An interior point method for solving systems of linear equations and inequalities

Markku Kallio * , Seppo Salo
Helsinki School of Economics, Runeberginkatu 14-16, Helsinki 00100, Finland

Received 1 March 1999; received in revised form 1 March 2000

Abstract

A simple interior point method is proposed for solving a system of linear equations subject to nonnegativity constraints. The direction of update is defined by projection of the current solution on a linear manifold defined by the equations. Infeasibility is discussed and extension for free and bounded variables is presented. As an application, we consider linear programming problems and a comparison with a state-of-the-art primal–dual infeasible interior point code is presented. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Interior point algorithms; Linear inequalities; Linear programming

1. The problem and a method of solution

An inspiration to our article arises from the vast literature on interior point methods for solving mathematical programming problems (see e.g. [2]). In particular, we are motivated by the success of primal–dual infeasible interior point algorithms for linear programming. Our aim is to solve systems of linear equations subject to nonnegativity constraints, for which optimality conditions for linear programming is a special case. We propose a simple interior point algorithm which can be initiated with any positive vector. It always converges, and if the problem is feasible, the limit point is a solution.

Let \( q \in \mathbb{R}^m \) and \( Q \in \mathbb{R}^{m \times n} \) with \( m < n \) and \( \text{rank}(Q) = m \). Our problem is to find \( z \in \mathbb{R}^n \) such that

\[
Qz = q, \quad \text{z} \geq 0. \tag{1.1}
\]

Denote \( M = \{ z \in \mathbb{R}^n \mid Qz = q \} \), \( Z = \text{diag}(z) \), and for \( d \in \mathbb{R}^n \), define \( Z^{-2} \)-norm by \( \|d\|_{Z^{-2}} = (d^T Z^{-2} d)^{1/2} \). Our method starts with \( z^0 \in \mathbb{R}^n \), \( z^0 \geq 0 \). For iteration \( k, k = 0, 1, 2, \ldots \), denote \( Z = \text{diag}(z^k) \), suppressing \( k \) in \( Z \). The direction \( d^k \) of update is defined by \( d^k = z - z^k \) where \( z \) is the point of \( M \) which is closest to \( z^k \) in \( Z^{-2} \)-norm; i.e., \( z \) is the unique solution to the following problem:

\[
\min_{z \in M} \|z - z^k\|_{Z^{-2}}. \tag{1.3}
\]

A step in this direction is taken. If \( z^k + d^k \geq 0 \) a solution for (1.1)–(1.2) is found; otherwise the step size \( \tau^k < 1 \) is applied so that \( z^{k+1} = z^k + \tau^k d^k > 0 \).
The steps of the algorithm are formalized as follows:

0. **Initialization.** Set the iteration count $k$ to 0, pick any $z^0 \in \mathbb{R}^n$ such that $z^0 > 0$, and let the step size parameter $\theta \in (0, 1)$.

1. **Termination test.** If $z^k$ satisfies (1.1)–(1.2), then stop.

2. **Direction.** Find $d^k$ according to (1.3).

3. **Step size.** Let $\bar{\tau}^k$ be the largest step size for which $z^k + \bar{\tau}^k d^k \geq 0$. If $\bar{\tau}^k \geq 1$, set $\tau^k = 1$; otherwise set $\tau^k = \theta^{\bar{\tau}^k}$.

4. **Update.** Compute

$$z^{k+1} = z^k + \tau^k d^k.$$  

Increment $k$ by one and return to Step 1.

### 2. Convergence

To begin, we show that the sequence generated by the algorithm converges. Thereafter we prove that the limit point solves (1.1)–(1.2) if a feasible solution exists. Several preliminary results are employed. The following three results characterize the update direction and convergence of the residual vector for (1.1).

**Lemma 1.** The direction vector defined by (1.3) is given by

$$d^k = Z^2 Q^T (Q Z^2 Q^T)^{-1} \tilde{q}^k,$$  

where $\tilde{q}^k$ is the residual vector defined by $\tilde{q}^k = q - Q z^k$, and $Z = \text{diag}(z^k)$.

