Dynamics of non-homogeneous systems. Proceedings of ISA RAS, 1997, pp. 211–234

EVTUSHENKO Yu.G., ZHADAN V.G.
Computing Centre of the RAS, Moscow, Russia

INTERIOR POINT METHODS BASED ON SPACE TRANSFORMATION TECHNIQUE

Revised version 22 December 2003

1. Introduction

In the past twenty five years quite general and effective space transformation technique has been developed for solving LP and NLP problems. The idea of this approach commonly occurs in the optimization literature, it came from nonlinear programming and projective geometry. Using a space transformation, the original problem with equality and inequality constraints is reduced to a problem with equality constraints only. Continuous and discrete versions of the stable gradient projection method are applied to the reduced problem. After an inverse transformation to the original space, a class 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. Vector fields described by these equations define flows leading to the optimal solution. As a result of the space transformation, the vector fields are changed and additional terms are introduced which serve as a barrier preventing the trajectories from leaving the feasible set. Therefore, we call these methods "barrier-projection" methods. In our algorithms we use the multiplicative barrier functions which are continuous and equal to zero on a boundary. We do not introduce conventional singular barriers and this feature provides a high rate of convergence. In this paper we give a survey of principal results which were published in the last two decades [6]–[17].

In Section 2 we describe a unified methodology for finding necessary and sufficient optimality conditions in extremal problems with functional equality constraints and nonfunctional inequality constraints. We show how numerous families of algorithms can be developed using various space transformations.

In Section 3, choosing an exponential space transformation, we obtain the Dikin algorithm [5] from the family of primal barrier-projection methods. This algorithm, however, does not posses the local convergence properties and, as a result, it converges only if starting points belong to relative interior of the feasible set. Furthermore, the convergence rate of a discrete version of the algorithms proves to be weaker than a linear one.

In 1984 N. Karmarkar [24] proposed a special sophisticated step-length rule in the method similar to discrete version of the Dikin affine scaling algorithm. Based on this rule the polynomial complexity was theoretically attained. After this publication an impressive number of papers have been published devoted to further modifications and improvements of the Dikin and Karmarkar algorithms. Many authors were trying to modify and explain these algorithms as classical methods. Various methods were obtained along this direction and the first of our algorithms published in seventies [8, 6, 9] were reinvented. Later on in eighties-nineties, we developed more efficient versions of these methods which are discussed here. These asymptotically stable methods are such that a feasible set is an attractor of the vector fields. They preserve feasibility, but a starting point can be infeasible. They belong to a class of interior-infeasible point algorithms. In Section 3 we show that among the barrier-projection algorithms there is a method which converges locally and exponentially fast to the optimal solution (in discrete case it converges locally with a linear rate).

In subsequent sections we apply our approach to primal and dual linear programming problems. For the sake of simplicity, we assume that these problems have unique non-degenerate solutions. In Section 4 we use a nonconventional representation of the dual linear programming problem and we propose a set of algorithms. Upon simplifying the problem and choosing a particular exponential space transformation we arrive at the dual affine scaling method proposed by I. Adler, N. Karmarkar, M. Resende and G. Veiga [1].

2. Outline of a space transformation technique

Define the following NLP problem:

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

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.

Important particular cases of (1) are a linear programming problem given in standard form

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

and its dual problem

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

where v is the n-vector of slack variables; $A \in \mathbb{R}^{mn}$ $(m < n)$; $c, x \in \mathbb{R}^n$; $b, u \in \mathbb{R}^m$ and $\operatorname{rank}(A) = m$.

We define the relative interior set of X and the interior set of U as:

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

and assume that these sets are nonempty. We also introduce the following sets:

$$\begin{array}{rcl} \mathbb{R}^n_+ & = & \{x \in \mathbb{R}^n : x \geq 0_n\}, \quad \mathbb{R}^n_{++} = \operatorname{int} \mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x > 0_n\}, \\ V & = & \{v \in \mathbb{R}^n : \;\text{ there exists }\; u \in \mathbb{R}^m \;\text{ such that }\; v = c - A^\top u\}, \\ V_U & = & \{v \in \mathbb{R}^n : \;\text{ there exists }\; u \in U \;\text{ such that }\; v = c - A^\top u\}. \end{array}$$

The set V_U is the image of U under the mapping $v(u) = c - A^\top u$. Therefore, $V_U = V \cap \mathbb{R}^n_+$.

For convenience, assume that the primal feasible set is bounded and both primal and dual problems are non-degenerate, which together imply that optimal solutions x_*, u_* exist and are unique. We split the vectors x_* and $v_* = v(u_*)$ in basic and nonbasic components. Without loss of generality we assume that

$$x_* = \begin{bmatrix} x_*^B \\ x_*^N \end{bmatrix}, \quad v_* = \begin{bmatrix} v_*^B \\ v_*^N \end{bmatrix}, \qquad x_*^B > 0_m, \;\; x_*^N = 0_d, \;\; v_*^B = 0_m, \;\; v_*^N > 0_d,$$

where $d = n - m$.

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

In order to construct a family of computational methods for solving the Problems (1) – (3) we use an approach based on space transformation. 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 $\mathbb{R}^n$ onto P or $\operatorname{int} P$, i.e. $P = \overline{\xi(\mathbb{R}^n)}$, where $\overline{B}$ is the topological closure of B. Let $\tilde{J}(y) = dx/dy$ be the Jacobian matrix of the transformation $x = \xi(y)$ with respect to y.

Consider the transformed minimization problem

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

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

Define the Lagrangian functions $L(x, u)$, $\tilde{L}(y, u)$ associated with the Problems (1) and (4), respectively

$$L(x, u) = f(x) + u^\top g(x), \qquad \tilde{L}(y, u) = \tilde{f}(y) + u^\top \tilde{g}(y).$$

Then the first-order necessary conditions for a local minimum for the Problem (4) in the transformed space are

$$\tilde{L}_y(y, u) = \tilde{f}_y(y) + \tilde{g}_y^\top(y) u = 0_n, \qquad \tilde{g}(y) = 0_m, \tag{5}$$

where $\tilde{f}_y = \tilde{J}^\top f_x$, $\tilde{g}_y = g_x \tilde{J}$.

If $\tilde{J}$ is a nonsingular, 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. Using this substitution, we rewrite expressions (5) in terms of the variable x. They take the form

$$J^\top(x) L_x(x, u) = 0_n, \qquad g(x) = 0_m, \qquad x \in P. \tag{6}$$

Some properties of the nonlinear systems, which are obtained after space transformations, were investigated in [15].

Let $K(x \mid P)$ and $K^*(x \mid P)$, respectively, denote the cone of feasible directions at the point x relative to the set P and its dual:

$$K(x \mid P) = \{z \in \mathbb{R}^n : \exists \lambda(z) > 0 \ \text{ such that } \ x + \lambda z \in P, \quad 0 < \lambda \leq \lambda(z)\},$$
$$K^*(x \mid P) = \{z \in \mathbb{R}^n : z^\top y \geq 0 \ \ \forall y \in K(x \mid P)\}.$$

Let $S(x \mid P)$ be a linear hull of the cone $K^*(x \mid P)$. The set of all vectors orthogonal to $S(x \mid P)$ is called orthogonal complement of $S(x \mid P)$ and is denoted by $S^\perp(x \mid P)$.

