On improved Choi–Goldfarb solution-containing ellipsoids in linear programming

Igor S. Litvinchev

Computing Center, Russian Academy of Sciences, Vavilov 40, 117967 Moscow, Russia

Received 1 May 1998; received in revised form 1 February 2000

Abstract
Ellipsoids that contain all optimal primal solutions, those that contain all optimal dual slack solutions, and primal–dual ellipsoids are derived. They are independent of the algorithm used and have a smaller size than the Choi–Goldfarb ellipsoids [J. Optim. Theory Appl. 80 (1994) 161–173]. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Linear programming; Interior point methods; Containing ellipsoids

1. Introduction
Unlike the simplex method for linear programming, which moves from one basic feasible solution to another basic feasible solution, interior algorithms create a sequence of interior points converging to an optimal feasible solution. To identify optimal basic and non-basic variables during the course of the algorithm several criteria have been proposed, associated with the interior point techniques used. These criteria were obtained by constructing ellipsoids which contain all optimal primal solutions and/or ellipsoids which contain all optimal dual slack solutions, and are all algorithm dependent (see e.g. [2,3,5,7–9]). Choi and Goldfarb [1] have further developed this approach, deriving solution containing ellipsoids which are independent of the interior point algorithm used, but larger than algorithm-dependent ellipsoids. Concerning solution-containing ellipsoids we also note that the first two-sided ellipsoidal approximation of polyhedra based on primal barrier functions is due to Sonnevend [4].

In this paper we propose the approach to diminish the size of the solution-containing ellipsoids derived in [1]. To get the tighter estimation of the ellipsoid size a certain localization of the optimal solution set is

E-mail address: igor@ccas.ru (I.S. Litvinchev).

1 The research supported in part by the grants from FAPESP, Brazil (No. 1997/7153-2) and RFBR, Russia (No. 99-01-01071).
It is shown that the proper choice of the localization results in a smaller containing ellipsoid. A similar approach is used to derive the primal–dual containing ellipsoids.

The paper is organized as follows. In Section 2, some basic notations and results, associated with Choi–Goldfarb ellipsoids are stated. In Section 3, the improved primal ellipsoids are constructed. In Section 4, the primal–dual ellipsoids are considered, and a new criterion for identifying the optimal basis is stated and compared with the previous results.

2. Definitions and notation

Consider the primal–dual pair of linear programs:

\[
\begin{align*}
\min & \{ c^t x \mid Ax = b, \; x \geq 0 \}, \\
\max & \{ b^t y \mid A^t y + z = c, \; z \geq 0 \},
\end{align*}
\]

where \( A \in \mathbb{R}^{m \times n} \) and \( b, c, x, y \) and \( z \) are suitably dimensioned. Let \( x^* \) and \((z^*, y^*)\) be optimal solutions to (1) and (2), respectively. It is assumed that the optimal solution sets for both (1) and (2) are bounded, or equivalently, strictly feasible primal and dual solutions exist, and that \( \text{rank}(A) = m \). The upper case letters \( X \) and \( Z \) are used to denote diagonal matrices whose elements are the elements of the vectors \( x \) and \( z \), and \( \| \cdot \| \) denotes the Euclidean norm.

It was established in [1] that if \( \bar{x} \) and \((\bar{z}, \bar{y})\) are primal and dual feasible solutions, then the ellipsoid

\[
E_p = \{ x \in \mathbb{R}^n \mid \| \tilde{Z}(x - \bar{x}) \| \leq \varepsilon \},
\]

where

\[
\varepsilon^2 = \| \tilde{Z} \bar{x} \|^2 + \Delta^2, \quad \Delta = \bar{x}^t \bar{x}
\]

contains all optimal primal solutions \( x^* \). Here the “ellipsoid” is understood in a general sense allowing also unbounded sets.

Using this ellipsoid the lower bound \( x_i^- (\varepsilon) \) for \( x_i^* \) was derived by the solution of the problem

\[
\min \{ x_i \mid \| \tilde{Z}(x - \bar{x}) \| \leq \varepsilon, \; Ax = b, \; c^t x \leq c^t \bar{x} \}.
\]

If \( \tilde{z} > 0 \) (which will be assumed valid throughout the paper) then the optimal solution \( x_i^- (\varepsilon) \) of this problem takes the following form. Let \( e_i \) denote the \( i \)th unit vector and

\[
R(\tilde{Z}) = \tilde{Z}^{-1} [I - \tilde{Z}^{-1} A^t (A \tilde{Z}^{-2} A^t)^{-1} A \tilde{Z}^{-1}] \tilde{Z}^{-1}.
\]

