ON SENSITIVITY OF CENTRAL SOLUTIONS IN SEMIDEFINITE PROGRAMMING

J.F. Sturm[1] and S. Zhang[2]

March 1998. Revised: May 1999 and September 2000.

ABSTRACT

In this paper we study the properties of the analytic central path of a semidefinite programming problem under perturbation of the right hand side of the constraints, including the limiting behavior when the central optimal solution, namely the analytic center of the optimal set, is approached. Our analysis assumes the primal-dual Slater condition and the strict complementarity condition. Our findings are as follows. First, on the negative side, if we view the central optimal solution as a function of the right hand side of the constraints, then this function is not continuous in general, whereas in the linear programming case this function is known to be Lipschitz continuous. On the positive side, compared with the previous conclusion we obtain a (seemingly) paradoxical result: on the central path any directional derivative with respect to the right hand side of the constraints is bounded, and even converges as the central optimal solution is approached. This phenomenon is possible due to the lack of a uniform bound on the derivatives with respect to the right hand side parameters. All these results are based on the strict complementarity assumption. Concerning this last property we give an example. In that example the set of right hand side parameters for which the strict complementarity condition holds is neither open nor closed. This is remarkable since a similar set for which the primal-dual Slater condition holds is always open.

Key words: analytic central path, semidefinite programming, condition number.

AMS subject classification: 15A48, 90C25, 90C31.

[1] Department of Quantitative Economics, Maastricht University, The Netherlands. Research performed at Communications Research Laboratory, McMaster University, Hamilton, Canada. Supported by Netherlands Organization for Scientific Research (NWO).

[2] Department of Systems Engineering & Engineering Management, Chinese University of Hong Kong.

1 Introduction

In recent years, semidefinite programming (abbreviated as SDP hereafter) has emerged as an important subject in optimization. It turns out that SDP has many applications; see e.g. [3, 4, 6, 10, 13, 35, 36]. Recent intensive research indicates that the interior point method can be an efficient means to solve SDPs and related problems. One may consult the interior point homepage at URL `http://www.mcs.anl.gov/home/otc/InteriorPoint/` and the related links for historical developments of SDP. Extensive treatments on the primal-dual interior point methods for SDP can be found in De Klerk [7] and Sturm [32].

These developments can be traced back to a few pioneering works. Among them are Nesterov and Nemirovsky [26] and Nesterov and Todd [27], who introduced the concepts of self-concordant barriers for convex cones and self-scaled barriers for symmetric cones, respectively. This gives rise to the possibility of solving SDP by interior point methods. A key observation is that an SDP can be seen as a linear program (LP) over the cone of positive semidefinite matrices. Indeed, many similarities between LP and SDP have been revealed ever since, and are used in the design of solution methods. Recognizing their similarities, SDP is also known as "LP for the 90's and 00's", [38]. However, there still exist some fundamental differences between the two problem classes. For a discussion on this subject, we refer to Luo, Sturm and Zhang [21]. In particular, LP possesses a very important property which is combinatorial in nature, i.e. if an LP problem has an optimal solution then it always has a basic optimal solution. Although some attempts are made to generalize concepts like basic solutions from LP to SDP (cf. Pataki [28]), there seem to be lots of complications still. As a consequence, classical results for sensitivity analysis and parametric programming for LP cannot be carried over easily to the SDP case, since most of these classical results are based on the analysis of a basic optimal solution. However, there is a general, well developed, perturbation theory for optimization problems; see the recent survey by Bonnans and Shapiro [5]. Under regularity and nondegeneracy conditions, Shapiro [32] showed that the then unique optimal solution is differentiable in the perturbation parameter; this is an application of the inverse function theorem. Bonnans and Shapiro [5] also gave an example with unique primal and dual optimal solutions, but violating the strict complementarity condition, in which the optimal solution is not Lipschitz stable. Without imposing nondegeneracy conditions, the directional derivatives of the optimal value function exist, in any perturbation direction; see Rockafellar [30]. Recently, Goldfarb and Scheinberg [12] investigated how these directional derivatives can be computed. Just as in linear programming, this amounts to computing dual multipliers on the boundary of the dual optimal face. Helmberg [15] demonstrated how to use dual solutions to estimate the optimal objective value when new constraints are added.

As is well known, each iteration of the interior point method involves the resolution of a scaled system of normal equations, typically of the form

$$\text{find } y \text{ such that } (ADA^{\mathrm{T}})y = b,$$

where A is an $m \times N$ matrix of rank m. The matrix D is an iteratively updated positive definite matrix; in linear programming it is further known to be diagonal. If the optimization problem is degenerate, then the matrix ADA^{T} gets increasingly ill–conditioned as the iteration process continues. Fortunately, this ill–conditioning does not result in any numerical problems (in the context of linear programming). In fact, computational experience has shown that the interior point method produces highly accurate solutions for linear programming problems, with or without degeneracy.

Now that degeneracy is no longer a major concern, one may look at another sort of nasty situation: What if the problem is almost infeasible, unbounded or unsolvable? If a tiny perturbation in the data can make the optimal solution set empty, then this may certainly be a point of concern, and this leads to the "distance to ill–posedness" measure, as introduced by Renegar [29]. The "distance to ill–posedness" has been used in global convergence analysis [29] and also in the context of sensitivity analysis; see Nunez and Freund [25].

Other approaches are based on the classical perturbation theory for linear systems of equations. Recall that the (matrix) condition number $\kappa(A)$ measures the sensitivity of the least squares solution x_{LS} of "$Ax = b$", with respect to changes in the right hand side b. Hoffman's error bound provides a possible way to generalize this condition number to linear equality and inequality systems. Namely, consider a polyhedral set

$$\mathcal{P} = \{x \mid Ax = b, x \geq 0\},$$

with A an $m \times N$ matrix and $\mathcal{P} \neq \emptyset$. The result of Hoffman states that there exists some (Lipschitz) constant $K > 0$ such that

$$\operatorname{dist}(x, \mathcal{P}) \leq K(\|Ax - b\| + \|[-x]_+\|) \text{ for all } x \in \Re^N,$$

were $[\cdot]_+$ denotes the non-negative part of a vector. Thus, the Lipschitz constant K is a measure of sensitivity to the right hand sides in "$Ax = b$" and "$x \geq 0$". Since the optimal solution set of a linear program can be described as a set of linear equalities and inequalities (viz. by using LP duality), we obtain a sensitivity measure for linear programming.

In the context of interior point methods, the constant in Hoffman's error bound may be a very conservative measure of sensitivity. Namely, interior point methods typically trace the central path towards the analytic center of the optimal solution set. In the following, a 'central solution' can be either a solution on the central path, or the analytic center of the optimal solution set (which is the limit of the central path) depending on the context. We are thus interested in the sensitivity of central solutions. We will encounter many similarities between the concept of central solutions in linear and semidefinite programming and the concept of least squares solutions in systems of linear equalities. For linear programming, it was shown by Holder, Sturm and Zhang [16] that all directional derivatives of the central optimal solution with respect to right hand side perturbations exist. An upper bound on the norm of the derivatives can be estimated as $\chi_{\rm LP}(A)/\|A\|_2$, where

$$\chi_{\rm LP}(A) = \sup\left\{ \|A\|_2 \|DA^{\rm T}(ADA^{\rm T})^{-1}\|_2 \,\middle|\, D \text{ is a positive diagonal matrix} \right\}.$$

The quantity $\chi_{\rm LP}(A)$ is always finite, and is called the condition number for linear programming [34]. Notice in particular that if A is invertible, then $\chi_{\rm LP}(A)$ reduces to the standard condition number of A for linear equation systems (and any associated linear program also reduces to a system of linear equations). The condition number for linear programming is an important quantity in the complexity theory of linear programming under the real-number computational model; see Vavavis and Ye [34].

The concept of centrality is important in perturbation theory for semidefinite programming. If we have a pair of primal and dual solutions (X, Z) with duality gap $\operatorname{tr} XZ$, then the distance to the optimal solution set can be of the same order as $\sqrt{\operatorname{tr} XZ}$, even if a pair of strictly complementary solutions exists. However, if (X, Z) is on the central path, and the strict complementarity condition holds, then the distance to the central optimal solution is at most $O(\operatorname{tr} XZ)$ (see Luo, Sturm and Zhang [20]). Thus, solutions on the central path are significantly less sensitive to perturbations in the duality gap. In this paper, we will show that a similar phenomenon occurs when analyzing the sensitivity of central solutions with respect to perturbations in the right hand side.

Organization of the paper and the contents of each section.

The organization of the paper is as follows. In the next section we introduce two equivalent formulations of the standard semidefinite programming problem. The first formulation is conventional and directly uses a matrix linear subspace in defining the domain of the problem. The second one, however, explicitly uses a vector subspace to indicate the domain. These two forms are, of course, equivalent. An advantage of using the second form is to make the formulae to be developed in the paper easy comparable to their linear programming counterparts. There is a crucial difference between the existence of an interior point solution (the Slater condition) and the existence of strictly complementary primal-dual optimal solutions. In fact, it is possible that a semidefinite program satisfying strict complementarity may loose this property with whatever small perturbations on the right hand side vector. In contrast, this can never happen in the case of Slater type properties. This and other related issues will also be discussed in Section 2, with some illustrating examples presented in the Appendix.