We will impose the following condition on the space transformation $\xi(y)$.

Condition 1. *At each point $x \in P$ the matrix $J(x)$ is defined and the null-space of $J^\top(x)$ coincides with the set $S(x \mid P)$.*

In particular, it follows from this condition that at all interior points $x \in \operatorname{int} P$ the matrix $J(x)$ is non-degenerate, becoming singular only on the boundary of the set P.

Definition 1. *Any pair $[x, u]$ is a weak KKT pair for the Problem* (1) *if it satisfies the conditions* (6).

It follows from this definition and condition 1 that $L_x(x_*, u_*) \in S(x_* \mid P)$ at any weak KKT pair. Define the Gram matrix $G(x) = J(x)J^\top(x)$. Since the null-spaces of $G(x)$ and $J^\top(x)$ coincide, conditions (6) for the pair $[x_*, u_*]$ can be rewritten in the form

$$G(x)L_x(x,u) = 0_n, \qquad g(x) = 0_m, \qquad x \in P. \tag{7}$$

Definition 2. *A weak KKT pair $[x, u]$ is a KKT pair for the Problem* (1) *if $L_x(x,u) \in \in K^*(x \mid P)$.*

Let $\mathrm{ri}\, B$ denote a relative interior of the set B.

Definition 3. *A KKT pair $[x, u]$ is a strong KKT pair if $L_x(x_*, u_*) \in \mathrm{ri}\, K^*(x_* \mid P)$.*

Definition 4. *The constraint qualification (CQ) for the Problem* (1) *holds at a point $x \in P$ if all vectors $g_x^i(x)$, $1 \le i \le m$, and any nonzero vector $p \in S(x \mid P)$ are linearly independent. We say that x is a regular point for the Problem* (1) *if the CQ holds at x.*

Theorem 1. *Let a regular point x_* be a solution of the Problem* (1)*. Then there exists a vector $u_* \in \mathbb{R}^m$ such that $[x_*, u_*]$ forms a weak KKT pair for the Problem* (1)*.*

The space transformation described above can be used to derive the second-order sufficient conditions for a point x_* to be an isolated minimum in the Problem (1). Introduce a null-space $N(x) = \{z \in \mathbb{R}^n : g_x(x)J(x)z = 0_m\}$.

Theorem 2. *Assume that f and g are twice-differentiable functions and the space transformation $\xi(y)$ satisfies condition* 1*. Sufficient conditions for a point $x_* \in P$ to be an isolated local minimum of the Problem* (1) *are that there exists a strong KKT pair $[x_*, u_*]$ such that*

$$z^\top J^\top(x_*)L_{xx}(x_*, u_*)J(x_*)z > 0 \tag{8}$$

for every $z \in N(x_*)$ such that $\|J(x_*)z\| \neq 0$.

If $P = \mathbb{R}^n$ (in other words, the condition $x \in P$ is missing), we can take the trivial space transformation $x = y$. In this case we have

$$J(x) = I_n, \qquad N(x) = \{z \in \mathbb{R}^n : g_x(x)z = 0_m\},$$
$$K(x \mid \mathbb{R}^n) = \mathbb{R}^n, \qquad S(x \mid \mathbb{R}^n) = K^*(x \mid \mathbb{R}^n) = \mathrm{ri}\, K^*(x \mid \mathbb{R}^n) = 0_n.$$

The Theorem 2 reduces to the well-known second-order sufficient conditions for an isolated local minimum (see, for example, [19]).

Suppose that the function $\xi(y)$ is such that the matrix $G(x)$ is continuously differentiable. Let $p \in \mathbb{R}^n$ and $G_x(x;p)$ denote a square matrix of order n whose (i,j)-element equals to

$$G_x^{ij}(x;p) = \sum_{k=1}^{n} \frac{\partial G^{ik}(x)}{\partial x^j}\, p^k.$$

We impose two additional conditions on the space transformation $\xi(y)$:

Condition 2. *At each point $x \in P$ for any vector $p \in \mathrm{ri}\, K^*(x \mid P)$ the matrix $G_x(x;p)$ is symmetric and its null-space coincides with $S^\perp(x \mid P)$.*

Condition 3. *If $x \in P$, then $z^\top G_x(x;p)z > 0$ for any non-zero vector $z \in S(x \mid P)$ and any vector $p \in \mathrm{ri}\, K^*(x \mid P)$.*

Let us consider an important particular case of the Problem (1), where $P = \mathbb{R}^n_+$:

$$\text{minimize } f(x) \text{ subjectto } x \in X = \{x \in \mathbb{R}^n : g(x) = 0_m,\ x \geq 0_n\}. \tag{9}$$

It is convenient for this set to use a component-wise space transformation:

$$x^i = \xi^i(y^i), \qquad 1 \leq i \leq n. \tag{10}$$

For such a transformation the inverse transformation $y = \delta(x)$ is also component-wise type $y^i = \delta^i(x^i)$, $1 \leq i \leq n$, and the corresponding matrices $J(x)$ and $G(x)$ are diagonal: $J(x) = D(\gamma(x))$, $G(x) = D(\theta(x))$, where

$$\begin{array}{lcllcl} \gamma^\top(x) & = & [\gamma^1(x^1), \ldots, \gamma^n(x^n)], & \gamma^i(x^i) & = & \dot{\xi}(\delta^i(x^i)), \\ \theta^\top(x) & = & [\theta^1(x^1), \ldots, \theta^n(x^n)], & \theta^i(x^i) & = & [\gamma^i(x^i)]^2, \end{array} \qquad 1 \leq i \leq n.$$

Let $\sigma(x) = \{i : x^i = 0\}$ be a set of active indices at the point $x \in \mathbb{R}^n_+$. In this case

$$\begin{array}{lcl} K^*(x \mid \mathbb{R}^n_+) & = & \{z \in \mathbb{R}^n_+ : \text{ if } i \notin \sigma(x), \text{ then } z^i = 0,\ 1 \leq i \leq n\}, \\ S(x \mid \mathbb{R}^n_+) & = & \{z \in \mathbb{R}^n : \text{ if } i \notin \sigma(x), \text{ then } z^i = 0,\ 1 \leq i \leq n\}, \end{array}$$

Condition 1 reduces to the following

Condition 4. *The vector function $\gamma(x)$ is defined at each point $x \in \mathbb{R}^n_+$ and $\gamma^i(x^i) = 0$ if and only if $i \in \sigma(x)$.*

In order to insure Conditions 2 and 3 we impose the

Condition 5. *The vector function $\theta(x)$ is differentiable in some neighborhood of $\mathbb{R}^n_+$ and $\dot{\theta}^i(0) > 0$, $1 \leq i \leq n$.*

As a rule, we perform the following quadratic and exponential transformations

$$x^i = \xi^i(y^i) = \frac{1}{4}(y^i)^2, \quad J(x) = D^{1/2}(x), \quad G(x) = D(x), \tag{11}$$

$$x^i = \xi^i(y^i) = e^{y^i}, \quad J(x) = D(x), \quad G(x) = D^2(x). \tag{12}$$

In these two cases the Jacobian matrix is singular on the boundary of the set $P = \mathbb{R}^n_+$. These transformations satisfy Condition 4. The Condition 5 holds only for the quadratic transformation (11).