**Proof.** By straightforward optimization in (1.3). \(\square\)

**Lemma 2.** There exists a monotonically decreasing sequence $\{x^k\}$ of scalars converging to $\tilde{z} \geq 0$ such that the residual vectors $\tilde{q}^k$, $k = 1, 2, \ldots$, satisfy

$$\tilde{q}^k = q - Qz^k = (1 - \tau^{k-1}) \tilde{q}^{k-1} = x^k \tilde{q}^0.$$  

**Proof.** Employing (1.4)–(2.1), $\tilde{q}^{k+1} = q - Q z^{k+1} = \tilde{q}^k - \tau^k Q d^k = (1 - \tau^k) \tilde{q}^k$, for $k \geq 0$. Observing that $0 < \tau^k \leq 1$, the sequence $\{x^k\}$ is nonnegative and monotonically decreasing and hence it converges to some $\tilde{z} \geq 0$. \(\square\)

**Lemma 3.** Let $H \in \mathbb{R}^{n \times (n-m)}$ be a matrix whose columns form a basis for the null space of $Q$, and let $h \in \mathbb{R}^n$ such that $Q h = q^0$. Then

$$d^k = z^k [h - H (H^T Z^{-2} H)^{-1} H^T Z^{-2} h],$$  

where $d^k$ is the direction given by (2.1), the scalars $z^k$ are defined by (2.2), and $Z = \text{diag}(z^k)$.

**Proof.** By Lemma 2, $q^k = q - Q z^k = z^k q^0 = z^k Q h$. Hence, by definition of $H$, any solution $z$ for $Qz = q$ is given by $z = z^k + \tilde{z}^k h - H v$ so that $d^k = z^k h - H v$, for some $v \in \mathbb{R}^{n-m}$. To determine $v$, problem (1.3) can be stated as $\min_v \in \mathbb{R}^{n-m} (h z^k - H v)^T Z^{-2} (h z^k - H v)$. Optimization yields $v = z^k (H^T Z^{-2} H)^{-1} H^T Z^{-2} h$ implying (2.3). \(\square\)

Lemma 5 below implies that the set of all possible direction vectors is bounded. To prove this result we employ a preliminary result and the following notation. Let $\mathcal{B} = \{j_1, \ldots, j_m\}$, $1 \leq j_1 < \cdots < j_m \leq n$. For $B \in \mathcal{B}$, let $Q_B \in \mathbb{R}^{m \times m}$ be the submatrix defined by columns (indexed by) $B$ of $Q$. Similarly, for $I \in \mathbb{R}^{n \times n}$, let $I_B \in \mathbb{R}^{m \times m}$ be the submatrix defined by columns $B$ of $I$. Denote $\mathcal{B} = \{B \in \mathcal{B} | \text{rank}(Q_B) = m\}$.

**Lemma 4.** Let $Q \in \mathbb{R}^{m \times n}$, rank($Q$) = $m$ and $S = \text{diag}(s)$, $s \in \mathbb{R}^n$, $s > 0$. Then

$$S Q^T (Q S Q^T)^{-1} = \sum_{B \in \mathcal{B}} \lambda_B I_B Q_B^{-1},$$  

where

$$\lambda_B = \Lambda_B / A,$$  

$$\Lambda_B = |Q_B|^2 \prod_{j \in B} s_j$$  

and

$$A = |Q S Q^T| = \sum_{B \in \mathcal{B}} \Lambda_B = \sum_{B \in \mathcal{B}} \Lambda_B.$$  

**Proof.** For $s \in \mathbb{R}^n$ and $B \in \mathcal{B}$, replacing components $s_j$, $j \notin B$, by zero, we obtain a vector denoted by $s_B$. Let $S_B = \text{diag}(s_B)$ and $I_B = I_B^2 S_B I_B$ (the diagonal matrix of elements $s_j$, $j \in B$). Then $A = \Lambda(s)$, the determinant of $Q S Q^T$, is a polynomial of $s$ with the following properties: (i) each term of $\Lambda(s)$ is of degree $m$ because each element of $Q S Q^T$ is a linear function of $s$; (ii) for all $B \in \mathcal{B}$, $\Lambda(s_B) = |Q_B S_B Q_B^T| = |Q_B S_B Q_B^T|$. 


\(|Q_B|^2|S_B| = \Delta_B\); (iii) if \(s\) has less than \(m\) nonzero components, then \(\text{rank}(Q) < m\) and \(\Delta(s) = 0\).

Consider a term \(A\prod_{j \in B} s_j^m\) of \(A(s)\) and suppose \(m_j > 1\), for some \(j\). Then (i) implies \(|B| < m\). Hence (iii) implies \(\Delta(s) = 0\), for all \(s\) such that \(s_j = 0\), for \(j \notin B\). It is elementary to show, that the polynomial \(\Delta(S)\) with such a property must have \(a = 0\). Consequently, in a term \(A\prod_{j \in B} s_j^m\), \(a \neq 0\) implies \(m_j = 1\), \(j \notin B\), and furthermore, \(B \in \mathcal{B}\) by (i). Hence by (ii), \(\Delta(s) = \sum_{B \in \mathcal{B}} \Delta_B\). But \(\Delta(s_B) = \Delta_B = 0\), for \(B \in \mathcal{B} \setminus \mathcal{B}\), so that (2.7) follows.

