RUTCOR
RESEARCH
REPORT

Primal-Dual Interior Point Algorithms for Convex Quadratically Constrained and Semidefinite Optimization Problems

I. Adler[a] F. Alizadeh[b]

RRR 46-95, December 1995

[a]University of California–Berkeley
[b]RUTCOR, Rutgers University, P.O. Box 5062, New Brunswick, NJ 08903. email: alizadeh@rutcor.rutgers.edu. Research supported in part by NSF grant CCR-9501941.

RUTCOR • Rutgers Center for Operations Research • Rutgers University • P.O. Box 5062 • New Brunswick New Jersey • 08903-5062
Telephone: 908-445-3804
Telefax: 908-445-5472
Email: rrr@rutcor.rutgers.edu
http://rutcor.rutgers.edu/~rrr

Rutcor Research Report
RRR 46-95, December 1995

Primal-Dual Interior Point Algorithms for Convex Quadratically Constrained and Semidefinite Optimization Problems

I. Adler F. Alizadeh

Abstract. In this paper we examine primal–dual interior point methods for optimization problems over the semidefinite cone and the p-cones (in particular the "ice cream cone"). We study similarities and differences of such interior point methods with the analogous algorithms for linear programming.

1 Introduction

It has been observed that many of the techniques used to derive and analyze interior point methods for linear programming may be extended, in a sense "word by word", to more general domains. For instance in [Ali91] Ye's methods based on Todd–Ye potential reduction function were extended to semidefinite programming (SDP). In [MN93] Karmarkar's original algorithm was extended to optimization problems over the "ice cream cone" (see below for definition); optimization over such cones is equivalent to convex quadratically constrained quadratic programming. On the other hand Nesterov and Nemirovskii have laid out a general theory of interior point methods based on the concept of *p–self–concordant* barrier functions [NN94]. In this important work the authors have shown that one can maximize a linear function over any convex set endowed with such barriers in time proportional to $O(\sqrt{p})$ iterations.

In this work we are concerned with extension of primal–dual methods of the type originally proposed by Kojima et al [KMY89]. Such methods are proposed for linear programming and define the primal–dual central path as

$$\{(\mathbf{x}, \mathbf{y}, \mathbf{z}) \,|\, A\mathbf{x} = b,\ A^T\mathbf{y} + \mathbf{z} = \mathbf{c},\ \text{and } x_i z_i = \mu\}$$

Monteiro and Adler in [MA89] presented a simple analysis of such methods and showed that the number of iterations is proportional to $\sqrt{n}$. It turns out that "word–by–word" extension of this analysis to semidefinite programming or to quadratically constrained problems (QCQP) is more challenging. For instance Monteiro and Adler implicitly use the fact that multiplication of diagonal matrices is commutative. However, analogous analysis for SDP or QCQP does not involve diagonal matrices. We will elaborate on some of these difficulties in the following section. Our goal is to derive an $O(\sqrt{n})$ iteration primal-dual algorithm for SDP and QCQP which is a direct extension of Kojima et al and Monteiro Adler algorithms.

Notation:

When in the context of QCQP we use boldface lower case letters for column vectors $\mathbf{x}$etc. If we need to refer to the j^{th} entry of a vector we use parenthesis: $(\mathbf{x_i})_j$ means the j^{th} entry of the i^{th} vector. Also, for reasons to become clear shortly, in QCQP unknown primal $\mathbf{x_i}$ and dual $\mathbf{z_i}$ vectors along with the objective vectors $\mathbf{c_i}$ are indexed from zero, with zeroth entry playing a special role. For these vectors we have occasion to use the subvector indexed from 1; for this purpose we use $\overline{\mathbf{x}}$ etc. Thus $\mathbf{x} = ((\mathbf{x})_0, \overline{\mathbf{x}})^T$. In addition we define

$$\gamma(\mathbf{x}) \stackrel{\text{def}}{=} (\mathbf{x})_0^2 - \overline{\mathbf{x}}^T\overline{\mathbf{x}}$$

Also define the reflection matrix R and vector $\mathbf{u}$:

$$R \stackrel{\text{def}}{=} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 \end{pmatrix} \text{ and } \mathbf{u} \stackrel{\text{def}}{=} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

both of which are indexed from zero. Thus,

$$\gamma(\mathbf{x}) = \mathbf{x}^T R \mathbf{x}$$

In semidefinite programming we use capital letter for symmetric matrices: X, Z, C, etc. But sometimes we need to treat these matrices as vectors in n^2 dimensional Euclidean space. We use the operator **vec** for this purpose, for instance $\mathbf{vec}\, X, \mathbf{vec}\,(ABC)$ etc. The inverse of the **vec** operator is the Mat : Thus for an n^2 vector $\mathbf{v}\mathrm{Mat}\,\mathbf{v}$ is an $n \times n$ matrix V with $\mathbf{vec}\, V = \mathbf{v}$. Also, $A \succeq B$ (respectively $A \succ B$) means that $A - B$ is positive semidefinite (respectively positive definite.) Finally we use the following well-known facts:

$$\mathbf{vec}\,(ABC) = (C^T \otimes A)\mathbf{vec}\, B \text{ and } (A \otimes B)(C \otimes D) = (AC) \otimes (BD)$$

2 Complimentary slackness

In general every convex programming problem can be laid out in the following "cone optimization" format:

$$\begin{array}{ll} \min & \mathbf{c}^T\mathbf{x} \\ \text{s.t.} & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq_K \mathbf{0} \end{array} \tag{1}$$

Here K is a closed, convex and pointed (i.e $K \cap (-K) = \{\mathbf{0}\}$) cone and $\mathbf{x} \geq_K \mathbf{y}$ (respectively $\mathbf{x} >_K \mathbf{y}$ means that $\mathbf{x} - \mathbf{y} \in K$ (respectively $\mathbf{x} - \mathbf{y} \in \mathrm{Int}\, K$). Let K^* be the polar cone of K, that is:

$$K^* = \{\mathbf{z} : \ \mathbf{x}^T\mathbf{z} \geq 0 \text{ for all } \mathbf{x} \in K\}$$

Then the following problem is dual to (1):

$$\begin{array}{ll} \max & \mathbf{b}^T\mathbf{y} \\ \text{s.t.} & A^T\mathbf{y} + \mathbf{z} = \mathbf{c} \\ & \mathbf{z} \geq_{K^*} \mathbf{0} \end{array} \tag{2}$$

For our purposes we assume the following:

1. $\exists \mathbf{x} >_K \mathbf{0}$ where $A\mathbf{x} = \mathbf{b}$.

2. $\exists \mathbf{y}, \mathbf{z} >_{K^*}$ such that $A^T\mathbf{y} + \mathbf{z} = \mathbf{c}$.

3. A is full rank.

These conditions are sufficient to guarantee that the optimal values of (1) and (2) are finite and equal to each other.

Theorem 2.1 Strong duality: *Let z_1 be the minimum value of (1) and z_2 the maximum value in (2). Then assuming conditions 1–3 above implies $z_1 = z_2$.*

Observe that the so–called weak duality theorem, $z_1 \geq z_2$ is quite easy:

$$z_1 - z_2 = \mathbf{c}^T\mathbf{x} - \mathbf{b}^T\mathbf{y} = \mathbf{c}^T\mathbf{x} - (A\mathbf{x})^T\mathbf{y} = \mathbf{x}^T(\mathbf{c} - A^T\mathbf{y}) = \mathbf{x}^T\mathbf{z} \geq 0$$

The last inequality is the direct consequence of the definition of polar. As a result we have shown that the quantity $\mathbf{x}^T\mathbf{z}$ always indicates the duality gap for a primal and dual feasible pair $(\mathbf{x}, \mathbf{z})$.

Let $\mathbf{x}$, $\mathbf{z}$ be two vectors. For each pair of polar cones (K, K^*) the three statements:

i. $\mathbf{x} \in K$

ii. $\mathbf{z} \in K^*$,

iii. $\mathbf{x}^T\mathbf{z} = 0$

impose n degrees of restriction on the pair $\mathbf{x}, \mathbf{z}$ in general; these n restrictions–sometimes referred to as *complementarity conditions*– along with linear equations $A\mathbf{x} = \mathbf{b}$, and $A^T\mathbf{y} + \mathbf{z} = \mathbf{c}$ should completely determine the primal and dual solutions $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ under appropriate nondegeneracy assumptions. However, there is no general formula for complementarity conditions for all cones[1]. However, for specific cones one can determine these conditions, often by elementary methods. In many cases one can find a function $\mathbf{F} : \Re^n \times \Re^n \rightarrow \Re^n$ expressing the complementary conditions; Also, $\mathbf{F}$ may be differentiable. Notice that in order to extend Kojima et al and Monteiro–Adler interior points methods we must have access to the complementarity conditions explicitly. We now examine some concrete cones.

2.1 The positive orthant

Let $K \stackrel{\text{def}}{=} \{\mathbf{x} : \; x_i \geq 0\}$. This is the positive orthant and optimization over K is linear programming. In this case the complementarity conditions are precisely the complementary slackness theorem which states that primal feasible $\mathbf{x}$, and dual feasible $\mathbf{z}$, are optimal iff $x_i z_i = 0$.