Let e^i denote the n-th order unit vector whose i-th component is equal to one. The CQ for the Problem (9) holds at a point x, if all the vectors $g^i(x)$, $1 \leq i \leq m$, and all e^j, such that $j \in \sigma(x)$, are linearly independent. The cone $N(x)$ takes the form $N(x) = \{z \in \mathbb{R}^n : g_x(x)D(\gamma(x))z = 0_m\}$.

The strict complementary condition (SCC) holds at a pair $[x, u]$, if $L^i_x(x, u) > 0$ for all $i \in \sigma(x)$. From the Theorem 2 the following second-order sufficient optimality conditions is obtained.

Theorem 3. *Sufficient conditions for a point $x_* \in X$ to be an isolated local minimum of the Problem (9) are that there exists a Lagrange multiplier vector u_* such that*

$$D(\gamma(x_*))L_x(x_*, u_*) = 0_n,$$

that the SCC holds at $[x_*, u_*]$ and that

$$z^\top D(\gamma(x_*))L_{xx}(x_*, u_*)D(\gamma(x_*))z > 0$$

for all $z \in N(x_*)$, *satisfying* $\|D(\gamma(x_*))z\| \neq 0$.

Now we will construct numerical methods for solving the Problem (1). We use the stable version of the gradient projection method for solving the Problem (4). The numerical method is stated as an initial-value problem involving the following system of ordinary differential equations:

$$\frac{dy}{dt} = -\tilde{L}_y(y, \tilde{u}(y)), \quad \tilde{L}_y(y, \tilde{u}) = \tilde{f}_y(y) + \tilde{g}_y^\top(y)\tilde{u}, \quad y(0, y_0) = y_0 \in \mathbb{R}^n. \tag{13}$$

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

$$\frac{d\tilde{g}}{dt} = \tilde{g}_y \frac{dy}{dt} = -\tau \tilde{g}(y), \qquad \tau > 0. \tag{14}$$

From this condition we obtain the system of linear algebraic equations

$$\tilde{g}_y(y)\tilde{g}_y^\top(y)\tilde{u}(y) + \tilde{g}_y(y)\tilde{f}_y(y) = \tau\tilde{g}(y),$$

where $\tilde{f}_y = \tilde{J}^\top f_x$, $\tilde{g}_y = g_x\tilde{J}$. By differentiating $\xi(y)$ with respect to y and taking into account (13) and (14), we have

$$\frac{dx}{dt} = \frac{d\xi}{dy}\frac{dy}{dt} = J(x)\frac{dy}{dt} = -G(x)L_x(x, u(x)), \qquad x(0, x_0) = x_0 \in P, \tag{15}$$

$$\Gamma(x)u(x) + g_x(x)G(x)f_x(x) = \tau g(x), \tag{16}$$

where $\Gamma(x) = g_x(x)G(x)g_x^\top(x)$.

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

Proof. We will show that the rank of the matrix $B(x) = J^\top(x)g_x^\top(x)$ is equal to m. If $x \in \operatorname{int} P$, it is obvious, since the matrix $J(x)$ is non-singular, and by the CQ the rank of the matrix $g_x^\top(x)$ is equal to m.

Now let $x \in \operatorname{fr} P$. If the rank of the matrix $B(x)$ is less then m, there is a non-zero vector $z \in \mathbb{R}^n$ such that $B(x)z = J^\top(x)g_x^\top(x)z = 0$. According to the Condition 1 the non-zero vector $p = g_x^\top(x)z$ belongs to the space $S(x \mid P)$. Thus the vectors $g_x^i(x)$, $1 \leq i \leq m$, and the vector p are linearly dependent which contradicts to the CQ. Therefore, $B(x)$ is the full rank matrix. □

Lemma 2. *Let the space transformation* $\xi(y)$ *satisfy Condition* 1. *Then the regular point* x_* *is an equilibrium state of system* (15) *if and only if the pair* $[x_*, u_*]$, *where* $u_* = u(x_*)$, *is a weak KKT pair for the Problem* (1).

Proof. Suppose that x_* is an equilibrium state of the system (15). Then

$$G(x_*)L_x(x_*, u_*) = 0,$$

and, therefore, $L_x(x_*, u_*) \in \ker G(x_*)$. But the rank of both matrices $J^\top(x_*)$ and $G(x_*) = J(x_*)J^\top(x_*)$ is equal to m. Thus, $L_x(x_*, u_*) \in \ker J^\top(x_*)$ and, by the Condition 1, $L_x(x_*, u_*) \in K^*(x_* \mid P)$. The equality $g(x_*) = 0$ follows from (16). □

Hence, corresponding to any regular point x_* we can define a corresponding Lagrange multiplier $u(x_*)$ by solving linear algebraic equations (16). If a local solution of the original problem (1) occurs at a regular point $x_* \in P$, then $[x_*, u(x_*)]$ forms a weak KKT pair for the Problem (1) and x_* is an equilibrium state of (15).

Let W be a $m \times n$ rectangular matrix whose rank is m. We introduce the pseudo-inverse matrix $W^+ = W^\top(WW^\top)^{-1}$ and the orthogonal projector $\pi(W) = I_n - W^+W$. If at a regular point x we define $u(x)$ and substitute it into the right-hand side of (15), then (15) can be rewritten in the following projective form:

$$\frac{dx}{dt} = -J(x)\left\{\pi[g_x(x)J(x)]J^\top(x)f_x(x) + \tau[g_x(x)J(x)]^+g(x)\right\}. \tag{17}$$

Let $x(t, x_0)$ denote the solution of the Cauchy Problem (15) with an initial condition $x_0 = x(0, x_0)$. In what follows, we assume that the initial-value problem under consideration is always uniquely solvable. A trajectory $x(t, x_0)$ can be continued as long as points on it are regular. By differentiating $f(x(t, x_0))$ with respect to t we arrive at

$$\frac{df}{dt} = -\left\|J^\top(x)L_x(x, u(x))\right\|^2 + \tau u^\top(x)g(x). \tag{18}$$

Hence the objective function $f(x(t, x_0))$ monotonically decreases either on the feasible set X or when the trajectory is close to X, i.e. $\|g(x(t, x_0))\|$ is sufficiently small.

