Dynamics of non-homogeneous systems. Proceedings of ISA RAS, 1999, v. 2, pp. 86–108.

Evtushenko Yu.G.

Computing Centre of Russian Academy of Sciences
Moscow, Russia

Moretti A.

Applied Mathematics Department
IMECC-UNICAMP, CP 6065, CEP 13081-970
Campinas — SP — Brazil

Zhadan V.G.

Computing Centre of Russian Academy of Sciences
Moscow, Russia

NEWTON'S STEEPEST DESCENT FOR LINEAR PROGRAMMING[1]

Revised version 28 May 2003

1. Introduction and Preliminaries

Newton's method has been widely used for solving systems of nonlinear equations, nonlinear programming problems, optimal control problems and, recently, for linear programming (LP). We mention only a few references [23, 10, 32, 11]. Due to extensive activity in this area, the list is not complete. Many publications are devoted to path-following interior-point algorithms for linear programming. Extensive theoretical investigations, convergence analysis, and practical implementation are given in numerous papers and reports, e.g., [3]-[9], [33, 34, 36, 12].

In this paper, Newton's method is applied to a nonlinear system of equations derived from the optimality conditions for the LP problem. The method is stated as an initial-value problem involving a system of ordinary differential equations. We consider continuous and discrete versions of the primal-dual Newton method for solving the LP problem. We present some results on convergence rate and give a description of the algorithm that was implemented.

The paper is organized as follows. In the remainder of this section, we formulate the linear program in standard form and consider some properties of this program.

In Section 2, we describe the continuous version of Newton's method. The right-hand side of the ordinary differential equation defines the Newton search directions that are determined from the solution of a linear system with $2n$ unknowns. In Section 3 we present various special cases and consequences of proposed numerical method. In Section 4, we reduce this system to 2×2-block system with n unknowns and to a system with symmetric well-defined matrix which does not have any singular terms in the vicinity of the optimal solution. In this linear system, in addition to the matrix of constraints, we introduce an additional matrix whose range space coincides with the null space of the matrix of constraints.

[1]Research was supported by Grants N° 99-01-01186 and N° 96-15-96124 from the Russian Foundation for Basic Research and by FAPESP Grant N° 96-6631-5.

In Section 5, we present some new results concerning the convergence rate of the method. In Section 6, we consider the discrete version of the method. We use the steepest descent approach for choosing the step lengths. These step lengths are determined from the solution of the auxiliary problem.

Let A be $m \times n$ full-rank matrix, $m < n$, $b, u \in \mathbb{R}^m$, and $c, x \in \mathbb{R}^n$. Consider the standard form of the linear program

$$
\textbf{(P)} \qquad \begin{cases} \text{minimize} & c^\top x \\ \text{subject to} & Ax = b, \\ & x \geq 0_n, \end{cases}
$$

and its dual

$$
\textbf{(D)} \qquad \begin{cases} \text{maximize} & b^\top u \\ \text{subject to} & c - A^\top u \geq 0_n, \end{cases}
$$

where 0_s is an s-dimensional null vector, 0_{sk} is an $s \times k$ rectangular null matrix. Subscripts will be used to distinguish values of quantities at a particular iteration while superscripts will indicate components of vectors.

For a matrix A of full row rank, the associated right $n \times m$ pseudoinverse matrix is defined as

$$
A^+ = A^\top (AA^\top)^{-1}.
$$

The orthogonal projector onto the range space of the matrix $A^\top$ is defined as

$$
(A^\top)^{\parallel} = A^+ A. \tag{1}
$$

The orthogonal projector onto the null space of matrix A is given by

$$
(A^\top)^{\perp} = I_n - (A^\top)^{\parallel},
$$

where I_n is the unit square matrix of order n.

Let K be a matrix whose rows span the null space of A. By choosing a basis for this space we can assume that K is an $\ell \times n$ matrix with the row rank ℓ, with $m + \ell = n$. This definition yields the orthogonality condition:

$$
KA^\top = 0_{\ell m}. \tag{2}
$$

We partition A and K as

$$
A = [B \mid N], \qquad K = [K_B \mid K_N], \tag{3}
$$

where B is an arbitrary nonsingular square matrix of order m, K_N is a nonsingular square matrix of order ℓ, N and $K_B^\top$ are $m \times \ell$ matrices. Popular choice for K is

$$
K = [-N^\top (B^\top)^{-1} \mid I_\ell]. \tag{4}
$$

By multiplying this matrix on the right by $A^\top$, we obtain that in this case the condition (2) holds.

Since the maximum rank of K is ℓ, we can define the pseudo-inverse matrix

$$
K^+ = K^\top (KK^\top)^{-1}
$$

of K and use the orthogonal projector $(K^\top)^{\perp}$ onto the null space of the matrix K and the orthogonal projector $(K^\top)^{\parallel}$ onto the range space of the matrix $K^\top$. The relations between the projectors are as follows:

$$
(A^\top)^{\perp} = (K^\top)^{\parallel}, \qquad (K^\top)^{\perp} = (A^\top)^{\parallel}. \tag{5}
$$

Define the vector of dual slacks $v \in \mathbb{R}^n$ and consider the overdetermined linear system of n equations with m unknowns

$$A^\top u = c - v. \tag{6}$$

The vector u can be regarded as an implicit function of v. In general, this equation does not have a solution. Therefore, we define the unique pseudosolution

$$u(v) = (AA^\top)^{-1}A(c - v). \tag{7}$$

Substituting this expression into (6), we obtain $A^\top u(v) = (A^\top)^{\|}(c - v)$. Clearly, the pseudosolution $u(v)$ solves (6) if and only if $(A^\top)^{\perp}(c - v) = 0_n$, i.e., the vector $c - v$ belongs to the row space of matrix A. According to (5), in this case $K(A^\top)^{\perp}(c-v) = K(K^\top)^{\|}(c-v) = d - Kv = 0_\ell$, where $d = Kc$.

Let q be an arbitrary n-dimensional vector (some of its components can be negative) that satisfies the following condition:

$$Aq = b.$$

Define n-dimensional functions $w(v)$ and $\tilde{w}(x)$ associated with the primal and dual problems:

$$w(v) = c - v - A^\top u(v), \qquad \tilde{w}(x) = q - x - K^\top \tilde{u}(x),$$

where $\tilde{u}(x)$ is a pseudosolution

$$\tilde{u}(x) = (KK^\top)^{-1}K(q - x) \tag{8}$$

of the equation

$$K^\top \tilde{u} = q - x. \tag{9}$$

The projections of a vector z onto the null space and row space of matrix A are denoted, respectively, by $z^{\perp}$ and $z^{\|}$. Taking into account (5), (7), (8) and (9), we obtain

$$w(v) = (A^\top)^{\perp}(c - v) = c^{\perp} - v^{\perp}, \qquad \tilde{w}(x) = (A^\top)^{\|}(q - x) = q^{\|} - x^{\|}, \tag{10}$$

$$A^\top u(v) = (A^\top)^{\|}(c - v) = c^{\|} - v^{\|}, \qquad K^\top \tilde{u}(x) = (A^\top)^{\perp}(q - x) = q^{\perp} - x^{\perp}. \tag{11}$$

Below $D(z)$ denotes the diagonal matrix whose entries are the components of a vector z. We denote component-wise operations on vectors by the usual notation for real numbers. Thus, given two vectors z, h of the same dimension, $D(z/h)$ denotes the diagonal matrix, whose i-th diagonal element is z^i/h^i.

Let $\mathbb{R}^n_+$ and $\mathbb{R}^n_{++}$ denote nonnegative and, respectively, strictly positive orthants in $\mathbb{R}^n$. Introduce feasible sets for the primal variable x and for the dual slacks v:

$$X = \{x \in \mathbb{R}^n_+ : b - Ax = 0_m\}, \qquad V = \{v \in \mathbb{R}^n_+ : Kv - d = 0_\ell\}.$$

Their relative interior parts are defined as intersections with the positive orthant

$$X_+ = X \cap \mathbb{R}^n_{++}, \qquad V_+ = V \cap \mathbb{R}^n_{++}.$$

We assume that the sets X_+ and V_+ are nonempty and that primal and dual nondegeneracy hold. In this case, both problems have unique solutions x_* and u_*, respectively, strict complementarity holds (i.e., $|x^i_*| + |v^i_*| \neq 0$ for all i), and the following inequalities are valid:

$$\|b\| \neq 0, \qquad \|d\| \neq 0. \tag{12}$$

In the sequel, the optimal dual slacks $v_* = c - A^\top u_*$ will be also referred to as a solution of the dual problem **(D)**.