2.2 Ice Cream Cones

Let Q be the "ice cream" cone, that is

$$Q = \left\{ \mathbf{x} \in \Re^{n+1} | (\mathbf{x})_0 \geq \sqrt{\left(\sum_{i=1}^{n} (\mathbf{x})_i^2 \right)} \right\} \tag{3}$$

[1]Actually if we have any smooth barrier function for the cone K we can derive the n explicit equations defining the complementarity relations; see for example the section on the logarithmic barrier functions in the sequel. On the other hand Nesterov and Nemirovskii show that any closed, pointed and convex cone is endowed with an n-self-concordant barrier function. They construct a universal barrier $b(\mathbf{x})$ by considering logarithm of volume of the set-polar of K centered at $\mathbf{x}$. Thus in a sense we have a general method of deriving the complementarity relations in the most general case. Nevertheless such a procedure is not computationally tractable in general.

We say $\mathbf{x} \geq_{\mathcal{Q}} \mathbf{y}$ iff $\mathbf{x} - \mathbf{y} \in \mathcal{Q}$ and $\mathbf{x} >_{\mathcal{Q}} \mathbf{y}$ iff $\mathbf{x} - \mathbf{y} \in \operatorname{Int} \mathcal{Q}$. ($\mathbf{x} \leq_{\mathcal{Q}} \mathbf{y}$ and $\mathbf{x} <_{\mathcal{Q}} \mathbf{y}$ are defined similarly.)

Consider the following pair of primal and dual problems:

$$\begin{array}{llr@{\qquad}lll}
\min & \mathbf{c_1}^T\mathbf{x_1} + \cdots + \mathbf{c_n}^T\mathbf{x_n} & & \max & \mathbf{b}^T\mathbf{y} & \\
\text{s.t.} & A_1\mathbf{x_1} + \cdots + A_n\mathbf{x_n} = \mathbf{b} & & \text{s.t.} & A_i^T\mathbf{y} + \mathbf{z_i} = \mathbf{c_i} & \text{for } i = 1\cdots n \\
 & \mathbf{x_i} \geq_{\mathcal{Q}} \mathbf{0} & \text{for } i = 1\cdots n & & \mathbf{z_i} \geq_{\mathcal{Q}} \mathbf{0} & \text{for } i = 1\cdots n
\end{array} \tag{4}$$

Note that each $\mathbf{x_i}, \mathbf{z_i} \in \Re^{n_i}$ may have different size for each i. Assuming there is some vector $\mathbf{y}$ such that $\mathbf{c_i} - A_i\mathbf{y} >_{\mathcal{Q}} \mathbf{0}$ for each $i = 1\cdots n$ (Slater condition) strong duality holds and the values of the two objective optimization problems are equal. (We assume that both are bounded and have finite solution.)

Since our intent is to derive a strict primal-dual interior point algorithm *à la* Monteiro–Adler, first we need a complementary slackness theorem.

Observe that at the optimum, we have:

$$\sum_i \mathbf{x_i}^T\mathbf{z_i} = 0 \tag{5}$$

But for each $\mathbf{x_i} \geq_{\mathcal{Q}} \mathbf{0}$ and $\mathbf{z_i} \geq_{\mathcal{Q}} \mathbf{0}$ we have (by self-polarity of the ice cream cone) $\mathbf{x_i}^T\mathbf{z_i} \geq 0$. Thus, (5) implies that

$$\mathbf{x_i}^T\mathbf{z_i} = 0 \text{ for } i = 1\cdots n \tag{6}$$

Now an explicit complimentary slackness theorem can be derived easily from the following lemma:

Lemma 2.1 *Let $\mathbf{x} \geq_{\mathcal{Q}} \mathbf{0}$ and $\mathbf{z} \geq_{\mathcal{Q}} \mathbf{0}$. Then $\mathbf{x}^T\mathbf{z} = 0$ iff*

i.

$$\frac{(\mathbf{x})_1}{(\mathbf{z})_1} = \cdots = \frac{(\mathbf{x})_n}{(\mathbf{z})_n} = -\frac{(\mathbf{x})_0}{(\mathbf{z})_0} \textit{ or equivalently } (\mathbf{x})_i(\mathbf{z})_0 + (\mathbf{z})_i(\mathbf{x})_0 = 0 \textit{ for } i = 1, \cdots, n$$

and

ii.

$$\gamma(\mathbf{x}) = (\mathbf{x})_0^2 - (\mathbf{x})_1^2 + \cdots + (\mathbf{x})_n^2 = 0 \textit{ or equivalently } \gamma(\mathbf{z}) = (\mathbf{z})_0^2 - (\mathbf{z})_1^2 + \cdots + (\mathbf{z})_n^2 = 0$$

Proof: This lemma is essentially a rewording of the Cauchy-Schwartz-Boniakovsky inequality (see the more general lemma 2.3.) Here is a simple proof: We have

$$(\mathbf{x})_0^2 \geq \sum_{i=1}^n (\mathbf{x})_i^2 \tag{7}$$

and

$$(\mathbf{z})_0^2 \geq \sum_{i=1}^n (\mathbf{z})_i^2$$

Multiplying both sides of the latter inequality by $((\mathbf{x})_0/(\mathbf{z})_0)^2$ we have:

$$(\mathbf{x})_0^2 \geq \sum_{i=1}^{n} \left((\mathbf{x})_i \frac{(\mathbf{x})_0}{(\mathbf{z})_0} \right)^2 \tag{8}$$

Also $\mathbf{x}^T\mathbf{z} = 0$ is equivalent to

$$-2\mathbf{x}_0^2 = 2\sum_{i=1}^{n} \left((\mathbf{x})_i (\mathbf{z})_i \frac{(\mathbf{x})_0}{(\mathbf{z})_0} \right) \tag{9}$$

Adding (7), (8) and (9) we get:

$$0 \geq \sum_{i=1}^{n} \left((\mathbf{x})_i + (\mathbf{z})_i \frac{(\mathbf{x})_0}{(\mathbf{z})_0} \right)^2$$

which immediately implies part i. To prove part ii observe that the ratio can be written as:

$$\frac{(\mathbf{x})_0}{(\mathbf{z})_0} = -\frac{(\mathbf{x})_1}{(\mathbf{z})_1} = \cdots = -\frac{(\mathbf{x})_n}{(\mathbf{z})_n} = \frac{\gamma(\mathbf{x})}{(\mathbf{x})_0(\mathbf{z})_0 + \cdots + (\mathbf{x})_n(\mathbf{z})_n} = \frac{(\mathbf{x})_0(\mathbf{z})_0 + \cdots + (\mathbf{x})_n(\mathbf{z})_n}{\gamma(\mathbf{z})}$$

Therefore

$$\sum_{j=0}^{n} (\mathbf{x})_j (\mathbf{z})_j = 0 \text{ iff } \gamma(\mathbf{x}) = \gamma(\mathbf{z}) = 0.$$

which proves the lemma. The converse is obvious. ■

Corollary 2.1 *Let vectors $\mathbf{x_i}$, $\mathbf{z_i}$ (for $i = 1, \cdots, n$) and $\mathbf{y}$ be feasible for the primal-dual pair (4). Then (assuming Slater condition) they are optimal iff for $i = 1, \cdots, n$*

$$-\frac{(\mathbf{x})_0}{(\mathbf{z})_0} = \frac{(\mathbf{x_i})_1}{(\mathbf{z_i})_1} = \cdots = \frac{(\mathbf{x_i})_{n_i}}{(\mathbf{z_i})_{n_i}} \tag{10}$$

$$(\mathbf{x_i})_0^2 = (\mathbf{x_i})_1^2 + \cdots + (\mathbf{x_i})_{n_i}^2 \tag{11}$$

Proof: This is immediate from previous lemma. It is useful however to derive it by applying KKT conditions on the primal problem in (4). For simplicity we assume there is only one inequality constraint '$\geq_{\mathcal{Q}}$' (i.e $n = 1$); the proof below immediately extends to the case where there are k inequalities. Define the Lagrangian functions:

$$\mathcal{L}(\mathbf{x}, \mathbf{y}, \alpha) \stackrel{\text{def}}{=} \mathbf{c}^T\mathbf{x} + (\mathbf{b} - A\mathbf{x})^T\mathbf{y} + \alpha\gamma(\mathbf{x}) \tag{12}$$

Now taking gradient of $\mathcal{L}$ with respect to $\mathbf{x}$, $\mathbf{y}$and α, we get:

$$\begin{array}{l} \nabla_{\mathbf{x}}\mathcal{L} = \mathbf{c}^T - \mathbf{y}^T A - 2\alpha\left((\mathbf{x})_0, -(\mathbf{x})_1, \cdots, (\mathbf{x})_n\right) = \mathbf{0} \\ \nabla_{\mathbf{y}}\mathcal{L} = \mathbf{b} - A\mathbf{x} = \mathbf{0} \\ \frac{d}{d\alpha}\mathcal{L} = \gamma(\mathbf{x}) = 0 \end{array}$$

Renaming $\mathbf{z} \stackrel{\text{def}}{=} \mathbf{c}^T - \mathbf{y}^T A$ we get the result. (Thus, the equations arising from $\nabla_{\mathbf{x}}\mathcal{L} = \mathbf{0}$ give rise to same equations as given in Lemma 1.1.) ■

A related problem

Conn and Overton in [CO94] have worked on the following variant (**note:** vectors are indexed from 1):

$$
\begin{array}{ll}
\min & \|\mathbf{z_1}\| + \cdots + \|\mathbf{z_n}\| \\
\text{s.t.} & \sum_i A_i^T \mathbf{y} + \mathbf{z_i} = \mathbf{c_i} \quad \text{for } i = 1, \cdots n
\end{array}
\qquad
\begin{array}{lll}
\max & \mathbf{c_1}^T \mathbf{x_1} + \cdots + \mathbf{c_n}^T \mathbf{x_n} & \\
\text{s.t.} & A_i \mathbf{x_i} = 0 & \text{for } i = 1, \cdots, n \\
 & \|\mathbf{x_i}\| \leq 1 & \text{for } i = 1, \cdots, n
\end{array}
\tag{13}
$$

They derive the complementary slackness theorem

$$
\|\mathbf{z_i}\| (\mathbf{x_i})_j = (\mathbf{z_i})_j \tag{14}
$$

Notice that his problem can be easily transformed to the standard problem (4). Simply introduce new variables $z_1, \cdots, z_k n$, and replace objective of the min problem by $z_1 + \cdots + z_k$ and add the constraint that vectors $(z_i, \mathbf{z_i}^T)^T \geq_{\mathcal{Q}} \mathbf{0}$. Then Conn and Overton's dual problem as well as their complementary slackness theorem are derivable from ours.