To prove (2.4), we employ the following known facts: if \(A\) and \(B\) are square matrices of equal size and \(A^\ast\) and \(B^\ast\) denote their adjoint matrices, then \(A^\ast = |A| l\) and \((AB)^\ast = B^\ast A^\ast\). Denote \(P = SQ^T(QSQ^T)^\ast\). Elements of \((QSQ^T)^\ast\) are minors of order \(m - 1\) of \(|Q|S^2|Q^\ast|\). Therefore, each element of the matrix \(P(s)\), is a polynomial of \(s\) and each term of such polynomials is of degree \(m\). Using the facts above, for all \(B \in \mathcal{B}\), \(P(s_B) = S_B^2 Q^T(QSQ^T)^\ast = I_B S_B Q^T(Q_B S_B Q^T)^\ast = |S_B| |Q_B| I_B Q_B^\ast\). Hence, \(P(s) = 0\) for all \(s\) with \(s = s_B\), \(B \in \mathcal{B} \setminus \mathcal{B}\), or with \(s\) having less than \(m\) nonzero components. Repeating the same arguments as with \(A(s)\), properties of \(P(s)\), with \(Q_B = |Q_B|Q_B^{-1}\), for \(B \in \mathcal{B}\), imply
\[
P = \sum_{B \in \mathcal{B}} \Delta_B I_B Q_B^{-1}, \tag{2.8}
\]
where \(\Delta_B\) is given by (2.6). Finally, observing that \(SO^T(QSQ^T)^{-1} = A^{-1}P\), (2.8) yields (2.4).

**Corollary 1.** For \(B \in \mathcal{B}\), let \(d_B^k\) be the basic solution to \(Qd = q^k\) with \(d_j, j \in B\) as basic variables, i.e. \(d_B^k = I_B Q_B^{-1} q^k\). Then the direction vector \(d^k\) in Lemma 1 is the convex combination
\[
d^k = \sum_{B \in \mathcal{B}} \lambda_B d_B^k, \tag{2.9}
\]
where \(\lambda_B\)’s are defined by (2.5)-(2.7) and \(S = Z^2\), with \(Z = \text{diag}(z^k)\).

**Proof.** By direct application of Lemma 4. □

**Lemma 5.** Let \(H\) and \(h\) be defined as in Lemma 3, let \(s \in \mathbb{R}^n\) and define \(S = \text{diag}(s)\) so that \(Se = s\). For \(0 < s < \infty\), define \(d_s = h - H(H^T S^{-2} h)^{-1}H^T S^{-2} h\). Then there is a scalar \(\rho < \infty\) such that \(\text{sup}_s \|d_s\| < \rho\).

**Proof.** By direct application of Lemma 4. □

**Theorem 1.** The sequence \(\{z^k\}\) generated by the algorithm converges.

**Proof.** If the sequence \(\{z^k\}\) is finite, the algorithm stops in a solution \(\hat{z}\) satisfying (1.1)-(1.2). Assume next that the sequence \(\{z^k\}\) is infinite. By Lemma 5, there is \(\rho' < \infty\) such that in (2.3), \(d^k = x^k(h-r)\) for some \(r\) such that \(Qr = 0\) and \(\|r\| < \rho'\). Define a closed cone of directions \(D = \{d \in \mathbb{R}^n | d = \mu(h-r), \mu > 0, Qr = 0, \|r\| < \rho\}\) so that \(d^k \in D\) for all \(k\).

By Lemma 2, the residuals \(\hat{q}^k\) converge to \(\hat{q}^0\). Define \(M = \{z \in \mathbb{R}^n | q - Qz = \hat{q}^0, z > \hat{x}\}\). For \(z \in \mathbb{R}^n\), denote \(z + D = \{w \in \mathbb{R}^n | w = z + d, d \in D\}\). Let \(C^k = \{z^k + D\} \cap M\) so that \(z^k \in C^k\). For some \(d \in D\), \(z^{k+1} = z^k + d \in C^k\), because \(z^{k+1} \in M\). Because \(D\) is a cone, we have \(C^{k+1} = \{z^{k+1} + D\} \cap M = \{z^k + d + D\} \cap M \subset \{z^k + D\} \cap M = C^k\). Consequently, \(z^l \in C^k\) for \(l \geq k\).