Then

\[
x_i^- (\varepsilon) = \bar{x}_i - \varepsilon \sqrt{c_i^t R(\tilde{Z}) e_i - (c_i^t R(\tilde{Z}) e_i)^2 / c_i^t R(\tilde{Z}) c} \quad \text{for} \; c_i^t R(\tilde{Z}) e_i < 0
\]

and

\[
x_i^- (\varepsilon) = \bar{x}_i - \varepsilon \sqrt{c_i^t R(\tilde{Z}) e_i} \quad \text{for} \; c_i^t R(\tilde{Z}) e_i \geq 0.
\]

Respectively, the following criterion for identifying optimal basic variables was obtained: if \( x_i^- (\varepsilon) > 0 \), then the \( i \)th primal variable is an optimal basic variable, i.e., \( x_i^* > 0 \) and \( z_i^* = 0 \).

We see that the smaller the size \( \varepsilon \) of the ellipsoid \( E_p \), the larger is the value of \( x_i^- (\varepsilon) \). Below we consider the approach to diminish the size of the solution-containing ellipsoid.
3. Improved primal solution-containing ellipsoid

Denote by $P^*$ the optimal solution set of (1) and let $W$ be a closed bounded subset of $R^n$ such that $W \supset P^*$. We will refer to such a $W$ as a primal localization. Let $S \in R^{n \times n}$ be a diagonal matrix having positive diagonal entries $s_i$.

For any $x \in P^*$ we have

$$
\|S\bar{Z}(x - \bar{x})\|^2 = \|S\bar{Z}\bar{x}\|^2 + \|S\bar{Z}x\|^2 - 2\bar{x}^i(S\bar{Z})^2x
\leq \|S\bar{Z}\bar{x}\|^2 + \max_{w \in W} \{\|S\bar{Z}w\|^2 - 2\bar{x}^i(S\bar{Z})^2w\}
\equiv \varepsilon^2(S, W),
$$

where the inequality follows from the condition $x \in P^* \subseteq W$.

Then the ellipsoid

$$E_p(S, W) = \{x \in R^n \mid \|S\bar{Z}(x - \bar{x})\| \leq \varepsilon(S, W)\}
$$

contains all optimal primal solutions.

Using the ellipsoid $E_p(S, W)$ the lower bound $x_\ast$ can be derived similarly as in Section 2. It is not hard to verify that we need only to substitute $R(\bar{Z})$ for

$$R(S\bar{Z}) = (S\bar{Z})^{-1}[I - (S\bar{Z})^{-1}A'(A(S\bar{Z})^{-2}A')^{-1}A(S\bar{Z})^{-1}](S\bar{Z})^{-1}.$$

and use $\varepsilon(S, W)$ instead of $\varepsilon$.

The reason for using the scaled ellipsoid is as follows. Suppose for example that the pair $(\bar{x}, \bar{z})$ was produced by a primal–dual method [6]. Let $S = (X\bar{Z})^{-1/2}$. For this case $R(S\bar{Z})$ takes the form

$$R(D^{-1}) = D[I - DA'(AD^2A')^{-1}AD]D,$$

where $D = X^{1/2}Z^{-1/2}$. Note that the matrix $AD^2A'$ is computed and decomposed into its Cholesky factor in most of the primal–dual methods.

To define a localization $W$ we can use some constraints of the primal problem or a linear combination of them. Below we consider the localization which results in a closed form expression for $\varepsilon(S, W)$.

Proposition 1. Let the primal localization be defined by $W_1 = \{w \in R^n \mid \bar{z}^iw \leq A, w \geq 0\}$. Then

$$\varepsilon^2(S, W_1) = \|S\bar{Z}\bar{x}\|^2 + A\max_i \{\varepsilon^2_i(A - 2\bar{z}_i\bar{x}_i)\}.$$

Proof. For any $x \in P^*$ we have

$$\bar{z}^i x = (c - A'\bar{y})x = c^i x - \bar{y}^i b \leq c^i \bar{x} - \bar{y}^i b = \bar{z}^i \bar{x} = A$$

and hence $P^* \subseteq W_1$. Moreover since $\bar{z} > 0$ then $W_1$ is bounded, i.e., $W_1$ is a localization. To estimate $\varepsilon(S, W_1)$ in (3) we should solve the problem

$$f^* = \max_w \{\|S\bar{Z}w\|^2 - 2\bar{x}^i(S\bar{Z})^2w\}.$$

The objective function of this problem is strictly convex and hence the optimal solution is attained at a vertex of the non-degenerate simplex $W_1$. The non-zero component of the $i$th vertex is $A/\bar{z}_i$ and thus

$$f^* = A\max_i \{\varepsilon^2_i(A - 2\bar{z}_i\bar{x}_i)\},$$

which by (3) gives the required expression for $\varepsilon(S, W_1)$.