The system of ordinary differential equations (15), where $u(x)$ is given by (16), has the first integral

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

This means that if $\tau > 0$, (15) 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 (see [7, 10, 27]). Therefore, we can call the method (17) "the stable version of the barrier-projection method". If $x_0 \in X$, then the trajectory $x(t, x_0)$ of (15) remains on this manifold because $g(x(t, x_0)) \equiv 0_m$ for all $t \geq 0$ and the trajectories of (15) coincide with the trajectories of the following system:

$$\frac{dx}{dt} = -J(x)\pi[g_x(x)J(x)]J^\top(x)f_x(x), \tag{19}$$

which can be obtained from (17) if we put $\tau = 0$. But in contrast to (17) this system is neutrally stable with respect to equality constraints. It means that if $g(x_0) = c$, then $g(x(t, x_0)) \equiv c$ for all $t \geq 0$ and we have to introduce a correction procedure to remove the violation of constraints. If the condition $x \in P$ is missing, then we can put $\xi(y) = y$, hence $J(x) = I_n$ and (19) coincides with well-known gradient projection method [25], which is also neutrally stable.

Note that, according to the Condition 1, the subspace $S^\perp(x \mid P)$ coincides with the space of the columns of the matrix $J(x)$, and, since the vector $G(x)L_x(x, u(x))$ belongs to this space, the velocity vector $\dot{x}$ always lies in $S^\perp(x \mid P)$. Thus, if x is a boundary point of P, then the vector $\dot{x}$ belongs to the eigen subspace of the space $\mathbb{R}^n$, which coincides with the space $M(x) - x$, where $M(x)$ is the intersection of the support planes of the set P at the point x. If the cone $K^*(x \mid P)$ has a non-empty interior, then this subspace degenerates to a single point (the origin of coordinates).

Let us show that trajectories $x(t, x_0)$ of (15) cannot cross the boundary of P. To the contrary, suppose this is not true and a trajectory $x(t, x_0)$ starting inside P leaves P at $t_1 > 0$. Then there exists a vector $p \in K^*(x(t_1, x_0) \mid P)$ such that $p^\top \dot{x}(t_1, x_0) < 0$. But the vector $\dot{x}(t, x_0)$ belongs to the orthogonal complement of the subspace $S(x \mid P)$ and the vector p belongs to $S(x \mid P)$. Hence $p^\top \dot{x}(t_1, x_0) = 0$ and, consequently, $x(t, x_0) \in P$ for all $t \geq 0$.

Thus the matrix $G(x)$ plays the role of a "barrier" preventing $x(t, x_0)$ from intersecting the boundary of P. The trajectory $x(t, x_0)$ can approach the boundary points only in the limit as $t \to +\infty$. If the initial point x_0 lies on the boundary, then the entire trajectory of system (15) belongs to the boundary of P.

By applying the Euler method for solving system (17), we obtain

$$x_{k+1} = x_k - \alpha_k J(x_k) \left\{ \pi[g_x(x_k)J(x_k)]J^\top(x_k)f_x(x_k) + \tau[g_x(x_k)J(x_k)]^+ g(x_k) \right\}, \tag{20}$$

where a step-size $\alpha_k > 0$.

Each equilibrium point x_* of the system (17) is a fixed point of iterations (20), i.e. $x_k = x_*$ implies $x_{k+1} = x_*$, and if iterates (20) converge to a regular point x_*, then the pair $[x_*, u(x_*)]$ satisfies conditions (7).

Theorem 4. *Let $[x_*, u_*]$ be a weak KKT pair of the Problem* (1)*, where the CQ and the second-order sufficiency conditions of the Theorem* 2 *hold. Let the space transformation $\xi(y)$ satisfy the Conditions* 1 – 3 *and $\tau > 0$. Then x_* is an asymptotically stable equilibrium state of the system* (17)*; there exists a positive number α_* such that for any fixed $0 < \alpha_k < \alpha_*$ the sequence $\{x_k\}$, generated by* (20)*, converges locally with a linear rate to x_* while the corresponding sequence $\{u_k\}$, where $u_k = u(x_k)$, converges to u_*.*

Proof. The proof is based on the Lyapunov linearization principle. The conditions for asymptotic stability to be valid are expressed in terms of eigenvalues determined by a matrix arisen from the right-hand side of (17) linearized about the equilibrium point x_*. The asymptotic stability of the point x_* implies the local, exponentially fast convergence of a trajectory $x(t, x_0)$ to the optimal solution x_*.

Denote $\delta x(t) = x(t, x_0) - x_*$ and linearize system (17) in the neighborhood of the point x_*. Then we obtain the equation of the first approximation of (17) about the equilibrium point x_*:

$$\delta \dot{x} = -Q(x_*, u_*)\delta x, \tag{21}$$

where

$$Q(x, u) = \tilde{M}(x) \left[G(x)L_{xx}(x, u) + G_x(x; L_x(x, u)) \right] + \tau G(x)\tilde{P}(x),$$
$$\tilde{M}(x) = I_n - G(x)\tilde{P}(x), \qquad \tilde{P} = G g_x^\top (g_x G g_x^\top)^{-1} g_x.$$

The stability of system (21) is determined by the properties of the roots of the characteristic equation

$$\det\ (Q(x_*, u_*) - \lambda I_n) = 0. \tag{22}$$

Suppose that the point x_* is such that the rank of the matrix $G(x_*)$ is equal to s, where $s < n$. Since $G(x_*)$ is the Gram matrix, there exists an orthogonal matrix U such that $G(x_*) = UHU^\top$ and the matrix H has the form

$$H = \left[\begin{array}{cc} H^B & 0_{s,n-s} \\ 0_{n-s,s} & 0_{n-s,n-s} \end{array} \right].$$

Here H^B is a diagonal matrix of order s, whose diagonal elements are the positive eigenvalues of $G(x_*)$. Moreover, since the matrix $G_x(x_*; L(x_*, u_*))$ is symmetric and the matrices $G(x_*)$ and $G_x(x_*; L(x_*, u_*))$ commutate, the matrix U can be chosen in such a way that the matrix $Y = U^\top G_x U$ will also be diagonal. Thus $Q(x_*, u_*)$ can be represented in the form

$$Q(x_*, u_*) = UWU^\top, \qquad W = (I_n - HU^\top \tilde{P} U)(HU^\top L_{xx} U + Y) + \tau HU^\top \tilde{U}$$

and, therefore, its eigenvalues are equal to the eigenvalues of the matrix W.

Let the V and $V^\perp$ be the null-space of the matrix H and its orthogonal complement, respectively:

$$\begin{array}{rcl} V & = & \ker H = \{y \in \mathbb{R}^n : y^1 = \ldots = y^s = 0\}, \\ V^\perp & = & \{y \in \mathbb{R}^{s+1} : y^{s+1} = \ldots = y^n = 0\}. \end{array}$$

The following relations between the subspaces $S(x_* \mid P)$, $S^\perp(x_* \mid P)$ and V, $V^\perp$ take place:

$$S(x_* \mid P) = UV, \qquad S^\perp(x_* \mid P) = UV^\perp.$$

If $z \in S^\perp(x_* \mid P)$, then, according to Condition 2, $G_x z = 0$. Thus, $G_x U y = 0$ for all $y \in V^\perp$. It follows from here that the matrix Y has the form

$$Y = [0_{n,s}, B], \qquad B^\top = [0_{n-s,s}, C],$$

where C is a diagonal non-degenerate matrix of order $n - s$.