In the interior of the feasible region of a semidefinite program, solutions on the analytic central path are less sensitive to the changes of the right hand side vector. Section 3 is devoted to the sensitivity analysis of this nature. In other words, we study how a point on the central path, with a fixed $\mu > 0$, would react to the changes on the right hand side vector b. The rate of change, viz. the Jacobian matrix, turns out to be in the form of a pseudo projection matrix. This type of projection matrices do not have a uniform bound in general. The geometric meanings and properties of the projection matrices, including the limiting property, are given in Section 3.

In Section 4 we show that the Jacobian matrix as introduced in Section 3 is in fact convergent to a finite limit as $\mu \to 0$. The limiting matrix, in return, completely characterizes the central optimal solutions of the primal-dual problems. This limiting matrix itself, unfortunately, may not be continuous with respect to the right hand side b.

We conclude the paper with a discussion in Section 5.

2 Semidefinite programming

Let $\mathcal{S}^{n\times n}$ denote the linear space spanned by all $n \times n$ real symmetric matrices. Let $\mathcal{S}^{n\times n}_{+}$ denote the cone of positive semidefinite matrices, and $\mathcal{S}^{n\times n}_{++}$ be the cone of positive definite matrices.

Consider the following semidefinite program in standard form:

$$\begin{array}{lll} (P) & \text{minimize} & C \bullet X \\ & \text{subject to} & A[i] \bullet X = b_i, \qquad \text{for } i = 1, ..., m \\ & & X \succeq 0 \end{array}$$

where C, $A[i] \in \mathcal{S}^{n\times n}$ for $i = 1, ..., m$. The standard inner product for two symmetric matrices is $A \bullet B = \operatorname{tr} AB = \sum_{i,j} A_{ij} B_{ij}$, and $A \succeq B$ means that $A - B \in \mathcal{S}^{n\times n}_{+}$.

The dual of (P) reads

$$\begin{array}{lll} (D) & \text{maximize} & \sum_{i=1}^{m} y_i b_i \\ & \text{subject to} & \sum_{i=1}^{m} y_i A[i] + Z = C \\ & & Z \succeq 0. \end{array}$$

In this paper, however, we choose to use the vector representation of semidefinite programming problems. This means that we will restrict ourselves to semidefinite programming in the $N = n(n+1)/2$ dimensional Euclidean space of real symmetric $n \times n$ matrices. Let $\{U[1], U[2], \ldots, U[N]\}$ be an orthonormal basis of the the space of $n \times n$ real symmetric matrices $\mathcal{S}^{n\times n}$. We can then represent any real symmetric matrix by its N coordinates with respect to this basis. Obviously, under a nonsingular linear mapping the matrix space and its vector representation are equivalent. In this paper, we will use these two representations interchangeably when there is no confusion.

We consider the image of the cone $\mathcal{S}_+^{n\times n}$ in $\Re^N$:

$$\left\{ x \in \Re^N \,\middle|\, \sum_{i=1}^N x_i U[i] \succeq 0 \right\}. \tag{2.1}$$

Its interior is the image of $\mathcal{S}_{++}^{n\times n}$

$$\left\{ x \in \Re^N \,\middle|\, \sum_{i=1}^N x_i U[i] \succ 0 \right\}.$$

If no confusion is possible, we will also denote these two images by $\mathcal{S}_+^{n\times n}$ and $\mathcal{S}_{++}^{n\times n}$ respectively. Note that the cone $\mathcal{S}_+^{n\times n}$ is self–dual.

In the vector form we can reformulate the primal-dual semidefinite programming problems as

$$\begin{array}{lll} (P_b) & \text{minimize} & c^{\mathrm{T}}x \\ & \text{subject to} & Ax = b, \\ & & x \in \mathcal{S}_+^{n\times n}, \end{array}$$

where $c \in \Re^N$, $b \in \Re^m$, $A \in \Re^{m\times N}$, and the decision variable is $x \in \Re^N$. Without loss of generality, we assume that A has full row rank. Associated with (P_b) is its dual problem

$$\begin{array}{lll} (D_b) & \text{maximize} & b^{\mathrm{T}}y \\ & \text{subject to} & A^{\mathrm{T}}y + z = c \\ & & z \in \mathcal{S}_+^{n\times n}. \end{array}$$

Since we shall frequently use the matrix and vector form of the problem formulation in an interchangeable way, we shall first explain our notation.

Let $a[i] \in \Re^N$ denote the i-th row of A, i.e.

$$A^{\mathrm{T}} = \left[\begin{array}{ccc} a[1], & \cdots, & a[m] \end{array} \right].$$

We adopt the convention that lower case symbols $x, c, a[i]$ denote vectors in $\Re^N$, and the corresponding upper case symbols $X, C, A[i]$ denote their symmetric matrix representations. Under isomorphism, we have

$$X = \sum_{k=1}^N x_k U[k], \;\; C = \sum_{k=1}^N c_k U[k]$$

and

$$A[i] = \sum_{k=1}^N a_k[i] U[k].$$

Conversely, we have

$$x_k = U[k] \bullet X, \;\; c_k = U[k] \bullet C \tag{2.2}$$

and

$$a_k[i] = U[k] \bullet A[i], \qquad \text{for all } k = 1, 2, \ldots, N.$$

In some places we will use the convention to write $x = \operatorname{vec}_{\mathcal{S}} X$ to state explicitly that $x_k = U[k] \bullet X$, $k = 1, 2, \ldots, N$, when $U[k]$'s are the usual unit symmetric matrices. More discussions on the notion of $\operatorname{vec}_{\mathcal{S}}$ and vec will be given in Section 3.

Denote the primal feasible set to be

$$\mathcal{F}_P(b) = \{x \mid Ax = b\} \cap \mathcal{S}_+^{n\times n}$$

and the dual feasible set

$$\mathcal{F}_D = \{z \mid A^{\mathrm{T}} y + z = c \text{ for some } y \in \Re^m\} \cap \mathcal{S}_+^{n\times n}.$$

We will only consider those values of b for which (P_b) is feasible, i.e. $b \in A\mathcal{S}_+^{n\times n}$, and we assume that $\mathcal{F}_D$ contains a positive definite solution:

Assumption 1 $\mathcal{F}_D \cap \mathcal{S}_{++}^{n\times n} \neq \emptyset$.

The above condition is known as the dual Slater condition. This condition is equivalent to stating that the set of primal optimal solutions is nonempty and bounded for all $b \in A\mathcal{S}_+^{n\times n}$, see [21, 30]. Therefore, Assumption 1 holds if and only if the primal has a *central optimal solution* which is sometimes also called the *analytic center of the optimal face*, denoted by $x^*(b)$. For linear programming, the central optimal solution is known to be Lipschitz continuous; see Holder, Sturm and Zhang [16]. Moreover, it is well known that the primal and dual central optimal solutions always satisfy the so called strict complementarity condition. Unfortunately, neither of these properties carry over to semidefinite programming. In particular, the set of b values for which the strict complementarity condition holds can be neither open nor closed. This is in contrast with the primal-dual Slater condition. The set of b values for which the primal and dual problems both satisfy the Slater condition obviously forms an open set.

Proposition 2.1 *The set of b vectors for which (P_b) and (D_b) satisfy the strict complementarity condition is neither closed nor open in general.*

Proof. See Example 1 in Appendix. **Q.E.D.**

Furthermore, a rather counter-intuitive result is that even in the favorable situation that both the Slater condition and the strict complementary condition hold, we still do not necessarily have continuity of the central solution. Formally we have the following result.

Proposition 2.2 *The central optimal solution for (P_b) is in general not continuous in b, even if the primal–dual Slater condition and strict complementarity condition are satisfied.*

Proof. See Example 2 in Appendix. **Q.E.D.**

Remark that unlike the above situation, the central optimal solution to a linear programming problem must always be Lipschitz continuous in b; see Holder, Sturm and Zhang [16].

3 Sensitivity on the central path

In this section, we study the sensitivity of solutions on the central path, with respect to changes in the right hand side vector b. In Section 4, we will study the limiting behavior if the central optimal solution is approached.

Denote the family of central paths for the parametic problem pair (P_b), (D_b) by

$$\{(X(b,\mu), Z(b,\mu)) \in \mathcal{F}_P(b) \times \mathcal{F}_D \mid X(b,\mu)Z(b,\mu) = \mu I,\ \mu > 0,\ b \in A\mathcal{S}_{++}^{n\times n}\}. \tag{3.1}$$

Thus, we have by definition that $X(b,\mu)Z(b,\mu) = \mu I$. Moreover, $(x(b,\mu), y(b,\mu), z(b,\mu))$ is uniquely defined by (b,μ); see [26]. Using the classical inverse function theorem, it follows that

$$(X(b,\mu), y(b,\mu), Z(b,\mu))$$

is a smooth function of b on the domain $A\mathcal{S}_{++}^{n\times n}$.