The first-order necessary and sufficient optimality conditions for problems **(P)** and **(D)** can be written as a system of $2n$ equalities:

$$\begin{array}{rcl} D^\lambda(x)v &=& 0_n, \\ Kv - d &=& 0_\ell, \\ Ax - b &=& 0_m, \end{array} \tag{13}$$

and $2n$ inequalities:

$$x \geq 0_n, \qquad v \geq 0_n, \tag{14}$$

where λ is a positive scalar. In (13), we introduced the expression $D^\lambda(x)v$ instead of the commonly used $D(x)v$. It was motivated by our previous work on space transformation techniques [5], [7]-[17]. The expression will allow us to obtain different step sizes for the primal and dual variables.

In system (13), we have $2n$ scalar equality equations and $2n$ unknown scalar variables. If we find solutions x_*, v_* of (13) and (14), then we define $u_* = u(v_*)$ from (7). Clearly, the pair $[x_*, u_*]$ coincides with the unique solutions of **(P)** and **(D)**, respectively.

2. Continuous Version of the Method

We use the continuous version of Newton's method for solving the system of equalities (13). The computation is described by the system of ordinary differential equations:

$$D^{\lambda-1}(\Xi)W(x,v)\begin{bmatrix} \dot{x} \\ \dot{v} \end{bmatrix} = -D^{\lambda-1}(\Xi)D(\gamma)R(x,v), \tag{15}$$

where W is a square $2n$ matrix, R is a $2n$-dimensional vector:

$$W(x,v) = \begin{bmatrix} \lambda D(v) & D(x) \\ 0_{\ell n} & K \\ A & 0_{mn} \end{bmatrix}, \qquad R(x,v) = \begin{bmatrix} D(x)v \\ Kv - d \\ Ax - b \end{bmatrix},$$

the diagonal matrix $D(\gamma)$ has the first $n + \ell$ diagonal components equal to α and m other diagonal components equal to τ; the first n components of the vector Ξ coincide with the vector x and the other n components are equal to one. In the sequel we will simplify (15) and investigate the following system of $2n$ differential equations:

$$W(x,v)\begin{bmatrix} \dot{x} \\ \dot{v} \end{bmatrix} = -D(\gamma)R(x,v). \tag{16}$$

We have introduced three auxiliary parameters λ, α, and τ. For the sake of simplicity, we assume that $\lambda = \alpha/\tau$.

By following the trajectories satisfying (16), we can, in theory, obtain a solution of the system of nonlinear equations (13). In practice, we build the iterative procedures using a discretization of these differential equations. The vectors $\dot{x}$ and $\dot{v}$, defined by (16), are called the Newton directions.

From the system (16), we have

$$K\dot{v} = \alpha(d - Kv), \tag{17}$$

$$A\dot{x} = \tau(b - Ax). \tag{18}$$

Let $x(t)$, $v(t)$ denote the solutions of the Cauchy problem (16) with initial conditions $x_0 = x(0)$, $v_0 = v(0)$. Using orthogonal decomposition of the vectors $x(t)$ and $v(t)$, we will seek a solution of the system (16) in the following form:

$$x(t) = x^{\|}(t) + x^{\perp}(t), \qquad v(t) = v^{\|}(t) + v^{\perp}(t). \tag{19}$$

We substitute these representations in (17) and (18). By multiplying both sides of differential equation (17) on the left by the matrix K^+ and multiplying (18) on the left by A^+, we obtain

$$\begin{aligned} \dot{v}^{\perp} &= \alpha[c^{\perp} - v^{\perp}] = \alpha w, && (20)\\ \dot{x}^{\|} &= \tau[q^{\|} - x^{\|}] = \tau \tilde{w}. && (21) \end{aligned}$$

Differentiating (10) and using the above equations produces

$$\dot{w} = -\alpha w, \qquad \dot{\tilde{w}} = -\tau \tilde{w}. \tag{22}$$

The equations (20), (21) and (22) can be easily integrated:

$$\begin{aligned} w(v(t)) &= c^{\perp} - v^{\perp}(t) = [c^{\perp} - v^{\perp}(0)]e^{-\alpha t}, && (23)\\ \tilde{w}(x(t)) &= q^{\|} - x^{\|}(t) = [q^{\|} - x^{\|}(0)]e^{-\tau t}. && (24) \end{aligned}$$

By multiplying the solution (23) on the left by $-K$ and the solution (24) on the left by $-A$, we obtain

$$\begin{aligned} Kv(t) - d &= [Kv_0 - d]e^{-\alpha t}, && (25)\\ Ax(t) - b &= [Ax_0 - b]e^{-\tau t}. && (26) \end{aligned}$$

Differentiating (19) and taking into account (10), (20), (21), we have:

$$\dot{x}(t) = \tau \tilde{w}(x(t)) + \dot{x}^{\perp}(t), \qquad \dot{v}(t) = \alpha w(v(t)) + \dot{v}^{\|}(t). \tag{27}$$

We introduce an m-dimensional vector function $g(t)$ and an ℓ-dimensional vector function $\tilde{g}(t)$ and use the following representation:

$$\dot{x}^{\perp}(t) = -\tau K^{\top}\tilde{g}(t), \qquad \dot{v}^{\|}(t) = -\alpha A^{\top} g(t). \tag{28}$$

The pseudosolutions $u(v(t))$ and $\tilde{u}(x(t))$, defined by (7) and (8), are composite functions of the independent variable t. Considering (28), these functions satisfy the following differential equations:

$$\dot{u} = \alpha g, \qquad \dot{\tilde{u}} = \tau \tilde{g}.$$

Substituting (27) and (28) in (16), we observe that the vectors g and $\tilde{g}$ satisfy the following linear algebraic system of n equations in n unknowns:

$$D(x)A^{\top} g + D(v)K^{\top}\tilde{g} = D(x)[c^{\perp} - v^{\perp}] + D(v)[q^{\|} - x^{\|}] + D(x)v. \tag{29}$$

The initial conditions for the system of ordinary equations (28) are as follows:

$$x^{\perp}(0) = x_0 - x^{\|}(0), \qquad v^{\|}(0) = v_0 - v^{\perp}(0).$$

We introduce an n-dimensional indicator vector:

$$y = D^{-1}(v)(A^{\top} g - w(v)). \tag{30}$$

This vector can be also represented in the following form:

$$y = e_n + D^{-1}(x)(\tilde{w}(x) - K^\top \tilde{g}), \tag{31}$$

where e_s denotes the vector of ones in $\mathbb{R}^s$.

Using this notation, we can rewrite (28) as follows:

$$\dot{x}^\perp(t) = -\tau[D(x)(e_n - y) + \tilde{w}], \qquad \dot{v}^\parallel(t) = -\alpha[D(v)y + w].$$

The vectors $\dot{x}^\perp(t)$ and $\dot{v}^\parallel(t)$ are orthogonal. Therefore, at each t we have:

$$\tilde{w}^\top D(v)y + w^\top D(x)(e_n - y) = y^\top D(xv)(y - e_n). \tag{32}$$

The system (16) can be written in original variables as follows:

$$\dot{x} = \tau D(x)[y - e_n], \qquad \dot{v} = -\alpha D(v)y. \tag{33}$$

The vector y satisfies the following conditions:

$$AD(y)x = b, \tag{34}$$
$$KD(e_n - y)v = d. \tag{35}$$

As the vectors $D(y)x - q$ and $D(e_n - y)v - c$ belong to the null spaces of A and K, respectively, they are orthogonal, i.e.,

$$c^\top D(x)y + q^\top D(v)(e_n - y) = c^\top q + y^\top D(xv)(e_n - y).$$

The first integral of the system (16) is

$$D^\lambda(x(t))v(t) = D^\lambda(x_0)v_0 e^{-\alpha t}. \tag{36}$$

From this expression we conclude that all components of the vectors $x(t)$ and $v(t)$ do not change their signs along the trajectories of the system (16). Therefore, if the starting points are $x_0 \in X_+$, $v_0 \in V_+$, the solution of (16) exists for all $0 \le t < \infty$ and $x(t) \ge 0_n$, $v(t) \ge 0_n$. If $x_0^i = 0$, $v_0^j = 0$, then $x^i(t) \equiv 0$, $v^j(t) \equiv 0$ for all $t \ge 0$. Hence we can say that the system (16) has an adhesion property.