2.3 Semidefinite Programming

Now Let us consider the space of $n \times n$ symmetric matrices ($\Re^{\frac{n \times n}{2}}$) and consider the cone of positive semidefinite matrices. In this case we use $\mathcal{P}$ instead of K as our cone:

$$
\mathcal{P} \stackrel{\text{def}}{=} \{X \in \Re^{\frac{n \times n}{2}} : \ \mathbf{a}^T X \mathbf{a} \geq 0 \text{ for all } \mathbf{a} \in \Re^n\}
$$

This cone is also self–polar and the following pair of primal and dual optimization problems are referred to as the standard *semidefinite programming problem*

$$
\begin{array}{lll}
\underline{\textbf{Primal}} & & \\
\min & C \bullet X & \\
\text{s.t.} & A_i \bullet X = b_i & \text{for } i = 1, \cdots, m \\
 & X \succeq 0 &
\end{array}
\qquad
\begin{array}{ll}
\underline{\textbf{Dual}} & \\
\max & \mathbf{b}^T \mathbf{y} \\
\text{s.t.} & C - \sum_{i=1}^m y_i A_i \succeq 0.
\end{array}
\tag{15}
$$

Here $C \bullet X$ is the inner product of two matrices:

$$
C \bullet X \stackrel{\text{def}}{=} \sum C_{ij} X_{ij} = \text{trace } (CX)
$$

The relation $U \succeq V$ means that $U - V \in \mathcal{P}$; "$\succeq$" is the Löwner partial order.

The complementarity condition follows from the following lemma:

Lemma 2.2 *Let $X \succeq 0$ and $Z \succeq 0$. then $X \bullet Z = 0$ iff $XZ = 0$.*

Therefore, in semidefinite programming, our system of equations has the form:

$$
\begin{array}{ll}
A_i \bullet X = b_i & \text{for } i = 1, \cdots m \\
\sum y_i A_i + Z = C & \\
XZ = 0 &
\end{array}
\tag{16}
$$

2.4 p–cones

Instead of the ice cream cone we could use the p-cones. Let $p \geq 2$ be any real number and define $1 \leq q \leq 2$ as follows:

$$\frac{1}{p} + \frac{1}{q} = 1$$

Define the "p-cone" to be

$$\mathcal{P} \stackrel{\text{def}}{=} \left\{ \mathbf{x} \middle| \, (\mathbf{x})_0 \geq \left(|(\mathbf{x})_1|^p + \cdots + |(\mathbf{x})_n|^p\right)^{1/p} \right\} \tag{17}$$

We say $\mathbf{x} \geq_p \mathbf{y}$ iff $\mathbf{x} - \mathbf{y} \in \mathcal{P}$ and $\mathbf{x} >_p \mathbf{y}$ iff $\mathbf{x} - \mathbf{y} \in \operatorname{Int} \mathcal{P}$. ($\mathbf{x} \leq_p \mathbf{y}$ and $\mathbf{x} <_p \mathbf{y}$ are defined similarly.)

It turns out that the polar of p-cones are q-cones, where, $1/p + 1/q = 1$. Thus p-cones are self polar iff $p = q = 2$. Now the analog of primal dual pair (4) is the following pair:

$$\begin{array}{lll} \min & \mathbf{c}^T\mathbf{x} & \\ \text{s.t.} & A\mathbf{x} = \mathbf{b} & \\ & \mathbf{x} \geq_p \mathbf{0} & \end{array} \qquad \begin{array}{ll} \max & \mathbf{b}^T\mathbf{y} \\ \text{s.t.} & A^T\mathbf{y} + \mathbf{z} = \mathbf{c} \\ & \mathbf{z} \geq_q \mathbf{0} \end{array} \tag{18}$$

(Notice that unlike the QCQP problem, the case $n = 1$, is not trivial.) It turns out that the complementary slackness analog for p-cones is given by the following lemma:

Lemma 2.3 *Let $\mathbf{x} \geq_p \mathbf{0}$ and $\mathbf{z} \geq_q \mathbf{0}$. Then, $\mathbf{x}^T\mathbf{z} = 0$ iff*

1.

$$\frac{|(\mathbf{x})_1|^p}{|(\mathbf{z})_1|^q} = \cdots = \frac{|(\mathbf{x})_n|^p}{|(\mathbf{z})_n|^q} = \frac{(\mathbf{x})_0^p}{(\mathbf{z})_0^q}$$

and equivalently

$$(\mathbf{z})_0|(\mathbf{x})_i|^{p-1} - |(\mathbf{z})_i|(\mathbf{x})_0^{p-1} = 0 \text{ for } i = 1, \cdots, n$$

and

2.

$$(\mathbf{x})_0 = \left(|(\mathbf{x})_1|^p + \cdots + (|(\mathbf{x})_n|^p\right)^{1/p} \text{ equivalently } (\mathbf{z})_0 = \left(|(\mathbf{z})_1|^q + \cdots + (|(\mathbf{z})_n|^q\right)^{1/q}$$

Proof: This lemma is essentially Hölder's inequality in disguise. We have,

$$\begin{aligned} -(\mathbf{x})_1(\mathbf{z})_1 - \cdots - (\mathbf{x})_n(\mathbf{z})_n &= (\mathbf{x})_0(\mathbf{z})_0 \\ &\geq \left(|(\mathbf{x})_1|^p + \cdots + |(\mathbf{x})_n|^p\right)^{1/p} \left(|(\mathbf{z})_1|^q + \cdots + |(\mathbf{z})_n|^q\right)^{1/q} \\ &\quad \text{(by Hölder's inequality)} \\ &\geq |(\mathbf{x})_1(\mathbf{z})_1| + \cdots + |(\mathbf{x})_n(\mathbf{z})_n| \end{aligned}$$

But since $-(\mathbf{x})_j(\mathbf{z})_j \leq |(\mathbf{x})_j(\mathbf{z})_j|$ we must have equality throughout, and in particular we must have $(\mathbf{x})_j(\mathbf{z})_j \leq 0$. Also, Hölder's inequality is satisfied with equality iff

$$\frac{|(\mathbf{x})_1|^p}{|(\mathbf{z})_1|^q} = \cdots = \frac{|(\mathbf{x})_n|^p}{|(\mathbf{z})_n|^q} = \frac{(\mathbf{x})_0^p}{(\mathbf{z})_0^q}$$

The second part of the lemma is proved as in lemma 1.1. ■

Corollary 2.2 *Let vectors* $\mathbf{x}$*,* $\mathbf{z}$ *and* $\mathbf{y}$ *be feasible for the primal and dual problems in (18).Then they are optimal (assuming Slater condition) iff*

$$\frac{(\mathbf{x})_0^p}{(\mathbf{z})_0^p} = \frac{|(\mathbf{x})_1|^p}{|(\mathbf{z})_1|^q} = \cdots = \frac{|(\mathbf{x})_n|^p}{|(\mathbf{z})_n|^q} \tag{19}$$

$$(\mathbf{x})_0^p = |(\mathbf{x})_1|^p + \cdots + |(\mathbf{x})_n|^p \textit{ Or equivalently} \tag{20}$$

$$(\mathbf{z})_0^q p = |(\mathbf{z})_1|^q + \cdots + |(\mathbf{z})_n|^q \tag{21}$$

For optimization problem over $p$cones the system to be solved now is given by:

$$\begin{array}{l} A\mathbf{x} = \mathbf{b} \\ A^T\mathbf{y} + \mathbf{z} = \mathbf{c} \\ \sum_j^n (\mathbf{x})_j(\mathbf{z})_j = \mu \\ |(\mathbf{x})_j|^{p-1}(\mathbf{z})_0 - (\mathbf{x})_0^{p-1}|(\mathbf{z})_j| = 0 \quad \text{for } j = 1\cdots n \end{array} \tag{22}$$

Deriving Newton's direction for this system is quite similar to the case of the ice cream cone. However, proving polynomiality seems to be more challenging. The case of p-cones therefore, will not be discussed in the sequel.

3 Primal-dual interior point algorithms

We now study the primal–dual interior point algorithms which are extensions of [KMY89] and [MA89]. We focus on the ice cream cone and the positive semidefinite cone, omitting treatment of p-cones as they present further difficulties. We also look at the well-known LP case as a blueprint based on which the other problems are developed.