$\square$
Note that if \( S = I \) and \( \bar{x} > 0 \), then
\[
\tilde{c}^2(I, W_1) = \|\tilde{Z}\bar{x}\|^2 + A^2 - 2A \min_i \tilde{z}_i < \varepsilon^2
\]
and hence \( E_p(I, W_1) \subset E_p \).

The other choice of the scaling matrix \( S \) associated with the primal–dual algorithms is \( S = (XZ)^{-1/2} \). For this case, we have
\[
\tilde{c}^2((XZ)^{-1/2}, W_1) = \frac{\Delta^2}{\min_i \tilde{z}_i} - \Delta.
\]

A similar approach can be used to derive a dual ellipsoid that contains all optimal dual slack solutions, i.e.
\[
E_d(S, V) = \{ z \in \mathbb{R}^n \mid \|S\tilde{X}(z - \bar{z})\| \leq \sigma(S, V) \},
\]
where
\[
\sigma^2(S, V) = \|S\tilde{X}\bar{z}\|^2 + \max_{\tau \in V} \{ \|S\tilde{X}\tau\|^2 - 2\bar{z}^t(S\tilde{X})^2\tau \}.
\]

A dual localization \( V \) is a closed bounded subset of \( \mathbb{R}^n \) such that \( z^* \in V \) for any optimal dual slack solution \( z^* \). It is not hard to verify that if \( \bar{x} > 0 \), then \( V_1 = \{ \tau \in \mathbb{R}^n \mid \tilde{z}^t\tau \leq \Delta, \tau \geq 0 \} \) is the dual localization and \( \sigma(S, V_1) = \varepsilon(S, W_1) \). The ellipsoid \( E_d(I, V_1) \) is smaller than the dual-containing ellipsoid \( E_d = \{ z \in \mathbb{R}^n \mid \|\tilde{X}(z - \bar{z})\| \leq \varepsilon \} \) considered in [1] and it can be used similarly to identify optimal non-basic primal variables.

4. Primal–dual solution-containing ellipsoid

Denote by \( F^* \subseteq \mathbb{R}^{2n} \) the set of all optimal primal–dual pairs \((x^*, z^*)\) and let \( \Omega \) be a closed bounded subset of \( \mathbb{R}^{2n} \) such that \( \Omega \supseteq F^* \). We will refer to such an \( \Omega \) as a primal–dual localization. In this section we assume \((\bar{x}, \bar{z}) > 0\).

For any \((x, z) \in F^*\) we have similar to (3):
\[
\|S\tilde{X}(x - \bar{x})\|^2 + \|S\tilde{X}(z - \bar{z})\|^2 \leq 2\|S\tilde{X}\bar{z}\|^2 + \max_{(w, \tau) \in \Omega} \{ \|S\tilde{Z}w\|^2 + \|S\tilde{X}\tau\|^2 \\
- 2\bar{z}^t(S\tilde{Z})^2w - 2\bar{z}^t(S\tilde{X})^2\tau \} \equiv \delta^2(S, \Omega)
\]
and hence the ellipsoid
\[
E_{pd}(S, \Omega) = \{ (x, z) \in \mathbb{R}^{2n} \mid \|S\tilde{X}(x - \bar{x})\|^2 + \|S\tilde{X}(z - \bar{z})\|^2 \leq \delta^2(S, \Omega) \}
\]
contains all optimal primal–dual pairs \((x^*, z^*)\).

**Proposition 2.** Let the primal–dual localization be defined by \( \Omega_1 = \{ (w, \tau) \in \mathbb{R}^{2n} \mid \tilde{z}^t w + \tilde{x}^t \tau = \Delta, (w, \tau) \geq 0 \} \). Then
\[
\delta^2(S, \Omega_1) = 2\|S\tilde{X}\bar{z}\|^2 + \Delta \max_i \{ s_i^*(\Delta - 2\bar{z}_i) \}.
\]

**Proof.** For any \((x, z) \in F^*\) we have \( \tilde{z}_i^1 + \tilde{x}_i^1 z = \Delta \) since \((x - \bar{x})^t(z - \bar{z}) = 0\) and \( x^t z = 0 \) by the complementarity condition. Hence \( F^* \subseteq \Omega_1 \) and \( \Omega_1 \) is a localization. Then the required expression for \( \delta(S, \Omega_1) \) is obtained similar to the proof of Proposition 1. \( \square \)

Now we use \( E_{pd} \) to identify, if possible, the optimal basis. Consider the problem
\[
q^t x \rightarrow \min,
\]
(4)
\[ Ax = b, \quad A^t y + z = c, \quad (5) \]
\[ \| S\hat{z}(x - \hat{x}) \|^2 + \| S\hat{x}(z - \hat{z}) \|^2 \leq \delta^2, \quad (6) \]
\[ \hat{x}^t x + \hat{x}^t z = \Delta, \quad (7) \]
\[ \hat{x}^t z \leq \Delta, \quad (8) \]
\[ \hat{x}^t x \leq \Delta. \quad (9) \]

Here \( q \) is such that the objective of this problem coincides with \( x_i \) up to a positive scale factor. In [9] the problem similar to (4)–(6) was discussed for \( q = D e_i \).