Let $\lambda = 1$, $x_0 \in X_+$, $v_0 \in V_+$ and let the pairwise products $x_0^i v_0^i$ be identical for all i. It follows from (36) that the subsequent products $x^i(t)v^i(t)$ are also identical for all i. Hence in this particular case, the trajectory of the system (16) belongs to so-called central path which plays a vital role in the theory of primal-dual algorithms.

3. Special Cases

We introduce two scaled matrices

$$\bar{A} = AD^{1/2}\left(\frac{x}{v}\right), \qquad \tilde{K} = KD^{1/2}\left(\frac{v}{x}\right).$$

Many special cases of the system (16) have been studied extensively in the literature. We mention only few of them.

$\boxed{\mathbf{1.}}$ Let $v_0 \in V_+$. Then the vector $v_0 - c$ belongs to the row space of the matrix A, and, according to (23), we have $w(0) = w(t) = 0_n$ for all t. Hence this property holds on the remainder of the trajectory. In this case, the formula for y is simplified, that is,

$$y = D^{-1}(v)A^\top g = D^{-1/2}(xv)\bar{A}^+ b.$$

Expressions (33) yield

$$\dot{x} = \tau \left[D^{1/2} \left(\frac{x}{v} \right) \bar{A}^{+} b - x \right], \qquad \dot{v} = -\alpha D^{1/2} \left(\frac{v}{x} \right) \bar{A}^{+} b. \tag{37}$$

This case was investigated in [1, 2, 17].

$\boxed{\mathbf{2.}}$ Let $x_0 \in X_+$ and $v_0 \in V_+$. According to (23) and (24), $w(0) = w(t) = 0_n$, $\tilde{w}(0) = = \tilde{w}(t) = 0_n$, all trajectories $x(t)$ and $v(t)$ will remain in the sets X_+, V_+, respectively, for all $t \geq 0$. Setting $b = Ax$, (30), (31), (32), and (33) yield

$$y = D^{-1}(v) A^\top (\bar{A} \bar{A}^\top)^{-1} b = e_n - D^{-1}(x) K^\top (\tilde{K} \tilde{K}^\top)^{-1} d,$$

$$y^\top D(xv)(y - e_n) = 0, \qquad \dot{v}^{\|} = \dot{v}, \qquad \dot{x}^{\perp} = \dot{x}, \tag{38}$$

$$\dot{x} = \tau \left[D \left(\frac{x}{v} \right) A^\top (\overline{AA^\top})^{-1} Ax - x \right], \qquad \dot{v} = -\alpha A^\top (\bar{A} \bar{A}^\top)^{-1} Ax. \tag{39}$$

This case was investigated in [2, 11].

$\boxed{\mathbf{3.}}$ Let $\alpha = \tau$. Then from (39) we have:

$$D(v)\dot{x} + D(x)\dot{v} = -\tau D(x) v.$$

If $\tau = 1$, we have the pure Newton method.

$\boxed{\mathbf{4}}$. Setting $\alpha = 0$, $\tau = 1$ and $v = e_n$ in (33), we obtain the following method:

$$\dot{x} = D^{1/2}(x) [\hat{A}^{\perp} \hat{c} + \hat{A}^{+} (Ax - b)], \tag{40}$$

where $\hat{A} = AD^{1/2}(x)$. This method was proposed and investigated in [8], where it was shown to converge exponentially to an optimal solution point x_* which is an asymptotically stable attractor, while all other vertices are nonstable stationary points for system (40). This variant is considered in [17].

$\boxed{\mathbf{5.}}$ Simplifying the method (40) to consider an interior variant, where $x(t) \in X_+$ for all t, yields the following from (40):

$$\dot{x} = D(x) [A^\top (AD(x)A^\top)^{-1} AD(x) - e_n] c. \tag{41}$$

This method also converges exponentially but, in contrast to (40), it does not have the asymptotic stability property with respect to the equality constrain $Ax = b$.

This method was proposed, studied and implemented in 1977 (see [8]). It was called a "barrier-projection method". Method (41) resembles Dikin's algorithm [4], sometimes called the "variation on Karmarkar's algorithm" [3]. It has better local convergence properties than Dikin's method. The differences between these methods are analyzed in [17]. Recently this method was reinvented in [19, 22] and [24].

$\boxed{\mathbf{6.}}$ Setting $\alpha = 1$, $\tau = 0$ and $x = e_n$ in (33), we have:

$$\dot{u} = (AD^{-1}(v)A^\top)^{-1} b, \qquad \dot{v} = -A^\top (AD^{-1}(v)A^\top)^{-1} b.$$

This method was introduced in [8]. It is similar to the affine-scaling algorithm proposed in [1]. Comparative analysis of the local convergence properties is given in [16, 17, 18].

4. Computation of Newton's Directions

Most of the computational effort in implementation of Newton's method is spent in solving the linear system (29). Here we focus on this system. Let G be an n-dimensional vector, $G^\top = [g^\top, \tilde{g}^\top]$. We rewrite (29) as

$$M(x,v)G = F(x,v), \tag{42}$$

where

$$\begin{array}{rcl} M(x,v) & = & [D(x)A^\top \mid D(v)K^\top], \\ F(x,v) & = & D(x)v + D(x)w(v) + D(v)\tilde{w}(x). \end{array} \tag{43}$$

This system is uniquely solvable in $\mathbb{R}^n_{++} \times \mathbb{R}^n_{++}$. Consider the two simplest cases where this system can be easily solved by eliminating either g or $\tilde{g}$. Assume that $v \in \mathbb{R}^n_{++}$. Then multiplying both sides of equation (42) on the left by $AD^{-1}(v)$ and taking into account the orthogonality condition (2), we obtain the linear system in m unknowns:

$$AD\left(\frac{x}{v}\right)A^\top g = AD\left(\frac{x}{v}\right)w(v) + b. \tag{44}$$

This approach is very popular in publications devoted to interior point techniques. Suppose now that $x \in \mathbb{R}^n_{++}$. Since the diagonal of $D(x)$ is strictly positive, we can rearrange the system (42). Multiplying on the left by $KD^{-1}(x)$, we find another linear system in ℓ unknowns:

$$KD\left(\frac{v}{x}\right)K^\top \tilde{g} = KD\left(\frac{v}{x}\right)\tilde{w}(x) + d. \tag{45}$$

It is sufficient to solve only one system — either (44) or (45) — because on substituting the computed solution in (29), we obtain all the expressions necessary for determining the Newton's directions (27). We usually choose the system with the smallest dimensionality.

