ISRN KTH/OPT SYST/FR 02/12 SE
Coden: TRITA/MAT-02-OS12
ISSN 1401-2294

Characterization of the limit point of the central path in semidefinite programming

by

Göran Sporre and Anders Forsgren

Optimization and Systems Theory

DEPARTMENT OF MATHEMATICS
ROYAL INSTITUTE OF TECHNOLOGY
SE-100 44 STOCKHOLM, SWEDEN

ISSN 1401-2294

CHARACTERIZATION OF THE LIMIT POINT OF THE CENTRAL PATH IN SEMIDEFINITE PROGRAMMING

Göran SPORRE* and Anders FORSGREN†

Technical Report TRITA-MAT-2002-OS12
Department of Mathematics
Royal Institute of Technology
June 2002

Abstract

In linear programming, the central path is known to converge to the analytic center of the set of optimal solutions. Recently, it has been shown that this is not necessarily true for linear semidefinite programming in the absence of strict complementarity. The present paper deals with the formulation of a convex problem whose solution defines the limit point of the central path. This problem is closely related to the analytic center problem for the set of optimal solutions. In the strict complementarity case the problems are shown to coincide.

Key words. Semidefinite programming, interior method, primal-dual interior method, central path, analytic center.

AMS subject classifications. 90C22, 90C51, 65K05.

1. Introduction

We consider the linear semidefinite programming problem

$$\begin{array}{ll} \underset{X \in \mathcal{S}_+^n}{\text{minimize}} & \operatorname{trace}(CX) \\ \text{subject to} & \operatorname{trace}(A^{(i)}X) = b_i, \quad i = 1, \dots, m, \end{array} \tag{1.1}$$

where $\mathcal{S}_+^n$ is the set of symmetric positive semidefinite $n \times n$ matrices. Similarly, $\mathcal{S}_{++}^n$ denotes the set of symmetric positive definite $n \times n$ matrices. In (1.1), $C, A^{(1)}, \dots, A^{(m)}$ are symmetric $n \times n$ matrices and $b \in I\!R^m$. Throughout, for convenience, we make the following linear independence assumption.

Assumption 1. *The matrices $A^{(1)}, \dots, A^{(m)}$ are linearly independent.*

*Optimization and Systems Theory, Department of Mathematics, Royal Institute of Technology, SE-100 44 Stockholm, Sweden (goran.sporre@math.kth.se). Research supported by the Swedish Research Council (VR).

†Optimization and Systems Theory, Department of Mathematics, Royal Institute of Technology, SE-100 44 Stockholm, Sweden (anders.forsgren@math.kth.se). Research supported by the Swedish Research Council (VR).

For general discussions on linear semidefinite programming and examples of applications, see, e.g., Vandenberghe and Boyd [22] and Todd [21]. For a comprehensive discussion on various aspects of semidefinite programming, see Wolkowicz, Saigal and Vandenberghe [23]. The standard dual of (1.1) can be formulated as

$$\begin{array}{ll} \underset{y \in I\!R^m,\, S \in \mathcal{S}_+^n}{\text{maximize}} & b^T y \\ \text{subject to} & \displaystyle\sum_{i=1}^m y_i A^{(i)} + S = C, \quad i = 1, \dots, m. \end{array} \tag{1.2}$$

Unlike linear programming, there is in general no strong duality theorem for the pair (1.1) and (1.2). However, throughout this paper we make the following assumption which guarantees that there is no duality gap; see Rockafellar [20, Thm. 31.4].

Assumption 2. *There exists a triple* $(X^0, y^0, S^0) \in \mathcal{S}_{++}^n \times I\!R^m \times \mathcal{S}_{++}^n$ *such that* $\operatorname{trace}(A^{(i)} X^0) = b_i$, $i = 1, \dots, m$, *and* $\sum_{i=1}^m y_i^0 A^{(i)} + S^0 = C$.

It is possible to formulate a duality theory with nice complexity properties and with no duality gap without any assumption like Assumption 2, see Ramana [19]. However, due to the size of the dual it is generally not tractable for computational purposes.

The primal barrier trajectory, $X(\mu)$, associated with (1.1) is given by the symmetric positive definite solution to the problem

$$\begin{array}{lll} \underset{X \in \mathcal{S}_{++}^n}{\text{minimize}} & \operatorname{trace}(CX) - \mu \ln \det X & \\ \text{subject to} & \operatorname{trace}(A^{(i)} X) = b_i, & i = 1, \dots, m, \end{array} \tag{1.3}$$

for positive values of μ. Since (1.3) has a strictly convex objective, see, e.g., Monteiro and Todd [13, p. 273], and a convex feasible region, the optimal solution is unique given that it exists. We return to the issue of existence after formulating the so-called primal-dual equations below.

