (2) are defined by $L(x,u) = f(x) + u^\top g(x)$, $\tilde{L}(y,u) = \tilde{f}(y) + u^\top \tilde{g}(y)$, respectively. To obtain the numerical solution of Problem (2) we seek the limit points of the solutions of the system described by the following vector differential equation:

$$\frac{dy}{dt} = -\tilde{L}_y(y, u(y)), \tag{3}$$

where $\tilde{L}_y(y,u) = \tilde{f}_y(y) + \tilde{g}_y^\top(y)u$, $\tilde{f}_y = \tilde{J}^\top f_x$, $\tilde{g}_y = g_x \tilde{J}$, $g_x(x)$ is the $m \times n$ Jacobian matrix of $g(x)$ with respect to x, $\tilde{J} = dx/dy$ is the Jacobian matrix of the transformation $x = \xi(y)$ with respect to y.

The function $u(y)$ is chosen to satisfy the following condition:

$$\frac{d\tilde{g}}{dt} = -\tilde{g}_y(y)\tilde{L}_y(y, u(y)) = -\tau \tilde{g}(y), \qquad \tau > 0. \tag{4}$$

If $\tilde{J}(y)$ is a nonsingular matrix, then there exists an inverse transformation $y = \delta(\mathrm{x})$, 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 $x(t)$ with respect to t, we obtain from (3) and (4)

$$\frac{dx}{dt} = \frac{d\xi}{dy}\frac{dy}{dt} = J(x)\frac{dy}{dt} = -G(x)L_x(x, u(x)), \qquad x_0 \in P, \tag{5}$$

$$\Gamma(\mathrm{x})u(\mathrm{x}) + g_x(x)G(x)f_x(x) = \tau g(x), \tag{6}$$

where we have introduced the two Gram matrices:

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

Let W be a $m \times n$ rectangular matrix whose rank is m. We introduce the pseudoinverse matrix $W^+ = W^\top(WW^\top)^{-1}$ and the orthogonal projector $\pi(W) = I_n - W^+W$, where I_n is the $n \times n$ identity matrix. The operator $\pi(W)$ projects any n-dimensional vector onto the nullspace $\ker W = \{z \in R^n : Wz = 0_m\}$.

If at a point x the matrix $g_x(x)$ has full rank, then we can find from (6) the function $u(x)$, substitute it into the right-hand side of (5) and write (5) in the following projective form:

$$\frac{dx}{dt} = -J(x)\left\{\pi[g_x(x)J(x)]J^\top(x)f_x(x) + \tau[g_x(x)J(x)]^+g(x)\right\}. \tag{7}$$

Let $x(t, x_0)$ denote the solution of the Cauchy problem (7) with initial condition $x(0, x_0) = x_0$, $x_0 \in P$.

If the condition $x \in P$ is absent in Problem (1), if $x_0 \in X$ and/or $\tau = 0$, then method (7) coincides with the gradient-projection method which has been used by many authors (see, for example, [16, 17]). In [10] we proved under standard assumptions that the solution of Problem (1) could be found as limit points of the trajectories $x(t, x_0)$ as $t \to \infty$.

The right-hand side of system (5) is well-defined for all $x \in P$. Sometimes $G(x)$ can be extended to an open set containing P so that system (5) is defined also for x such that they do not belong to P.

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. $P = R^n_+$. It is convenient for this set P to use a component-wise differentiable space transformation $\xi(y)$

$$x^i = \xi^i(y^i), \qquad 1 \le i \le n. \tag{8}$$

For such transformation the corresponding Jacobian matrix is diagonal and

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

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

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

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

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

C$_1$. The matrix $G(x)$ is defined, continuous at each point $x \in P$ and it is singular only on the boundary of P.

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

$\mathbf{C}_3$. The space transformation $\xi(y)$ satisfies condition $\mathbf{C}_2$, the map $\theta(x)$ is defined and differentiable in a neighborhood of R^n_+, $\dot{\theta}^i(0) > 0,\ 1 \le i \le n$.

$\mathbf{C}_4$. There exists $\alpha > 0$ such that

$$\theta^i(x^i) = (x^i)^\alpha + O((x^i)^{\alpha+1}), \qquad 1 \le i \le n. \tag{9}$$

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

$$x^i = \xi^i(y^i) = \frac{1}{4}(y^i)^2, \qquad J(x) = D^{1/2}(x), \qquad G(x) = D(x), \tag{10}$$

$$x^i = \xi^i(y^i) = e^{y^i}, \qquad J(x) = D(x), \qquad G(x) = D^2(x). \tag{11}$$

In these two cases the Jacobian matrix is singular on the boundary of the set P. These transformations satisfy $\mathbf{C}_1$ and $\mathbf{C}_2$. Condition $\mathbf{C}_3$ holds only for transformation (10).