We can use the least squares approach to solve (42). We obtain the normal equations by multiplying both sides of (42) by the square matrix $M^\top$:

$$M_2(x,v)G = F_2(x,v), \tag{46}$$

where we introduced a symmetric positive definite $n \times n$ matrix with 2×2 block matrices

$$M_2(x,v) = M^\top(x,v)M(x,v) = \begin{bmatrix} AD^2(x)A^\top & AD(xv)K^\top \\ KD(xv)A^\top & KD^2(v)K^\top \end{bmatrix}$$

and n-dimensional vector

$$F_2(x,v) = M^\top(x,v)F(x,v).$$

Lemma 1. *If $x \in \mathbb{R}^n_{++}$, $v \in \mathbb{R}^n_{++}$, then the matrices $M(x,v)$, $M_2(x,v)$ are nonsingular and there is a one-to-one correspondence between their inverses:*

$$M_2^{-1} = M^{-1}(M^{-1})^\top, \qquad M^{-1} = M_2^{-1}M^\top, \tag{47}$$

where

$$M^{-1}(x,v) = \begin{bmatrix} \left(AD\left(\frac{x}{v}\right)A^\top\right)^{-1} AD^{-1}(v) \\ \left(KD\left(\frac{v}{x}\right)K^\top\right)^{-1} KD^{-1}(x) \end{bmatrix}. \tag{48}$$

Proof. Multiplying (48) on the right by $M(x, v)$, we obtain the expression

$$M^{-1}(x, v)M(x, v) = I_n,$$

which proves (48). Solving the systems (42) and (46), we have

$$G = M^{-1}F = M_2^{-1}F_2 = M_2^{-1}M^\top F. \tag{49}$$

From these conditions we get the formulas (47). $\square$

Using (48), we find explicit expressions for vectors g and $\tilde{g}$. On substituting in (28), we obtain

$$\begin{array}{rcl} \dot{x}^\perp & = & -\tau K^\top \left(KD\left(\frac{v}{x}\right)K^\top\right)^{-1} KD^{-1}(x)F(x, v), \\ \dot{v}^\parallel & = & \alpha A^\top \left(AD\left(\frac{x}{v}\right)A^\top\right)^{-1} AD^{-1}(v)F(x, v). \end{array} \tag{50}$$

All formulas (42) – (50) can be used for practical computations only if the vectors x and v are not very close to the solutions of problems **(P)** and **(D)**. All these formulas become increasingly ill-conditioned when the trajectories approach the optimal solutions with some components of the vectors x_* and v_* being zero. Matrices $AD(x/v)A^\top$ and $KD(v/x)K^\top$ may be ill-conditioned or singular when the elements of the diagonal matrices $D(x/v)$ or $D(v/x)$ take on both very large and very small values. This severe drawback is analyzed in many publications. See, for example, [11]. Various modifications of Cholesky codes were proposed in order to cope with these shortcomings.

We propose at later stages of the computation to refrain from solving the systems (44) or (45) and, instead, to focus on the system (42). We will use the partition (3), where B is $m \times s$ current matrix (not obligatory optimal basis). Let ε be a nonnegative number. We mention three different rules for the partition:

1. If $x^i > \varepsilon$, then the i-th column of A is included in the matrix B, else it is included in N.
2. If $v^i > \varepsilon$, then the i-th column of A is included in the matrix N, else it is included in B.
3. If $x^i > v^i$, then the i-th column of A is included in the matrix B, else it is included in N.

In the course of computation, the columns of the current matrices B and N will be altered. It is desirable for the sequence of matrices B to converge to the primal optimal basis. If $\varepsilon = 0$, the vectors x and v satisfy the conditions of complementarity and the strict complementarity (i.e., $x^iv^i = 0$, $|x^i| + |v^i| \neq 0$ for all i), then all three rules yield the same partition.

We say that a pair $x \in \mathbb{R}^n_+$, $v \in \mathbb{R}^n_+$ satisfies *ε-strict complementarity condition*, if

$$|x^i| + |v^i| \geq 2\varepsilon$$

for all i.

Let us use the first partition rule. Without loss of generality, we assume that a current point x is such that only the first s components of the vector x belong to the matrix B. If ε — strict complementarity condition holds, then we have

$$x = \begin{bmatrix} x^B \\ x^N \end{bmatrix}, \quad x^B \geq \varepsilon e_s, \qquad v = \begin{bmatrix} v^B \\ v^N \end{bmatrix}, \quad v^N \geq \varepsilon e_{\ell_1},$$

where $\ell_1 = n - s$. Using this notation, we can rewrite (29) as follows:

$$\begin{array}{rcl} D(x^B)B^\top g & + & D(v^B)K_B^\top \tilde{g} = D(x^B)v^B + \\ & + & D(x^B)\left[(c^\perp)^B - (v^\perp)^B\right] + D(v^B)\left[(q^\parallel)^B - (x^\parallel)^B\right], \end{array} \tag{51}$$

$$
\begin{aligned}
D(x^N)N^\top g \quad &+ \quad D(v^N)K_N^\top \tilde{g} = D(x^N)v^N + \\
&+ \quad D(x^N)\left[(c^\perp)^N - (v^\perp)^N\right] + D(v^N)\left[(q^\|)^N - (x^\|)^N\right].
\end{aligned} \tag{52}
$$

Let us denote

$$
\bar{K}_B = K_B D\left(\frac{v^B}{x^B}\right), \qquad \bar{N} = ND\left(\frac{x^N}{v^N}\right).
$$

Multiplying (51) on the left by the nonsingular matrix $D^{-1}(x^B)$ and (52) by $D^{-1}(v^N)$, we obtain the equivalent system:

$$
\begin{aligned}
B^\top g + \bar{K}_B^\top \tilde{g} &= (v^\|)^B + (c^\perp)^B + D\left(\frac{v^B}{x^B}\right)[(q^\|)^B - (x^\|)^B], \\
\bar{N}^\top g + K_N^\top \tilde{g} &= (x^\perp)^N + (q^\|)^N + D\left(\frac{x^N}{v^N}\right)[(c^\perp)^N - (v^\perp)^N].
\end{aligned}
$$

We rewrite this system as follows:

$$
\bar{M}G = \bar{F}, \tag{53}
$$

where $\bar{M} = \bar{D}M$, $\bar{F} = \bar{D}F$ and

$$
\bar{D} = \begin{bmatrix} D^{-1}(x^B) & 0_{s\ell_1} \\ 0_{\ell_1 s} & D^{-1}(v^N) \end{bmatrix}, \quad \bar{M} = \begin{bmatrix} B^\top & \bar{K}_B^\top \\ \bar{N}^\top & K_N^\top \end{bmatrix}.
$$

The unique solution of (53) coincides with the solution of the following normal system:

$$
\bar{M}_2(x,v)G = \bar{F}_2(x,v). \tag{54}
$$

Here we introduced a symmetric positive definite $n \times n$ matrix

$$
\bar{M}_2 = \bar{M}^\top \bar{M} = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}.
$$

The matrices Q_{ij} and the vector $\bar{F}_2 = \bar{M}^\top \bar{F}$, $\bar{F}_2^\top = [(f_1)^\top, (f_2)^\top]$ are as follows:

$$
\begin{aligned}
Q_{11} &= BB^\top + \bar{N}\bar{N}^\top, & Q_{12} &= B\bar{K}_B^\top + \bar{N}K_N^\top, \\
Q_{21} &= \bar{K}_B B^\top + K_N \bar{N}^\top, & Q_{22} &= \bar{K}_B \bar{K}_B^\top + K_N K_N^\top,
\end{aligned}
$$

$$
\begin{aligned}
f_1 &= B \left[(v^\|)^B + (c^\perp)^B + D\left(\frac{v^B}{x^B}\right)[(q^\|)^B - (x^\|)^B]\right] + \\
&+ \bar{N} \left[(x^\perp)^N + (q^\|)^N + D\left(\frac{x^N}{v^N}\right)[(c^\perp)^N - (v^\perp)^N]\right], \\
f_2 &= \bar{K}_B \left[(v^\|)^B + (c^\perp)^B + D\left(\frac{v^B}{x^B}\right)[(q^\|)^B - (x^\|)^B]\right] + \\
&+ K^N \left[(x^\perp)^N + (q^\|)^N + D\left(\frac{x^N}{v^N}\right)[(c^\perp)^N - (v^\perp)^N]\right].
\end{aligned}
$$

Lemma 2. *If $x \in \mathbb{R}^n_{++}$, $v \in \mathbb{R}^n_{++}$, then the solution of (42) is given by*

$$
G = \bar{M}_2^{-1}\bar{F}_2 = \bar{M}_2^{-1}\bar{M}^\top \bar{F} = \bar{M}^{-1}\bar{F},
$$

where the inverse of $\bar{M}_2$ can be defined in the following form:

$$
\bar{M}_2^{-1} = \begin{bmatrix} Q_{11}^{-1} + Q_{11}^{-1}Q_{12}H^{-1}Q_{21}Q_{11}^{-1} & -Q_{11}^{-1}Q_{12}H^{-1} \\ -H^{-1}Q_{21}Q_{11}^{-1} & H^{-1} \end{bmatrix}, \tag{55}
$$

where we introduced a square matrix $H = Q_{22} - Q_{21}Q_{11}^{-1}Q_{12}$.