If \(z \in C^k\), then there is \(r \in \mathbb{R}^n\), with \(Qr = 0\) and \(\|r\| < \rho\), and \(\mu > 0\) such that \(z = z^k + \mu(h-r)\). Hence, employing Lemmas 2 and 3, \(Qr - Qz = q - Qz^k\), \(z^k - \hat{x} \geq 0\). Thus, \(\hat{x} - \hat{z} \geq 0\). By Lemma 2, \(z^k \to \hat{z}\) converges to \(\hat{z}\), and therefore, the diameter of \(C^k\) converges to \(0\). Hence, \(\{z^k\}\) is a Cauchy sequence, wherefore it converges. □

**Theorem 2.** If a feasible solution \(\hat{z}\) for (1.1)-(1.2) exists, then \(\{z^k\}\) converges to a feasible solution \(\hat{z}\). Furthermore, if \(\{z^k\}\) is infinite, then \(\hat{z}_j = 0\) implies \(\hat{z}_j = 0\) for all feasible solutions \(\hat{z}\).

**Proof.** Assume that a feasible solution \(\hat{z}\) exists. By Theorem 1, \(z^k\) converges to \(\hat{z}\). If \(n = m\), Lemma 1 implies \(z^l = \hat{z} \geq 0\). If convergence is finite, \(\hat{z}\) is feasible and there is nothing to prove. Otherwise, after possible permutation of the variables, partition \(z^1 = (z_1^1, z_2^1)\), with \(z_i \in \mathbb{R}^n, i = 1, 2\), such that \(z_1 > 0, z_2 > 0\). Suppose \(n_1 = 0\); i.e. \(\hat{z} > 0\). If \(\hat{z} < 1\) for all \(k\), then (1.4) yields \(z^{k+1} = z^k + \hat{q}^k d^k\), with \(\hat{z} > 0\) and \(\lim inf_k \|z^k d^k\| > 0\). This contradicts convergence of \(z^k\). Hence, convergence is finite and \(\hat{z} > 0\) is feasible.
From now on we assume that $n > m$, $n_1 > 0$ and \{z^k\} is infinite. For all $k$ and $j$ ($j = 1, \ldots, n$), define $v^j_k = \sup |s_k| z^j_k$, which converge to zero. Consider a collection of diagonal scaling matrices $S = \text{diag}(s) \in R^{n \times n}$, where $s \in V^k = \{s \in R^n \mid s \geq 0, |s_j - z^j_k| \leq v^j_k \}$ for all $j$. Then $z^k \in V^k$ and $Z = \text{diag}(z^k)$ belongs to this collection. Because $v^j_k$ is nonincreasing, for all $j$, $z^l \in V^l \subset V^k$, for all $l \geq k$. For all $k$, observing (2.3), define a closed cone of directions
\[
D^k = \{d \in R^n \mid d = \mu S^2 u, \mu \geq 0, S^2 u = h - Hv, \n H^T u = 0, \ s \in V^k \}
\]
so that $d^l \in D^l \subset D^k$, for $l \geq k$. Define $C^k = z^k + D^k$ so that $C^k$ is a closed cone with vertex at $z^k$. Hence $z^l \in C^l \subset C^k$, for all $l \geq k$. Because $C^k$ is closed, $\bar{z} \in C^k$.

We aim to show that, if $\bar{x} > 0$ or $\bar{z} = 0$ and $\bar{z} \neq 0$, then $\bar{z} \notin C^k$, for some $k$, wherefore $\bar{x} = 0$ and $\bar{z} = 0$.

We proceed with the following decomposition:
\[
d = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}, \quad S = \begin{bmatrix} S_1 & 0 \\ 0 & S_2 \end{bmatrix},
\]
where $d_i \in R^n$ and $S_i \in R^{n \times n}$, for $i = 1, 2$.

Pick an index $v \leq n_1$ so that $z_v = 0$. For all $k$, there is $l \geq k$, such that $z^l_v = u_v^l$. Because $\bar{z} \in C^l$, $\bar{z} = z^l + d$, for some $d \in D^l$. Hence, $-d_1 = z^l_v - \bar{z}_v = z^l_v - \mu S^2 u_1 > 0$ with $\mu > 0$ and the diagonal of $S_1$ strictly positive. Consequently, for $k$ large enough, the diagonal of $S$ is strictly positive and $d = \mu (h - Hv)$, with $\mu = x^l - \bar{x}$ and $H^T u = H^T S^{-2} (h - Hv) = 0$. In the latter equation, premultiplication by $v^T$ yields
\[
v^T H^T S^{-2} (h - Hv) = 0. \quad (2.10)
\]