The first-order optimality conditions for (1.3) state that there exists a vector $y(\mu) \in I\!R^m$ such that $X(\mu$ and $y(\mu)$ satisfy

$$\begin{array}{rcll} C - \mu X(\mu)^{-1} & = & \sum_{i=1}^m y_i(\mu) A^{(i)}, & \\ \operatorname{trace}(A^{(i)} X(\mu)) & = & b_i, & i = 1, \dots, m, \\ X(\mu) & \in & \mathcal{S}_{++}^n. & \end{array} \tag{1.4}$$

Letting $S(\mu) = \mu X(\mu)^{-1}$, this can alternatively be written as

$$S(\mu) + \sum_{i=1}^m y_i(\mu) A^{(i)} = C, \tag{1.5a}$$

$$\operatorname{trace}(A^{(i)} X(\mu)) = b_i, \qquad i = 1, \dots, m, \tag{1.5b}$$

$$X(\mu) S(\mu) = \mu I, \tag{1.5c}$$

$$X(\mu), S(\mu) \in \mathcal{S}_{++}^n. \tag{1.5d}$$

Equations (1.5a)–(1.5c) are known as the primal-dual equations in semidefinite programming and the set of unique solutions $(X(\mu), y(\mu), S(\mu))$ to (1.5) for positive μ

is known as the central path. For a proof of existence and uniqueness of the central path, see, e.g., Kojima, Shindoh and Hara [11, Thm 3.1]. Note that (1.5a) along with $S(\mu) \in \mathcal{S}^n_{++}$ implies that $y(\mu), S(\mu)$ is feasible with respect to (1.2). Also, note that (1.5) may be viewed as a perturbation of the optimality conditions, since these are obtained by setting $\mu = 0$ in (1.5c) and replacing $X(\mu), S(\mu) \in \mathcal{S}^n_{++}$ with $X(\mu), S(\mu) \in \mathcal{S}^n_{+}$.

The central path is a key concept in many so-called interior methods for solving linear semidefinite programs, see, e.g., Helmberg *et al.* [10], Monteiro [14], Alizadeh, Haeberly and Overton [3], Potra and Sheng [18], Zhang [24], Monteiro and Zhang [16], Monteiro and Tsuchiya [15], and de Klerk *et al.* [5]. See also Kojima, Shindoh and Hara [11] who treat the more general case of monotone linear semidefinite complementarity problems. Much of the work done on semidefinite programming in recent years has been inspired by the work of Nesterov and Nemirovskii which gives a foundation for polynomial time interior methods for convex programming; see Nesterov and Nemirovskii [17] and in particular Chap. 6.4 which treats semidefinite programming.

The central path in linear semidefinite programming converges as $\mu \to 0^+$; see Halická, de Klerk and Roos [9, Thm A.3]. See also Graña Drummond and Peterzil [8] for a more general result concerning convergence of the central path for a class of convex semidefinite programs including linear semidefinite programming. Let the limit point of the central path be denoted by $(\bar{X}, \bar{y}, \bar{S})$ and let $p = \operatorname{rank} \bar{X}$ and $q = n - \operatorname{rank} \bar{S}$. From (1.5c) and (1.5d) it is clear that $\bar{X}\bar{S} = \bar{S}\bar{X} = 0$. Therefore, there is an orthonormal $n \times n$ matrix Q, partitioned as

$$Q = \begin{pmatrix} Q_1 & Q_2 & Q_3 \end{pmatrix}$$

with $Q_1 \in I\!R^{n\times p}$, $Q_2 \in I\!R^{n\times(q-p)}$ and $Q_3 \in I\!R^{n\times(n-q)}$, such that

$$Q^T\bar{X}Q = \begin{pmatrix} \bar{X}_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad Q^T\bar{S}Q = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \bar{S}_{33} \end{pmatrix}, \tag{1.6}$$

where $\bar{X}_{11}$ and $\bar{S}_{33}$ are positive definite and diagonal matrices. Note that Q is associated with $(\bar{X}, \bar{y}, \bar{S})$, but this dependency is omitted for notational convenience.

Let q_i denote the ith column of Q. In consistency with the partition of (1.6), for any symmetric $n \times n$ matrix Y we use the notation

$$\begin{aligned} Y_{ij} &= Q_i^T Y Q_j, \quad i, j \in \{1, 2, 3\} \quad \text{and} \\ y_{ij} &= q_i^T Y q_j, \quad\;\; i, j \in \{1, \dots, n\}. \end{aligned} \tag{1.7}$$

This notation is used to avoid formulae cluttered with the rotation induced by Q. We feel that the main ideas of this paper are easiest understood if Q is thought of as being the identity matrix. This means that no rotation is needed to bring $\bar{X}$ and $\bar{S}$ to the form (1.6). Further, Y_{ij} and y_{ij} can then be interpreted in the usual way, i.e., as block (i, j) and component (i, j) of Y. Therefore, we *recommend* the reader to adopt this view. The example in Section 1.2 and 3 has $Q = I$. A suitable Q can be found by letting the columns of Q_1 and Q_3 be the normalized eigenvectors corresponding to nonzero eigenvalues of $\bar{X}$ and $\bar{S}$ respectively and letting the columns of Q_2 be any orthonormal basis for the null space of $(\bar{X}\ \bar{S})^T$.