Proof. Matrices Q_{11} and Q_{22} can be represented as

$$Q_{11} = AS_1A^\top, \qquad Q_{22} = KS_2K^\top,$$

$$S_1(x,v) = \begin{bmatrix} I_s & 0_{s\ell_1} \\ 0_{\ell_1 s} & D\left(\dfrac{x^N}{v^N}\right) \end{bmatrix}, \qquad S_2(x,v) = \begin{bmatrix} D\left(\dfrac{v^B}{x^B}\right) & 0_{s\ell_1} \\ 0_{\ell_1 s} & I_\ell \end{bmatrix}.$$

We conclude that the matrices Q_{11} and Q_{22} are nonsingular. Clearly,

$$\mathrm{Det}[\bar{M}_2] = \mathrm{Det}[Q_{11}]\mathrm{Det}[H].$$

Therefore, the matrix H is also nonsingular. Now we obtain that all matrices in the expression (55) are well defined and the inverse matrix $\bar{M}_2^{-1}$ exists. $\square$

Similarly to (47), we have:

$$\bar{M}_2^{-1} = M^{-1}\bar{D}^{-2}(M^{-1})^\top, \qquad M^{-1} = \bar{M}_2^{-1}M^\top\bar{D}^2. \tag{56}$$

Consider the equation (42) in the vicinity of the optimal solutions x_*, v_* of the problems **(P)** and **(D)**. We use the previous partition (3) and the representation (4), but here we assume that B is the optimal basis. In this case

$$x_* = \begin{bmatrix} x_*^B \\ x_*^N \end{bmatrix}, \qquad x_*^B > 0_m, \quad x_*^N = 0_\ell,$$

$$v_* = \begin{bmatrix} v_*^B \\ v_*^N \end{bmatrix}, \qquad v_*^B = 0_m, \quad v_*^N > 0_\ell,$$

$$x_*^\perp = \begin{bmatrix} (x_*^\perp)^B \\ (x_*^\perp)^N \end{bmatrix}, \quad v_*^\parallel = \begin{bmatrix} (v_*^\parallel)^B \\ (v_*^\parallel)^N \end{bmatrix}, \quad q = \begin{bmatrix} q^B \\ q^N \end{bmatrix}, \quad c = \begin{bmatrix} c^B \\ c^N \end{bmatrix}.$$

It is easy to see that the following result holds.

Lemma 3. *Let* $[x_*, v_*]$ *be a non-degenerate solution pair of the primal and dual problems* **(P)**, **(D)**. *Then the matrix* $M(x_*, v_*)$ *is nonsingular and*

$$F(x_*, v_*) = 0_n,$$

$$M^{-1}(x_*, v_*) = \begin{bmatrix} (B^\top)^{-1}D^{-1}(x_*^B) & 0_{m\ell} \\ 0_{\ell m} & (K_N^\top)^{-1}D^{-1}(v_*^N) \end{bmatrix}, \tag{57}$$

$$M_2^{-1}(x_*, v_*) = \begin{bmatrix} (B^\top)^{-1}D^{-2}(x_*^B)B^{-1} & 0_{m\ell} \\ 0_{\ell m} & (K_N^\top)^{-1}D^{-2}(v_*^N)K_N^{-1} \end{bmatrix}. \tag{58}$$

The solutions of **(P)** and **(D)** can be represented as

$$x_*^B = B^{-1}b, \quad u_* = (B^\top)^{-1}c^B, \quad v_*^N = c^N - N^\top u_*,$$

or, in equivalent form,

$$x_*^B = q^B - K_B^\top\tilde{u}_*, \quad \tilde{u}_* = (K_N^\top)^{-1}q^N, \quad v_*^N = (K_N)^{-1}d.$$

In view of (1), (2), (10) and (11) we obtain: $w(v_*) = \tilde{w}(x_*) = 0_n$,

$$x_*^\parallel = A^\parallel x_* = A^\top(AA^\top)^{-1}Ax_* = A^\top(AA^\top)^{-1}Aq = q^\parallel,$$

$$v_*^\perp = K^\| v_* = K^\top (KK^\top)^{-1} K v_* = K^\top (KK^\top)^{-1} d = c^\perp,$$
$$v_*^\| = v_* - c^\perp, \quad x_*^\perp = x_* - q^\|, \quad (x_*^\perp)^N = -(q^\|)^N, \quad (v_*^\|)^B = -(c^\perp)^B.$$

Define two scalars

$$\varepsilon_* = \min_{1\le i\le m} x_*^i > 0, \qquad \kappa_* = \min_{m+1\le j\le n} v_*^j > 0.$$

Substitute x_* and v_* in (29). If strict complementarity holds, then using the third partition rule, we form the matrix B which coincides with the primal optimal basis. The first and second rules also provide correct results if

$$\varepsilon_* \ge \varepsilon > 0, \qquad \kappa_* \ge \varepsilon > 0,$$

respectively. Now we obtain the exact inverse matrices (57) and (58) from (55) and (56). The system (42) has a unique zero solution at the optimal solution point $[x_*, v_*]$, and there exists a neighborhood about this point, where the system is well-defined and can be solved using Gaussian elimination, Gauss–Jordan algorithm, and numerous other methods of linear algebra. The use of formulas (47) and (48) is not convenient in a neighborhood of the solution pair. Indeed, the formula (57) can not be obtained directly from (48) by working with the pair $[x, v]$ to $[x_*, v_*]$. Thus instead of solving the systems (44) or (45) we should solve either the linear algebraic system (53) with a nonsymmetric matrix, or the system (54) with a symmetric matrix. Both systems do not have any singular terms in the vicinity of the optimal, pair $[x_*, v_*]$.

5. Convergence Properties

Newton's method has very attractive convergence properties near a solution. We study convergence using methods of stability analysis proposed by Lyapunov in 1892 and refined by many contributors in this field. We introduce the discrete version of Newton's method:

$$z_{k+1} = z_k - W^{-1}(x_k, v_k) D(\gamma_k) R(x_k, v_k), \tag{59}$$

where $z_k^\top = [x_k^\top, v_k^\top]$.

Consider the local convergence properties of the system (16) in a neighborhood of the solution of (13). We start from the simplest case of fixed coefficients λ, α, τ. Let $\gamma_* = \max[\alpha, \tau]$.

Theorem 1. *Suppose that the system of equations* (13) *has a solution pair* $[\bar{x}, \bar{v}]$, *such that the* m *columns of the matrix* A *corresponding to nonzero components of the vector* $\bar{x}$ *and the* ℓ *columns of* K *corresponding to nonzero components of* $\bar{v}$ *form two square nonsingular matrices. Then for any* $\alpha > 0$, $\tau > 0$, *the pair* $[\bar{x}, \bar{v}]$ *is an asymptotically stable equilibrium point of the system* (16). *If the starting points* $x(0)$, $v(0)$ *are sufficiently close to* $\bar{x}$ *and* $\bar{v}$, *respectively, then the following estimates hold:*

$$\lim_{t\to\infty} \sup_t \frac{\ln \|x(t) - \bar{x}\|}{t} = -\tau, \quad \lim_{t\to\infty} \sup_t \frac{\ln \|v(t) - \bar{v}\|}{t} = -\alpha, \tag{60}$$

$$\lim_{t\to\infty} y^i(t) = \begin{cases} 1, & \text{if } \bar{x}^i \ne 0, \\ 0, & \text{if } \bar{x}^i = 0, \end{cases} \qquad 1 \le i \le n. \tag{61}$$

If $0 < \gamma_* < 2$, *then the discrete version* (59) *converges locally to the pair* $[\bar{x}, \bar{v}]$ *at least linearly. If* $\alpha = \tau = 1$, *then the sequence* $[x_k, v_k]$ *converges quadratically to* $[\bar{x}, \bar{v}]$.