It is now useful to work with the so-called Kronecker products of matrices for ease of deriving various formulae later. A comprehensive discussion on the linear algebra of Kronecker products can be found in Chapter 4 of Horn and Johnson [18] (see also Horn and Johnson [17] for related background information). Below we will sketch some basic facts to serve our discussion.

For matrices $A \in \Re^{m\times n}$ and $B \in \Re^{p\times q}$, the standard "vec" operator on A is defined as

$$\mathrm{vec}\,(A) = (a_{11}, \cdots, a_{m1}, a_{12}, \cdots, a_{m2}, \cdots, a_{1n}, \cdots, a_{mn})^T$$

and the Kronecker product between A and B is defined as

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}.$$

The following lemma proves to be useful in many applications; its proof can be found in Chapter 4 of Horn and Johnson [18].

Lemma 3.1 *It holds that*

- $(A \otimes B)(C \otimes D) = AC \otimes BD$
- $\mathrm{vec}\,(AB) = (I \otimes A)\mathrm{vec}\, B = (B^T \otimes I)\mathrm{vec}\, A$
- $\mathrm{vec}\,(AXB) = (B^T \otimes A)\mathrm{vec}\, X.$

Notice that for a symmetric matrix $X \in \mathcal{S}^{n\times n}$, we have that $\mathrm{vec}\,(X)$ is an n^2-dimensional vector, whereas $\mathrm{vec}_{\mathcal{S}}\,(X)$ is an N-dimensional vector, with $N = n(n+1)/2$; see (2.2).

Differentiating with respect to b on both sides of the identity

$$\mu I = X(b,\mu)Z(b,\mu)$$

yields (with abuse of notation)

$$0 = X(b,\mu)\left(\frac{\partial}{\partial b_i} Z(b,\mu)\right) + \left(\frac{\partial}{\partial b_i} X(b,\mu)\right) Z(b,\mu) \text{ for } i = 1,2,\ldots,m.$$

Applying Lemma 3.1 we obtain

$$0 = (I \otimes X(b,\mu))\left(\frac{\partial}{\partial b^{\mathrm{T}}} \mathrm{vec}\, Z(b,\mu)\right) + (Z(b,\mu) \otimes I)\left(\frac{\partial}{\partial b^{\mathrm{T}}} \mathrm{vec}\, X(b,\mu)\right).$$

Pre-multiplying the above identity with the matrix $(X(b,\mu) \otimes I)$, and using Lemma 3.1, yields

$$\begin{aligned} 0 &= (X(b,\mu) \otimes X(b,\mu))\left(\frac{\partial}{\partial b^{\mathrm{T}}} \mathrm{vec}\, Z(b,\mu)\right) + (X(b,\mu)Z(b,\mu) \otimes I)\left(\frac{\partial}{\partial b^{\mathrm{T}}} \mathrm{vec}\, X(b,\mu)\right) \\ &= (X(b,\mu) \otimes X(b,\mu))\left(\frac{\partial}{\partial b^{\mathrm{T}}} \mathrm{vec}\, Z(b,\mu)\right) + \mu\left(\frac{\partial}{\partial b^{\mathrm{T}}} \mathrm{vec}\, X(b,\mu)\right). \end{aligned} \tag{3.2}$$

We remark here (again using Lemma 3.1) that for any $n \times n$ symmetric matrix W, it holds that

$$(X(b,\mu) \otimes X(b,\mu))\mathrm{vec}\, W = \mathrm{vec}\,(X(b,\mu)WX(b,\mu)),$$

which is a linear transformation within the space of symmetric $n \times n$ matrices. Notice further that if $W \neq 0$ then

$$W \bullet (X(b,\mu)WX(b,\mu)) > 0.$$

Therefore, we can find a positive definite matrix $D(b,\mu) \in \Re^{N\times N}$ such that

$$D(b,\mu)w = \mathrm{vec}_{\mathcal{S}}\,(X(b,\mu)WX(b,\mu)) \text{ for all } W \in \mathcal{S}^{n\times n},$$

where $w = \mathrm{vec}\, W$; see also [2].

Differentiating the feasibility relations

$$Ax(b,\mu) = b, \quad A^{\mathrm{T}}y(b,\mu) + z(b,\mu) = c,$$

and using (3.2), we conclude that $(\partial x(b,\mu)/\partial b^{\mathrm{T}}, \partial y(b,\mu)/\partial b^{\mathrm{T}}, \partial z(b,\mu)/\partial b^{\mathrm{T}})$ satisfies the system (3.3)–(3.5):

$$A\frac{\partial x(b,\mu)}{\partial b^{\mathrm{T}}} = I \tag{3.3}$$

$$A^{\mathrm{T}}\frac{\partial y(b,\mu)}{\partial b^{\mathrm{T}}} + \frac{\partial z(b,\mu)}{\partial b^{\mathrm{T}}} = 0 \tag{3.4}$$

$$D(b,\mu)\frac{\partial z(b,\mu)}{\partial b^{\mathrm{T}}} + \mu\frac{\partial x(b,\mu)}{\partial b^{\mathrm{T}}} = 0. \tag{3.5}$$

Since $D(b,\mu)$ is positive definite, and A has full row rank, it follows that the system (3.3)–(3.5) is invertible. In other words, the Jacobian with respect to b is nonsingular for the central path equations.

Combining (3.4) and (3.5), we have

$$\frac{\partial x(b,\mu)}{\partial b^{\mathrm{T}}} = \frac{1}{\mu}D(b,\mu)A^{\mathrm{T}}\frac{\partial y(b,\mu)}{\partial b^{\mathrm{T}}}. \tag{3.6}$$

Substituting this expression into (3.3) yields

$$AD(b,\mu)A^{\mathrm{T}}\frac{\partial y(b,\mu)}{\partial b^{\mathrm{T}}} = \mu I. \tag{3.7}$$

Since $D(b,\mu)$ is nonsingular, this further implies that

$$\frac{\partial y(b,\mu)}{\partial b^{\mathrm{T}}} = \mu(AD(b,\mu)A^{\mathrm{T}})^{-1}.$$

Now substitute the above expression into (3.6) to obtain

$$\frac{\partial x(b,\mu)}{\partial b^{\mathrm{T}}} = D(b,\mu)A^{\mathrm{T}}(AD(b,\mu)A^{\mathrm{T}})^{-1}. \tag{3.8}$$

We define for $D \in \mathcal{S}_{++}^{N\times N}$ a matrix $Q(D)$ as

$$Q(D) := DA^{\mathrm{T}}(ADA^{\mathrm{T}})^{-1}A, \tag{3.9}$$

which is sometimes called a pseudo projection matrix. Remark that $Q(D)$ is well defined, since A has full row rank and D is positive definite. Furthermore, we see from (3.8) that

$$\frac{\partial x(b,\mu)}{\partial b^{\mathrm{T}}} = Q(D(b,\mu))A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}, \tag{3.10}$$

so that

$$\|A\|_2\left\|\frac{\partial x(b,\mu)}{\partial b^{\mathrm{T}}}\right\|_2 \le \|Q(D(b,\mu))\|_2\kappa(A), \tag{3.11}$$

where

$$\kappa(A) = \|A\|_2 \cdot \|A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}\|_2$$

is the (matrix) condition number of A. We are thus naturally led to investigate the norm of the pseudo projection $Q(D(b,\mu))$. Note that in the case when $D(b,\mu)$ is a positive diagonal matrix, then the norm of $Q(D(b,\mu))$ is known to be uniformly bounded. This is an important fact used in the complexity analysis of interior point methods for linear programming. We will come back to this point in Proposition 3.1.

In the following, we let $\angle(x,z)$ denote the angle between two nonzero vectors $x,z \in \Re^N$, i.e.

$$\angle(x,z) := \arccos \frac{x^{\mathrm{T}} z}{\|x\|_2 \|z\|_2}.$$

Clearly, for any given positive definite matrix D, the range space of $Q(D)$ intersects the null space of A only at the origin. It is unclear, however, to what extend the range space of $Q(D)$ can lean towards the null space of A. This issue is addressed in the next theorem.