Denote by U^B and U^N submatrices of the matrix U consisting of the first s and last $n - s$ columns of U, respectively. Denote also by H^B the left-hand upper square submatrix of order s of the matrix H. Let

$$\tilde{P}^B = (U^B)^\top g_{x_B}^\top [g_{x_B} U^B H^B (U^B)^\top g_{x_B}^\top]^{-1} g_{x_B} U^B, \qquad L_{xx}^B = (U^B)^\top L_{xx} U^B.$$

Then the matrix W can be decomposed into the following blocks:

$$W = \left[\begin{array}{cc} W_1 & W_3 \\ 0_{n-s,s} & W_2 \end{array} \right],$$

where the matrix W_3 is not essential and

$$W_1 = (I_s - H^B \tilde{P}^B) H^B L_{xx}^B + \tau H^B \tilde{P}^B, \qquad W^2 = (U^N)^\top G_x U^N.$$

The characteristic equation for the matrix W splits into two equations:

$$|W_1 - \lambda_i I_s| = 0, \quad |W_2 - \lambda_j I_{n-s}| = 0, \qquad 1 \leq i \leq s, \quad s+1 \leq j \leq n. \tag{23}$$

At first we will find the solution of the second equation. Let $z_j \in \mathbb{R}^{n-s}$ be the eigenvector, corresponding to the eigenvalue λ_j. Then the following equalities take place:

$$(U^N)^\top G_x U^N z_j = \lambda_j z_j, \qquad s+1 \leq j \leq n. \tag{24}$$

Multiplying both sides of (24) by $z_j^\top$, we have

$$z_j^\top (U^N)^\top G_x U^N z_j = \lambda \|z_j\|^2, \qquad s+1 \leq j \leq n.$$

The vectors $h_j = U^N z_j$ belong to the subspace $S(x_* \mid P)$. Therefore, according to Condition 3, the eigenvalues

$$\lambda_j = z_j^\top (U^N)^\top G_x U^N \frac{z_j}{\|z_j\|^2} = \frac{h_j^\top G_x h_j}{\|z_j\|^2} \tag{25}$$

are real and strictly positive. Denote

$$\hat{\lambda}_1 = \min_{s+1 \leq j \leq n} \lambda_j > 0.$$

Let A be the square root of the matrix H, and let Λ^B be the upper left-hand square submatrix of Λ of the order s. Instead of finding the roots of the first equation (23) we will find the eigenvalues of the matrix $\Omega_1 = (\Lambda^B)^{-1} W_1 \Lambda^B$ which is similar to the matrix W_1. We have the following characteristic equation:

$$|\Omega_1 - \lambda_i I_s| = |\hat{M}\hat{L}^B_{xx} + \tau\hat{P} - \lambda_i I_s| = 0, \qquad 1 \le i \le s,$$

where

$$\hat{M} = I_s - \hat{P}, \qquad \hat{P} = \Lambda^B \tilde{P}^B \Lambda^B, \qquad \hat{L}^B_{xx} = \Lambda^B L^B_{xx} \Lambda^B.$$

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

$$\hat{K}(x_*) = \{\bar{x} \in \mathbb{R}^s : g_{x_B}(x_*) U^B \Lambda^B \bar{x} = 0_m\}$$

and its orthogonal complement $\hat{K}^\perp(x_*)$, respectively. Furthermore we have

$$\hat{P}\hat{M} = 0_{ss}, \qquad \hat{P}\hat{P} = \hat{P}, \qquad \hat{M}\hat{M} = \hat{M}. \tag{26}$$

Let λ_i and z_i be an eigenvalue and a corresponding nonzero eigenvector of the matrix Ω_1. Then

$$(\hat{M}\hat{L}^B_{xx} + \tau\hat{P}) z_i = \lambda_i z_i, \qquad z_i \in \mathbb{R}^s. \tag{27}$$

If $\hat{P} z_i \ne 0_s$, then premultiplying (27) by the matrix $\hat{P}$ and taking into account (26), we obtain $\lambda_i = \tau$. If $\hat{P} z_i = 0_s$, then $z_i \in \hat{K}(x_*)$ and multiplying (27) by $z_i^\top$ we have

$$\lambda_i = z_i^\top \Lambda^B L^B_{xx}(x_*, u_*) \Lambda^B \frac{z_i}{\|z_i\|^2}. \tag{28}$$

Now we will take into account that $U^B \Lambda^B z_i = U \Lambda h_i = J(x_*) h_i$ for some vector $h_i \in \mathbb{R}^s$, whose first s components ate equal to the corresponding components of the vector z_i. Then (28) can be rewritten in the form

$$\lambda_i = \frac{h_i^\top J^\top(x_*) L_{xx} J(x_*) h_i}{\|z_i\|^2}. \tag{29}$$

It follows from $z_i \in \hat{K}(x_*)$, $z_i \ne 0_s$, that $h_i \in N(x_*)$ and $J(x_*) h_i \ne 0_n$. Taking into account (8) we conclude that any eigenvalue λ_i corresponding to the eigenvector z_i from the tangent manifold $\hat{K}(x_*)$ is strictly positive and

$$\hat{\lambda}_2 = \min_{i \in \Delta(x_*)} \lambda_i > 0, \qquad \Delta(x_*) = \{i : z_i \in \hat{K}(x_*)\}.$$

These results imply that all roots of the characteristic equation for the matrix Q are real and the smallest root $\lambda_* = \min[\hat{\lambda}_1, \hat{\lambda}_2, \tau]$ is positive. Hence, according to Lyapunov's linearization principle [2], the equilibrium point x_* is asymptotically stable and the following estimation holds:

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

Denote $\alpha_* = \dfrac{2}{\lambda^*}$, where $\lambda^* = \max\limits_{1 \le i \le n} \lambda_i$. If the step-size $\alpha_k < \alpha_*$, then by Theorem 2.3.7 in [7] the linear convergence of the discrete version (20) follows from the proof given above. $\square$

The Theorem 4, being applied to the Problem (9), gives the the following statement.

Theorem 5. *Let $[x_*, u_*]$ be a weak KKT pair of the Problem (9), where the CQ and the second-order sufficiency conditions of the Theorem 3 hold. Let the component-wise space transformation (10) satisfy the Conditions 4, 5 and $\tau > 0$. Then x_* is an asymptotically stable equilibrium state of system (17); there exists a positive number α_* such that for any fixed $0 < \alpha_k < \alpha_*$ the sequence $\{x_k\}$, generated by (20), converges locally with a linear rate to x_* while the corresponding sequence $\{u_k\}$ converges to u_*.*

This theorem was proved in [10, 12].