Proof. The pair $[\bar{x}, \bar{v}]$ solves the system (13), therefore it is an equilibrium point of the system (16). We linearize formulas obtained in the previous sections. Define the deviations from the pair $[\bar{x}, \bar{v}]$:

$$\delta x(t) = x(t) - \bar{x}, \qquad \delta v(t) = v(t) - \bar{v}.$$

The equations of the first approximation about the equilibrium are

$$\delta \dot{x}(t) = -\tau \delta x(t), \qquad \delta \dot{v}(t) = -\alpha \delta v(t). \tag{62}$$

Integrating these equations, we obtain

$$\delta x(t) = \delta x(0) e^{-\tau t}, \qquad \delta v(t) = \delta v(0) e^{-\alpha t}. \tag{63}$$

Formulas (19), (28) and (63) enable us to obtain the following approximate representations:

$$K^\top \tilde{g}(t) = \delta x^\perp(t), \qquad A^\top g(t) = \delta v^\parallel(t). \tag{64}$$

From (10) and (64) we have

$$A^\top g(t) - w(v(t)) = \delta v(t), \qquad K^\top \tilde{g}(t) - \tilde{w}(x(t)) = \delta x(t).$$

Substituting these expressions in (30) and (31), we obtain

$$y(t) = D^{-1}(\bar{v} + \delta v(t))\delta v(t) = e_n - D^{-1}(\bar{x} + \delta x(t))\delta x(t).$$

Taking into account (63), we obtain from the last two expressions that the statement (61) holds.

Due to Lyapunov linearization principle, we conclude from (62) that the state $[\bar{x}, \bar{v}]$ is asymptotically stable in the sense of Lyapunov, and property (60) holds, the trajectories of (16) converge locally exponentially to the pair $[\bar{x}, \bar{v}]$.

The statement about linear convergence of the discrete version (59) follows from [6, Theorem 2.3.7]. The proof of quadratic convergence is nearly identical to the proof of convergence of Newton's method. $\square$

In the theorem above we did not impose any conditions on the signs of x and v. Therefore, if a pair $[\bar{x}, \bar{v}]$ satisfies all conditions of the theorem 1 and $\bar{x} \geq 0_n$, $\bar{v} \geq 0_n$, then the solutions of (16) locally converge to the optimal solutions of **(P)** and **(D)**. The pairs $[\bar{x}, \bar{v}]$ that satisfy (13), but not (14), are referred to as a "spurious solutions". In [11] we read: "Spurious solutions abound, and none of them gives any useful information about solutions of **(P)** or **(D)**, so it is best to exclude them altogether from the region of search". In order to justify these pairs we mention that if $\bar{x} \geq 0_n$, $\bar{v} \leq 0_n$, then the trajectories $x(t)$ locally converge to a point that maximizes $c^\top x$, subject to the constraint $x \in X$; the trajectories $v(t)$ converge to the corresponding optimal dual slack vector that is nonpositive in this case. In problem **(P)**, instead of condition $x \geq 0_n$, we can have the condition $x \leq 0_n$. Such a problem will not be worse than **(P)** and the pair x, v will give us the solutions of the maximization or minimization problems if $v \leq 0_n$ or $v \geq 0_n$, respectively.

Define the nonnegative Lyapunov function

$$L(x, v) = D^\lambda(x)v + \|Ax - b\| + \|Kv - d\|$$

and introduce two level sets:

$$\Omega_0 = \left\{ [x, v] : L(x, v) \leq L(x_0, v_0), \;\; x \in \mathbb{R}^n_+, \;\; v \in \mathbb{R}^n_+ \right\},$$

$$\tilde{\Omega}_0 = \left\{ [x, v] \in \Omega_0 : \;\; x \in \mathbb{R}^n_{++}, \;\; v \in \mathbb{R}^n_{++} \right\},$$

where x_0, v_0 are fixed arbitrary vectors from $\mathbb{R}^n_{++}$.

Theorem 2. *Suppose that the problems* **(P)** *and* **(D)** *have a unique optimal pair* $[x_*, v_*]$. *Assume that the vectors* x_0, v_0 *are such that the set* $\tilde{\Omega}_0$ *is nonempty and bounded. Then all trajectories of* (16), *starting from* $\tilde{\Omega}_0$, *converge to the attractor pair* $[x_*, v_*]$.

Proof. The Lyapunov function $L(x, v) > 0$ for all $[x, v]$ from Ω_0, except the pair $[x_*, v_*]$, where $L(x_*, v_*) = 0$. The first integrals of the system (16) are (25), (26) and (38). The solution of (16) belong to Ω_0 and are, therefore, bounded. The right-hand side of (38) is strictly positive and tends to zero only as $t \to \infty$. By moving along the trajectories of (16) we do not violate nonnegativity of x and v as the Lyapunov function $L(x(t), v(t))$ monotonically decreases. Therefore, the trajectories do not cross the boundary of the set Ω_0. All trajectories that emanate from $\tilde{\Omega}_0$ remain in the interior of Ω_0. According to La Salle's Invariance Principle [2], the solutions $x(t)$, $v(t)$ can be extended as $t \to \infty$, their positive limit set is a compact connected attractor contained in Ω_0 and it coincides with the equilibrium pair $[x_*, v_*]$, which is unique on Ω_0. $\square$

Consider the global convergence properties of the discrete variant of the method. For this case, we define the deviation from the optimal solution of problems **(P)** and **(D)** as

$$\Delta x_k = x_k - x_*, \quad \Delta v_k = v_k - v_*, \quad \Delta z_k^\top = [\Delta x_k^\top, \Delta v_k^\top].$$

We use the partition (3), where B is an optimal basis. Therefore, we have $\Delta x_k^N = x_k^N$, $\Delta v_k^B = v_k^B$. We rewrite (59) as follows:

$$W(x_k, v_k)\Delta z_{k+1} = W(x_k, v_k)\Delta z_k - D(\gamma_k)R(x_k, v_k).$$

In detail, we have

$$\begin{array}{lcl} \lambda_k D(v_k)\Delta x_{k+1} & + & D(x_k)\Delta v_{k+1} = (1 + \lambda_k\alpha_k)D(\Delta x_k)\Delta v_k + \\ & + & (1 - \alpha_k)D(x_*)v_k + (\lambda_k - \alpha_k)D(v_*)x_k, \\ K\Delta v_{k+1} & = & (1 - \alpha_k)K\Delta v_k, \\ A\Delta x_{k+1} & = & (1 - \tau_k)A\Delta x_k. \end{array} \tag{65}$$

From (30), (31) and (33), we obtain

$$x_{k+1} = D(x_k)[e_n + \tau_k(y_k - e_n)], \quad v_{k+1} = D(v_k)[e_n - \alpha_k y_k], \tag{66}$$

where

$$y_k = D^{-1}(v_k)\left(A^\top g_k - w(v_k)\right) = e_n + D^{-1}(x_k)\left(\tilde{w}(x_k) - K^\top \tilde{g}_k\right).$$

Let $x_k > 0$, $v_k > 0$. In order to guarantee the nonnegativity of the vectors x_{k+1}, v_{k+1}, the steps α_k, τ_k must satisfy the conditions:

$$e_n \geq \alpha_k y_k, \qquad e_n \geq \tau_k(e_n - y_k).$$

It is easy to see that these conditions hold, if

$$\alpha_k \leq \alpha_k^* = \frac{1}{[y_k^*]_+}, \qquad 0 < \tau_k \leq \tau_k^* = \frac{1}{[1 - y_*^k]_+}, \tag{67}$$

where $[\alpha]_+ = \max[0, \alpha]$; y_k^* and y_*^k are, respectively, the maximum and minimum components of the vector y_k.

The numbers α_k^* and τ_k^* determine the largest possible steps with respect to the primal and dual variables along the Newton's directions for which all the components of the vectors x and v remain nonnegative at the k-th iteration.