Theorem 3.1 *Let $A \in \Re^{m\times N}$ have full row rank. Then for any*

$$Q \in \mathrm{cl}\,\{Q(D) \mid D \in \mathcal{S}_{++}^{N\times N}\}$$

it holds that

$$\min\{\sin\angle(x,Qs) \mid Ax = 0,\ x \neq 0,\ s \neq 0\} = \frac{1}{\|Q\|_2} > 0.$$

Moreover,

$$AQ = A, \quad \mathrm{rank}\,(Q) = m.$$

Proof. Let P denote the orthogonal projection onto the image of A^{T}, i.e.

$$P := A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}A.$$

Since $Q(D)P = Q(D)$, $AQ(D) = A$ and $PQ(D) = P$, for all positive definite D, we must also have

$$QP = Q, \quad AQ = A \ \text{ and } PQ = P. \tag{3.12}$$

Therefore, $Qs = QPs$ for all $s \in \Re^N$. Thus, for any $s \in \Re^N$ there is $y \in \Re^m$ such that

$$Qs = QPs = QA^{\mathrm{T}}y, \quad \|A^{\mathrm{T}}y\|_2 = \|Ps\|_2 \leq \|s\|_2.$$

By definition of the matrix 2-norm, we further get

$$\|Q\|_2 = \max_{s\neq 0} \frac{\|Qs\|_2}{\|s\|_2} = \max_{y\neq 0} \frac{\|QA^{\mathrm{T}}y\|_2}{\|A^{\mathrm{T}}y\|_2}. \tag{3.13}$$

From (3.12) we know that $AQ = A$, and so $A^{\mathrm{T}}y \perp (Q - I)A^{\mathrm{T}}y$. Therefore, we can apply the basic relations of a rectangular triangle to conclude that

$$\sin \angle((Q - I)A^{\mathrm{T}}y, QA^{\mathrm{T}}y) = \frac{\|A^{\mathrm{T}}y\|_2}{\|QA^{\mathrm{T}}y\|_2}.$$

Together with (3.13), this implies further that

$$\min\{\sin \angle(x, Qs) \mid Ax = 0,\ s \in \Re^N\} \leq \frac{1}{\|Q\|_2}. \tag{3.14}$$

Conversely, for any $y \in \Re^m$ and x with $Ax = 0$, we have

$$\sin \angle(x, QA^{\mathrm{T}}y) = \frac{\|QA^{\mathrm{T}}y - tx\|}{\|QA^{\mathrm{T}}y\|} \tag{3.15}$$

where $t \in \Re$ is chosen such that $QA^{\mathrm{T}}y - tx \perp x$. However,

$$\|QA^{\mathrm{T}}y - tx\| \geq \|P(QA^{\mathrm{T}}y - tx)\| = \|PQA^{\mathrm{T}}y\|. \tag{3.16}$$

Recall now from (3.12) that

$$PQA^{\mathrm{T}} = PA^{\mathrm{T}} = A^{\mathrm{T}}. \tag{3.17}$$

Combining (3.15)–(3.17), we obtain

$$\sin \angle(x, QA^{\mathrm{T}}y) \geq \frac{\|PQA^{\mathrm{T}}y\|}{\|QA^{\mathrm{T}}y\|} = \frac{\|A^{\mathrm{T}}y\|}{\|QA^{\mathrm{T}}y\|} \geq \frac{1}{\|Q\|_2}, \tag{3.18}$$

where the last step follows from (3.13). The lemma follows from (3.18) and (3.14) by inclusion.

Q.E.D.

Remark that for given $D \in \mathcal{S}_{++}^{N\times N}$, the matrices $Q(D)$ and DA^{T} have the same range space, since $Qs = DA^{\mathrm{T}}y$ with $y = (ADA^{\mathrm{T}})^{-1}As$. Remark also that if $DA^{\mathrm{T}}y \neq 0$ then $ADA^{\mathrm{T}}y \neq 0$ and hence $DA^{\mathrm{T}}y$ cannot be in the null space of A. If we restrict D to be a positive diagonal matrix, then $\angle(x, DA^{\mathrm{T}}y)$ with $Ax = 0$ is bounded from below by a positive constant, independent of x, y and D:

Proposition 3.1 *Let $A \in \Re^{m\times N}$ have full row rank. Then the set*

$$\{Q(D) \mid D \textit{ is a positive diagonal matrix}\}$$

is bounded. Therefore, the angle $\angle(x, DA^{\mathrm{T}}y)$ with $Ax = 0$, $x \neq 0$, $y \neq 0$ and D a positive diagonal matrix is bounded below from zero by a positive constant, which is independent of x, y and D.

The above result is well-known (cf. [8, 9, 31, 33]) and is widely used for the analysis of interior point methods in linear programming. The minimum angle between $\operatorname{Ker} A$ and $\operatorname{Img} DA^{\mathrm{T}}$ is also used as a condition number for linear programs whose constraint matrix is A; see Vavasis and Ye [34].

If Q is a limit point of the (bounded) set $\{Q(D) \mid D \text{ is a positive diagonal matrix}\}$, one might expect that $\operatorname{Img} Q = \operatorname{Img} DA^{\mathrm{T}}$ for some (possibly non-positive) diagonal matrix D. However, this is not true; consider for instance the case

$$A = \begin{bmatrix} -1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

Writing $D = \operatorname{diag}(d_1, d_2, d_3)$, we see that

$$AD = \begin{bmatrix} -d_1 & d_2 & 0 \\ 0 & d_2 & d_3 \end{bmatrix}.$$

Hence, if we choose $d_1 = d_2 = \epsilon > 0$ and $d_3 = 1$ then for $\epsilon \downarrow 0$ we have

$$\operatorname{Img} \operatorname{diag}(\epsilon, \epsilon, 1) A^{\mathrm{T}} \to \operatorname{Img} \begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{\mathrm{T}}.$$

But it is clear that this limiting image cannot be written as $\operatorname{Img} DA^{\mathrm{T}}$ for any diagonal matrix D.

In the setting of semidefinite programming however, the scaling matrix is a linear transformation of the form $W \to XWX$ with X positive definite. Proposition 3.1 is thus only applicable if we restrict to a fixed eigenvector basis for X. Still, one may wonder whether Proposition 3.1 can be extended to positive definite transformations D of the form $D(W) = XWX$ for general positive definite X. This turns out not to be the case.

Consider an example with

$$A[1] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad A[2] = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

and

$$(X_\epsilon)^{-1} = \begin{bmatrix} 1 & 1/\epsilon \\ 1/\epsilon & 2/\epsilon^2 \end{bmatrix}, \quad Z_\epsilon = \begin{bmatrix} 1 & 0 \\ 0 & -\epsilon^2 \end{bmatrix}.$$

We see that

$$X_\epsilon^{-1} Z_\epsilon X_\epsilon^{-1} = \begin{bmatrix} 0 & -1/\epsilon \\ -1/\epsilon & -3/\epsilon^2 \end{bmatrix} = -\frac{1}{\epsilon} A[1] - \frac{3}{\epsilon^2} A[2].$$

But,

$$Z_\epsilon \to \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \in \{X \mid A[1] \bullet X = 0,\ A[2] \bullet X = 0\}.$$

However, if we consider the scaling matrices on the central path, then the result holds true again for those $b \in \mathcal{A}\mathcal{S}^{n\times n}_{++}$ where a strictly complementary solution exists. We will make this claim precise in Lemma 4.1. In order to prove this result, we first need to study the limiting behavior of the central path and its derivatives. All these will be topics of discussion for the next section.

4 Limiting behavior

Throughout this section, we consider a fixed right hand side vector $b \in \mathcal{A}\mathcal{S}^{n\times n}_{++}$ for which a strictly complementary solution exists, and we shall study the limiting behavior of the analytic central path as μ approaches zero. Properties of the analytic central path in the context of linear programming or linear complementarity problems have been well studied in the literature; see, e.g., [1, 19, 22, 23, 24, 37].

Since b is fixed, we can drop it as an argument in $(X(b,\mu), Z(b,\mu))$ and $D(b,\mu)$, arriving at the central path

$$\{(X(\mu), Z(\mu)) \in \mathcal{F}_P(b) \times \mathcal{F}_D \mid X(\mu)Z(\mu) = \mu I, \mu > 0\},$$

(see (3.1)). In our setting, the primal-dual analytic central path is a well defined smooth curve [26], and it converges to the central optimal solution $(X^*, Z^*) = \lim_{\mu\downarrow 0}(X(\mu), Z(\mu))$; see [20, 11]. Luo, Sturm and Zhang [20] showed that

$$\|X(\mu) - X^*\| = O(\mu) \text{ and } \|Z(\mu) - Z^*\| = O(\mu) \quad \text{with } 0 < \mu < 1. \tag{4.1}$$

The existence of strictly complementary solutions is essential for the above relation to hold.