Let \( \tilde{e} = S^{-1} e, \quad \tilde{q} = (S\hat{z})^{-1} q, \quad \tilde{d}^2 = \| \tilde{e} \|^2 + \| \tilde{e} \|^2 - \| \tilde{e} u \|^2, \) where \( \tau_v, \tau_u \) are used to denote the projections of \( \tau \in \mathbb{R}^n \) onto the null space of the matrices \( A(S\hat{z})^{-1} \) and \( AS\hat{X} \), respectively. Denote also

\[ B = (\tilde{e}^t \tilde{q}_e) (\| \tilde{e} \|^2 - \| \tilde{e} u \|^2) \left[ \| \tilde{q}_e \|^4 \left( \| \tilde{q}_e \|^2 - (\tilde{e}^t \tilde{q}_e)^2 \right) \right]^{-1}. \]

**Proposition 3.** The optimal solution \( x^- \) of problem (4)–(9) is such that
(a) if \( \tilde{e}^t \tilde{q}_e \geq 0 \) and \( \Delta \| \tilde{q}_e \| < \delta(\tilde{e}^t \tilde{q}_e) \), then

\[ q^t x^- = q^t \tilde{x} - B \frac{\| \tilde{e} \|^2 - \| \tilde{e} u \|^2}{\| \tilde{e} \|^2 - \| \tilde{e} u \|^2}, \]

(b) if \( \tilde{e}^t \tilde{q}_e \leq 0 \) and \( \Delta^2 < B(\| \tilde{e} \|^2 - \| \tilde{e} u \|^2) - \tilde{d}^2 \), then

\[ q^t x^- = q^t \tilde{x} - \sqrt{B(\| \tilde{e} \|^2 - \| \tilde{e} u \|^2) - \tilde{d}^2}, \]

(c) else

\[ q^t x^- = q^t \tilde{x} - (\tilde{e}^t \tilde{q}_e) \Delta \tilde{d}^2 - \sqrt{(\tilde{d}^2 - \Delta^2 \tilde{d}^2)(\| \tilde{q}_e \|^2 - \Delta^2 (\tilde{e}^t \tilde{q}_e)^2)}. \]

**Proof.** The KKT optimality conditions for (4)–(9) give

\[ q + A^t u + 2\lambda(S\hat{z})^2(x - \hat{x}) - (\theta - \beta)\hat{z} = 0, \quad (10) \]
\[ v + 2\lambda(S\hat{x})^2(z - \hat{z}) - (\theta - \alpha)\hat{x} = 0, \quad (11) \]
\[ Av = 0, \quad \lambda > 0, \quad x \geq 0, \quad \beta \geq 0, \]

where \( \theta, \alpha, \) and \( \beta \) are the Lagrange multipliers for the restrictions (7)–(9).

Multiplying (10) by \( A(S\hat{z})^{-2} \) and utilizing the condition \( A(x - \hat{x}) = 0 \) we get

\[ u = [A(S\hat{z})^{-2} A^t]^{-1} A(S\hat{z})^{-1} ((\theta - \beta)\hat{e} - \tilde{q}). \]

Multiplying (10) by \( (S\hat{z})^{-2} \) and substituting the expression for \( u \) we obtain

\[ x - \hat{x} = \frac{1}{2\lambda} (S\hat{z})^{-1} ((\theta - \beta)\hat{e} - \tilde{q}). \]

Now multiplying (11) by \( A \) and using the conditions \( Av = 0, \) \( A^t (y - \hat{y}) + (z - \hat{z}) = 0 \) we get

\[ y - \hat{y} = -\frac{\theta - \alpha}{2\lambda} [A(S\hat{z})^{-2} A^t]^{-1} A \hat{x}. \]
and hence
\[ z - \tilde{z} = -A^1(y - \tilde{y}) = \frac{\theta - x}{2\lambda} (S\tilde{X})^{-1}(\tilde{e} - \tilde{e}_u). \]