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

$$\text{minimize } f(x) \text{ subject to } x \in X = \{x \in \mathbb{R}^n : g(x) = 0_m, \;\; h(x) \leq 0_c\}, \tag{30}$$

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

Space transformation 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:

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

Then the original Problem (30) is transformed into the equivalent problem

$$\text{minimize } f(x) \text{ subject to } z \in Z = \{z \in \mathbb{R}^{n+c} : \Phi(z) = 0_q, \;\; p \in \mathbb{R}^c_+\}. \tag{31}$$

Here P is a positive orthant in $\mathbb{R}^c$. This problem is similar to (9). In order to take into account the constraint $p \geq 0_c$ we introduce a surjective component-wise differentiable mapping $\xi : \mathbb{R}^c \to \mathbb{R}^c_+$ and make the space transformation $p = \xi(y)$, where $y \in \mathbb{R}^c$, $\overline{\xi(\mathbb{R}^c)} = \mathbb{R}^c_+$. Let $\xi_y(y)$ be the square $c \times c$ Jacobian matrix of the mapping $\xi(y)$ 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 diagonal matrices:

$$J(p) = \xi_y(y)|_{y=\psi(p)} = D(\gamma(p)), \qquad G(p) = J(p)J^\top(p) = D(\theta(p)).$$

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

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

The Problem (31) can be formulated as follows:

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

In the last Problem we have only equality constraints, therefore we can use the numerical method described above. After an inverse transformation to the space of x and p we obtain from (15), (16)

$$\frac{dz}{dt} = -\tilde{G}(p)L_z(z, w(z)). \tag{33}$$

Here

$$\begin{gathered} L(z, w) = f(x) + w^\top \Phi(z), \qquad L_z(z, w) = f_z(z) + \Phi_z^\top(z) w, \\ \Phi_z(z)\tilde{G}(p)L_z(z, w(z)) = \tau\Phi(z), \end{gathered} \tag{34}$$

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

System (33) can be rewritten in the more detailed form

$$\frac{dx}{dt} = -L_x(x, w(z)), \qquad \frac{dp}{dt} = -G(p)v(z), \tag{35}$$

where the function $w^\top(z) = [u^\top(z), v^\top(z)]$ is found from the following linear system of q equations:

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

Condition (34) can be written as

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

Therefore, system (35) has two first integrals:

$$g(x(t,z_0)) = g(x_0)e^{-\tau t}, \quad h(x(t,z_0)) + p(t,z_0) = (h(x_0)+p_0)e^{-\tau t}, \qquad z_0^\top = [x_0^\top, p_0^\top]. \tag{36}$$

Similarly to (18) we obtain

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

Consider the simplified version of method (35). Suppose that along the trajectories of system (35) 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 (35) the additional vector p and integrate the system which does not employ this vector:

$$\frac{dx}{dt} = -L_x(x, w(x)), \tag{37}$$

where

$$\Gamma(x)w(x) + \Phi_x(x)f_x(x) = \tau \begin{bmatrix} g(x) \\ 0_c \end{bmatrix}, \quad \Gamma(x) = \Phi_x(x)\Phi_x^\top(x) + \begin{bmatrix} 0_{mm} & 0_{mc} \\ 0_{cm} & G(-h(x)) \end{bmatrix}. \tag{38}$$

Along the trajectories of (37) we have

$$\begin{aligned} &\frac{dg}{dt} = -\tau g(x), \qquad \frac{dh}{dt} = -G(-h(x))v(x), \\ &\frac{df}{dt} = -\|L_x(x,w(x))\|^2 - \|J^\top(-h(x))v(x)\|^2 + \tau u^\top(x)g(x). \end{aligned} \tag{39}$$

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 to the contrary that this is not true and $h^i(x(t,x_0)) > 0$ for some $t > 0$. Then there is an earlier instant $0 < t_1 < t$ such that $h^i(x(t_1,x_0)) = 0$ and $\dot{h}^i(x(t_1,x_0)) > 0$. This contradicts (39) since $\theta^i(0) = 0$. Hence $x(t,x_0) \in X$ for all $t \geq 0$. Thus the matrix $G(-h(x))$ plays the role of a "barrier" preventing $x(t,x_0)$ from intersecting the hypersurface $h^i(x) = 0$.

The method (37) is closely related to (15). Let us consider the Problem (9). We have the two alternatives ensuing from (15) or (37). The main body of computational work required when using any numerical integration method is to evaluate the right-hand sides of equations for various values of x. This could be done by solving the linear system (16) of m equations or the system (38) of $m+n$ equations, respectively. One might expect that introducing a slack variable p increases the computational work considerably. However, with allowance made for the the simple structure of (38), we can reduce the computational time by using the Frobenius formula for an inverse matrix Upon some transformations we find that (37), (38) can be written in the form of (15), (16), respectively, if we put

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

Therefore, the performances of both methods which seem at a glance to be unrelated are in fact very similar.

3. Primal barrier-projection methods for solving LP problem

Applying the stable barrier-projection method (15) for solving the Problem (2), we obtain the following continuous and discrete versions:

$$\frac{dx}{dt} = -G(x)[c - A^\top u(x)], \qquad x(0, x_0) = x_0, \tag{40}$$

$$x_{k+1} = x_k - \alpha_k G(x_k)[c - A^\top u_k], \qquad u_k = u(x_k), \tag{41}$$

where the function $u(x)$ is found from the linear equation (16) which can be rewritten as

$$AG(x)A^\top u(x) - AG(x)c = \tau(b - Ax). \tag{42}$$

By differentiating the objective function with respect to t we obtain

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

The system of ordinary differential equations (40) has the first integral

$$Ax(t, x_0) = b + (Ax_0 - b)e^{-\tau t}. \tag{43}$$

Under the non-degeneracy assumption all feasible points are regular, and each weak KKT pair $[x, u(x)]$ is such that x is an equilibrium state of system (40). The pair $[x, u(x)]$ is a strong KKT pair if and only if $x = x_*$.

Denote $\alpha_* = 2/\mu^*$,

$$\mu^* = \max\left[\tau, \max_{m+1 \le i \le n} \dot{\theta}^i(x_*^i)v_*^i\right], \qquad \mu_* = \min\left[\tau, \min_{m+1 \le i \le n} \dot{\theta}^i(x_*^i)v_*^i\right].$$

Here μ^* and μ_* are, respectively, the largest and smallest eigenvalues of the first approximation matrix of the right-hand side of (40) at the optimal solution x_*.

Theorem 6. *Let x_*, u_* be unique non-degenerate solutions of the Problems* (2) *and* (3), *respectively. Assume that the component-wise space transformation $\xi(y)$ satisfy the Conditions* 4, 5 *and $\tau > 0$. Then the following statements are true:*

1) The point x_* is an asymptotically stable equilibrium state of system (40)*.*

2) The solutions $x(t, x_0)$ of system (40) *converge locally exponentially fast to the optimal point x_*.*

3) For any fixed $0 < \alpha_k < \alpha_*$ the sequence $\{x_k\}$, generated by (41)*, converges locally with a linear rate to x_* while the corresponding sequence $\{u_k\}$ converges to u_*.*

4) All extreme feasible points of X are unstable equilibrium points of (40) *and* (41)*, except the optimal solution x_*.*

The exponential rate of convergence of the solution $x(t, x_0)$ of (40) to the equilibrium point x_* implies that there exist a neighborhood $\Delta(x_*)$ of x_* and a constant $C > 0$ such that

$$\|x(t, x_0) - x_*\| \leq C\|x_0 - x_*\|e^{-\mu_* t}$$

for all $x_0 \in \Delta(x_*)$, $t > 0$.