Applying a basis transformation if necessary, we may assume that the central optimal solutions X^* and Z^* can be partitioned as

$$X^* = \begin{bmatrix} X^*_{\mathcal{B}} & 0 \\ 0 & 0 \end{bmatrix}, \quad Z^* = \begin{bmatrix} 0 & 0 \\ 0 & Z^*_{\mathcal{N}} \end{bmatrix},$$

where $X^*_{\mathcal{B}}$ and $Z^*_{\mathcal{N}}$ are both positive definite. The subscripts $\mathcal{B}$ and $\mathcal{N}$ are reminiscent of linear programming, where they stand for "basic" and "nonbasic" respectively. Using this notation we can partition a given $n \times n$ symmetric matrix M as

$$\begin{bmatrix} M_{\mathcal{B}} & M_{\mathcal{U}} \\ M_{\mathcal{U}}^T & M_{\mathcal{N}} \end{bmatrix}.$$

In particular, we have

$$\mu I = X(\mu)Z(\mu) = \begin{bmatrix} X_{\mathcal{B}}(\mu) & X_{\mathcal{U}}(\mu) \\ X_{\mathcal{U}}(\mu)^T & X_{\mathcal{N}}(\mu) \end{bmatrix} \cdot \begin{bmatrix} Z_{\mathcal{B}}(\mu) & Z_{\mathcal{U}}(\mu) \\ Z_{\mathcal{U}}(\mu)^T & Z_{\mathcal{N}}(\mu) \end{bmatrix}. \tag{4.2}$$

Let $\dot{X}(\mu)$ denote the derivative of $X(\mu)$ with respect to the parameter μ, and $\dot{Z}(\mu)$ denote the derivative of $Z(\mu)$. Recently, Halická [14] has established, based on (4.1) and several other key observations made in [20], the analyticity of the central path solutions $(X(\mu), y(\mu), Z(\mu))$ for all $\mu \geq 0$, including the boundary case $\mu = 0$. Formally, we have

Theorem 4.1 *Suppose that the primal and dual have interior solutions and also a strictly complementary solution pair. Then, $(X(\mu), y(\mu), Z(\mu))$ is an analytic function on the doamin $\mu \in [0, \infty)$. In particular, the following limits exist:*

$$\dot{X}(0) := \lim_{\mu \downarrow 0} \dot{X}(\mu), \quad \dot{Z}(0) := \lim_{\mu \downarrow 0} \dot{Z}(\mu).$$

For the case that there is no strictly complementary solution, it is known from Goldfarb and Scheinberg [11] that $\|\dot{X}(\mu)\| \to \infty$ and $\|\dot{Z}(\mu)\| \to \infty$ for $\mu \downarrow 0$. Therefore, the existence of a strictly complementary solution pair is crucial in establishing Theorem 4.1. We will now use the existence of $\dot{X}(0)$ and $\dot{Z}(0)$ to prove that the pseudo projection matrix $Q(D(\mu))$ is bounded under strict complementarity.

Lemma 4.1 *Suppose that the primal and dual have interior solutions and also a strictly complementary solution pair. Then*

$$\{Q(D(\mu)) \mid 0 < \mu \leq 1\}$$

is bounded. Therefore, the angle $\angle(x, D(\mu)A^{\mathrm{T}}y)$ with $Ax = 0$, $x \neq 0$, $y \neq 0$ and $0 < \mu \leq 1$ is bounded below from zero by a positive constant, which is independent of x, y and μ.

Proof. Recall that $X(\mu) \succ 0$ and consequently $D(\mu) \succ 0$ for all $\mu > 0$. Hence, we know from Theorem 3.1 that

$$\min\{\sin \angle(x, Q(D(\mu))s) \mid Ax = 0,\, x \neq 0,\, s \neq 0\} = \frac{1}{\|Q(D(\mu))\|_2}.$$

Since $Q(D(\mu))$ is continuous in μ, this means that we need only to be concerned with the situation when $\mu \downarrow 0$.

Suppose to the contrary that $Q(D(\mu))$ is not bounded for $\mu \downarrow 0$. It then follows from Theorem 3.1 that there exists a sequence $\{\mu_k \mid k = 1, 2, \ldots\}$ with $\mu_k \to 0$ and a sequence

$$s^{(k)} = D(\mu_k)A^{\mathrm{T}}u^{(k)}, \quad k = 1, 2, \ldots,$$

and that the sequence of the corresponding matrices $\{S^{(k)} \mid k = 1, 2, \ldots\}$ is convergent with limit

$$S^{\infty} = \lim_{k \to \infty} S^{(k)} \quad \text{and } As^{\infty} = 0, \quad S^{\infty} \neq 0. \tag{4.3}$$

Since $X(\mu_k)Z(\mu_k) = \mu_k I$, we have $\mathrm{vec}_{\mathcal{S}}\,(Z(\mu_k)S^{(k)}Z(\mu_k)) = \mu_k^2 D(\mu_k)^{-1}s^{(k)} = \mu_k^2 A^{\mathrm{T}}u^{(k)}$, whereas $As^\infty = 0$. This implies the identity

$$S^\infty \bullet (Z(\mu_k)S^{(k)}Z(\mu_k)) = 0 \text{ for } k = 1, 2, \ldots, \tag{4.4}$$

which will play a crucial role throughout the proof. Letting $k \to \infty$ in (4.4) yields

$$0 = S^\infty \bullet (Z^* S^\infty Z^*) = S_{\mathcal{N}}^\infty \bullet (Z_{\mathcal{N}}^* S_{\mathcal{N}}^\infty Z_{\mathcal{N}}^*),$$

implying that $S_{\mathcal{N}}^\infty = 0$.

Since $\|Z(\mu_k) - Z^*\| = O(\mu_k)$ (see (4.1)), we have

$$Z(\mu_k)S^{(k)}Z(\mu_k) = Z^*S^{(k)}Z^* + (Z(\mu_k) - Z^*)S^{(k)}Z^* + Z^*S^{(k)}(Z(\mu_k) - Z^*) + O(\mu_k^2). \tag{4.5}$$

Notice further that $S^\infty \bullet (Z^*S^{(k)}Z^*) = 0$, because $S_{\mathcal{N}}^\infty = 0$. Hence, we get from (4.4) and (4.5) that

$$\begin{aligned} 0 &= \lim_{k\to\infty} \frac{S^\infty \bullet (Z(\mu_k)S^{(k)}Z(\mu_k))}{\mu_k} \\ &= 2 \lim_{k\to\infty} \mathrm{tr}\, \frac{S^\infty (Z(\mu_k) - Z^*)S^{(k)}Z^*}{\mu_k} \\ &= 2\, \mathrm{tr}\, S^\infty \dot{Z}(0) S^\infty Z^*, \end{aligned}$$

where we used Theorem 4.1. As $Z_{\mathcal{B}}^* = 0$, $Z_{\mathcal{U}}^* = 0$ and $S_{\mathcal{N}}^\infty = 0$, we can simplify the above expression to

$$0 = \mathrm{tr}\, S^\infty \dot{Z}(0) S^\infty Z^* = \mathrm{tr}\, (S_{\mathcal{U}}^\infty)^{\mathrm{T}} \dot{Z}_{\mathcal{B}}(0) S_{\mathcal{U}}^\infty Z_{\mathcal{N}}^*.$$

Since $Z_{\mathcal{N}}^* \succ 0$ and $\dot{Z}_{\mathcal{B}}(0) = (X_{\mathcal{B}}^*)^{-1} \succ 0$ we conclude that $S_{\mathcal{U}}^\infty = 0$.

Having established that $S_{\mathcal{U}}^\infty = 0$ and $S_{\mathcal{N}}^\infty = 0$, we obtain from (4.4) that

$$\begin{aligned} 0 &= S_{\mathcal{B}}^\infty \bullet (Z(\mu_k)S^{(k)}Z(\mu_k))_{\mathcal{B}} \\ &= S_{\mathcal{B}}^\infty \bullet (Z_{\mathcal{B}}(\mu_k)S_{\mathcal{B}}^{(k)}Z_{\mathcal{B}}(\mu_k) + Z_{\mathcal{U}}(\mu_k)S_{\mathcal{N}}^{(k)}Z_{\mathcal{U}}(\mu_k)^{\mathrm{T}}) + 2\mathrm{tr}\, S_{\mathcal{B}}^\infty Z_{\mathcal{B}}(\mu_k)S_{\mathcal{U}}^{(k)}Z_{\mathcal{U}}(\mu_k)^{\mathrm{T}}. \end{aligned} \tag{4.6}$$

Now divide (4.6) by μ_k^2 on both sides and let $k \to \infty$. Noticing that $S_{\mathcal{U}}^{(k)} \to 0$ as $k \to \infty$, we obtain

$$0 = S_{\mathcal{B}}^\infty \bullet \dot{Z}_{\mathcal{B}}^* S_{\mathcal{B}}^\infty \dot{Z}_{\mathcal{B}}^*$$

and so $S_{\mathcal{B}}^\infty = 0$. This finally implies that $S^\infty = 0$, which contradicts (4.3). The theorem is thus proven. **Q.E.D.**

Below, we show that the pseudo projection $Q(D(\mu))$ is not only bounded, but in fact it converges to a certain matrix Q^* if $\mu \downarrow 0$.

Theorem 4.2 *Suppose that the primal and dual semidefinite programs* (P) *and* (D) *have interior solutions and also have a strictly complementary solution pair. Then the following limit exists:*

$$Q^* := \lim_{\mu \downarrow 0} Q(D(\mu)).$$

Proof. From Lemma 4.1 we know that $Q(\mu)$ is bounded, and hence it has a cluster point for $\mu \to 0$. Let $\bar{Q}$ and $\tilde{Q}$ denote two such cluster points. Let $r \in \Re^N$ be arbitrary, and define

$$w(\mu) := Q(\mu)r, \quad \bar{w} := \bar{Q}r, \quad \tilde{w} := \tilde{Q}r.$$

Notice that $\bar{w}$ and $\tilde{w}$ are both cluster points of $w(\mu)$ for $\mu \to 0$. We will prove the theorem by showing that $\bar{w} = \tilde{w}$, irrespective of the choice of r, $\bar{Q}$ and $\tilde{Q}$.