By assumption, $Q \bar{z} = q$ so that $Q (\bar{z} - z^l) = x^l q^0$. Let $h = (\bar{z} - z^l) / x^l$ so that $Q h = q^0$. Denote $\beta = \mu x^l - \bar{x} / x^l$ so that $\beta \in (0, 1]$. Then $d = z^l - \bar{z} = \mu (h - Hv) = \beta (z^l - z^l) - \mu Hv$. Solving for $\mu Hv$ yields
\[
\mu Hv = \bar{z} - \bar{\bar{z}}, \quad (2.11)
\]
where $\bar{\bar{z}} \equiv \bar{z} + (1 - \beta) z^l$, and
\[
\mu (h - Hv) = \bar{z} - z^l. \quad (2.12)
\]

With (2.11)–(2.12) we have $-\mu^2 v^T H^T S^{-2} (h - Hv) = (z^l - \bar{z})^T S^{-2} (z^l - \bar{z})$. Decomposing the latter expression and observing (2.10) yields
\[
\bar{z}_1^T S_1^{-2} \bar{z}_1^1 + (z_2 - \bar{z}_2)^T S_2^{-2} (z_2^1 - z_2) = 0. \quad (2.13)
\]
Let $k$ and $l \geq k$ increase without limit. Then, if $n_1 < n$, $z^k_1$ converges to $\bar{z}_2 > 0$ and $S_2$ converges to diag($\bar{z}_2$) so that $(\bar{z}_2 - z^k_2)^T S_2^{-2} (z^k_2 - \bar{z}_2)$ converges to zero. For the first term in (2.13), observing that $z^k_1 - \bar{z}_1 > 0$, we have
\[
z_1^k S_1^{-2} z^k_1 > z_1 z_1 / s^k_1 \geq z_1 / v^1. \quad (2.14)
\]

Suppose $\bar{z} > 0$. Then $\beta$ converges to zero and $z^k_1/v^1$ in (2.14) converges to one. Therefore, for some $k$ large enough and $l \geq k$ such that $z^l_1 = \bar{z}_1^l$, (2.13) does not have a solution, implying $\bar{z} \notin C^l$. This is a contradiction, wherefore $\bar{z} = 0$.

For the second part of the assertion, suppose $\bar{x} = 0$ and $\bar{z} > 0$. Then $\mu = x^l$, $\beta = 1$ and $\bar{z} = z^l$ so that $\bar{z}_v / v^1 = \bar{z}_v / v^1$, which diverges as $v^1$ converges to zero. Therefore, for some $k$ and $l \geq k$, (2.13) does not have a solution, implying $\bar{z} \notin C^l$. This is a contradiction, wherefore $\bar{z}_v = 0$. \hfill \Box

3. Infeasibility

Detection of infeasibility may take place in alternative ways. First, if (1.1)–(1.2) does not have a feasible solution, then by Theorem 1 the sequence \{z^k\} converges to a point $\bar{z}$. By Lemma 2, the corresponding residual vector is given by $q - Q \bar{z} = \bar{q} q^0$, where $\bar{q} q^0$ is the residual vector at the initial point $z^0$ and $\bar{q} > 0$ is the limit point of the nonnegative and monotone decreasing sequence \{z^k\} defined by Lemma 2. Thus, an indication for infeasibility is that $x^k$ converges to a strictly positive level. However, in a finite number of iterations, it may be difficult to judge when to conclude infeasibility and stop.

Alternatively, we may modify the method of Section 1 (in a way to be discussed shortly) to find $z \in R^n$ and $\kappa \in R$ to satisfy the homogeneous system
\[
Qz - \kappa q = 0, \quad (3.1)
\]
\[
z \geq 0, \quad \kappa \geq 0, \quad (3.2)
\]
which is always feasible.