In (40) we can take a starting vector x_0 such that $\|Ax_0 - b\| \neq 0$. Moreover, if the Condition 5 holds then the components of x_0 corresponding to nonbasic components of the vector x_* may be negative. In this case the entire trajectory $x(t, x_0)$ is infeasible. Nevertheless, owing to the local convergence property, the trajectory $x(t, x_0)$ converges to x_*, if $\|x_0 - x_*\|$ is sufficiently small. It follows from (43) that, if initially $Ax_0 = b$, then $Ax(t, x_0) \equiv b$ for all $t \geq 0$. Still further if $x_0 \geq 0_n$, then the trajectory $x(t, x_0)$ of (40) remains in the feasible set X and the objective function monotonically decreases along $x(t, x_0)$. (40) preserves feasibility, hence it is an interior point method. If Condition 5 holds, then this method allows to start the computation from an infeasible point. Therefore, we can call it an interior-infeasible method.

If we use the quadratic and exponential space transformations, then (40) and (41) can be cast in the form

$$\frac{dx}{dt} = D^\beta(x)(A^\top u(x) - c), \qquad AD^\beta(x)A^\top u(x) = AD^\beta(x)c + \tau(b - Ax), \tag{44}$$

$$x_{k+1} = x_k + \alpha_k D^\beta(x_k)[A^\top u_k - c], \tag{45}$$

where $\beta = 1$ for (11), and $\beta = 2$ for (12).

The continuous version of (44) had been studied by Smirnov [26]. He proved that the following estimate:

$$\|x^N(t, x_0)\| = \begin{cases} O(e^{-\mu_* t}), & \text{if } \beta = 1, \\ O(t^{1/(1-\beta)}), & \text{if } \beta > 1 \end{cases}$$

holds for the vector of nonbasic components. Hence the trajectories of (44) derived with the help of the quadratic space transformation locally converge faster than the trajectories of (44), derived with the help of the exponential transformation. Therefore, in our codes we have mainly used the quadratic space transformation.

It was shown that if β is odd, then x_* is an asymptotically stable point of (44) and the local convergence takes place. If β is even, then the optimal point x_* is unstable and trajectories converge to x_* provided that $x_0 \in \Delta(x_*) \cap X_0$.

If we set $\beta = 2$, $\tau = 0$, then (44) leads to the Dikin affine scaling method

$$\frac{dx}{dt} = D^2(x)(A^\top u(x) - c), \qquad AD^2(x)A^\top u(x) = AD^2(x)c. \tag{46}$$

Discrete and continuous versions of (46) were investigated in numerous papers (see, for example, [3, 4, 5, 22, 23, 28, 29]). It is worthy of note that the optimal solution x_* proves to be unstable equilibrium point and, therefore, (46) does not possess the local convergence property. In the latter case a starting point x_0 should meet a constraint $x_0 \in \Delta(x_*) \cap X_0$ for the convergence to take place.

If we set $\beta = 1$, $\tau = 0$, then (44) yields

$$\frac{dx}{dt} = D(x)(A^\top u(x) - c), \qquad AD(x)A^\top u(x) = AD(x)c. \tag{47}$$

Since the quadratic space transformation is used here, the Condition 5 holds and we have the exponential rate of convergence (in discrete case a linear rate). However owing to the simplification $\tau = 0$ a starting point x_0 must be such that $x_0 \in \Delta(x_*)$, $Ax_0 = b$. The convergence property of the method (47) is more attractive, than (46). Therefore, the method (47) is considered in the West as essential improvement of Dikin–Karmarkar method. The book [21] claims that (47) "is identical with the gradient flow for linear programming proposed and studied extensively by Faybusovich (1991). Independently this flow was also studied by Herzel, Recchioni and Zirilli (1991)... Starting from the seminal work of Khachian and Karmarkar, there has been a lot of progress in developing ulterior point algorithms for linear programming and nonlinear programming, due to Bayer, Lagarias and Faybusovich".

In point of fact, (47) has been proposed and analyzed as far back as 1977 in [9], and later on a comprechensive exposition of this method was described in numerous our papers and in the book [7] appeared in English in 1985 and three years before in Russian.

Still further, the nonlocal convergence analysis of (47) was carried out in [14]. We assumed that a condition $\sum_{i=1}^{n} x^i = 1$ was introduced among the other equality constraints. In this case the Lyapunov function

$$V(x) = \sum_{i=1}^{m} x_*^i [\ln x_*^i - \ln x^i] \tag{48}$$

decreases along the trajectories of the system (47) insofar as

$$\frac{dV}{dt} = c^\top (x_* - x) \leq 0 \tag{49}$$

everywhere on the feasible set X. Using (49) we evaluated a number of iterations required for finding ε-solution of (2), using step-sizes α_k determined by the steepest descent approach.

Numerous papers have been published in the West on the interior point techniques [3, 4, 18, 20, 22, 23, 28]. Some remarks deserve to be made concerning the main differences between the methodology adopted in these papers and our original approach:

1. Along with the LP problems we considered the NLP problems as well.

2. We developed asymptotically stable interior-infeasible-point algorithms. Our analysis was not confined to the interior point technique. For this reason the current points are in general allowed to be infeasible, however if the initial or current points are feasible, then the corresponding trajectory remains in the feasible set.

3. In all proposed methods multiplicative barriers were used and we did not resort to singular penalties.

4. The steepest descent approach was employed in computations. The trajectory could move along the boundary of the feasible set.

All these items can be considered as advantages of presented approach.

4. Dual barrier-projection methods for solving LP problem

By extension of the space and by converting the inequality constraints to equalities, we transform the original dual problem (3) into the following equivalent problem:

$$\text{maximize } b^\top u \text{ with respect to } u \text{ and } y \text{ subject to } \xi(y) + A^\top u - c = 0_n, \tag{50}$$

where $v = \xi(y) \in \mathbb{R}^n_+$, $\overline{\xi(\mathbb{R}^n)} = \mathbb{R}^n_+$.