Applying the Euler method for solving system (5), we obtain

$$x_{k+1} = x_k - h_k G(x_k) L_x(x_k, u(x_k)), \qquad x_0 \in P, \tag{12}$$

where $h_k > 0$ is a stepsize.

In [10] we proved the local linear convergence of algorithm (12) if stepsize h_k is fixed and sufficiently small and transformation (10) is used.

We say that x is a regular point for Problem (1) if the vectors $g^i_x(x),\ 1 \le i \le m$, are linearly independent. The equation (6) can be rewritten as

$$g_x(x) G(x) L_x(x, u(x)) = \tau g(x). \tag{13}$$

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

$$G(x) L_x(x, u(x)) = 0_n, \tag{14}$$

then $[x, u(x)]$ is a Kuhn–Tucker point of Problem (1). We say that x is an equilibrium point of system (5) if the right-hand side evaluated at x is a null vector. The right-hand side of system (5) defines a vector field which vanishes at equilibrium points. At every regular point this field is nonvanishing except points x such that $[x, u(x)]$ forms a Kuhn–Tucker pair.

Now we apply Newton's method for finding a solution x of nonlinear equation (14). The continuous version of Newton's method leads to the initial value problem for the following system of ordinary differential equations:

$$\Lambda(x)\frac{dx}{dt} = -\gamma G(x) L_x(x, u(x)), \qquad x(t, x_0) = x_0, \tag{15}$$

where $\gamma \in R^1$ is a scalar, $\Lambda(x)$ is the Jacobian matrix of the mapping $G(x)L_x(x, u(x))$ with respect to x:

$$\Lambda(x) = \dot{G} D(L_x) + G L_{xx} + G g_x^\top \frac{du}{dx}. \tag{16}$$

Here all matrices and vectors are evaluated at a point x and the function $u(x)$ is defined from (13); we took into account that the transformation $\xi(y)$ satisfies (8), therefore G is a diagonal matrix and $\dot{G} = D(\dot{\theta})$. By differentiating equality (13) with respect to x, we obtain

$$g_x \left[\dot{G} D(L_x) + G L_{xx} + G g_x^\top \frac{du}{dx} \right] = \tau g_x. \tag{17}$$

Here for the sake of simplicity we assume that $g(x)$ is a linear function of x.

We find du/dx from (17) and after substituting it into (16), we obtain

$$\Lambda = [I_n - H][\dot{G}D(L_x) + GL_{xx}] + \tau H,$$

where $H = Gg_x^\top(g_x Gg_x^\top)^{-1} g_x$.

3. Barrier-projection method for linear programming

In this section we apply barrier-projection method (7) to a linear programming problem. In (1) we set $f(x) = c^\top x$, $g(x) = b - Ax$, $P = R^n_+$, where c $\in R^n$, $b \in R^m$, and A is an $m \times n$ real matrix with rank m, $m < n$. Now Problem (1) is stated in the standard LP form:

$$\text{minimize } c^\top x \text{ subject to } X = \{x \in R^n, \quad b - Ax = 0_m, \quad x \geq 0_n\}. \tag{18}$$

We introduce the dual LP problem

$$\text{maximize } b^\top u \text{ subject to } u \in U = \{u \in R^m : v = c - A^\top u \geq 0_m\}, \tag{19}$$

where $v = L_x(x, u)$ is a vector of dual slack variables.

We define a relative interior set of X and an interior set of U:

$$X_0 = \{x \in R^n : Ax = b, \quad x > 0_n\}, \qquad U_0 = \{u \in R^m : v = \mathrm{c} - A^\top u > 0_m\}.$$

We assume that the set X_0 and U_0 are nonempty, the primal and dual nondegeneracies hold. In this case both problems have unique solutions x_* and u_*, respectively.

Applying methods (5) and (12) for solving Problem (18), we obtain the following continuous and discrete versions

$$\frac{dx}{dt} = -G(x)[c - A^\top u(x)], \qquad x(0, x_0) = x_0 > 0_n, \tag{20}$$

$$x_{k+1} = x_k - h_k G(x_k)[c - A^\top u(x_k)], \qquad x_0 > 0_n, \tag{21}$$

where the function $u(x)$ is found from linear equation (6) which can be rewritten as follows:

$$AG(x)A^\top u(x) - AG(x)c = \tau(b - Ax). \tag{22}$$

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

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

Hence $c^\top x(t, x_0)$ is a monotonically decreasing function of $t > 0$, if $x(t, x_0) \in X$ or the trajectory is close to X, i.e. $\|Ax(t, x_0) - b\|$ is sufficiently small.