Direct application of the method in Section 1 may not work as illustrated by the following example. Let (1.1) be given by $z_1 - z_2 = 5$ so that $\bar{z}_1 = 5$, $\bar{z}_2 = 0$ and $\kappa = 1$ solve (3.1)–(3.2). Initializing our method for (3.1)–(3.2) with $z^0_1 = z^0_2 = \kappa = 1$ yields in one iteration $z^1_1 = 12$, $z^1_2 = 0.6$ and $\kappa = 0$, which is feasible for (3.1)–(3.2) so that the iterations stop. However, this solution provides a homogeneous solution, but not a feasible solution for (1.1)–(1.2).
To prevent such an early termination, we modify the step size rule of Step 3 in Section 1 by setting \( \tau^k = 1 \) if \( \tau^k > 1 \) (instead of the earlier condition \( \tau^k \geq 1 \)). \(^1\) Otherwise as before, \( \tau^k = \theta \tilde{\tau}^k \), where \( \theta \) is the step size parameter. Then all results of Section 2 remain valid, but convergence is finite only if a feasible interior point is encountered. Hence by Theorem 2, the modified algorithm applied to (3.1)–(3.2) converges to a feasible solution \( \tilde{z} \) and \( \tilde{k} \). If the original system (1.1)–(1.2) has a feasible solution \( \tilde{z} \), then \( \tilde{z} \) and \( \tilde{k} = 1 \) solve (3.1)–(3.2). In such a case, if convergence is finite, we obtain an interior solution \( \tilde{z} > 0 \) and \( \tilde{k} > 0 \), and if convergence is infinite, \( \tilde{k} > 0 \) implies \( \tilde{k} > 0 \) by Theorem 2. Consequently in both cases, \( \tilde{z}/\tilde{k} \) solves (1.1)–(1.2). In the example above, with \( \theta = 0.99 \), finite convergence occurs in iteration two, and rounded figures are \( \tilde{z}_1 = 1.228, \tilde{z}_2 = 0.589 \) and \( \tilde{k} = 0.00999 \). Conversely, if (1.1)–(1.2) is infeasible, then the solution \( \tilde{z} \) and \( \tilde{k} \) for the homogeneous system (3.1)–(3.2) obtained by the method of Section 1, employing the modified step size rule, satisfies \( \tilde{k} = 0 \).

4. Free and bounded variables

We consider next a modification of the system (1.1)–(1.2), where some of the variables are free and some have simple bounds. A free variable \( z_j \) has an upper bound equal to \( \infty \) and a lower bound equal to \( -\infty \). If a variable \( z_j \) is nonfree, but its lower bound is \( -\infty \), we replace \( z_j \) by \(-z_j\), so that the revised variable has a finite lower bound. Thereafter, we translate each nonfree variable so that its lower bound becomes 0; i.e., if a variable \( z_j \) has a lower bound \( l_j > -\infty \), we replace \( z_j \) by \( z_j + l_j \). Thereby, if the revised upper bound is less than \( \infty \), it is denoted by \( u_j \). Finally, we subdivide the set of indices \( J = \{1, 2, \ldots, n\} \) into \( J = J_F \cup J_B \cup J_P \) so that

\[
-\infty \leq z_j \leq \infty \quad \text{for } j \in J_F, \\
0 \leq z_j \leq u_j \quad \text{for } j \in J_B, \\
0 \leq z_j \leq \infty \quad \text{for } j \in J_P.
\]

Hence, in the resulting modified system, (1.2) is replaced by 0 \( \leq z_j \leq u_j \), for \( j \in J_B \), and 0 \( \leq z_j \), for \( j \in J_P \).

Introducing new variables in the usual way, an equivalent system can be reformulated employing nonnegative variables only. For free variables \( j \in J_F \), substitute \( z_j \) by \( z^+_j - z^-_j \) with \( z^+_j \geq 0 \) and \( z^-_j \geq 0 \). For bounded variables \( j \in J_B \), introduce a slack variable \( v_j \) so that the requirement \( z_j \leq u_j \) becomes \( z_j + v_j = u_j \) with \( v_j \geq 0 \). Denoting by \( Q_j \) the column vector of \( Q \), the system is now revised to

\[
\sum_{j \in J_F} Q_j(z^+_j - z^-_j) + \sum_{j \in J_B} Q_j z_j + \sum_{j \in J_F} Q_j v_j = q, \quad (4.1)
\]

\[
z_j + v_j = u_j \quad \text{for } j \in J_B, \quad (4.2)
\]

\[
z^+_j, z^-_j \geq 0 \quad \text{for } j \in J_F, \quad z_j, v_j \geq 0 \quad \text{for } j \in J_B, \quad (4.3)
\]

Clearly, the method of Section 1 applies to solve system (4.1)–(4.3). Straightforward algebra shows, that the following steps can be employed to compute the direction vector of update in each iteration \( k \). Suppressing the iteration count \( k \) for variables \( z^+_j, z^-_j, v_j \), and \( z_j \), define a diagonal scaling matrix \( S = \text{diag}(s) \in R^{n \times n} \) as follows:

\[
s_j = \begin{cases}
(z^+_j)^2 + (z^-_j)^2 & \text{for } j \in J_F, \\
(v_j z_j)/(v_j^2 + z_j^2) & \text{for } j \in J_B, \\
z_j^2 & \text{for } j \in J_P.
\end{cases}
\]