Recall from Theorem 3.1 that $A\bar{Q} = A\tilde{Q} = A$, and hence

$$A(\bar{w} - \tilde{w}) = 0. \tag{4.7}$$

As $X(\mu)Z(\mu) = \mu I$, we also have

$$\mathrm{vec}_{\mathcal{S}}\,(Z(\mu)W(\mu)Z(\mu)) = \mu^2 D(\mu)^{-1}Q(\mu)r = \mu^2 A^{\mathrm{T}}(AD(\mu)A^{\mathrm{T}})^{-1}Ar$$

which together with (4.7) implies that

$$(\bar{W} - \tilde{W}) \bullet (Z(\mu)W(\mu)Z(\mu)) = 0 \text{ for all } \mu > 0. \tag{4.8}$$

Letting $\mu \to 0$, we further obtain that

$$0 = (\bar{W} - \tilde{W}) \bullet (Z^*\bar{W}Z^*) = (\bar{W} - \tilde{W}) \bullet (Z^*\tilde{W}Z^*),$$

and thus

$$\mathrm{tr}\,(\bar{W} - \tilde{W})Z^*(\bar{W} - \tilde{W})Z^* = 0.$$

Hence, $(\bar{W} - \tilde{W})_{\mathcal{N}} = 0$.

Similar as in (4.5), we have

$$Z(\mu)W(\mu)Z(\mu) = Z^*W(\mu)Z^* + (Z(\mu) - Z^*)W(\mu)Z^* + Z^*W(\mu)(Z(\mu) - Z^*) + O(\mu^2),$$

and $(\bar{W} - \tilde{W}) \bullet (Z^*W(\mu)Z^*) = (\bar{W} - \tilde{W})_{\mathcal{N}} \bullet (Z^*W(\mu)Z^*)_{\mathcal{N}} = 0$ because $(\bar{W} - \tilde{W})_{\mathcal{N}} = 0$. Thus, dividing in (4.8) by μ and letting $\mu \downarrow 0$, we obtain that

$$0 = 2\,\mathrm{tr}\,(\bar{W} - \tilde{W})\dot{Z}(0)\bar{W}Z^* = 2\,\mathrm{tr}\,(\bar{W} - \tilde{W})\dot{Z}(0)\tilde{W}Z^*,$$

and so

$$0 = \mathrm{tr}\,(\bar{W} - \tilde{W})\dot{Z}(0)(\bar{W} - \tilde{W})Z^*.$$

As $Z^*_{\mathcal{B}} = 0$, $Z^*_{\mathcal{U}} = 0$ and $(\bar{W} - \tilde{W})_{\mathcal{N}} = 0$, we can simplify the above expression to

$$0 = \mathrm{tr}\,(\bar{W} - \tilde{W})\dot{Z}(0)(\bar{W} - \tilde{W})Z^* = \mathrm{tr}\,(\bar{W} - \tilde{W})_{\mathcal{U}}^{\mathrm{T}}\dot{Z}_{\mathcal{B}}(0)(\bar{W} - \tilde{W})_{\mathcal{U}}Z^*_{\mathcal{N}}.$$

Since $Z^*_{\mathcal{N}} \succ 0$ and $\dot{Z}_{\mathcal{B}}(0) = (X^*_{\mathcal{B}})^{-1} \succ 0$, this implies that $(\bar{W} - \tilde{W})_{\mathcal{U}} = 0$.

Having established that $\bar{W}_{\mathcal{U}} = \tilde{W}_{\mathcal{U}}$ and $\bar{W}_{\mathcal{N}} = \tilde{W}_{\mathcal{N}}$, we obtain from (4.8) that

$$\begin{aligned} 0 &= (\bar{W} - \tilde{W})_{\mathcal{B}} \bullet (Z(\mu)W(\mu)Z(\mu))_{\mathcal{B}} \\ &= (\bar{W} - \tilde{W})_{\mathcal{B}} \bullet (Z_{\mathcal{B}}(\mu)W_{\mathcal{B}}(\mu)Z_{\mathcal{B}}(\mu) + Z_{\mathcal{U}}(\mu)W_{\mathcal{N}}(\mu)Z_{\mathcal{U}}(\mu)^{\mathrm{T}}) \\ &\quad +2\,\mathrm{tr}\,(\bar{W} - \tilde{W})_{\mathcal{B}}Z_{\mathcal{B}}(\mu)W_{\mathcal{U}}(\mu)Z_{\mathcal{U}}(\mu)^{\mathrm{T}}. \end{aligned} \tag{4.9}$$

Dividing (4.9) by μ^2 on the both sides and letting $\mu \downarrow 0$ yields

$$\begin{cases} 0 = (\bar{W} - \tilde{W})_{\mathcal{B}} \bullet (\dot{Z}_{\mathcal{B}}(0)\bar{W}_{\mathcal{B}}\dot{Z}_{\mathcal{B}}(0) + \dot{Z}_{\mathcal{U}}(0)\bar{W}_{\mathcal{N}}\dot{Z}_{\mathcal{U}}(0)^{\mathrm{T}}) + 2\,\mathrm{tr}\,(\bar{W} - \tilde{W})_{\mathcal{B}}\dot{Z}_{\mathcal{B}}(0)\bar{W}_{\mathcal{U}}\dot{Z}_{\mathcal{U}}(0)^{\mathrm{T}} \\ 0 = (\bar{W} - \tilde{W})_{\mathcal{B}} \bullet (\dot{Z}_{\mathcal{B}}(0)\tilde{W}_{\mathcal{B}}\dot{Z}_{\mathcal{B}}(0) + \dot{Z}_{\mathcal{U}}(0)\bar{W}_{\mathcal{N}}\dot{Z}_{\mathcal{U}}(0)^{\mathrm{T}}) + 2\,\mathrm{tr}\,(\bar{W} - \tilde{W})_{\mathcal{B}}\dot{Z}_{\mathcal{B}}(0)\bar{W}_{\mathcal{U}}\dot{Z}_{\mathcal{U}}(0)^{\mathrm{T}}, \end{cases}$$

where we used that $\bar{W}_{\mathcal{U}} = \tilde{W}_{\mathcal{U}}$ and $\bar{W}_{\mathcal{N}} = \tilde{W}_{\mathcal{N}}$. Substracting the above two identities from each other yields

$$0 = (\bar{W} - \tilde{W})_{\mathcal{B}} \bullet \dot{Z}_{\mathcal{B}}(0)(\bar{W} - \tilde{W})_{\mathcal{B}}\dot{Z}_{\mathcal{B}}(0),$$

and so $(\bar{W} - \tilde{W})_{\mathcal{B}} = 0$. This finally implies that $\bar{W} = \tilde{W}$ and therefore $\bar{Q} = \tilde{Q}$, since $r \in \Re^N$ is arbitrary. Hence, $\bar{Q} = \tilde{Q} = \lim_{\mu\downarrow 0} Q(\mu)$. **Q.E.D.**

We remark that similar to (3.3)–(3.5), it holds that

$$A\dot{x}(\mu) = 0 \tag{4.10}$$

$$A^{\mathrm{T}}\dot{y}(\mu) + \dot{z}(\mu) = 0 \tag{4.11}$$

$$D(\mu)\dot{z}(\mu) + \mu\dot{x}(\mu) = x(\mu). \tag{4.12}$$

Thus, we easily arrive at the identities

$$\dot{x}(\mu) = \frac{1}{\mu}\left(x(\mu) - D(\mu)A^{\mathrm{T}}(AD(\mu)A^{\mathrm{T}})^{-1}b\right), \quad \dot{z}(\mu) = A^{\mathrm{T}}(AD(\mu)A^{\mathrm{T}})^{-1}b. \tag{4.13}$$

Recalling the definition of $Q(D)$ and using the fact that $Ax(\mu) = b$, we further get

$$\dot{x}(\mu) = \frac{1}{\mu}\left(x(\mu) - Q(D(\mu))A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}b\right). \tag{4.14}$$

Similarly, since $D(\mu)z(\mu) = \mu x(\mu)$, $z(\mu) = c - A^{\mathrm{T}}y(\mu)$ and $AQ(D(\mu)) = A$,

$$\dot{z}(\mu) = \frac{1}{\mu}Q(D(\mu))^{\mathrm{T}}z(\mu) = \frac{1}{\mu}\left(Q(D(\mu))^{\mathrm{T}}c - A^{\mathrm{T}}y(\mu)\right) = \frac{1}{\mu}\left(z(\mu) - [I - Q(D(\mu))]^{\mathrm{T}}c\right). \tag{4.15}$$

Since $\dot{X}(\mu)$ and $\dot{Z}(\mu)$ are bounded, it follows from (4.14) and (4.15) that

$$\|x(\mu) - Q(D(\mu))A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}b\| = O(\mu), \quad \|z(\mu) - [I - Q(D(\mu))]^{\mathrm{T}}c\| = O(\mu). \tag{4.16}$$

The solutions $x(\mu)$ and $z(\mu)$ are interior solutions with a positive duality gap $x(\mu)^{\mathrm{T}}z(\mu) = n\mu$. On the other hand, the solutions $\hat{x}(\mu) := Q(D(\mu))A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}b$ and $\hat{z}(\mu) := [I - Q(D(\mu))]^{\mathrm{T}}c$ satisfy the linear constraints of the primal and the dual problems respectively. Moreover, they also satisfy the optimality condition $(\hat{x}(\mu))^{\mathrm{T}}\hat{z}(\mu) = 0$, but may violate the semidefiniteness condition by an amount less than $\mu\|\dot{X}(\mu)\|$ and $\mu\|\dot{Z}(\mu)\|$ respectively.