If the space transformation $\xi(y)$ satisfies (8), and conditions $\mathbf{C}_1$ and $\mathbf{C}_2$ hold, then system (22) has a unique solution for all $x \geq 0_n$ and the trajectories of (20) do not leave the positive orthant R^n_+. Suppose not: let $x^i(T, x_0) < 0$ for some $T > 0$. Then there exists a time $T_* < T$ such that $x^i(T_*, x_0) = 0$ and $dx^i(T_*, x_0)/dt < 0$. This contradicts (20) since, according to $\mathbf{C}_2$, $\theta^i(x^i(T_*, x_0)) = 0$. Thus a transformation function plays the role of a "barrier", preventing the trajectory $x(t, x_0)$ from passing through the boundary of P. Therefore, we call (7) and (20) a "barrier-projection method".

The system of ordinary differential equations (20) has the first integral

$$Ax(t, x_0) = b + (Ax_0 - b)e^{-\tau t}. \tag{23}$$

This means that if $\tau > 0$, then method (20) has a remarkable property: all its trajectories approach the feasible set as t tends to infinity and the polyhedra X is an asymptotically stable attractor for the system (see [6, 11, 18]). Therefore, we call method (20) "the stable version of the barrier-projection method". On the contrary the well-known gradient-projection method is neutrally stable with respect to the equality constraints. It means that, if $Ax_0 - b = \beta$, $\|\beta\| \neq 0$, then $Ax(t, x_0) - b \equiv \beta$ for all $t \geq 0$ and we have to introduce an additional correction procedure in order to satisfy feasibility. This procedure increases the computation time.

From (23) it follows that, if $A\mathrm{x}_0 = b$, then $Ax(t, x_0) \equiv b$ for all $t \geq 0$ and the trajectory $x(t, x_0)$ of (20) remains in the feasible set X, the objective function monotonically decreases along the trajectories. The gradient-projection method and the method described above can be considered as particular cases of the interior point techniques. But we do not restrict ourselves to only interior point techniques. Methods (7) and (20) belong to the more general family of algorithms. In our methods the current points are often infeasible with respect to equality constraint, but if the starting points or the current points are feasible, then the subsequent trajectory remains in the feasible set, i.e. the feasibility is preserved.

Theorem 3.1. *Let x_*, u_* be unique solutions of Problems* (18) *and* (19)*, respectively. Let the space transformation $\xi(y)$ satisfy conditions $\mathbf{C}_2$ and $\mathbf{C}_3$. Then the system* (20) *with $\tau > 0$ is asymptotically stable at the isolated solution point x_*. There exists $h_* > 0$ such that for any fixed $0 < h_k < h_*$ the sequence $\{x_k\}$, generated by* (21)*, converges locally with a linear rate to x_* while the corresponding sequence u_k converges to u_*.*

Proof. Denote $\delta x(t) = x(t, x_0) - x_*$ and linearize system in the neighborhood of the point x_*. Then we obtain the equation of the first approximation of (20) about the equilibrium point x_*:

$$\delta \dot{x} = -Q\delta x, \tag{24}$$

where $Q = MD(\dot{\theta})D(v) + \tau P$, $M = I_n - P$, $P = GA^\top(AGA^\top)^{-1}A$. Here all functions are evaluated at the points $x = x_*$, $u = u_* = u(x_*)$, $v = v_* = c - A^\top u_*$.

The stability of system (20) is determined by the properties of the roots of the characteristic equation

$$\det(Q - \lambda I_n) = 0. \tag{25}$$

For proof we split the vectors x_* and v_* in two vectors

$$x_* = \begin{bmatrix} x_*^B \\ x_*^N \end{bmatrix}, \qquad v_* = \begin{bmatrix} v_*^B \\ v_*^N \end{bmatrix}, \tag{26}$$

where x_*^B, $v_*^B \in R^m$; x_*^N, $v_*^N \in R^d$; $d = n - m$. All components of the vectors x_*^N and v_*^B are equal to zero and all components of x_*^B and v_*^N are interior, i.e. $x_*^B > 0$, $v_*^N > 0$. In a similar way we represent vector $\theta(x_*)$ and matrices:

$$\begin{aligned} \theta(x_*) &= \begin{bmatrix} \theta_*^B \\ \theta_*^N \end{bmatrix}, \qquad A = [B \,|\, N], \qquad P = \begin{bmatrix} P^B & P^{BN} \\ P^{NB} & P^N \end{bmatrix}, \\ J &= \begin{bmatrix} J^B & 0_{md} \\ 0_{dm} & J^N \end{bmatrix}, \\ G^B &= D(\theta_*^B) = J^B J^B. \end{aligned} \tag{27}$$