If in Theorem 1, we substitute the optimal pair $[x^*, v^*]$ for $[\bar{x}, \bar{v}]$, then in addition to the discrete analog of (61), we obtain that

$$\lim_{k\to\infty} \alpha_k^* = \lim_{k\to\infty} \tau_k^* = 1.$$

Therefore, if k is sufficiently large, the behavior of Newton's method is similar to the "pure Newton's" variant, where $\alpha_k = \tau_k = \lambda_k = 1$. From (65), we obtain for this case

$$\begin{array}{rcl} D(x_k)\Delta v_{k+1} + D(v_k)\Delta x_{k+1} & = & D(\Delta x_k)\Delta v_k, \\ K\Delta v_k & = & 0_\ell, \\ A\Delta x_k & = & 0_m. \end{array} \tag{68}$$

We conclude that, if $k \geq 1$, then all vectors Δx_k belong to the null space of A and all vectors Δv_k belong to the range space of $A^\top$. Therefore, we can represented these vectors as

$$\Delta v_k = -A^\top g_k, \qquad \Delta x_k = -K^\top \tilde{g}_k, \tag{69}$$

where g_k is an m-dimensional vector and $\tilde{g}_k$ is an ℓ-dimensional vector. Substituting these expressions in (68), we obtain a system of linear equations for determining the vector $G_{k+1}^\top = [g_{k+1}^\top, \tilde{g}_{k+1}^\top]$:

$$M(x_k, v_k)G_{k+1} + D(\Delta x_k)\Delta v_k = 0_n,$$

where M is defined in (43).

The inverse of M computed for various cases is given by the formulas (48), (56), (57). If $x_k > 0_n$, $v_k > 0_n$, we can use (48). Taking into account (69), we have

$$\begin{array}{rcl} \Delta v_{k+1} & = & A^\top \left(AD\left(\dfrac{x_k}{v_k}\right)A^\top\right)^{-1} AD^{-1}(v_k)D(\Delta x_k)\Delta v_k, \\ \Delta x_{k+1} & = & K^\top \left(KD\left(\dfrac{v_k}{x_k}\right)K^\top\right)^{-1} KD^{-1}(x_k)D(\Delta x_k)\Delta v_k. \end{array}$$

If the pair $[x_k, v_k]$ is sufficiently close to $[x_*, v_*]$, we can use the formulas (57), which yield

$$\begin{array}{rcl} v_{k+1}^B & \simeq & D\left(\dfrac{x_k^B}{x_*^B} - e_m\right)v_k^B, \\ \Delta v_{k+1}^N & \simeq & N^\top (B^\top)^{-1} v_{k+1}^B, \\ x_{k+1}^N & \simeq & D\left(\dfrac{v_k^N}{v_*^N} - e_\ell\right)x_k^N, \\ \Delta x_{k+1}^B & \simeq & K_B^\top (K_N^\top)^{-1} x_{k+1}^N. \end{array}$$

These results establish the second order of convergence if the current pair $[x_k, v_k]$ is sufficiently close to the optimal pair.

6. Steepest Descent Approach

From the formulas (66), we obtain

$$\begin{array}{rcl} Kv_{k+1} - d & = & (1-\alpha_k)(Kv_k - d), \\ A(x_{k+1} - q) & = & (1-\tau_k)A(x_k - q), \\ \Phi_{k+1} & = & D\left([e_n + \tau_k(y_k - e_n)][e_n - \alpha_k y_k]\right)\Phi_k, \\ \varphi_{k+1} & = & (1-\tau_k)\varphi_k + (\tau_k - \alpha_k)y_k^\top \Phi_k + \alpha_k\tau_k y_k^\top [I_n - D(y_k)]\Phi_k, \end{array}$$

where

$$\Phi_k = D(x_k)v_k, \qquad \varphi_k = x_k^\top v_k.$$

Lemma 4. *If* $-\infty < y_*^k \le y_k^* < \infty$, *then* $0 < \alpha_k^*$, $0 < \tau_k^*$.

Proof. From (67), we have the equality

$$\frac{1}{\alpha_k^*} + \frac{1}{\tau_k^*} = \begin{cases} 1 + y_k^* - y_*^k, & \text{if } y_k^* \ge 0, \ 0 \le y_*^k \le 1; \\ y_k^*, & \text{if } y_*^k > 1; \\ 1 - y_*^k, & \text{if } y_k^* < 0. \end{cases} \tag{70}$$

It follows from (70) that, if the maximal and the minimal components of the vector y_k are bounded, then α_k^* and τ_k^* are bounded away from zero. □

It is possible to obtain more precise lower estimates for α_k^* and τ_k^* directly from (67). Let σ be a positive number satisfying the following inequality:

$$\|y_k\|_\infty \le \sigma, \tag{71}$$

where $\|y\|_\infty = \max\limits_{1 \le i \le n} |y^i|$. In this case we have

$$\alpha_k^* \ge \frac{1}{\sigma}, \qquad \tau_k^* \ge \frac{1}{1+\sigma}.$$

Lemma 5. *Let* $x_k \in X_+$ *and* $v_k \in V_+$. *Then* $\alpha_k^* < \infty$, $\tau_k^* < \infty$.

Proof. It suffices to show that $y_k^* > 0$, $y_*^k < 1$. Suppose the contrary, i.e., let $y_k^* < 0$, then $y_k \le 0_n$. From (32), (38), we have

$$\sum_{i=1}^n x_k^i v^i y_k^i (y_k^i - 1) = 0.$$

Since $x_k > 0$ and $v_k > 0$, this equality is possible only if $y_k = 0_n$. But, according to (34), this implies that $b = 0_m$, which contradicts (12). Hence $y_k^* > 0$.

If $y_*^k > 1$, $y_k \ge e_n$. By the same process as above, we have $y_k = e_n$ and from (35), we obtain that $d = 0_\ell$. We arrive at a contradiction with (12), therefore, $y_*^k < 1$. □

In order for x_{k+1} and v_{k+1} to be interior points, we must have

$$0 \le \alpha_k \le \omega\alpha_k^* = \bar{\alpha}_k, \qquad 0 \le \tau_k \le \omega\tau_k^* = \bar{\tau}_k, \tag{72}$$

where $0 < \omega < 1$.

The steps α_k, τ_k are best chosen so as to minimize the absolute values of all components of the three vectors Φ_{k+1}, $Kv_{k+1} - d$, and $A(x_{k+1} - q)$ under the conditions (72). Thus, we have a multicriteria minimization problem. The simplest way to solve it is to use the following linear convolution function:

$$\vartheta_k(\alpha, \tau) = \varphi_{k+1} + |1-\tau| \, \|A(x_k - q)\| + |1 - \alpha| \, \|Kv_k - d\|. \tag{73}$$

The auxiliary problem thus obtained is: to find

$$\vartheta_k(\tilde{\alpha}_k, \tilde{\tau}_k) = \min_{\substack{0 \le \alpha \le \bar{\alpha}_k \\ 0 \le \tau \le \bar{\tau}_k}} \vartheta_k(\alpha, \tau). \tag{74}$$

The solution of this problem is trivial. It is enough to compare the values of the goal function $\vartheta(\alpha, \tau)$ in nine points $[\alpha_i, \tau_j]$, where

$$\begin{aligned} \alpha_i &= 0, \bar{\alpha}_k \text{ and } 1, \quad \text{if } \bar{\alpha}_k > 1, \\ \tau_j &= 0, \bar{\tau}_k \text{ and } 1, \quad \text{if } \bar{\tau}_k > 1, \end{aligned}$$

and to choose the smallest value.

Finally we outline the algorithm.

DATA: Initial pair $x_0 > 0_n$, $v_0 > 0_n$, the safety factor $0 < \omega < 1$, and the stopping tolerance $\varepsilon > 0$.

For $k = 0, 1, \dots,$ do:

Step 1. Compute $\vartheta_k(0, 0)$ and, if this number is less than ε, then stop.

Step 2. Define the maximal step lengths α_k^*, τ_k^*. Solve the auxiliary minimization problem (74) and define $\tilde{\alpha}_k$, $\tilde{\tau}_k$.