By (7) we have \( \tilde{x}^1(z - \tilde{z}) + \tilde{z}^1(x - \tilde{x}) = -A. \) Substituting \( x - \tilde{x} \) and \( z - \tilde{z} \) in this expression we obtain the equation for \( \theta \) which gives
\[ \theta = d^2[\tilde{e}^1\tilde{q}_e + \beta ||\tilde{e}_v||^2 + \alpha(||\tilde{e}||^2 - ||\tilde{e}_u||^2) - 2\lambda A]. \]

If \( \alpha > 0 \), then by the complementarity condition for (8) we have \( \tilde{x}^1z = A. \) Hence by (7) it follows \( 0 = \tilde{z}^1x < A \) and \( \beta = 0 \) by the complementarity condition for (9). Similarly if \( \beta > 0 \), then \( \alpha = 0 \). We can write both these conditions in the form \( \alpha \beta = 0. \)

Remembering that \( A = \tilde{x}^1\tilde{z} \) the complementarity conditions for (8) and (9) take the form \( \alpha \tilde{x}^1(z - \tilde{z}) = 0 \) and \( \beta \tilde{z}^1(x - \tilde{x}) = 0. \) Substituting the expressions for \( x - \tilde{x} \) and \( z - \tilde{z} \) we obtain
\[ \alpha(\theta - x) = 0, \]
\[ \beta((\theta - \beta)||\tilde{e}_v||^2 - \tilde{e}^1\tilde{q}_e) = 0. \]

Utilizing the expression for \( \theta \) and the condition \( \alpha \beta = 0 \) we get
\[ \alpha d^2(\tilde{e}^1\tilde{q}_e - \alpha ||\tilde{e}_v||^2 - 2\lambda A) = 0, \]
\[ \beta \left[ -\tilde{e}^1\tilde{q}_e \frac{||\tilde{e}||^2 - ||\tilde{e}_u||^2}{||\tilde{e}_v||^2} - \beta d^2(||\tilde{e}||^2 - ||\tilde{e}_u||^2) - 2\lambda d^2 \right] = 0. \]

If \( \tilde{e}^1\tilde{q}_e \leq 0 \) then the term in the brackets in (12) is strictly negative and hence \( \alpha = 0. \) By (13) we have that if \( \tilde{e}^1\tilde{q}_e > 0 \) then \( \beta = 0. \)

Since the quadratic constraint (6) must be active, then
\[ \delta^2 = \left( \frac{1}{2\lambda} \right)^2 ||(\theta - \beta)\tilde{e}_v - \tilde{q}_e||^2 + \left( \frac{\theta - x}{2\lambda} \right)^2 ||\tilde{e} - \tilde{e}_u||^2. \]

Substituting the expression for \( \theta \) we obtain after the algebraic manipulations:
\[ \delta^2 = A^2d^2 + \left( \frac{1}{2\lambda} \right)^2 (||\tilde{q}_e||^2 - d^2(\tilde{e}^1\tilde{q}_e)^2) - \left( \frac{1}{2\lambda} \right)^2 (\tilde{e}^1\tilde{q}_e)(\theta - \beta)(||\tilde{e}||^2 - ||\tilde{e}_u||^2) + 2\lambda A[\beta ||\tilde{e}_v||^2 + \alpha(||\tilde{e}||^2 - ||\tilde{e}_u||^2)]. \]

Now we are in a position to derive the main results of the proposition.
Suppose that \( \tilde{e}^1\tilde{q}_e > 0 \) and \( A||\tilde{q}_e|| < \delta(\tilde{e}^1\tilde{q}_e) \). By (13) it follows that \( \beta = 0. \) Let \( \alpha > 0. \) Then by (12) we have
\[ \alpha = \frac{\tilde{e}^1\tilde{q}_e - 2\lambda A}{||\tilde{e}_v||^2}. \]

Substituting this \( \alpha \) in (14) with \( \beta = 0 \) we obtain
\[ \left( \frac{1}{2\lambda} \right)^2 \frac{\delta^2||\tilde{e}_v||^2 - A^2}{||\tilde{q}_e||^2||\tilde{e}_v||^2 - (\tilde{e}^1\tilde{q}_e)^2}. \]

Note that since \( A||\tilde{q}_e|| < \delta(\tilde{e}^1\tilde{q}_e) \), then the numerator of the above expression is strictly positive. Using the expression for \( \lambda \) it is not hard to verify that \( \alpha \) in (15) is strictly positive. Now utilizing the expressions for
\( x - \bar{x}, \alpha \) and \( \theta \) we have
\[
q'(x - \bar{x}) = \frac{1}{2\lambda}[\theta \tilde{e}_x \tilde{q}_x - \|\tilde{q}_x\|^2] = -\frac{\tilde{e}_x^T \tilde{q}_x}{\|\tilde{e}_x\|^2} \Delta - \frac{1}{2\lambda} \left( \|\tilde{q}_x\|^2 - \left(\tilde{e}_x^T \tilde{q}_x\right)^2 \right),
\]
which finally proves part (a) of the proposition.