Letting $\mu \to 0$ in (4.16) yields the following result.

Theorem 4.3 *Suppose that the primal and dual have interior solutions and also a strictly complementary solution pair. Then*

$$x^* = Q^*A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}b, \quad z^* = (I - Q^*)^{\mathrm{T}}c.$$

5 Discussion

We have derived a neat formula for the central optimal solution, viz. $x^*(b) = Q^*(b)x_{\mathrm{LS}}(b)$, where $x_{\mathrm{LS}}(b) := A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}b$ is the least squares solution to "$Ax = b$". This identity holds true if (P_b) and (D_b) have a strictly complementary solution pair and satisfy Slater's condition; see Theorem 4.3. Since the condition number $\kappa(A)$ of the matrix A measures the sensitivity of $x_{\mathrm{LS}}(b)$ with respect to perturbations in b, it is tempting to use $\|Q^*(b)\|_2\kappa(A)$ as a measure of sensitivity in $x^*(b) = Q^*(b)x_{\mathrm{LS}}(b)$; see also (3.11). Moreover, $\|Q^*(b)\|_2$ has an intuitively attractive geometric interpretation; see Theorem 3.1. Also, (3.10) implies that

$$\lim_{\mu \downarrow 0} \frac{\partial x(b,\mu)}{\partial b^{\mathrm{T}}} = Q^*(b)A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}.$$

However, the central optimal solution $x^*(b)$ is obviously not everywhere differentiable in b, and the matrix $Q^*(b)$ is not everywhere continuous in b, even in the case of linear programming. The mere existence of $Q^*(b)$ is therefore not sufficient to bound $\|Q^*(b)\|$ on a compact set of right hand side parameters. In fact, by the examples in the Appendix we see that $x^*(b)$ may even have discontinuities. Furthermore, $Q^*(\cdot)$ may fail to exist for arbitrarily small perturbations in b in which case strict complementarity must fail for such perturbations (see Theorem 4.2). The examples in the appendix show that this is in fact possible. This means that the issues involved in sensitivity analysis for semidefinite programming are far more delicate than in its linear programming counterpart. It remains a challenge to design theoretically sound methods for sensitivity and parametric analysis for the general SDP.

Acknowledgment: We would like to thank the associate editor and anonymous referees for their careful reading and valuable comments which helped to improve the presentation of the paper.

References

[1] Adler, I., and Monteiro, R.D.C., *Limiting behavior of the affine scaling continuous trajectories for linear programming problems*, Mathematical Programming 50 (1991) 29–51.

[2] Alizadeh, F., Haeberly, J.A. and Overton, M., *Primal-dual interior point metods for semidefinite programming: convergence rates, stability and numerical results*, SIAM Journal on Optimization 8 (1998) 746-768.

[3] Ben-Tal, A. and Bendsoe, M.P., *A new method for optimal truss topology design*, SIAM Journal on Optimization 3 (1993) 322-358.

[4] Ben-Tal, A. and Nemirovsky, A., *Optimal design of engineering structures*, Optima 47 (1995) 4-9.

[5] Bonnans, J.F. and Shapiro, A., *Optimization problems with perturbations, a guided tour*, SIAM Review 40 (1998) 2, 228–264.

[6] Boyd, S., L. El Ghaoui, Feron, E. and Balakrishnan, V., Linear matrix inequalities in system and control theory, Studies in Applied Mathematics, vol. 15, SIAM, Philadelphia, U.S.A., 1994.

[7] De Klerk, E, *Interior point methods for semidefinite programming*, Ph.D thesis, Delft University of Technology, The Netherlands, 1997.

[8] Dikin, I.I., *Iterative solution of problems of linear and quadratic programming*, Soviet Mathematics Doklady 8 (1967) 674-675.

[9] Forsgren, A., *On linear least-squares problems with diagonally dominant weight matrices*, SIMAX 17 (1996) 763–788.

[10] Goemans, M.X. and Williamson, D.P., *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*, Journal ACM 42 (1995) 1115-1145.

[11] Goldfarb, D. and Scheinberg, K., *Interior point trajectories in semidefinite programming*, SIAM Journal on Optimization 8 (1998) 871–886.

[12] Goldfarb, D. and Scheinberg, K., *On parametric semidefinite programming*, Applied Numerical Mathematics 29 (1999) 361–377.

[13] Grötschel, M., Lovász, L. and Schrijver, A., Geometric algorithms and combinatorial optimization, Springer-Verlag, New York, 1988.

[14] Halická, M., *Analyticity of the central path at the boundary point in semidefinite programming*, Working Paper, Department of Applied Mathematics, Comenius University, Bratislava, Slovakia, 2000. Paper presented at *Workshop IPM 2000*, Budapest, July 2000.

[15] Helmberg, C., *Fixing variables in semidefinite relaxations*, Preprint SC 96–43, Konrad–Zuse–Zentrum für Informationstechnik Berlin, Germany, 1996.

[16] Holder, A., Sturm, J.F. and Zhang, S., *Analytic central path, sensitivity analysis and parametric linear programming*, CCM Report 118, Center for Computational Mathematics, Mathematics Department, University of Colorado at Denver, USA, 1997.

[17] Horn, R.A. and Johnson, C.R., Matrix analysis, Cambridge University Press, Cambridge, New York, USA, 1985.

[18] Horn, R.A. and Johnson, C.R., Topics in matrix analysis, Combridge University Press, Cambridge, New York, USA, 1991.

[19] Kojima, M., Mizuno, S. and Noma, T., *Limiting behavior of trajectories by a continuation method for monotone complementarity problems*, Mathematics of Operations Research 15 (1990) 662–675.

[20] Luo, Z.-Q., Sturm, J.F. and Zhang, S., *Superlinear convergence of a symmetric primal-dual path following algorithm for semidefinite programming*, SIAM Journal on Optimization 8 (1998) 59–81.

[21] Luo, Z.-Q., Sturm, J.F. and Zhang, S., *Duality results for conic convex programming*, Report 9719/A, Econometric Institute, Erasmus University Rotterdam, The Netherlands, 1997.

[22] McLinden, L., *The complementarity problem for maximal monotone multifunctions.* In Cottle, R.W., Giannessi, F. and Lions, J.L., eds., *Varitional Inequalities and Complementarity Problems*, John Wiley and Sons, New York, pp. 251–270, 1980.

[23] Megiddo, N., *Pathways to the optimal set in linear programming.* In Megiddo, N. ed., *Progress in Mathematical Programming: Interior Point and Related Methods*, Springer Verlag, New York, pp. 131–158, 1989.

[24] Monteiro, R.D.C. and Tsuchiya, T., *Limiting behavior of the derivatives of certain trajectories associated with a montone horizontal linear complementarity problem*, Mathematics of Operations Research 21 (1996) 793–814.

[25] Nunez, M.A. and Freund, R.M., *Condition measures and properties of the central trajectory of a linear program*, working paper, 1996.

[26] Nesterov, Y. and Nemirovskii, A., Interior point polynomial methods in convex programming, Studies in Applied Mathematics, vol. 13, SIAM, Philadelphia, 1994.

[27] Nesterov, Y. and Todd, M.J., *Self–scaled barriers and interior point methods for convex programming*, Mathematics of Operations Research 22 (1997) 1–42.

[28] Pataki, G., *A simplex generalization for semidefinite programming*, talk presented at INFORMS Dallas Meeting, U.S.A., 1997.

[29] Renegar, J., *Linear programming, complexity theory and elementary functional analysis*, Mathematical Programming 70 (1995) 279–351.

[30] Rockafellar, R.T., Conjugate duality and optimization, Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, U.S.A., 1974.

[31] Stewart, G.W., *On scaled projections and pseudoinverses*, Linear Algebraa and Its Applications 112 (1989) 189-193.

[32] Sturm, J.F., *Primal-dual interior point approach to semidefinite programming*, Ph.D thesis, Tinbergen Institute Series 156, Erasmus University, The Netherlands, 1997.

[33] Todd, M.J., *A Dantzig-Wolfe-like variant of Karmarkar's interior-point linear programming algorithm*, Operations Research 38 (1990) 1006-1018.

[34] Vavasis, S.A. and Ye, Y., *A primal-dual interior point method whose running time depends only on the constraint matrix*, Mathematical Programming 74 (1996) 79-120.