3.1 logarithmic barrier for primal and dual problems

3.1.1 Ice cream cone:

Let us replace the primal problem with

$$\begin{array}{lll} (P_\mu) & \min & \sum_{i=1}^n \mathbf{c_i}^T\mathbf{x_i} - \mu\sum_{i=1}^n \ln\gamma(\mathbf{x_i}) \\ & \text{s.t.} & \sum_{i=1}^n A_i\mathbf{x_i} = \mathbf{b} \end{array} \tag{23}$$

Now in order to solve the primal-dual pair (4) we derive the KKT conditions on (23) and apply Newton's iteration to the resulting system of equations:

$$\mathbf{c_i} - A_i^T\mathbf{y} - \frac{2\mu}{\gamma(\mathbf{x_i})}\begin{pmatrix} (\mathbf{x_i})_0 \\ -(\mathbf{x_i})_1 \\ \vdots \\ -(\mathbf{x_i})_{n_i} \end{pmatrix} = \mathbf{0} \quad \text{for } i = 1, \cdots, n \tag{24}$$

$$\sum_{i=1}^{n} A_i\mathbf{x_i} = \mathbf{b} \tag{25}$$

Setting

$$\mathbf{z_i} \stackrel{\text{def}}{=} \mathbf{c_i} - A_i^T\mathbf{y}$$

We have

$$\frac{(\mathbf{x_i})_0}{(\mathbf{z_i})_0} = -\frac{(\mathbf{x_i})_1}{(\mathbf{z_i})_1} = \cdots = -\frac{(\mathbf{x_i})_{n_i}}{(\mathbf{z_i})_{n_i}} = \frac{\gamma(\mathbf{x_i})}{2\mu} = \frac{\gamma(\mathbf{x_i})}{(\mathbf{x_i})_0(\mathbf{z_i})_0 + \cdots + (\mathbf{x_i})_{n_i}(\mathbf{z_i})_{n_i}} \tag{26}$$

which implies

$$(\mathbf{x_i})_0(\mathbf{z_i})_0 + \cdots + (\mathbf{x_i})_{n_i}(\mathbf{z_i})_{n_i} = 2\mu \tag{27}$$

Similarly if we relax the dual problem we get

$$\begin{array}{lll} (D_\mu) & \max & \mathbf{b}^T\mathbf{y} - \nu\sum_{i=1}^{n}[\ln[(\mathbf{z_i})_0^2 - (\mathbf{z_i})_1^2 - \cdots - (\mathbf{z_i})_{n_i}^2] \\ & \text{s.t.} & A_i^T\mathbf{y} + \mathbf{z_i} = \mathbf{c_i} \end{array} \tag{28}$$

And the KKT conditions on the dual gives:

$$\mathbf{b} - \sum_{i=1}^{n} A_i\mathbf{x_i} = \mathbf{0} \tag{29}$$

$$\mathbf{c_i} - A_i\mathbf{y} - \mathbf{z_i} = \mathbf{0} \tag{30}$$

$$\frac{2\nu}{\gamma(\mathbf{z_i})}\begin{pmatrix} (\mathbf{z_i})_0 \\ -(\mathbf{z_i})_1 \\ \vdots \\ -(\mathbf{z_i})_{n_i} \end{pmatrix} = \begin{pmatrix} (\mathbf{x_i})_0 \\ (\mathbf{x_i})_1 \\ \vdots \\ (\mathbf{x_i})_{n_i} \end{pmatrix} \tag{31}$$

Which implies

$$\frac{(\mathbf{x_i})_0}{(\mathbf{z_i})_0} = -\frac{(\mathbf{x_i})_1}{(\mathbf{z_i})_1} = \cdots = -\frac{(\mathbf{x_i})_{n_i}}{(\mathbf{z_i})_{n_i}} = \frac{2\nu}{\gamma(\mathbf{z_i})} = \frac{(\mathbf{x_i})_0(\mathbf{z_i})_0 + \cdots + (\mathbf{x_i})_{n_i}(\mathbf{z_i})_{n_i}}{\gamma(\mathbf{z_i})} \tag{32}$$

and therefore

$$(\mathbf{x_i})_0(\mathbf{z_i})_0 + \cdots + (\mathbf{x_i})_{n_i}(\mathbf{z_i})_{n_i} = 2\nu \tag{33}$$

We see that by identifying $\mathbf{x_i}$, $\mathbf{y}$ and $\mathbf{z_i}$ in the primal with the corresponding variables in the dual we will have $\mu = \nu$; in other words applying KKT on logarithmic barrier relaxation of the primal is equivalent to KKT conditions on the logarithmic barrier relaxation of the dual.

3.1.2 positive semidefinite cone:

Using the function $\ln \det X$ and proceeding as before by replacing the primal problem with the relaxed barrier version we get:

$$\begin{array}{lll} \min & C \bullet X - \mu \ln \det X & \\ \text{s.t.} & A_i \bullet X = b_i & \text{for } i = 1, \cdots, n \end{array} \tag{34}$$

The Lagrangian function is given by

$$\mathcal{L}(X, \mathbf{y}) \stackrel{\text{def}}{=} C \bullet X - \mu \ln \det X - \sum_i y_i (b_i - A_i \bullet X)$$

and applying the KKT conditions we get the following system of equations:

$$\begin{array}{l} \nabla_X \mathcal{L} = C - \sum y_i A_i - \mu X^{-1} = 0 \\ \nabla_{y_i} \mathcal{L} = b_i - A_i \bullet X = 0 \text{ for } i = 1, \cdots, n \end{array} \tag{35}$$

Now introducing a new symmetric matrix Z we get the following system:

$$\begin{array}{ll} A_i \bullet X = b_i & \text{for } i = 1, \cdots, n \\ \sum y_i A_i + Z = C & \\ XZ = \mu I & \end{array} \tag{36}$$

where the last equation is equivalent to $Z = \mu X^{-1}$. Notice that this is equivalent to the relaxed from of the complementary slackness condition $XZ = 0$.

Now let us apply the same approach on the dual:

$$\begin{array}{ll} \max & \mathbf{b}^T \mathbf{y} - \nu \ln \det Z \\ \text{s.t.} & \sum y_i A_i + Z = C \end{array}$$

The dual Lagrangian now is (with the Lagrange multiplier a symmetric matrix):

$$\mathcal{L}(X, \mathbf{y}, Z) = \mathbf{b}^T \mathbf{y} - \nu \ln \det Z - X \bullet (C - \sum y_i A_i - Z)$$

Applying KKT we get

$$\begin{array}{l} \nabla_X \mathcal{L} = C - \sum y_i A_i - Z = 0 \\ \nabla_{y_i} \mathcal{L} = b_i - A_i \bullet X = 0 \text{ for } i = 1, \cdots, m \\ \nabla_Z \mathcal{L} = \nu Z^{-1} - X = 0 \end{array}$$

which is essentially the same system one gets from the primal problem.

3.2 The Newton Iteration

All three cones (the positive orthant, the ice cream cone and the positive semidefinite cone) have the property that the complementarity condition is of the form $\mathbf{F}(\mathbf{x}, \mathbf{z}) = \mathbf{0}$, where each $\mathbf{F}_i$ is a *bilinear form* in $\mathbf{x}$ and $\mathbf{z}$. Therefore the derivative of $\mathbf{F}$ with respect to $\mathbf{z}$is a matrix dependent only on $\mathbf{x}$, and with respect to $\mathbf{x}$a matrix entirely dependent on $\mathbf{z}$; let us call these matrices $\mathcal{Z}$ and $\mathcal{X}$ respectively.

Therefore, the process of deriving the Newton iteration in all three cones are quite similar. First notice that the relaxed complementary slackness theorem in all three cases has a matrix formulation:

$$\mathcal{X}\mathcal{Z}\mathbf{e} = \mathcal{X}\mathbf{z} = \mathcal{Z}\mathcal{X}\mathbf{e} = \mathcal{Z}\mathbf{x} = \mu\mathbf{e} \tag{37}$$

where matrices $\mathcal{X}$ and $\mathcal{Z}$ and vector $\mathbf{e}$are different in each of the three optimization problems. We will derive the specific expressions for each of these cones in the following subsections. However, notice that for all three, the newton direction $\Delta\mathbf{x}, \Delta\mathbf{y}, \Delta\mathbf{z})$ is given by the following system:

$$\begin{pmatrix} \mathcal{A} & 0 & 0 \\ 0 & \mathcal{A}^T & I \\ \mathcal{Z} & 0 & \mathcal{X} \end{pmatrix} \begin{pmatrix} \Delta\mathbf{x} \\ \Delta\mathbf{y} \\ \Delta\mathbf{z} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathcal{X}\mathcal{Z}\mathbf{e} - \mu\mathbf{e} \end{pmatrix} \tag{38}$$

where $\mathcal{A}$ is the coefficient matrix of the primal problem. The vector $\mathbf{e}$is sometimes referred to as the "center" of the cone. The relaxed complementarity condition can be written in the matrix form:

$$\mathcal{X}\mathcal{Z}\mathbf{e} = \mathcal{Z}\mathcal{X}\mathbf{e} = \mu\mathbf{e}$$

Solving (38) symbolically we can obtain the newton direction as follows:

$$\begin{array}{l} \Delta\mathbf{x} = [\mathcal{Z}^{-1} - \mathcal{Z}^{-1}\mathcal{X}\mathcal{A}^T(\mathcal{A}\mathcal{Z}^{-1}\mathcal{X}\mathcal{A})^{-1}\mathcal{A}\mathcal{Z}^{-1}](\mathcal{X}\mathcal{Z}\mathbf{e} - \mu\mathbf{e}) \\ \Delta\mathbf{y} = -[(\mathcal{A}\mathcal{Z}^{-1}\mathcal{X}\mathcal{A})^{-1}\mathcal{A}\mathcal{Z}^{-1}](\mathcal{X}\mathcal{Z}\mathbf{e} - \mu\mathbf{e}) \\ \Delta\mathbf{z} = \mathcal{A}^T[(\mathcal{A}\mathcal{Z}^{-1}\mathcal{X}\mathcal{A})^{-1}\mathcal{A}\mathcal{Z}^{-1}](\mathcal{X}\mathcal{Z}\mathbf{e} - \mu\mathbf{e}) \end{array} \tag{39}$$

3.2.1 In LP:

Here the relaxed complementarity conditions $x_i z_i = \mu$ for $i = 1, \cdots, n$ can be written in the matrix form (37), where

$$\mathcal{X} = \text{Diag }(\mathbf{x}), \mathcal{Z} = \text{Diag }(\mathbf{z}), \text{ and } \mathbf{e} = \mathbf{1}$$

($\mathbf{1}$is the vector of all ones.)