From $\mathbf{C}_2$ it follows that $\theta_*^N = 0_d$, and P^{BN}, P^{NB}, P^N are null matrices. Hence the matrix Q can be decomposed into the following blocks

$$Q = \begin{bmatrix} \tau I_m & Q_3 \\ 0_{dm} & D(\dot{\theta}_*^N)D(v_*^N) \end{bmatrix}, \qquad Q_2 = D(\dot{\theta}_*^N)D(v_*^N),$$

where the matrix Q_3 is not essential.

The characteristic equation (25) is equivalent to two equations:

$$|(\tau-\lambda)I_m| = 0, \qquad |Q_2-\lambda I_d| = 0.$$

The solutions of these equations are found explicitly: $\lambda_j = \tau$, $\lambda_i = \dot{\theta}^i(x_*^i)v_*^i$, $1 \le j \le m$, $m+1 \le i \le n$. From $\mathbf{C}_3$ and the strict complementary condition we obtain: $\tilde{\lambda} = \min_{s+1\le i\le n} \lambda_i > 0$.

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

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

Denote $h_* = 2/\lambda^*$, where $\lambda^* = \max_{m+1\le i\le n}[\tau, \lambda_i]$. If the stepsize $h_k < h_*$, then by Theorem 2.3.7 from [6], the linear convergence of the discrete version (21) follows from the proof given above. □

By introducing condition $\mathbf{C}_3$, we assume that the matrix $G(x)$ is differentiable at least in the neighborhood of the solution point x_*. In this case we proved a local convergence. If $G(x)$ is defined only on the set X, then the local convergence takes place, if $x_0 \in X_0$ and x_0 is sufficiently close to x_*. In the last case we say that the trajectories $x(t,x_0)$ converge locally on X_0. If we use the exponential space transformation (11) and set $\tau = 0$, then from (20), (22) we obtain

$$\frac{dx}{dt} = D^2(x)(A^\top u(x) - c), \qquad AD^2(x)A^\top u(x) = AD^2(x)c. \tag{28}$$

The discrete and continuous versions of this method were investigated in various papers (see, for example, [1, 3, 4, 13, 14, 19, 20]). In [1] the discrete version was called "a variation on Karmarkar's algorithm". We should remark that method (28) does not possess the local convergence property. Here the convergence takes place only if x_0 belongs to the relative interior of X. Theorem 3.1 cannot be used for the exponential space transformation (11) because this transformation does not satisfy condition $\mathbf{C}_3$. If we try to use the same approach, then we obtain that among the roots of the characteristic equation (25) there are zero roots and, therefore, Lyapunov's linearization principle can not be used. In this case the convergence was proved by G. Smirnov on the basis of the vector Lyapunov function. He investigated the transformation (9) and proved that, if $\alpha > 1$, then $\|x_N(t,x_0)\| \approx O(t^{-1/(\alpha-1)})$ as $t \to \infty$. If we use the quadratic space transformation (10), then $\|x_N(t,x_0)\| \approx O(e^{-\lambda_* t})$. Hence the trajectories of system (20) with the quadratic transformation converge locally faster than the trajectories of system (20) 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. $P = \{x \in R^n : a \le \le x \le b\}$. Here we use the following transformation:

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

The statement of Theorem 3.1 is generalized for this case.

For the sake of simplicity in this section we consider the case where the quadratic space transformation (10) is used, the starting point $x_0 \in X_0$ and there is a condition

$$\sum_{i=1}^{n} x^i = 1 \tag{29}$$

among the equality constraints. Under the given assumptions methods (20), (21) and condition (22) can be rewritten as follows:

$$\begin{aligned} \frac{dx}{dt} &= -D(x)[c - A^\top u(x)], \qquad x(0,x_0) = x_0 \in X_0, &(30)\\ x_{k+1} &= x_k - h_k D(x_k)[c - A^\top u(x_k)], \qquad x_0 \in X_0, &(31)\\ AD(x)[c - A^\top u(x)] &= 0_m. &(32)\end{aligned}$$

From (29) and (32) we have

$$\begin{aligned} c^\top x &= x^\top A^\top u(x) = b^\top u(x), \qquad Ax = b,\\ c^\top \frac{dx}{dt} &= b^\top \frac{du}{dt} = -\|D^{1/2}(x)(c - A^\top u(x))\|^2 \le 0. \end{aligned} \tag{33}$$

We observe that the objective functions of the primal and dual problems monotonically decrease along the trajectories of system (30). The duality gap is equal to zero along entire trajectories.