Define residuals, \( \hat{u}^k_j = u_j - z_j - v_j \), for \( j \in J_B \), and \( q^k = q^k - \sum_{j \in J_B} s_j \hat{u}^k_j Q_j \hat{u}^k_j \). Define an auxiliary vector \( \lambda^k \in R^m \) by \( \lambda^k = (Q^T S Q)^{-1} q^k \) and scalars \( \mu^k_j \), for \( j \in J_B \), by \( \mu^k_j = (u_j - z_j^2 Q_j^T \lambda^k_j)/(z_j^2 + v_j^2) \). Then the components of the direction vector, obtained by applying (2.1) to the revised system (4.1)–(4.3), are as follows:

\[
d^k_j = \begin{cases}
(z^+_j)^2 Q_j^T \lambda^k_j & \text{for } z^+_j, j \in J_F, \\
-(z^-_j)^2 Q_j^T \lambda^k_j & \text{for } z^-_j, j \in J_F, \\
v_j^2 \mu^k_j & \text{for } v_j, j \in J_B, \\
z_j^2 Q_j^T \lambda^k_j & \text{for } z_j, j \in J_P.
\end{cases}
\]

Note that the essential part of computational effort relates to the inverse of the matrix \( Q S Q^T \), which has the same dimension and structure as the matrix \( Q Z Q^T \) in (2.1) of the original system (1.1)–(1.2).

\(^1\) In an alternative modification we set \( \tau^k = 1 \) if \( \tau^k \geq 1 \), provided that the component \( \kappa \) does not decrease to zero in iteration \( k \). To shorten the discussion, we do not pursue this alternative further.
Table 1
The number of rows (Rows), columns (Cols) of test problems, as well as the number of iterations for Andersen (AA) and the method (KS) of Section 1

<table>
<thead>
<tr>
<th>Problem</th>
<th>Rows</th>
<th>Cols</th>
<th>AA</th>
<th>KS</th>
</tr>
</thead>
<tbody>
<tr>
<td>afiro</td>
<td>27</td>
<td>32</td>
<td>9</td>
<td>21</td>
</tr>
<tr>
<td>sc50b</td>
<td>50</td>
<td>48</td>
<td>11</td>
<td>18</td>
</tr>
<tr>
<td>sc50a</td>
<td>50</td>
<td>48</td>
<td>11</td>
<td>20</td>
</tr>
<tr>
<td>sc105</td>
<td>105</td>
<td>103</td>
<td>12</td>
<td>21</td>
</tr>
<tr>
<td>adlittle</td>
<td>56</td>
<td>97</td>
<td>13</td>
<td>24</td>
</tr>
<tr>
<td>scagr7</td>
<td>129</td>
<td>140</td>
<td>15</td>
<td>27</td>
</tr>
<tr>
<td>stocfor1</td>
<td>117</td>
<td>111</td>
<td>12</td>
<td>24</td>
</tr>
<tr>
<td>blend</td>
<td>74</td>
<td>83</td>
<td>10</td>
<td>24</td>
</tr>
<tr>
<td>sc205</td>
<td>295</td>
<td>203</td>
<td>12</td>
<td>23</td>
</tr>
<tr>
<td>share2b</td>
<td>96</td>
<td>79</td>
<td>11</td>
<td>27</td>
</tr>
<tr>
<td>lotfi</td>
<td>153</td>
<td>308</td>
<td>17</td>
<td>30</td>
</tr>
<tr>
<td>share1b</td>
<td>117</td>
<td>225</td>
<td>26</td>
<td>38</td>
</tr>
<tr>
<td>scorpion</td>
<td>388</td>
<td>358</td>
<td>14</td>
<td>21</td>
</tr>
<tr>
<td>scagr25</td>
<td>471</td>
<td>500</td>
<td>16</td>
<td>40</td>
</tr>
<tr>
<td>scatp1</td>
<td>300</td>
<td>480</td>
<td>19</td>
<td>32</td>
</tr>
<tr>
<td>brandy</td>
<td>220</td>
<td>249</td>
<td>14</td>
<td>30</td>
</tr>
<tr>
<td>israel</td>
<td>174</td>
<td>142</td>
<td>19</td>
<td>48</td>
</tr>
<tr>
<td>scsd1</td>
<td>77</td>
<td>760</td>
<td>10</td>
<td>25</td>
</tr>
<tr>
<td>agg</td>
<td>488</td>
<td>163</td>
<td>17</td>
<td>72</td>
</tr>
<tr>
<td>bandm</td>
<td>305</td>
<td>472</td>
<td>15</td>
<td>30</td>
</tr>
<tr>
<td>e226</td>
<td>223</td>
<td>282</td>
<td>16</td>
<td>35</td>
</tr>
<tr>
<td>scfxm1</td>
<td>330</td>
<td>457</td>
<td>18</td>
<td>38</td>
</tr>
<tr>
<td>beaconfd</td>
<td>173</td>
<td>262</td>
<td>12</td>
<td>25</td>
</tr>
<tr>
<td>degen2</td>
<td>444</td>
<td>534</td>
<td>10</td>
<td>27</td>
</tr>
<tr>
<td>agg2</td>
<td>516</td>
<td>302</td>
<td>18</td>
<td>34</td>
</tr>
<tr>
<td>agg3</td>
<td>516</td>
<td>302</td>
<td>17</td>
<td>35</td>
</tr>
<tr>
<td>scrs8</td>
<td>490</td>
<td>1169</td>
<td>25</td>
<td>59</td>
</tr>
<tr>
<td>scsd6</td>
<td>147</td>
<td>1350</td>
<td>12</td>
<td>27</td>
</tr>
<tr>
<td>ship04s</td>
<td>402</td>
<td>1458</td>
<td>17</td>
<td>25</td>
</tr>
<tr>
<td>bnl1</td>
<td>643</td>
<td>1175</td>
<td>26</td>
<td>78</td>
</tr>
<tr>
<td>scfxm2</td>
<td>660</td>
<td>914</td>
<td>20</td>
<td>41</td>
</tr>
</tbody>
</table>