[35] Vandenberghe, L. and Boyd, S., *Semidefinite programming*, SIAM Review 38 (1996) 49-95.

[36] Vandenberghe, L., Boyd, S. and Wu, S.-P., *Determinant maximization with linear matrix inequality constraints*, SIMAX 19 (1998) 499-533.

[37] Witzgall, C., Boggs, P.T. and Domich, P.D., *On the convergence behavior of trajectories for linear programming*. Lagarias, J.C. and Todd, M.J. eds., *Mathematical Developments Arising from Linear Programming: Proceedings of a Joint Summer Research Conference held at Bowdoin College, Brunswick, Maine, U.S.A., June/July 1988*, volume 114 of *Contemporary Mathematics*, American Mathematical Society, USA, pp. 161–187, 1990.

[38] Wolkowicz, H., *Applications of semidefinite programming (linear programming for the 90's and 00's)*, talk presented at Workshop on High Performance Optimization Techniques II, Rotterdam, The Netherlands, 1997.

Appendix

Example 1 *Consider the following SDP instance:*

$$\begin{array}{lll} (\bar{P}) & \textit{minimize} & C \bullet X \\ & \textit{subject to} & A[i] \bullet X = b_i, \quad \textit{for } i = 1,2 \\ & & X \in \mathcal{S}^{3\times 3} \textit{ and } X \succeq 0 \end{array}$$

where

$$C = \begin{bmatrix} 1 & 0 & 0 \\ & 0 & 0 \\ & & 0 \end{bmatrix}, \quad A[1] = \begin{bmatrix} 0 & 0 & -1/2 \\ & 1 & 0 \\ & & 0 \end{bmatrix},$$

$$A[2] = \begin{bmatrix} 0 & -1/2 & 0 \\ & 0 & 0 \\ & & 1 \end{bmatrix}.$$

Since all matrices are symmetric, we only indicate the upper-triangle part of the above matrices.

In fact, $(\bar{P})$ can also be explicitly written as

$$\begin{array}{lll} (\bar{P}) & \text{minimize} & x_{11} \\ & \text{subject to} & \begin{bmatrix} x_{11} & x_{33} - b_2 & x_{22} - b_1 \\ & x_{22} & x_{23} \\ & & x_{33} \end{bmatrix} \succeq 0. \end{array}$$

The dual of $(\bar{P})$ is

$$\begin{array}{lll} (\bar{D}) & \text{maximize} & b_1 y_1 + b_2 y_2 \\ & \text{subject to} & y_1 A[1] + y_2 A[2] \preceq C, \end{array}$$

which is equivalent to:

$$\begin{array}{lll} (\bar{D}) & \text{maximize} & b_1 y_1 + b_2 y_2 \\ & \text{subject to} & Z = \begin{bmatrix} 1 & y_2/2 & y_1/2 \\ & -y_1 & 0 \\ & & -y_2 \end{bmatrix} \succeq 0. \end{array}$$

First of all, it is obvious that for any values of b_1, b_2, the problems $(\bar{P})$ and $(\bar{D})$ always satisfy the Slater condition.

Moreover, if b_1 and b_2 are nonnegative, then the optimal value is zero, and any optimal solution to $(\bar{P})$ is of the form

$$X^* = \begin{bmatrix} 0 & 0 & 0 \\ & b_1 & x_{23} \\ & & b_2 \end{bmatrix}, \quad x_{23}^2 \le b_1 b_2.$$

If $b_1 > 0$, then y_1^* is zero for any optimal solution to $(\bar{D})$, and by the positive semi-definiteness of Z this implies that also $y_2^* = 0$. Hence

$$Z^* = \begin{bmatrix} 1 & 0 & 0 \\ & 0 & 0 \\ & & 0 \end{bmatrix} \text{ if } b_1 > 0,\ b_2 \geq 0.$$

We have the following cases:

1. If $b_1 = b_2 = 0$, then there is a strictly complementary solution, e.g.

$$X^* = 0,\ y_1^* = y_2^* = -1,\ Z^* = \begin{bmatrix} 1 & -1/2 & -1/2 \\ & 1 & 0 \\ & & 1 \end{bmatrix}.$$

2. If $b_1 > 0$ and $b_2 = 0$, then there is a unique optimal solution pair, and this pair violates the condition of strict complementarity, viz.

$$X^* = \begin{bmatrix} 0 & 0 & 0 \\ & b_1 & 0 \\ & & 0 \end{bmatrix},\ y_1^* = y_2^* = 0,\ Z^* = \begin{bmatrix} 1 & 0 & 0 \\ & 0 & 0 \\ & & 0 \end{bmatrix}.$$

3. If $b_1 > 0$ and $b_2 > 0$, then there is a strictly complementary solution pair, e.g.

$$X^* = \begin{bmatrix} 0 & 0 & 0 \\ & b_1 & 0 \\ & & b_2 \end{bmatrix},\ y_1^* = y_2^* = 0,\ Z^* = \begin{bmatrix} 1 & 0 & 0 \\ & 0 & 0 \\ & & 0 \end{bmatrix}.$$

Nevertheless, in our example the Slater condition always holds for any value of b_1 and b_2. Therefore, both the primal and the dual central optimal solutions exist. In fact, the primal central optimal solution happens to be Lipschitz continuous in b for the above example. In general however, this is not the case.

Example 2 *Consider the following SDP instance:*

$$\begin{array}{lll} (\tilde{P}) & \textit{minimize} & C \bullet X \\ & \textit{subject to} & A[i] \bullet X = b_i, \quad \textit{for } i = 1,2,3,4 \\ & & X \in \mathcal{S}^{4\times 4} \textit{ and } X \succeq 0 \end{array}$$

where

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 \\ & & 0 & 0 \\ & & & 1 \end{bmatrix},\quad A[1] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ & 0 & 0 & 0 \\ & & 0 & 0 \\ & & & 0 \end{bmatrix},\quad A[2] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ & 1 & 0 & 0 \\ & & 0 & 0 \\ & & & 0 \end{bmatrix},$$

$$A[3] = \begin{bmatrix} 0 & -1/2 & 0 & 0 \\ & 0 & 0 & 0 \\ & & 1 & 0 \\ & & & 0 \end{bmatrix}, \quad A[4] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 \\ & & 0 & 1/2 \\ & & & 0 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 1 \\ 0 \\ b_4 \end{bmatrix}.$$

In fact, $(\tilde{P})$ can also be explicitly written as

$$\begin{array}{lll} (\tilde{P}) & \text{minimize} & x_{44} \\ & \text{subject to} & \begin{bmatrix} 1 & x_{33} & x_{13} & x_{14} \\ & 1 & x_{23} & x_{24} \\ & & x_{33} & b_4 \\ & & & x_{44} \end{bmatrix} \succeq 0. \end{array}$$

The dual of $(\tilde{P})$ is

$$\begin{array}{lll} (\tilde{D}) & \text{maximize} & y_1 + y_2 + b_4 y_4 \\ & \text{subject to} & y_1 A[1] + y_2 A[2] + y_3 A[3] + y_4 A[4] \preceq C, \end{array}$$

which is equivalent to:

$$\begin{array}{lll} (\tilde{D}) & \text{maximize} & y_1 + y_2 + b_4 y_4 \\ & \text{subject to} & Z = \begin{bmatrix} -y_1 & y_3/2 & 0 & 0 \\ & -y_2 & 0 & 0 \\ & & -y_3 & -y_4/2 \\ & & & 1 \end{bmatrix} \succeq 0. \end{array}$$

First of all, it is obvious that for any $b_4 \in \Re$, the problems $(\tilde{P})$ and $(\tilde{D})$ always satisfy the Slater condition.

If $b_4 = 0$, then the optimal value is zero, and any primal optimal solution is of the form:

$$X^* = \begin{bmatrix} 1 & x_{33} & x_{13} & 0 \\ & 1 & x_{23} & 0 \\ & & x_{33} & 0 \\ & & & 0 \end{bmatrix}.$$

The central optimal solution is found by maximizing the determinant of the principal submatrix formed by the first 3 rows and columns of X^*, thus

$$X^*(0) = \begin{bmatrix} 1 & \sqrt{1/3} & 0 & 0 \\ & 1 & 0 & 0 \\ & & \sqrt{1/3} & 0 \\ & & & 0 \end{bmatrix}.$$

The dual optimal solution is unique:

$$Z^*(0) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 \\ & & 0 & 0 \\ & & & 1 \end{bmatrix}.$$

Suppose now that $b_4 \neq 0$. Then the optimal value is b_4^2 and

$$X^*(b_4) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ & 1 & 0 & 0 \\ & & 1 & b_4 \\ & & & b_4^2 \end{bmatrix}, \quad Z^*(b_4) = \begin{bmatrix} b_4^2/2 & -b_4^2/2 & 0 & 0 \\ & b_4^2/2 & 0 & 0 \\ & & b_4^2 & -b_4 \\ & & & 1 \end{bmatrix}.$$