Introduce the Lyapunov function

$$V(x) = \sum_{i \in J_B(x_*)} x_*^i [\ln(x_*^i) - \ln(x^i)],$$

where x_* is a solution of (18) and $J_B(x_*) = \{i \in [1:n] : x_*^i > 0\}$. The function $V(x)$ is well-defined everywhere on the set

$$X_1 = \{x \in X : x^i > 0 \text{ for } i \in J_B(x_*)\}.$$

Moreover, $V(x_*) = 0$ and $V(x) > 0$ for all $x \in X_1$ such that $x \neq x_*$. It follows from the following inequalities:

$$V(x) = -\sum_{i \in J_B(x_*)} x_*^i \ln \frac{x^i}{x_*^i} = -\ln \prod_{i \in J_B(x_*)} \left(\frac{x^i}{x_*^i}\right)^{x_*^i} > -\ln \left(\sum_{i \in J_B(x_*)} x^i\right) \ge 0.$$

Using (33) we obtain that the derivative of Lyapunov's function along the solution of (30) is

$$\frac{dV(x)}{dt} = V_x^\top(x)\dot{x} = c^\top x_* - x_*^\top A^\top u(x) = c^\top (x_* - x) < 0. \tag{34}$$

This inequality holds for all x such that $x \in X_1$, $x \neq x_*$.

For arbitrary $x_1 \in X_0$ define a Lebesque level set $Q(x_0) = \{x \in X_1 : V(x) \le V(x_0)\}$. This set is compact and does not contain any vertex from X_1 except x_*. By our choice $x_0 \in X_1$ and (34) implies that $V(x(t,x_0)) \le V(x_0)$ for all $t \ge 0$. Hence $x(t,x_0) \in Q(x_0)$.

Let us define

$$K_0 = \min_{x \in Q(x_0)} \frac{\langle c, x - x_*\rangle}{V(x)}, \qquad K^0 = \sup_{x \in Q(x_0)} \frac{\langle c, x - x_*\rangle}{V(x)}. \tag{35}$$

From these definitions we obtain directly

$$K_0 V(x) \le c^\top (x - x_*) \le K^0 V(x) \tag{36}$$

for all $x \in Q(x_0)$.

Let $\|a\|_\infty = \max\limits_{1 \le i \le n} |a^i|$ be the Chebyshev norm of a vector a.

Theorem 3.2. *Suppose the assumptions of Theorem* 3.1 *hold and* $x_0 \in X_1$, $x_0 \neq x_*$. *Then* $0 < \mathrm{K}_0 < K^0 = \|v_*\|_\infty < +\infty$ *and for all* $t \geq 0$ *the following estimates hold:*

$$V(x_0)e^{-K^0 t} \leq V(x(t, x_0)) \leq V(x_0)e^{-\mathrm{K}_0 t}. \tag{37}$$

Proof. Inequalities (37) follow from (34) and (36). Now we show that $\mathrm{K}_0 > 0$ and $K^0 = \|v_*\|_\infty < +\infty$.

Let $J_B(x_*) = \{1, 2, \ldots, m\}$. Then partitions (26) and (27) take place. The same partition we will use for arbitrary point $x \in Q(x_0)$.

Introducing new variables $z = [z^1, \ldots, z^d]$ such that $z = x^N$, we obtain

$$x^B = B^{-1}b - B^{-1}Nx^N = x_*^B - B^{-1}Nz. \tag{38}$$

The function $V(x)$ and the set $Q(x_0)$ can be rewritten, respectively, in the following forms:

$$\begin{aligned} \tilde{V}(z) &= -\sum_{i=1}^{m} x_*^i \ln \frac{1 - (B^{-1}Nz)^i}{x_*^i}, \\ \tilde{Q}(x_0) &= \{z \in R_+^m : B^{-1}Nz < x_*^B, \quad \tilde{V}(z) \leq V(x_0)\}. \end{aligned}$$

Since transformation (38) is linear, the function $\tilde{V}(z)$ and the set $\tilde{Q}(x_0)$ are convex. Also we have

$$\langle c, x - x_* \rangle = \langle c^N, x^N \rangle - \langle c^B, B^{-1}Nx^N \rangle = \langle v_*^N, z \rangle, \tag{39}$$

$$K_0 = \min_{z \in \tilde{Q}(x_0)} \frac{\langle v_*^N, z \rangle}{\tilde{V}(z)}, \qquad K^0 = \sup_{z \in \tilde{Q}(x_0)} \frac{\langle v_*^N, z \rangle}{\tilde{V}(z)}. \tag{40}$$

The dual nondegeneracy implies that $v_*^N > 0_d$.