5. Application to linear programming

Let \( A \in \mathbb{R}^{k \times l} \), \( c \in \mathbb{R}^l \) and \( b \in \mathbb{R}^k \), and consider the following linear programming problem:

\[
\begin{align*}
\min_{x \in \mathbb{R}^l} & \{ c^T x \mid Ax = b, \ x \geq 0 \} \\
\text{and its dual} & \\
\max_{y \in \mathbb{R}^k, s \in \mathbb{R}^l} & \{ b^T y \mid A^T y + s = c, \ s \geq 0 \}.
\end{align*}
\]

Optimality conditions for (5.1)–(5.2) may be stated by (1.1) with

\[
Q = \begin{bmatrix} 0 & A^T & 1 \\ A & 0 & 0 \\ -c^T & b^T & 0 \end{bmatrix}, \quad q = \begin{bmatrix} c \\ b \\ 0 \end{bmatrix}, \quad z = \begin{bmatrix} x \\ y \\ s \end{bmatrix},
\]

and by nonnegativity constraints \( x \geq 0 \) and \( s \geq 0 \). To deal with the free variables \( y \) and to solve the problem, we may proceed as described in Section 4.

A rudimentary code is set up to test the method for some linear programming test problems from the Netlib collection. With this code, we only aim to study the number of iterations needed to solve the problem. Otherwise we do not worry about computational efficiency and memory requirements at this stage. Consequently, we employ dense Cholesky factorization to deal with the inverse \((QSQ^T)^{-1}\) and we use high
enough accuracy to deal with ill conditioning of the matrix $QSQ^T$. We prescale the matrix $A$ with a procedure that aims to equalize the norms of all rows and columns of $A$ close to 1. Thereafter, to initialize the iterations, we set all variables equal to ten. Our experimental code does not deal with simple bounds; i.e. all variables are nonnegative or free. We compare the results with a state of the art interior point code by Andersen and Andersen [1], wherefore a similar stopping criterion is employed in both codes. In particular, the vital requirement is that the absolute value of the relative duality gap $|c^T x - b^T y|/(1 + |c^T x|)$ is less than $10^{-8}$. Besides, both codes employ primal and dual infeasibility tolerances, which differ, because initialization procedures differ. We require the norms of primal and dual infeasibility vectors relative to norms of $b$ and $c$, respectively, to be less than $10^{-6}$. For the step size parameter, we use $\theta = 0.99$.

The results are shown in Table 1 below. We obtain approximately twice as many iterations as compared with using MOSEK v1.0.beta [1]. Given that the latter code employs most known enhancements for primal–dual infeasible interior point methods, we may regard our results quite satisfactory.

Acknowledgements

The authors wish to thank Erling Andersen for providing them with the results obtained with MOSEK, as well as an anonymous referee for drawing their attention to infeasible problems. Financial support for this work is gratefully acknowledged from The Foundation for the Helsinki School of Economics.

References