3.2.2 In QCQP:

The relaxed complementary slackness conditions are given by

$$\begin{array}{ll} \sum_{j=0}^{n_i}(\mathbf{x_i})_j(\mathbf{z_i})_j = \mu & \text{for } i = 1\cdots n \\ (\mathbf{x_i})_j(\mathbf{z_i})_0 + (\mathbf{x_i})_0(\mathbf{z_i})_j = 0 & \text{for } i = 1\cdots n, \ j = 1\cdots n_i \end{array} \tag{40}$$

Notice that $\mathbf{x_i}, \mathbf{z_i} \in \Re^{n_i+1}$. To see what $\mathcal{X}$, $\mathcal{Z}$ and $\mathbf{e}$ are explicitly we temporarily assume that $n = 1$ and $n_1 = k$ and therefore we will drop the subscripts. We have:

$$(\mathbf{x})_0(\mathbf{z})_0 + \cdots + (\mathbf{x})_k(\mathbf{z})_k = \mu \tag{41}$$

and

$$(\mathbf{x})_j(\mathbf{z})_0 + (\mathbf{x})_0(\mathbf{z})_j = 0 \text{ for } j = 1, \cdots, k \tag{42}$$

Now replacing $(\mathbf{x})_j$ and $(\mathbf{z})_j$ by $(\mathbf{x})_j - (\Delta\mathbf{x})_j$ and $(\mathbf{z})_j - \Delta\mathbf{z})_j$, respectively, and dropping the nonlinear terms, (41) results in

$$(\mathbf{x})_0\Delta(\mathbf{z})_0 + \Delta(\mathbf{x})_0(\mathbf{z})_0 + \cdots + (\mathbf{x})_k(\Delta\mathbf{z})_k + (\Delta\mathbf{x})_k(\mathbf{z})_k = (\mathbf{x})_0(\mathbf{z})_0 + \cdots + (\mathbf{x})_n(\mathbf{z})_n - \mu \tag{43}$$

Similarly (42) results in:

$$(\mathbf{x})_j\Delta\mathbf{z})_0 + (\Delta\mathbf{x})_j(\mathbf{z})_0 + (\mathbf{z})_j(\Delta\mathbf{x})_0 + (\Delta\mathbf{z})_j(\mathbf{x})_0 = (\mathbf{x})_j(\mathbf{z})_0 + (\mathbf{x})_0(\mathbf{z})_j \tag{44}$$

From (43),(44) we see that

$$\mathcal{Z} = \text{Arw}\,(\mathbf{z}) \stackrel{\text{def}}{=} \begin{pmatrix} (\mathbf{z})_0 & (\mathbf{z})_1 & \cdots & (\mathbf{z})_k \\ (\mathbf{z})_1 & (\mathbf{z})_0 & & \\ \vdots & & \ddots & \\ (\mathbf{z})_k & & & (\mathbf{z})_0 \end{pmatrix} \tag{45}$$

Similarly

$$X = \text{Arw}\,(\mathbf{x}) \stackrel{\text{def}}{=} \begin{pmatrix} (\mathbf{x})_0 & (\mathbf{x})_1 & \cdots & (\mathbf{x})_k \\ (\mathbf{x})_1 & (\mathbf{x})_0 & & \\ \vdots & & \ddots & \\ (\mathbf{x})_k & & & (\mathbf{x})_0 \end{pmatrix} \tag{46}$$

Finally

$$\mathbf{e} = \mathbf{u} \stackrel{\text{def}}{=} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{47}$$

Thus the operator "Arw" maps a (zero indexed) vector $\mathbf{x}$on to a "arrow shaped" matrix where the first row and column are identical to $\mathbf{x}$, the diagonal elements are $\mathbf{x}_0$ and everywhere else we have zero. The "Arw " operator is the analog of "Diag " operator in LP.

In the context of QCQP let us also define:

$$\Delta X \stackrel{\text{def}}{=} \text{Arw}\,(\Delta\mathbf{x}) \text{ and } \Delta Z \stackrel{\text{def}}{=} \text{Arw}\,(\Delta\mathbf{z})$$

In the general case where $n \neq 1$ we simply have:

$$\mathcal{X} = \text{Arw}\,(\mathbf{x_1}) \oplus \cdots \oplus \text{Arw}\,(\mathbf{x_n}) \text{ and similarly } \mathcal{Z} = \text{Arw}\,(\mathbf{z_1}) \oplus \cdots \oplus \text{Arw}\,(\mathbf{z_n})$$

where $\oplus$ is the direct sum of matrices. In addition

$$\mathbf{u} = \begin{pmatrix} \mathbf{u_1} \\ \vdots \\ \mathbf{u_k} \end{pmatrix}$$

(Here each $\mathbf{u_i} = (1, 0, \cdots, 0)^T$ is an n_i-vector.)

Observe that $\mathbf{x} \geq_Q \mathbf{0}$ iff $\mathrm{Arw}(\mathbf{x}) \succeq 0$. This shows that QCQP is also a special semidefinite programming problem as was observed by Nesterov and Nemirovskii [NN94]. However, in terms of computational complexity, it is not useful to treat QCQP as an SDP, because in that case interior point methods imply complexity proportional to $\sqrt{\sum n_i k}$ rather than $\sqrt{n}$, see the development under semidefinite programming in the next section. Notice also that $\mathrm{Arw}(\mathbf{x_i})$ and $\mathrm{Arw}(\mathbf{z_i})$ are low rank perturbations of diagonal matrices (in fact multiples of identity), and inverting each one is quite straightforward.

The system of equations to be solved in QCQP is therefore given by

$$\begin{array}{ll} A_1\mathbf{x_1} + \cdots + A_n\mathbf{x_k} = \mathbf{b} & \\ A_i^T\mathbf{y} + \mathbf{z_i} = \mathbf{c_i} & \text{for } i = 1 \cdots n \\ \sum_{j=0}^{n_i} \mathbf{x_i})_j(\mathbf{z_i})_j = \mu & \text{for } i = 1 \cdots n \\ (\mathbf{x_i})_j(\mathbf{z_i})_0 + (\mathbf{x_i})_0(\mathbf{z_i})_j = 0 & \text{for } i = 1 \cdots n,\ j = 1 \cdots n_i \end{array} \tag{48}$$

Remembering that $\mathbf{x_i}, \mathbf{z_i} \in \Re^{n_i+1}$ and $A_i \in \Re^{m\times(n_i+1)}$ and setting $N \stackrel{\text{def}}{=} \sum_i n_i + n$ the system (48) has $2N + m$ equations in the same number of unknowns. Applying newton iteration to (48) we get:

$$\left(\begin{array}{ccc|c|ccc} A_1 & \cdots & A_n & 0 & 0 & \cdots & 0 \\ \hline 0 & \cdots & 0 & A_1^T & I & & \\ \vdots & \ddots & \vdots & \vdots & & \ddots & \\ 0 & \cdots & 0 & A_n^T & & & I \\ \hline Z_1 & & & 0 & X_1 & & \\ & \ddots & & \vdots & & \ddots & \\ & & Z_n & 0 & & & X_n \end{array}\right) \begin{pmatrix} \Delta\mathbf{x_1} \\ \vdots \\ \Delta\mathbf{x_n} \\ \Delta\mathbf{y} \\ \Delta\mathbf{z_1} \\ \vdots \\ \Delta\mathbf{z_n} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ X_1Z_1\mathbf{u_1} - \mu\mathbf{u_1} \\ \vdots \\ X_nZ_n\mathbf{u_n} - \mu\mathbf{u_n} \end{pmatrix} \tag{49}$$

This is analog of (38) where $\mathcal{A} = [A_1, \cdots, A_n]$. And each $X_i = \mathrm{Arw}(\mathbf{x_i})$, each $Z_i = \mathrm{Arw}(\mathbf{z_i})$

3.2.3 In SDP:

For the SDP problem the relaxed complementary slackness is given by the matrix relation:

$$XZ = \mu I$$

However, this relation is not symmetric, that is given symmetric matrices X and Z, the matrix XZ is not always symmetric. The result is that after applying the newton iteration, the direction ΔX may not be a symmetric matrix. To overcome this problem we instead use the following as the relaxed complementarity condition:

$$XZ + ZX = 2\mu I$$

Again replacing X by $X - \Delta X$ and Z by $Z - \Delta Z$ and removing nonlinear terms the complementary relations give us:

$$X\Delta Z + \Delta Z X + Z\Delta X + \Delta X Z = 2\mu I - XZ - ZX \tag{50}$$

Writing this in the vector notation we get:

$$(Z \otimes I + I \otimes Z)\textbf{vec}\,(\Delta X) + (X \otimes I + I \otimes X)\textbf{vec}\,(\Delta Z) = 2\mu\textbf{vec}\,I - \textbf{vec}\,(XZ + ZX) \tag{51}$$

Thus, in SDP

$$\mathcal{X} = X \otimes I + I \otimes X,\ \mathcal{Z} = Z \otimes I + I \otimes Z,\ \text{and } \mathbf{e} = \textbf{vec}\,I$$

Recall that $\otimes$ is the Kronecker product of matrices and the operation $A \otimes I + I \otimes B$ is the *Kronecker sum* of matrices A and B, see [HJ90].