Denote $\tilde{S}(x_0) = \{z \in \tilde{Q}(x_0) : \tilde{V}(z) = V(x_0)\}$. Since for any $\bar{z} \in \tilde{S}(x_0)$ the function $\tilde{V}(z)$ is convex on the closed interval connecting the origin and $\bar{z}$, we have $\tilde{V}(\alpha\bar{z}) \leq \alpha\tilde{V}(\bar{z})$, $0 \leq \alpha \leq 1$. Therefore,

$$K_0 = \min_{z \in \tilde{S}(x_0)} \frac{\langle v_*^N, z \rangle}{\tilde{V}(z)} = \frac{1}{V(x_0)} \min_{z \in \tilde{S}(x_0)} \langle v_*^N, z \rangle, \qquad K^0 = \lim_{z \to +0} \sup \frac{\langle v_*^N, z \rangle}{\tilde{V}(z)}. \tag{41}$$

Note that K^0 does not depend on the starting point x_0.

Let $Z = \{z \in R_+^d : \sum_{i=1}^{d} z^i = 1\}$, then using (41) we obtain

$$K^0 = \sup_{z \in Z} \lim_{\alpha \to +0} \sup \frac{\alpha \langle v_*^N, z \rangle}{\tilde{V}(\alpha z)} = \sup_{z \in Z} \frac{\langle v_*^N, z \rangle}{\langle B^{-1}Nz, e \rangle}, \tag{42}$$

where e is a vector of ones.

From (29) and (38) we have $1 - \langle x^N, e \rangle = \langle x^B, e \rangle = \langle x_*^B, e \rangle - \langle B^{-1}Nx^N, e \rangle = 1 - \langle B^{-1}Nx^N, e \rangle$. Therefore, $\langle B^{-1}Nx^N, e \rangle = \langle x^N, e \rangle$ and

$$K^0 = \sup_{z \in Z} \langle v_*^N, z \rangle = \max_{1 \leq i \leq d} (v_*^N)^i = \|v_*\|_\infty. \tag{43}$$

The solution of the dual problem is bounded; therefore, we conclude that $K^0 < +\infty$.

Now we will estimate K_0. The function $\langle v_*^N, z\rangle$ attains its minimal value on the set $\tilde{S}(x_0)$ at a point $\bar{z}_i = [0,\dots,0,\beta_i,0,\dots,0]^\top$, where $\beta_i > 0$, $1 \le i \le d$. Taking into account that $\bar{z}_i \in \tilde{S}(x_0)$ we obtain the following equation for determining the value β_i:

$$\sum_{j=1}^{m} x_*^j \ln\left[\frac{1-\beta_i(B^{-1}a_{m+i})^j}{x_*^j}\right] + V(x_0) = 0, \tag{44}$$

where a_s is the s-th column of A. Thus

$$K_0 = \frac{1}{V(x_0)} \min_{1\le i\le d} (v_*^N)^i \beta_i > 0 \tag{45}$$

because all β^i are strictly positive. $\square$

Obviously, that the solution β_i of equation (44) is greater than the solution $\bar{\beta}_i$ of the equation

$$\ln\left(1-\bar{\beta}_i\frac{(B^{-1}a_{m+i})^{j_i}}{x_*^{j_i}}\right) + V(x_0) = 0, \tag{46}$$

where

$$\frac{(B^{-1}a_{m+i})^{j_i}}{x_*^{j_i}} = \max_{j\in J_i^+} \frac{(B^{-1}a_{m+i})^{j}}{x_*^{j}}, \qquad J_i^+ = \{1 \le j \le m : (B^{-1}a_{m+i})^j > 0\}.$$

Hence

$$\beta_i \ge \bar{\beta}_i = \left(1-e^{-V(x_0)}\right)\frac{x_*^{j_i}}{(B^{-1}a_{m+i})^{j_i}}.$$

Substituting this inequality in (45) and denoting $J_N(x_*) = \{i \in [1:n] : x_*^i = 0\}$, we obtain in general case

$$K_0 \ge \frac{1-e^{-V(x_0)}}{V(x_0)} \min_{i\in J_N(x_*)}\left[v_*^i\frac{(B^{-1}b)^{j_i}}{(B^{-1}a_i)^{j_i}}\right].$$

Now we consider the discrete version (31) of the method.