Note that if \( \tilde{e}_x^T \tilde{q}_x \geq 0 \) and \( \|q_x\| \geq \delta \tilde{e}_x^T \tilde{q}_x \), then the case \( \alpha > 0 \) is impossible and we have \( \alpha = \beta = 0 \).

Suppose now that \( \tilde{e}_x^T \tilde{q}_x \leq 0 \) and \( \Delta^2 < B(\delta^2(\|\tilde{e}_x\|^2 - \|\tilde{e}_u\|^2) - \Delta^2) \). By (12) it follows that \( \alpha = 0 \). Let \( \beta > 0 \). Then by (13) we have
\[
\beta = -\frac{\tilde{e}_x^T \tilde{q}_x(\|\tilde{e}_x\|^2 - \|\tilde{e}_u\|^2) + 2\lambda \|\tilde{e}_x\| \|\tilde{e}_x\|}{\|\tilde{e}_x\|^2(\|\tilde{e}_x\|^2 - \|\tilde{e}_u\|^2)}. \tag{16}
\]
Substituting this \( \beta \) in (14) with \( \alpha = 0 \) we obtain
\[
\left( \frac{1}{2\lambda} \right)^2 = \frac{\delta^2(\|\tilde{e}_x\|^2 - \|\tilde{e}_u\|^2) - \Delta^2}{(\|\tilde{e}_x\|^2 - \|\tilde{e}_u\|^2)(\|\tilde{q}_x\|^2 - (\tilde{e}_x^T \tilde{q}_x)^2) \|\tilde{e}_x\|^2}.
\]
Note that under our assumption the numerator of this expression is strictly positive. Moreover, if we substitute this \( \lambda \) in (16), then \( \beta > 0 \).

Substituting \( x - \bar{x}, \beta \) and \( \theta \) we have
\[
q'(x - \bar{x}) = \frac{1}{2\lambda}[(\theta - \beta) \tilde{e}_x \tilde{q}_x - \|\tilde{q}_x\|^2] = -\sqrt{\delta^2 - \frac{\Delta^2}{\|\tilde{e}_x\|^2 - \|\tilde{e}_u\|^2}} \left( \|\tilde{q}_x\|^2 - \left(\tilde{e}_x^T \tilde{q}_x\right)^2 \right)
\],
which proves part (b) of the proposition.

Note that if \( \tilde{e}_x^T \tilde{q}_x \leq 0 \) and \( \Delta^2 \geq B(\delta^2(\|\tilde{e}_x\|^2 - \|\tilde{e}_u\|^2) - \Delta^2) \) then the case \( \beta > 0 \) is impossible and \( \alpha = \beta = 0 \).

If either (a) or (b) is not satisfied, then \( \alpha = \beta = 0 \). For this case we have \( \theta = d^2(\tilde{e}_x^T \tilde{q}_x - 2\lambda \Delta) \) and
\[
q'(x - \bar{x}) = \frac{1}{2\lambda}[(\theta - \beta) \tilde{e}_x \tilde{q}_x - \|\tilde{q}_x\|^2] = \frac{d^2(\tilde{e}_x^T \tilde{q}_x - 2\lambda \Delta)}{2\lambda} \tilde{e}_x \tilde{q}_x - \frac{1}{2\lambda} \|\tilde{q}_x\|^2
\]
\[
= -d^2 \Delta^2 \tilde{e}_x^T \tilde{q}_x - \frac{1}{2\lambda} \|\tilde{q}_x\|^2 - d^2(\tilde{e}_x^T \tilde{q}_x)^2.
\]
Obtaining \( \lambda \) from (14) with \( \alpha = \beta = 0 \) we get the expression for \( q't' \) required in part (c) of the proposition.

Since \( (x^*, z^*, y^*) \) is feasible to (4)–(9) for \( \delta = \delta(S, \Omega_I) \), we have the following criterion for identifying the optimal basis: if \( q't' > 0 \), then \( x^*_i > 0 \) and \( z^*_i = 0 \).

Note that for part (c) of Proposition 3 we have \( \alpha = \beta = 0 \) and hence constraints (8) and (9) can be relaxed without changing the optimal value. For parts (a) and (b), either (8) or (9) is active and hence \( q't' \leq \min\{q't', q't'_n\} \).