The Problem (50) is similar to (32). System (35) being applied to solving (50) can be rewritten in terms of u and v as follows:

$$\frac{du}{dt} = b - Ax(u,v), \qquad \frac{dv}{dt} = -G(v)x(u,v), \tag{51}$$

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

If $u_0 \in U$, then we can get rid of the equation for v and this way simplify systems (51). In this case, (51) can be expressed as

$$\frac{du}{dt} = b - Ax(u), \qquad \left(G(v(u)) + A^\top A\right) x(u) = A^\top b, \tag{52}$$

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:

$$\begin{gathered} u_{k+1} = u_k + \alpha_k(b - Ax_k), \qquad v_{k+1} = v_k - \alpha_k G(v_k) x_k, \\ \left(G(v_k) + A^\top A\right) x_k = A^\top b + \tau(v_k + A^\top u_k - c). \end{gathered} \tag{53}$$

Similarly for the system (52) we have

$$u_{k+1} = u_k + \alpha_k(b - Ax_k), \qquad \left(G(v_k) + A^\top A\right) x_k = A^\top b, \tag{54}$$

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

Denote

$$\alpha_* = \frac{2}{\lambda^*}, \qquad \lambda^* = \max\left[\tau, \max_{1 \leq i \leq m} \dot\theta(0) x^i_*\right], \qquad \lambda_* = \min\left[\tau, \min_{1 \leq i \leq m} \dot\theta(0) x^i_*\right],$$

where λ^* and λ_* are, respectively, maximum and minimum eigenvalues of the matrix of the equation of the first approximation about the optimal solution u_*.

Theorem 7. *Let x_* and u_* be unique non-degenerate solutions of the Problems* (2) *and* (3)*, respectively, and let $v_* = c - A^\top u_*$. Assume that the component-wise space transformation $\xi(y)$ satisfies Conditions* 4, 5 *and $\tau > 0$. Then the following statements are true:*

1) *The pair $[u_*, v_*]$ is an asymptotically stable equilibrium state of system* (51)*.*

2) *The solutions $u(t,z_0)$, $v(t,z_0)$ of system* (51) *converge locally exponentially fast 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) The point u_* is an asymptotically stable equilibrium state of system (52).

4) The solutions $u(t, u_0)$ of system (52) *converge locally exponentially fast to the optimal solution u_* of the dual Problem* (3). *The corresponding function $x(u(t, u_0))$ converges to the optimal solution x_* of the primal Problem* (2).

5) For any fixed $0 < \alpha_k < \alpha_*$ the sequence $\{u_k, v_k\}$, generated by (53), *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 (54), *converges locally with a linear rate to u_* while the corresponding sequence $\{x_k\}$ converges to x_*.*

The method (52) was proposed and studied in 1977 (see [9]). Method (51) was given in [11]. In [16] we describe nonlocal convergence analysis of dual method. We supposed the Problem (2) is such that $Ae = 0_m$, where e is a vector of ones in $\mathbb{R}^n$. We used the Lyapunov function similar to (48):

$$V(u) = \sum_{i=m+1}^{n} v_*^i [\ln v_*^i - \ln v^i(u)].$$

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 null-space of A. We partition A as $A = [B, N]$, where the square matrix B is nonsingular. We can now write the matrix P as

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

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

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

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 (3) can be substituted by the following equivalent minimization problem:

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

Applying the stable barrier-projection method (40) to this problem, we obtain

$$\frac{dv}{dt} = G(v)(P^\top x(v) - \bar{x}), \tag{55}$$

$$PG(v)P^\top x(v) = PG(v)\bar{x} + \tau P(c - v). \tag{56}$$

If a point v is such that the matrix $PG(v)P^\top$ is invertible, then we can solve the linear equation (56) and obtain

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

Let $H(v) = G^{l/2}(v)$ and introduce the pseudo-inverse and projection matrices

$$(PH)^+ = (PH)^\top (PGP^\top)^{-1}, \qquad (PH)^\# = (PH)^+ PH.$$

The system (55), (56) can be rewritten in the following projective form:

$$\frac{dv}{dt} = H[\tau(PH)^{+}P(c-v) - (I_n - (PH)^{\#})H\bar{x}]. \tag{57}$$

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,

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

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 (57) can be rewritten as follows:

$$\frac{dv}{dt} = -G(v)\left(I_n - P^\top(PG(v)P^\top)^{-1}PG(v)\right)\bar{x}, \qquad v_0 \in \mathrm{ri}\, V_U. \tag{58}$$

Theorem 8. *Suppose that the conditions of the Theorem 7 hold. Then:*

1) *the point v_* is an asymptotically stable equilibrium point of system (55);*

2) *the solutions $v(t,v_0)$ of (57) 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*

$$v_{k+1} = v_k - \alpha_k G(v_k)(\bar{x} - P^\top x_k), \qquad x_k = x(v_k), \tag{59}$$

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

Since for system (58) $P\dot{v} = 0_d$, 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

$$\dot{v} = A^\top \lambda. \tag{60}$$

If $v > 0_n$, then after left multiplying both sides of (60) with $AG^{-1}(v)$ and in view of (58) we obtain

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

Hence, on the set $\mathrm{ri}\, V_U$ the method (58) takes the form

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

In u-space this method can be written as

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

If we use the quadratic and exponential space transformations (11), (12) we obtain

$$\frac{du}{dt} = (AD^{-1}(v(u))A^\top)^{-1}b, \qquad u_0 \in U_0, \tag{61}$$

and

$$\frac{du}{dt} = (AD^{-2}(v(u))A^{\top})^{-1}b, \qquad u_0 \in U_0, \tag{62}$$

respectively. The system (62) 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 the Theorem 8, the solution of (55) converges locally with an exponential rate to the equilibrium point $v_* = v(u_*)$. Therefore, the solutions of (61) also converge to the point u_* on the set U_0.

The discrete version of (61) consists of the iteration

$$u_{k+1} = u_k + \alpha_k(AD^{-1}(v_k)A^{\top})^{-1}b, \qquad u_0 \in U_0, \tag{63}$$

where $v_k = v(u_k)$. Taking into account the [7, Theorem 2.3.7] we conclude that the exponential rate of convergence of (61) insures local linear convergence of the discrete variant (63) if the step length α_k is sufficient small.

References

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

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

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

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

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

6. *Evtushenko Yu.G.* Two numerical methods of solving nonlinear programming problems, *Sov. Mam. Dokl.* **215** (2), 420–423, (1974).

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

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

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

10. *Evtushenko Yu.G. and Zhadan V.G.* 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).

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

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

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

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

15. *Evtushenko Yu.G. and Zhadan V.G.* Barrier-projective methods for nonlinear programming, *Comput. Mam. and Math. Physics*, **34** (5), 579–590, (1994).

16. *Evtushenko Yu.G. and Zhadan V.G.* Dual Barrier-Projective Methods in Linear Programming, In: *Resent trends in Optimization Theory and Applications* (Edited by R.P. Argawal), World Scientific Series in Applicable Analysis, **5**, 51–66, (1995).

17. *Evtushenko Yu.G. and Zhadan V.G.* Dual Barrier-Projective and Barrier-Newton Methods for Linear Programming Problems, *Comput. Math. and Math. Physics*, **36** (7), 30–45, (1996), (in Russian).

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

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

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

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

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

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

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

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

26. *Smirnov G.* Convergence of barrier-projection methods of optimization via vector Lyapunov functions, *Optimization Methods and Software*, **3** (1-3), 153–162, (1994).

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

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

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