Theorem 3.3. *Let the stepsize h_k in* (31) *be chosen as*

$$h_k = \sigma/\mu(x_k), \tag{47}$$

where $0<\sigma<1$, $\mu(x) = \max\limits_{1\le i\le n} v^i(x)$. Then for any $x_0 \in X_0$ there exists $0<\sigma(x_0)<1$ such that the following inequality

$$V(x_{k+1}) \le V(x_k)\left(1-\frac{h_k K_0}{2}\right) \tag{48}$$

holds for any $0<\sigma<\sigma(x_0)$. Here K_0 is defined by (35).

The proof is given in [11].

Denote

$$B(x_0) = \max_{x\in Q(x_0)} \max_{1\le i\le n} v^i(x).$$

If h_k is chosen in accordance with (47), then $h_k \ge \sigma(x_0)/B(x_0)$ for any $k \ge 0$. Thus, using (48), we obtain

$$V(x_{k+1}) \le V(x_0)\left[1-\frac{\sigma(x_0)K_0}{2B(x_0)}\right]^k.$$

The total number of iterations performed by algorithm (31) is no greater than $\bar{k}(x_0) = \left\lceil \frac{2B(x_0)}{K_0\sigma(x_0)} \ln\left(\frac{V(x_0)}{\varepsilon}\right)\right\rceil$, where $\varepsilon > 0$ denotes the tolerance for the Lyapunov function.

The total number of arithmetic operations at each iteration of algorithm (31) is essentially due to computation of the matrix $AD(x)A^\top$ and solution of the linear system (32). These computations requires $\approx m^2n/2 + m^3/6$ elementary operations. Since $m < n$, we can conclude that the computational cost of one step is roughly $2n^3/3$ arithmetic operations.

4. Barrier-Newton method for linear programming

In this section we apply barrier-Newton method (15) for solving linear programming Problem (18). In this case we have

$$\begin{aligned} \Lambda(x) &= [I_n - H(x)]\dot{G}(x)D(c - A^\top u(x)) + \tau H(x), &(49)\\ H(x) &= G(x)A^\top(AG(x)A^\top)^{-1}A. &(50)\end{aligned}$$

As in the previous section the vector-function $u(x)$ is found from linear equation (22).

Introduce a Lebesque level set in R^n

$$\begin{aligned} \Omega &= \{x \in R^n : x \geq 0_n, \quad \|Ax - b\| \leq \|Ax_0 - b\|, \\ 0_n &\leq G(x)(c - A^\top u(x)) \leq G(x_0)v_0\}, \end{aligned}$$

where x_0 is an initial point in (15), $v_0 = c - A^\top u_0,\ u_0 = u(x_0)$.

Theorem 4.1. *Suppose that the set Ω is compact and contains a unique stationary point x_*. Assume that the space transformation $\xi(y)$ satisfies $\mathbf{C}_2$ and is such that the matrix $\Lambda(x)$ is nonsingular everywhere on Ω. If starting point x_0 is such that $x_0 > 0_n,\ v_0 > 0_n$, then*

$$\lim_{t\to\infty} x(t, x_0) = x_*, \qquad \lim_{t\to\infty} u(x(t, x_0)) = u_*, \tag{51}$$

where x_*, u_* are the solutions of Problem (18) *and* (19), *respectively.*

Proof. If the matrix $\Lambda(x)$ is nonsingular, then from (49) and (50) we find that

$$A\Lambda(x) = \tau A, \qquad A = \tau A\Lambda^{-1}(\mathrm{x}). \tag{52}$$

The pair $[x_*, u_*]$ forms the Kuhn–Tucker pair in Problem (18); therefore, $x_* \in \Omega$. The solution of (15) exists at least for $t \geq 0$ such that $x(t, x_0) \in \Omega$, where the matrix $\Lambda(x)$ is nonsingular. Let us show that $x(t, x_0)$ does not leave the set Ω for any $t \geq 0$.