Step 3. Update the current pair $[x_k, v_k]$ by setting

$$x_{k+1} = D(x_k)[e_n + \tilde{\tau}_k(y_k - e_n)], \qquad v_{k+1} = D(v_k)[e_n - \tilde{\alpha}_k y_k].$$

A special case of this algorithm, where $v_0 \in V_+$, was investigated in [2].

Theorem 3. *Let $x_0 \in X_+$, $v_0 \in V_+$. Assume that the sequences x_k, v_k generated by the method of steepest descent* (66), (74) *are such that the inequality* (71) *holds for all k. Then for any $\varepsilon > 0$ the function $\vartheta_k(\tilde{\alpha}_k, \tilde{\tau}_k)$ will become less than ε after not more than*

$$k_* = \left\lceil \frac{1+\sigma}{\omega} \ln \left\lceil \frac{\vartheta_0}{\varepsilon} \right\rceil \right\rceil,$$

iterations, where $\lceil a \rceil$ is the least integer approaching the number a from above.

The proof of this theorem is similar to the one given in [15].

Acknowledgement

All authors would like to thank Natalia Alexandrova for useful comments and editing this paper.

References

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

2. *Bacciotti A.* Local Stability of Nonlinear Control Systems // Series on Advances in Mathematics for Applied Sciences, vol.8, Singapore: World Scientific Publishing Co. Ptc. Ltd., 1992.

3. *Barnes E.* A Variation on Karmarkar's Algorithm for Solving Linear Programming Problems // Mathematical Programming, 1986, vol. 36, pp. 174–182.

4. *Dikin I.I.* Iterative Solution of Problems of Linear and Quadratic Programming // Sov. Math. Dokl., 1967, vol. 8, pp. 674–675.

5. *Evtushenko Yu.G.* Two Numerical Methods of Solving Nonlinear Programming Problems // Sov. Math. Dokl., 1974, vol. 215, N° 2, pp. 420–423.

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

7. *Evtushenko Yu.G. and Zhadan V.G.* Numerical Methods for Solving Some Operations Research Problems // U.S.S.R. Comput. Maths. Math. Phys., 1973, vol. 13, N° 3, pp. 56–77.

8. *Evtushenko Yu.G. and Zhadan V.G.* A Relaxation Method for Solving Problems of Non-Linear Programming // U.S.S.R. Comput. Maths. Math. Phys., 1977, vol. 17, N° 4, pp. 73–87.

9. *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).

10. *Evtushenko Yu.G. and Zhadan V.G.* Barrier-Projective and Barrier-Newton Numerical Methods in Optimization (the Linear Programming Case) // Computing Centre of the USSR Academy of Sciences. Reports in Comput. Math., 1992 (in Russian).

11. *Evtushenko Yu.G. and Zhadan V.G.* Stable Barrier-Projection and Barrier-Newton Methods in Nonlinear Programming // Optimization Methods and Software, 1994, vol. 3, N° 3, pp. 237–256.

12. *Evtushenko Yu.G. and Zhadan V.G.* Stable Barrier-Projection and Barrier-Newton Methods for Linear and Nonlinear Programming // Algorithms for Continuous Optimization. Spedicato E. (ed.). NATO ASI Series. Kluwer Academic Publishers, 1994, pp. 255–285.

13. *Evtushenko Yu.G. and Zhadan V.G.* Stable Barrier-Projection and Barrier-Newton Methods in Linear Programming // Computational Optimization and Applications, 1994, vol. 3, N° 4, pp. 289–303.

14. *Evtushenko Yu.G. and Zhadan V.G.* Barrier-Projective Methods for Nonlinear Programming // Comput. Math. and Math. Phys., 1994, vol. 34, N° 5, pp. 579–590.

15. *Evtushenko Yu.G., Zhadan V.G. and Cherenkov A.P.* Application of Newton's Method to Solving Linear Programming Problems // Comput. Math. and Math. Phys., 1995, vol. 35, N° 6, pp. 850–866.

16. *Evtushenko Yu.G. and Zhadan V.G.* Dual Barrier-Projective Methods in Linear Programming // WSSIAA, 1995, N° 5, pp. 51–66.

17. *Evtushenko Yu.G. and Zhadan V.G.* Space-Transformation Technique: the State of Art // Nonlinear Optimization and Applications. G. Di Pillo and F. Gianessi (eds.). Plenum Press, 1996, pp. 101–123.

18. *Evtushenko Yu.G. and Zhadan V.G.* Dual Barrier-Projective and Barrier-Newton Methods for Linear Programming Problems // Comput. Math. and Math. Phys., 1996, vol. 36, N° 7, pp. 847–859.

19. *Faybusovich L.E.* Hamiltonian Structure of Dynamical Systems which Solve Linear Programming Problems // Physica D, 1991, vol. 53, pp. 217–232.

20. *Gonzaga C.* Path Following Methods for Linear Programming // SIAM Review, 1992, vol. 34, N° 2, pp. 167–224.

21. *Gonzaga C. and Tapia R.* On the Convergence of the Mizuno–Todd–Ye Algorithm to the Analytic Center of the Solution Set // SIAM J. Optim., 1997, vol. 7, N° 1, pp. 47–65.

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

23. *Hertog D.* Interior Point Approach to Linear, Quadratic and Convex Programming. Algorithms and Complexity. Kluwer Academic Publishers, 1994.

24. *Herzel S., Recchioni M.C. and Zirilli F.* A Quadratically Convergent Method for Linear Programming // Linear Algebra and its Applications, 1991, vol. 152, pp. 255–289.

25. *Kojima M., Mizuno S. and Yoshise A.* A Primal-Dual Interior Point Method for Linear Programming // Progress in Mathematical Programming. Interior Point and Related Methods. Megiddo N. (ed.). Berlin: Springer-Verlag, 1989, pp. 29–47.

26. *Kojima M., Megiddo N. and Mizuno S.* A Primal-Dual Infeasible-Interior Point Algorithm for Linear Programming // Mathematical programming, 1993, vol. 61, pp. 263–280.

27. *Lustig I.J., Marsten R.E. and Shanno D.F.* Computational Experience with a Primal-Dual Interior Point Method for Linear Programming // Linear Algebra and its Applications, 1989, vol. 152, pp. 191–222.

28. *Lustig I.J., Marsten R.E. and Shanno D.F.* On Implementing Mehrotra's Predictor-Corrector Interior Point Method for Linear Programming // SIAM J. Optim., 1992, vol. 2, pp. 435–449.

29. *Mizuno S., Todd M.J. and Ye Y.* On Adaptive Step Primal-Dual Interior-Point Algorithms for Linear Programming // Technical Report 944, N.Y., Ithaca, Cornell University, School of Operations Research and Industrial Engineering, 1990.

30. *Monteiro R.C. and Adler I.* Interior Path-Following Primal-Dual Algorithms. Part I: Linear Programming // Mathematical Programming, 1989, vol. 44, pp. 27–41.

31. *Roos C., Terlaky T. and Vial J.Ph.* Theory and Algorithms for Linear Optimization. An Interior Point Approach. John Wiley & Sons, 1997.

32. *Saigal R.* Linear Programming. A Modem Integrated Analysis. Kluwer Academic Publishers, 1995.

33. *Stanevichus A.I.-A., Sherbak L.V.* New Variants of Barrier-Newton Method for Solving Linear Programming Problem // Comput. Math. and Math. Phys., 1995, vol. 35, N° 12, pp. 1443–1452.

34. *Todd M.J. and Ye Y.* A Centered Projective Algorithm for Linear Programming // Math. Oper. Res., 1990, vol. 15, pp. 508–529.

35. *Wright S.J.* Primal-Dual Interior-Point Methods. Philadelphia: SIAM, 1997.

36. *Wright S.J.* Modified Cholesky Factorization in Interior-Point Algorithms for linear programming // Preprint ANL/MCS-P600-0596, Mathematics and Computer Science Division, Argonne National Laboratory, 1996.

37. *Zang Y. and Tapia R.A.* Superlinear and Quadratic Convergence of Primal-Dual Interior-Point Methods for Linear Programming // SIAM J. Optim., 1993, vol. 3, pp. 118–133.