In [9] the primal–dual containing ellipsoid centered at the origin was considered for \( S = (XZ)^{-1/2} \) and \( (\bar{x}, \bar{z}, \bar{y}) \in N_{\gamma}(\gamma) \), where \( \gamma \in [0, 1) \),
\[
N_{\gamma}(\gamma) = \left\{ (x, y, z) \in F^0 \left\| XZe - \frac{A}{n} e \right\| \leq \gamma \frac{A}{n} \right\}
\]
and \( F^0 \) is the set of all strictly feasible primal–dual solutions.

For \( (\bar{x}, \bar{z}, \bar{y}) \in N_{\gamma}(\gamma) \) we have \( \bar{x}, \bar{z}, \bar{y} \geq (1 - \gamma) A/n, i = \bar{1}, n \) and by Proposition 2
\[
\delta^2((XZ)^{-1/2}, \Omega_I) = \frac{\Delta^2}{\min \bar{x}_i, \bar{z}_i} \leq \frac{\Delta n}{1 - \gamma} = \delta^2,
\]
such that the size of our ellipsoid is the same as in [9]. Similarly we can estimate the size of the ellipsoid for \( (\bar{x}, \bar{z}, \bar{y}) \in N_{-\infty}(\eta) = \left\{ (x, y, z) \in F^0 \left\| x, y, z, \eta A/n, i = \bar{1}, n \right\} \), \( \eta \in (0, 1) \).
The lower bound \( q^* x_{IT} \) obtained in [9, Theorem 3] by the solution of the problem similar to (4)–(6) for \( q = D^{-1} e_i \) is as follows:

\[
q^* x_{IT} = e_i^T \tilde{P} \tilde{e} - \sqrt{(\delta^2 - \Delta)(1 - \tilde{p}_{ii})},
\]

where \( \tilde{P} = D^T A^i (A D^T A^i)^{-1} A D \) and \( \tilde{p}_{ii} \) is the \( i \)th diagonal entry of \( \tilde{P} \).

Since \( \tau_u = \tau_e = (I - \tilde{P}) \tau \) and \( d^2 = \Delta^{-1} \) for \( S = (\tilde{XZ})^{-1/2} \), part (c) of Proposition 3 gives

\[
q^* x_c = e_i^T \tilde{P} \tilde{e} - \sqrt{(\delta^2 - \Delta) \left( 1 - \tilde{p}_{ii} - \frac{1}{\Delta} [e_i^T (I - \tilde{P}) \tilde{e}]^2 \right)} > q^* x_{IT}.
\]

Hence Proposition 3 results in the stronger criterion for identifying the optimal basis.

It is interesting to compare the lower bound obtained in Proposition 3 with the bound \( x_i^- (\epsilon) \) in Section 2. We have

\[
\delta^2 (I, \Omega_1) = 2 \| \tilde{Z} \tilde{x} \|^2 + A^2 - 2A \min_i \tilde{z} \tilde{x}_i = e^2 + \| \tilde{Z} \tilde{x} \|^2 - 2A \min_i \tilde{z} \tilde{x}_i.
\]

If \((\tilde{x}, \tilde{z}, \tilde{y}) \in N_{\delta}(\gamma)\), then \((1 - \gamma) A/n \leq \tilde{z} \tilde{x}_i \leq (1 + \gamma) A/n\) and

\[
\| \tilde{Z} \tilde{x} \|^2 - 2A \min_i \tilde{z} \tilde{x}_i \leq \delta^2 (\gamma^2 + 4 \gamma - 1)/n.
\]

For \( \gamma \in [0, \sqrt{3} - 2) \) the right-hand side of the latter inequality is strictly negative and \( \delta(I, \Omega_1) < \epsilon \).

Since \( R(\tilde{Z}) : \tilde{Z}^{-1} (\tilde{Z}^{-1} \tau) \), for any \( \tau \in R^e \) and \((\tilde{Z}^{-1} \epsilon) = e \), by \( A^i \tilde{y} + \tilde{z} = c \), it is not hard to verify that \( x_i^- (\epsilon) \) coincides with the optimal solution of the following problem:

\[
x_i^- (\epsilon) = \min \{ x_i | \| \tilde{Z} (x - \tilde{x}) \| \leq \epsilon, \ A x = b, \ \ c^T x \leq c^T \tilde{x} \}.
\]

Comparing (17) with the problem (4)–(9) for \( q = e_i \), \( S = I \), \((\tilde{x}, \tilde{Z}, \tilde{y}) \in N_{\delta}(\gamma), \gamma \in [0, \sqrt{3} - 2), \delta = \delta(I, \Omega_1) < \epsilon \), we obtain \( x_i^- (\epsilon) < e_i^T x^- \) that results in improved criterion for identifying the optimal basis.