For the SDP problem our system of equation is given by:

$$\begin{array}{ll} \mathcal{A}\textbf{vec}\,X & = 0 \\ \mathcal{A}^T\mathbf{y} + \textbf{vec}\,Z & = \textbf{vec}\,C \\ XZ + ZX & = 2\mu I \end{array} \tag{52}$$

Thus the newton system of equations can be written as:

$$\begin{array}{l} \mathcal{A}\textbf{vec}\,\Delta X = \mathbf{0} \\ \mathcal{A}^T\Delta\mathbf{y} + \textbf{vec}\,\Delta Z = 0 \\ X\Delta Z + \Delta Z X + Z\Delta X + \Delta X Z = 2\mu I - XZ - ZX \end{array} \tag{53}$$

Or in the vector/Kronecker notation:

$$\begin{pmatrix} \mathcal{A} & 0 & 0 \\ 0 & \mathcal{A}^T & I \\ Z \otimes I + I \otimes Z & 0 & X \otimes I + I \otimes X \end{pmatrix} \begin{pmatrix} \textbf{vec}\,\Delta X \\ \Delta\mathbf{y} \\ \textbf{vec}\,\Delta Z \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \textbf{vec}\,(XZ + ZX - 2\mu I) \end{pmatrix} \tag{54}$$

3.2.4 General results for all three cones:

We now prove some statements that are true regardless of which of the three cones we are working on. First notice that the complementarity condition is equivalent to

$$\mathcal{X}\mathbf{z} = \mathcal{Z}\mathbf{x} = 0$$

in all three cases. To obtain the general system of equations arising from applying the Newton method one replaces $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ with, respectively $\mathbf{x} - \Delta\mathbf{x}$, $\mathbf{y} - \Delta\mathbf{y}$ and $\mathbf{z} - \Delta\mathbf{z}$, and ends up with solving a system of equations which has the following generic form:

Lemma 3.1 *The solution of (38) satisfies*

$$\Delta \mathbf{x}^T \Delta \mathbf{z} = 0$$

Proof: Multiply both sides of the second set of equations

$$\mathcal{A}^T \Delta \mathbf{y} + \Delta \mathbf{z} = \mathbf{0}$$

by $\Delta \mathbf{x}$ and notice that by the first set $\mathcal{A}\Delta \mathbf{x} = \mathbf{0}$. ∎

Lemma 3.2 *For all three cones*

1. $\mathcal{Z}\mathbf{e} = \mathbf{z}$ *and* $\mathcal{X}\mathbf{e} = \mathbf{x}$.
2. *On the central path* $\mathcal{X}\mathcal{Z} = \mathcal{Z}\mathcal{X}$.
3. $\mathcal{X}\mathbf{z} = \mathcal{Z}\mathbf{x} = \mu\mathbf{e}$

Second, let us define the primal–dual central path as:

$$\{\mathbf{w}_\mu = (\mathbf{x}^T, \mathbf{y}^T, \mathbf{z}^T)^T : (\mathbf{x}^T, \mathbf{y}^T, \mathbf{z}^T)^T \text{ satisfies (38) for some } \mu \geq 0 \}$$

Suppose $\mathbf{x}$ and $\mathbf{z}$ are on the primal-dual-central path for some μ, and $\Delta \mathbf{x}$ and $\Delta \mathbf{z}$ are newton iterates. Let us assume that we take a step length of α in this direction and that

$$\begin{aligned} \hat{\mathbf{x}} &= \mathbf{x} - \alpha \Delta \mathbf{x} \\ \hat{\mathbf{y}} &= \mathbf{y} - \alpha \Delta \mathbf{y} \\ \hat{\mathbf{z}} &= \mathbf{z} - \alpha \Delta \mathbf{z} \end{aligned}$$

are the new points.

Lemma 3.3 *For all three cones the maximum α such that both $\mathbf{x} - \alpha\Delta\mathbf{x}$ and $\mathbf{z} - \alpha\Delta\mathbf{z}$ are feasible is given by*

$$\frac{1}{\alpha} = \max\{\lambda_1(\mathcal{X}^{-1}\Delta\mathcal{X}), \lambda_1(\mathcal{Z}^{-1}\Delta\mathcal{Z})\}$$

Where $\lambda_1(A)$ is the the largest eigenvalue of A.

Notice that since for all three cones $\mathcal{X} \succeq 0$ and $\Delta\mathcal{X}$ is symmetric the eigenvalues of (not necessarily symmetric) matrix $\mathcal{X}^{-1}\Delta\mathcal{X}$ are all real; the same holds for $\mathcal{Z}^{-1}\Delta\mathcal{Z}$.

Lemma 3.4 *Let $\mathbf{x}$and $\mathbf{z}$be on the central path, and $\Delta\mathbf{x}$ and $\Delta\mathbf{z}$ be the solution of (38) with $\mu = 0$. Then*

$$\hat{\mathbf{x}}^T\hat{\mathbf{z}} = (1-\alpha)\mathbf{x}^T\mathbf{z}$$

Proof:

$$\begin{aligned}\hat{\mathbf{x}}^T\hat{\mathbf{z}} &= (\mathbf{x}-\alpha\Delta\mathbf{x})^T(\mathbf{z}-\alpha\Delta\mathbf{z})\\ &= \mathbf{x}^T\mathbf{z}-\alpha(\mathbf{e}^T\mathcal{Z}\Delta\mathbf{x}\mathbf{e}+\mathbf{e}^T\mathcal{X}\Delta\mathbf{z}\mathbf{e})+\alpha\Delta\mathbf{x}^T\Delta\mathbf{z}\\ &= (1-\alpha)\mathbf{x}^T\mathbf{z}\end{aligned}$$

■

In this section we will show that if α is chosen as large as possible then one step of the newton direction will reduce the duality gap by a factor of $1-1/\sqrt{n}$. Since we have already shown the amount of reduction is exactly $1-\alpha$, we only need to show that $\alpha \geq 1/\sqrt{n}$.

In the remainder of this section We manipulate the relation

$$\mathcal{Z}\Delta\mathbf{x}+\mathcal{X}\Delta\mathbf{z}=\mathcal{X}\mathcal{Z}\mathbf{e}-\mu\mathbf{e}$$

in each of the three cones and show the following lemma:

Lemma 3.5 *Let* $\mathbf{x}$*and* $\mathbf{z}$*be on the primal-dual central path. Let also that* $\Delta\mathbf{x}$ *and* $\Delta\mathbf{z}$ *be the solution of the system (38) with* $\mu=0$*. Finally let* α *be the maximal number such that both* $\hat{\mathbf{x}}=\mathbf{x}-\alpha\Delta\mathbf{x}$ *and* $\hat{\mathbf{z}}=\mathbf{z}-\alpha\Delta\mathbf{z}$ *are feasible. Then*

$$\alpha \geq \frac{1}{\sqrt{n}}$$

3.3 Behavior on the central path

We will now set out to prove lemma 3.5.

In case of LP Lemma 3.5 is easy. One has

$$\alpha=\max\{\alpha:\ \mathrm{Diag}\ (\mathbf{x})-\alpha\mathrm{Diag}\ (\Delta\mathbf{x})\succeq 0 \text{ and } \mathrm{Diag}\ (\mathbf{z})-\alpha\mathrm{Diag}\ (\Delta\mathbf{z})\succeq 0\}$$

which is equivalent to

$$\frac{1}{\alpha^2}=\max_i\left(\frac{(\Delta x_i)^2}{x_i^2},\frac{(\Delta z_i)^2}{z_i^2}\right)\leq\sum_i\left(\frac{\Delta x_i^2}{x_i^2}+\frac{\Delta z_i^2}{z_i^2}\right)=\left[\sum_i\left(\frac{\Delta x_i}{x_i}+\frac{\Delta z_i}{z_i}\right)^2-2\sum_i\frac{\Delta x_i\Delta z_i}{x_iz_i}\right]$$

But the last sum equals zero because all denominators $x_iz_i=\mu$ and the numerators add up to zero. Also, from $z_i\Delta x_i+x_i\Delta z_i=x_iz_i$ we have

$$\frac{\Delta x_i}{x_i}+\frac{\Delta z_i}{z_i}=1$$

Thus, the right hand side above adds up to n and so we get

$$\alpha \geq \frac{1}{\sqrt{n}}.$$

3.4 Proof of Lemma 3.5 in QCQP

First we note some useful relations. We will develop most of our machinery with the assumption $n = 1$ and $n_1 = k$, in order to avoid tedious notation. Recall that

$$\mathbf{x} = \begin{pmatrix} (\mathbf{x})_0 \\ \overline{\mathbf{x}} \end{pmatrix}, \quad \mathbf{z} = \begin{pmatrix} (\mathbf{z})_0 \\ \overline{\mathbf{z}} \end{pmatrix}$$

and

$$X = \mathrm{Arw}\,(\mathbf{x}) \text{ and } Z = \mathrm{Arw}\,(\mathbf{z})$$

Then

$$X = \begin{pmatrix} (\mathbf{x})_0 & \overline{\mathbf{x}}^T \\ \overline{\mathbf{x}}^T & (\mathbf{x})_0 I \end{pmatrix} \text{ and } Z = \begin{pmatrix} (\mathbf{z})_0 & \overline{\mathbf{z}}^T \\ \overline{\mathbf{x}}^T & (\mathbf{z})_0 I \end{pmatrix}$$

and

$$X^{-1} = \frac{1}{\gamma(\mathbf{x})} \begin{pmatrix} (\mathbf{x})_0 & -\overline{\mathbf{x}}^T \\ -\overline{\mathbf{x}} & \frac{1}{(\mathbf{x})_0}(\gamma(\mathbf{x})I + \overline{\mathbf{x}}\overline{\mathbf{x}}^T) \end{pmatrix} \text{ and } Z^{-1} = \frac{1}{\gamma(\mathbf{z})} \begin{pmatrix} (\mathbf{z})_0 & -\overline{\mathbf{z}}^T \\ -\overline{\mathbf{z}} & \frac{1}{(\mathbf{z})_0}(\gamma(\mathbf{z})I + \overline{\mathbf{z}}\overline{\mathbf{z}}^T) \end{pmatrix} \tag{55}$$