By differentiating $g(x)$ along the solutions of (15) and taking into account (13) we obtain

$$\frac{dg}{dt} = \gamma A\Lambda^{-1}(x)G(x)L_x(x, u(x)) = \frac{\gamma AG(x)L_x(x, u(x))}{\tau} = -\gamma g(x).$$

Hence system (15) has two first integrals:

$$\begin{aligned} Ax(t, x_0) &= b + (Ax_0 - b)\mathrm{e}^{-\gamma t}, &(53)\\ G(x(t, x_0))v(t) &= G(x_0)v_0 e^{-\gamma t}, &(54)\end{aligned}$$

where $v(t) = c - A^\top u(x(t, x_0))$.

The solution $x(t, x_0)$ of system (15) belongs to the compact set Ω for all $t \geq 0$. Hence this solution can be prolonged as $t \to \infty$. Since condition $\mathbf{C}_2$ holds, all components of vectors

$x(t, x_0)$, $v(t)$ can not change their sign. Therefore, the trajectory $x(t, x_0)$ do not cross the boundary of the set Ω. The function θ(x) this way play the role of the multiplicative barriers preserving nonnegativity. All trajectory remains in the set Ω. According to La Salle's Invariance Principle [2] the limit set of the solution is a compact connected set contained in Ω and coincides with the equilibrium point x_*, which is unique on Ω. Taking the limit as $t \to \infty$, we obtain from (53) and (54) that $Ax_* = b,\ G(x_*)v_* = 0_n,\ v_* \geq 0_n,\ x_* \geq 0_n$.

Due to condition $\mathbf{C}_2$ we have the complimentarity condition $x_*^i v_*^i = 0,\ 1 \leq i \leq n$. Hence we conclude that the pair x_*, u_* defined by (51) forms Kuhn–Tucker point in Problem (18). □

Integrating (15) using the Euler method, we obtain the following iterative process:

$$x_{k+1} = x_k - h\Lambda^{-1}(x_k)G(x_k)(c - A^\top u(x_k)), \tag{55}$$

where $h > 0$ is a stepsize and function $u(x)$ is defined by (22).

Each equilibrium point x_* of system (15) is a fixed point of iterations (55), i.e. $x_k = x_*$ implies $x_{k+1} = x_*$, and if iterates (55) converge to a regular point x_*, then the pair $[x_*, u(x_*)]$ satisfies the Kuhn–Tucker conditions.

If the conditions of Theorem 3.1 hold, then the matrix $\Lambda(x_*)$ is nonsingular. Therefore, if the stepsize h is fixed and $0 < h < 2$, then the discrete versions (55) 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_*.

References

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

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

3. Bayer, D.A. and Lagarias, J.C., "The nonlinear geometry of linear programming. Affine and projective scaling trajectories", *Trans. Amer. Math. Soc.*, Vol.314, No.2 (1989), pp. 499–526.

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

5. Evtushenko, Yu., "Two numerical methods of solving nonlinear programming problems", *Sov. Math. Dokl.*, Vol.15, No.2 (1974), pp. 420–423

6. Evtushenko, Yu., *Numerical Optimization Techniques*, Optimization Software, Inc. Publications Division, New York, 1985.

7. Evtushenko, Yu. and Zhadan, V., "Numerical methods for solving some operations research problems", *USSR Comput. Maths. Math. Phys.*, Vol.13, No.3 (1973), pp. 56–77.

8. Evtushenko, Yu. and Zhadan, V., "A relaxation method for solving problems of non-linear programming", *USSR Comput. Maths. Math. Phys.*, Vol.17, No.4 (1978), pp. 73–87.

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

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

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

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

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

14. Kallio, M., “On gradient projection for linear programming”, *Working Paper 94*, Yale School of Organization and Management, 1986.

15. Karmarkar, N., “A new polynomial-time algorithm for linear programming”, *Combinatorica*, No.4 (1984), pp. 373–395.

16. Luenberger, D., *Introduction to Linear and Nonlinear Programming*, Addison-Wesley Publishing Company, Reading, 1973.

17. Rosen, J., “The gradient projection method for nonlinear programming. Part I: linear constraints”, *SIAM J. Applied Math.*, Vol.8 (1960), pp. 181–217.

18. Tanabe, K., “A geometric method in nonlinear programming”, *Journal of Optimization Theory and Applications*, Vol. 30, No.2 (1980), pp. 181–210.

19. Vanderbei, R., Meketon, M., and Freedman, B., “A modification of Karmarkar’s linear programming algorithm”, *Algorithmica*, 1 (1986), pp. 395–407.

20. Wei, Zi-Luan, “An interior point method for linear programming”, *Journal of Computing Mathematics*, Oct. (l987), pp. 342–350.