The ellipsoids considered in this paper are algorithm-independent, that is they do not depend on how the primal–dual pair \((\tilde{x}, \tilde{Z})\) was obtained. Meanwhile the size of the ellipsoid depends on the proximity of \((\tilde{x}, \tilde{Z})\) to the central path: the closer to the central path, the smaller the size. This is one more reason for maintaining \((\tilde{x}, \tilde{Z})\) in the neighborhood of the central path during the course of the interior algorithm.

Acknowledgements

The author gratefully acknowledges two anonymous referees for their helpful comments and remarks.

Appendix

Since our expression for \( x_i^- (\epsilon) \) is slightly different from the one used in [1, Criterion (C4), p. 170], we give here the complete derivation for \( x_i^- (\epsilon) \).

**Proposition A1.** The optimal solution \( x_i^- (\epsilon) \) of the problem

\[
\min \{ e_i^T x | (x - \tilde{x})^T \tilde{Z}^2 (x - \tilde{x}) \leq \epsilon^2, \ A x = b, \ \ c^T x \leq c^T \tilde{x} \}
\]

is as follows:

\[
x_i^- (\epsilon) = \tilde{x}_i - \epsilon \sqrt{e_i^T R(\tilde{Z}) e_i - (c^T R(\tilde{Z}) e_i)^2/c^T P_{AZ}^{-1} c} \quad \text{for } c^T R(\tilde{Z}) e_i < 0
\]
and

\[ x_i^- (\varepsilon) = \bar{x}_i - \varepsilon \sqrt{e_i^T R(\bar{Z}) e_i} \quad \text{for} \quad c_i^T R(\bar{Z}) e_i \geq 0. \]

**Proof.** The Lagrangian, associated with (A.1) is

\[ L = e_i^T x + \lambda [(x - \bar{x})^T \bar{Z}^2 (x - \bar{x}) - e_i^2] + u^T (Ax - b) + \theta (c_i^T x - c_i^T \bar{x}). \]

The KKT condition \( \partial L / \partial x = 0 \) gives

\[ e_i + 2 i \bar{Z}^2 (x - \bar{x}) + \theta c + A^T u = 0. \]  \( \text{(A.2)} \)

Multiplying (A.2) by \((x - \bar{x})\) and utilizing conditions \( \lambda [(x - \bar{x})^T \bar{Z}^2 (x - \bar{x}) - e_i^2] = 0, \theta (c_i^T x - c_i^T \bar{x}) = 0, A(x - \bar{x}) = 0 \)

we obtain

\[ \lambda = (\bar{x}_i - x_i) / 2 \varepsilon. \]

Multiplying (A.2) by \( A \bar{Z}^{-2} \) we get

\[ u = - (A \bar{Z}^{-2} A^T)^{-1} A \bar{Z}^{-2} (e_i + \theta c). \]

From (A.2) we have

\[ 2 \lambda (\bar{x} - x) = \bar{Z}^{-2} (A^T u + e_i + \theta c) = R(\bar{Z}) (e_i + \theta c). \]  \( \text{(A.3)} \)

Multiplying (A.3) by \( c_i^T \) we obtain

\[ 0 \leq 2 \lambda c_i^T (\bar{x} - x) = c_i^T R(\bar{Z}) e_i + \theta c_i^T R(\bar{Z}) c. \]  \( \text{(A.4)} \)

If \( c_i^T R(\bar{Z}) e_i < 0 \), then it follows from (A.4) that \( \theta > 0 \) and hence by the complementarity condition \( c_i^T (\bar{x} - x) = 0 \). From (A.4) we have \( \theta = - c_i^T R(\bar{Z}) e_i / c_i^T R(\bar{Z}) c \).

Substituting \( \lambda \) and \( \theta \) in (A.3) and multiplying by \( e_i \) we obtain the first part of Proposition A1.

Let \( c_i^T R(\bar{Z}) e_i \geq 0 \). Suppose that we have relaxed the restriction \( c_i^T x \leq c_i^T \bar{x} \) in (A.1) or, just the same, put \( \theta = 0 \) in the Lagrangian. Then similar to (A.3) we obtain for the solution of the relaxed problem

\[ 2 \lambda (\bar{x} - x) = R(\bar{Z}) e_i \]

and since \( \lambda > 0 \) and \( c_i^T R(\bar{Z}) e_i \geq 0 \) we have \( c_i^T x \leq c_i^T \bar{x} \). Since the relaxed solution is feasible to (A.1), then it is optimal to (A.1). This immediately gives the required expression for \( x_i^- (\varepsilon) \).

Note that in [1, Criterion (C4), p. 170] there is a little flaw. The first expression for \( x_i^- (\varepsilon) \) in Proposition A1 was associated there with the case \( c_i^T R(\bar{Z}) e_i > 0 \).

**References**