On the central path, for a given constant $\mu > 0$ the following relationships hold:

1. We have

   $$(\mathbf{x})_0(\mathbf{z})_0 + \cdots + (\mathbf{x})_k(\mathbf{z})_k = \mu$$

   and

   $$(\mathbf{x})_0(\mathbf{z})_i + (\mathbf{x})_i(\mathbf{z})_0 = 0$$

   which is equivalent to

   $$r_\mu \stackrel{\text{def}}{=} \frac{(\mathbf{x})_0}{(\mathbf{z})_0} = -\frac{(\mathbf{x})_1}{(\mathbf{z})_1} = \cdots = -\frac{(\mathbf{x})_k}{(\mathbf{z})_k} = \frac{\mu}{\gamma(\mathbf{z})} = \frac{\gamma(\mathbf{x})}{\mu}$$

2. We have

   $$\mathbf{x} = r_\mu R\mathbf{z}$$

   and

   $$Z = \frac{1}{r_\mu} \begin{pmatrix} (\mathbf{x})_0 & -\overline{\mathbf{x}}^T \\ -\overline{\mathbf{x}} & (\mathbf{x})_0 I \end{pmatrix}$$

3. Thus,

   $$XZ = \frac{1}{r_\mu} \begin{pmatrix} \gamma(\mathbf{x}) & \mathbf{0}^T \\ \mathbf{0} & (\mathbf{x})_0^2 I - \overline{\mathbf{x}}\overline{\mathbf{x}}^T \end{pmatrix}$$

   Therefore X and Z on the central path commute with each other. We also have

   $$(XZ)^{-1} = r_\mu \begin{pmatrix} \frac{1}{\gamma(\mathbf{x})} & \mathbf{0}^T \\ \mathbf{0} & \frac{1}{(\mathbf{x})_0^2}\left(I + \frac{\overline{\mathbf{x}}\overline{\mathbf{x}}^T}{\gamma(\mathbf{x})}\right) \end{pmatrix}$$

Now Suppose $\Delta\mathbf{x}$ and $\Delta\mathbf{y}$ and $\Delta\mathbf{z}$ are derived from the solution of the Newton system with starting points $\mathbf{x}$ and $\mathbf{z}$ on the central path and with parameter μ. Recall that

$$\frac{1}{\alpha} = \max\{\lambda_1(X^{-1}\Delta X), \lambda_1(Z^{-1}\Delta Z)\}$$

Observing that $XZ = ZX$ on the central path, the relation $X\Delta Z\mathbf{u} + Z\Delta X\mathbf{u} = XZ\mathbf{u}$ implies

Lemma 3.6 *For $\mathbf{x}$ and $\mathbf{z}$ on the central path we have*

1.
$$Z^{-1}\Delta Z\mathbf{u} + X^{-1}\Delta X\mathbf{u} = \mathbf{u} \tag{56}$$

2.
$$\Delta Z Z^{-1}\mathbf{u} + \Delta X X^{-1}\mathbf{u} = \mathbf{u} \tag{57}$$

Proof:

1. Simply multiply both sides by $X^{-1}Z^{-1}$.
2. We have

$$\begin{aligned}
\Delta Z Z^{-1}\mathbf{u} + \Delta X X^{-1}\mathbf{u} &= \Delta Z\frac{R\mathbf{z}}{\gamma(\mathbf{z})} + \Delta X\frac{R\mathbf{x}}{\gamma(\mathbf{x})} \\
&= \frac{1}{r_\mu\gamma(\mathbf{z})}\Delta Z\mathbf{x} + \frac{r_\mu}{\gamma(\mathbf{x})}\Delta X\mathbf{z} \\
&= \frac{1}{\mu}(\Delta Z X\mathbf{u} + \Delta X Z\mathbf{u}) \\
&= \mathbf{u}
\end{aligned}$$

∎

To show the next result we first introduce some notation: Let

$\mathbf{u}^T(X^{-1}\Delta X) = \frac{1}{\gamma(\mathbf{x})}(\theta_\mathbf{x}, \boldsymbol{\beta}_\mathbf{x}^T)$
$(X^{-1}\Delta X)\mathbf{u} = \frac{1}{\gamma(\mathbf{x})}(\theta_\mathbf{x}, \boldsymbol{\delta}_\mathbf{x}^T)^T$

where

$\theta_\mathbf{x} \stackrel{\text{def}}{=} (\mathbf{x})_0(\Delta\mathbf{x})_0 - \overline{\mathbf{x}}^T\Delta\overline{\mathbf{x}}$
$\boldsymbol{\beta}_\mathbf{x} \stackrel{\text{def}}{=} (\mathbf{x})_0\Delta\mathbf{x} - (\Delta\mathbf{x})_0\mathbf{x}$
$\boldsymbol{\delta}_\mathbf{x} \stackrel{\text{def}}{=} -(\Delta\mathbf{x})_0\overline{\mathbf{x}} + W_{\overline{\mathbf{x}}}\Delta\overline{\mathbf{x}}$

Here $W_\mathbf{x}$ for any vector $\mathbf{x}$ is defined as

$$W_\mathbf{x} \stackrel{\text{def}}{=} \frac{1}{(\mathbf{x})_0}\left(\gamma(\mathbf{x})I + \mathbf{x}\mathbf{x}^T\right)$$

Note, therefore that

$$\mathbf{x}^T W_{\mathbf{x}} = (\mathbf{x})_0 \mathbf{x}^T$$

Similarly, define $\theta_{\mathbf{z}}$, $\boldsymbol{\beta}_{\mathbf{z}}$ and $\boldsymbol{\delta}_{\mathbf{z}}$, by replacing $\mathbf{z}$ for $\mathbf{x}$.

Observe that by inner producting the relation in parts 1 and 2 in Lemma 56 above we get

$$\mathbf{u}^T(X^{-1}\Delta X)^2\mathbf{u} + \mathbf{u}^T(Z^{-1}\Delta Z)^2\mathbf{u} + \mathbf{u}^T(X^{-1}\Delta X Z^{-1}\Delta Z)\mathbf{u} + \mathbf{u}^T(Z^{-1}\Delta Z X^{-1}\Delta X)\mathbf{u} = 1 \quad (58)$$

(the right hand side of 58 would be n for $n \neq 1$.) However one can prove:

Proposition 3.1 *With the assumptions of Lemma 56*

1.

$$\mathbf{u}^T(Z^{-1}\Delta Z X^{-1}\Delta X)\mathbf{u} + \mathbf{u}^T(X^{-1}\Delta X Z^{-1}\Delta Z)\mathbf{u} = 0 \quad (59)$$

2.

$$\frac{\mathbf{u}^T(Z^{-1}\Delta Z)^2\mathbf{u}}{\lambda_1(Z^{-1}\Delta Z)^2} \geq \frac{1}{2}, \qquad \frac{\mathbf{u}^T(X^{-1}\Delta X)^2\mathbf{u}}{\lambda_1(X^{-1}\Delta X)^2} \geq \frac{1}{2} \quad (60)$$

Proof:

1. We have,

$$\begin{aligned} \theta_{\mathbf{x}}\theta_{\mathbf{z}} &= [(\mathbf{x})_0(\Delta\mathbf{x})_0 - \overline{\mathbf{x}}^T\Delta\overline{\mathbf{x}}][((\mathbf{z})_0(\Delta\mathbf{z})_0 - \overline{\mathbf{z}}^T\Delta\overline{\mathbf{z}}] \\ &= (\mathbf{x})_0(\mathbf{z})_0(\Delta\mathbf{x})_0(\Delta\mathbf{z})_0 + \overline{\mathbf{x}}^T\Delta\overline{\mathbf{z}}\,\overline{\mathbf{z}}^T\Delta\overline{\mathbf{x}} + (\mathbf{z})_0(\Delta\mathbf{x})_0\overline{\mathbf{x}}^T\Delta\overline{\mathbf{z}} + (\mathbf{x})_0(\Delta\mathbf{z})_0\overline{\mathbf{z}}^T\Delta\overline{\mathbf{x}} \end{aligned}$$

and,

$$\begin{aligned} \boldsymbol{\beta}_{\mathbf{x}}^T\boldsymbol{\delta}_{\mathbf{x}} &= ((\mathbf{x})_0\Delta\overline{\mathbf{x}} - (\Delta\mathbf{x})_0\overline{\mathbf{x}})^T(-(\Delta\mathbf{z})_0\overline{\mathbf{z}} + W_{\overline{\mathbf{z}}}\Delta\overline{\mathbf{z}}) \\ &= -(\mathbf{x})_0(\Delta\mathbf{z})_0\overline{\mathbf{z}}^T\Delta\overline{\mathbf{x}} - (\mathbf{z})_0(\Delta\mathbf{x})_0\overline{\mathbf{x}}^T\Delta\overline{\mathbf{z}} - (\Delta\mathbf{x})_0(\Delta\mathbf{z})_0(\mathbf{x})_0(\mathbf{z})_0 + \mu[(\Delta\mathbf{x})_0(\Delta\mathbf{z})_0 + \Delta\overline{\mathbf{x}}^T\Delta\overline{\mathbf{z}}] \\ &\quad -\overline{\mathbf{x}}^T\Delta\overline{\mathbf{z}}\,\overline{\mathbf{z}}^T\Delta\overline{\mathbf{x}} \end{aligned}$$

Therefore

$\theta_{\mathbf{x}}\theta_{\mathbf{z}} + \boldsymbol{\beta}_{\mathbf{x}}^T\boldsymbol{\delta}_{\mathbf{z}} = \mu\left((\Delta\mathbf{x})_0(\Delta\mathbf{z})_0 + \Delta\overline{\mathbf{x}}^T\Delta\overline{\mathbf{z}}\right)$, and by symmetry
$\theta_{\mathbf{x}}\theta_{\mathbf{x}} + \boldsymbol{\beta}_{\mathbf{z}}^T\boldsymbol{\delta}_{\mathbf{x}} = \mu\left((\Delta\mathbf{z})_0(\Delta\mathbf{x})_0 + \Delta\overline{\mathbf{z}}^T\Delta\overline{\mathbf{x}}\right)$.

Now,

$$\mathbf{u}^T\left(X^{-1}\Delta X Z^{-1}\Delta Z\right)\mathbf{u} + \mathbf{u}^T\left(Z^{-1}\Delta Z X^{-1}\Delta X\right)\mathbf{u} = \frac{1}{\gamma(\mathbf{x})\gamma(\mathbf{z})}\left[2\theta_{\mathbf{x}}\theta_{\mathbf{z}} + \boldsymbol{\beta}_{\mathbf{x}}^T\boldsymbol{\delta}_{\mathbf{x}} + \boldsymbol{\beta}_{\mathbf{z}}^T\boldsymbol{\delta}_{\mathbf{x}}\right] = 0$$

which proves *1.* (Again, it should be noted that the last equality holds also for the case $n \neq 1$ since $\gamma(\mathbf{x_i})\gamma(\mathbf{z_i}) = \mu^2$ for all i and, while $\Delta\mathbf{x_i}^T\Delta\mathbf{z_i}$ is not necessarily zero, the sum $\sum_i \Delta\mathbf{x_i}^T\Delta\mathbf{z_i} = 0$.)

2. Note that $\lambda_1(X^{-1}\Delta X)$ is a root of

$$\gamma(\mathbf{x})\lambda^2 - 2\theta_{\mathbf{x}}\lambda + \gamma(\Delta\mathbf{x}) = 0$$

so,

$$\lambda_1^2(X^{-1}\Delta X) = \frac{\left(\theta_{\mathbf{x}} \pm \sqrt{\theta_{\mathbf{x}}^2 - \gamma(\mathbf{x})\gamma(\Delta\mathbf{x})}\right)^2}{\gamma^2(\mathbf{x})}$$

Also

$$\mathbf{u}^T(X^{-1}\Delta X)^2\mathbf{u} = \frac{1}{\gamma^2(\mathbf{x})}(\theta_{\mathbf{x}}^2 + \boldsymbol{\beta}_{\mathbf{x}}^T\boldsymbol{\delta}_{\mathbf{x}})$$

But,

$$\boldsymbol{\beta}_{\mathbf{x}}^T\boldsymbol{\delta}_{\mathbf{x}} = \theta_{\mathbf{x}}^2 - \gamma(\mathbf{x})\gamma(\Delta\mathbf{x})$$

So we have

$$\mathbf{u}^T(X^{-1}\Delta X)^2\mathbf{u} = \frac{1}{\gamma^2(\mathbf{x})}\left[\theta_{\mathbf{x}}^2 + \theta_{\mathbf{x}}^2 + -\gamma(\mathbf{x})\gamma(\Delta\mathbf{x})\right]$$

Now observing that for all real numbers u and v

$$u^2 + v^2 \geq \frac{1}{2}(u+v)^2$$

we get

$$\mathbf{u}^T(X^{-1}\Delta X)^2\mathbf{u} \geq \frac{1}{2}\lambda_1^2(X^{-1}\Delta X)$$

and by symmetry

$$\mathbf{u}^T(Z^{-1}\Delta Z)^2\mathbf{u} \geq \frac{1}{2}\lambda_1^2(Z^{-1}\Delta Z)$$

and this completes proof of *2*.

∎

Remark. By inner producting (56) and (57) we get:

$$\mathbf{u}^T(\Delta Z Z^{-1}\Delta Z Z^{-1} + \Delta X X^{-1}\Delta X X^{-1})\mathbf{u} = 1 \tag{61}$$

Two other terms that is the result of the inner product equals zero by part *1* of Proposition 3.1. (Again the right hand side is generally equal to $\mathbf{u}^t\mathbf{u} = n$ for the case $n \neq 1$.)

It is now easy to prove Lemma (3.6). By part *2* of Proposition 3.1 we have

$$\frac{1}{\alpha^2} = \max\left(\lambda_1^2(\Delta Z Z^{-1}), \lambda_1^2(\Delta X X^{-1})\right) \leq 2$$

And for the general case $n \neq 1$ $1/\alpha^2 \leq 2n$ or

$$\alpha \geq \frac{1}{\sqrt{2n}}$$

3.5 Proof of Lemma 3.4 in SDP

We concentrate on the relation:

$$X\Delta Z + \Delta ZX + Z\Delta X + \Delta XZ = XZ + ZX = 2\mu I$$

Which is the case when $\mu = 0$. First notice that

$$\frac{1}{\alpha} = \max\{\lambda_1(X^{-1}\Delta X), \lambda_1(Z^{-1}\Delta Z)\}$$

This is consistent with the lemma 3.3. To see this observe that $\Delta\mathcal{Z} = \Delta Z \otimes I + I \otimes \Delta Z$, and

$$\lambda_1(Z^{-1}\Delta Z) = \lambda_1\left[(Z \otimes I + I \otimes Z)^{-1}\Delta Z \otimes I + I \otimes \Delta Z\right]$$

Lemma 3.7 *For X and Z on the central path with $XZ = \mu I$ we have:*

$$X\Delta Z + \Delta XZ = Z\Delta X + \Delta ZX = \mu I$$

Proof: Let $P \stackrel{\text{def}}{=} X\Delta Z + \Delta XZ$. We need to show that $P = P^T = \mu I$. We know that $P + P^T = 2\mu I$ and $P^T = X^{-1}PX$. Thus

$$P + X^{-1}PX = 2\mu I$$

which is equivalent to the Lyapunov equation:

$$XP + PX = 2\mu X$$

Since X is positive definite and thus nonsingular, the Lyapunov equation has a unique solution, which must be $P = \mu I$. ∎

Now to prove lemma 3.4 in case of semidefinite programming we first multiply both sides of $X\Delta Z + \Delta XZ = \mu I$ by $(XZ)^{-1}$ which is $(1/\mu)I$ and get:

$$Z^{-1}\Delta Z + \Delta XX^{-1} = I$$

After squaring and then taking trace we get:

$$\text{trace }(Z^{-1}\Delta X)^2 + \text{trace }(\Delta XX^{-1})^2 + \text{trace }(Z^{-1}\Delta Z\Delta XX^{-1}) + \text{trace }(\Delta XX^{-1}Z^{-1}\Delta Z) = n$$

But the last two terms of the left hand side each equals $\Delta X \bullet \Delta Z$ which equals zero. Thus we have

$$\text{trace }(Z^{-1}\Delta X)^2 + \text{trace }(\Delta XX^{-1})^2 = n$$

and thus

$$\frac{1}{\alpha^2} = \max\left(\lambda_1^2(Z^{-1}\Delta Z), \lambda_1^2(X^{-1}\Delta X)\right) \leq n$$

Which implies

$$\alpha \geq \frac{1}{\sqrt{n}}$$

References

[Ali91] F. Alizadeh. *Combinatorial Optimization with Interior Point Methods and Semi-Definite Matrices*. PhD thesis, Computer Science Department, University of Minnesota, Minneapolis, Minnesota, 1991.

[CO94] Andrew R. Conn and Michael Overton. A Primal–Dual Interior Point MEthod for Minimizing a Sum of Euclidean Distances. Presented in Mathematical Programming Symposium, Ann Arbor, 1994.

[HJ90] R. Horn and C. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1990.

[KMY89] M. Kojima, S. Mizuni, and A. Yoshise. A primal-dual interior-point algorithm for linear programming. In N. Megiddo, editor, *Progress in Mathematical Programming*, Berlin, 1989. Springer Verlag.

[MA89] R. C. Monteiro and I. Adler. Interior path following primal–dual algorithms. Part I: Linear programming. *Math. Programming*, 44:27–41, 1989.

[MN93] K. Mints and A. Nemirovskii. Extension of Karmarkar's Algorithm onto Convex Quadratically Constrained Quadratic Problems. to appear, 1993.

[NN94] Y. Nesterov and A. Nemirovskii. *Interior Point Polynomial Methods in Convex Programming: Theory and Applications*. Society for Industrial and Applied Mathematics, Philadelphia, 1994.