It is straightforward to verify that $\bar{X}$ is an optimal solution to (1.1) and that $\bar{y}, \bar{S}$ is an optimal solution to (1.2). In fact, it can be shown that for any optimal solution, $X_{ij} = 0$ for all i and j except $i = j = 1$. Similarly, for any dual optimal solution, $S_{ij} = 0$ for all i and j except $i = j = 3$. These facts rely on $(\bar{X}, \bar{y}, \bar{S})$ being maximally complementary; see de Klerk [4, Chap. 2.2.2 and 2.3].

In linear programming it is well known that the central path converges to the so-called analytic center of the set of optimal solutions, see, e.g., Adler and Monteiro [1]. This result was extended to the linear semidefinite case by Goldfarb and Scheinberg [6, Thm. 4.3] where they claim that $\bar{X} = Q_1 \bar{U} Q^T$, where $\bar{U}$ is the optimal solution to the problem

$$\begin{array}{ll} \underset{U \in \mathcal{S}^p_{++}}{\text{maximize}} & \ln\det U \\ \text{subject to} & \operatorname{trace}(A^{(i)}_{11} U) = b_i, \quad i = 1, \ldots, m. \end{array} \tag{1.8}$$

Subsequently, Halická, de Klerk and Roos [9] have shown that this characterization is not correct in general. However, in the special case of strict complementarity, i.e., when $p = q$, the characterization given by (1.8) is correct, see, e.g., Luo, Sturm and Zhang [12, Lemma 3.4] or de Klerk [4, Thm 2.3.2]. For a discussion on complementarity and related issues in semidefinite programming, see Alizadeh, Haeberly and Overton [2].

In this paper we give a characterization of $\bar{X}$ that does not require strict complementarity. It consists of a problem similar to (1.8), but with an extra term added to the objective function. This term vanishes in the strict complementarity case.

The paper is organized as follows. After a few notational comments we review an example by Halická, de Klerk and Roos [9] in Section 1.2. The example is used for illustrational purposes. In Section 2 we derive the problem whose solution characterizes the limit point of the central path. We then discuss the results of Goldfarb and Scheinberg [6] in light of our results in Section 3. The special case of strict complementarity is discussed in Section 4 before we finish the paper with a summary and a brief discussion in Section 5.

1.1. Notation

We denote the ith unit vector by e_i. For two real sequences $\{v_1(\mu) : \mu > 0\}$ and $\{v_2(\mu) : \mu > 0\}$, $v_1(\mu) = O(v_2(\mu))$ means that $v_1(\mu)/v_2(\mu)$ is bounded as $\mu \to 0^+$. Further, $v_1(\mu) = \Theta(v_2(\mu))$ means that $v_1(\mu) = O(v_2(\mu))$ and $v_2(\mu) = O(v_1(\mu))$. For matrices, the "O" symbol should be interpreted componentwise, e.g. $X(\mu) = O(v_1(\mu))$ should be interpreted as each component of $X(\mu)/v_1(\mu)$ is bounded as $\mu \to 0^+$. For a symmetric matrix A, we mean by $A \succeq 0$ that A is positive semidefinite. For a symmetric positive semidefinite matrix A, we denote the matrix square root of A by $A^{1/2}$; see, e.g., Golub and van Loan [7, p. 149].

1.2. An illustrative example

The following example was given by Halická, de Klerk and Roos [9] to show that in the absence of strict complementarity, the central path need not converge to the analytic center as defined by (1.8).

$$\begin{array}{ll} \text{minimize} & x_{44} \\ \text{subject to} & \begin{pmatrix} 1-x_{22} & x_{12} & x_{13} & x_{14} \\ x_{12} & x_{22} & -x_{44}/2 & -x_{33}/2 \\ x_{13} & -x_{44}/2 & x_{33} & 0 \\ x_{14} & -x_{33}/2 & 0 & x_{44} \end{pmatrix} \succeq 0. \end{array} \tag{1.9}$$

Since the form of the problem is such that $Q = I$ is admissible (see below), *indices can be interpreted the normal way* in this section. Let $X \in \mathcal{S}^4_+$ and note that adding the linear constraints $x_{11}+x_{22} = 1$, $x_{33}+2x_{24} = 0$, $x_{44}+2x_{23} = 0$, and $x_{34} = 0$ gives the positive semidefiniteness constraint of (1.9). Since the linear constraints can be written on trace form, (1.9) can be stated on the form (1.1). It is straightforward to verify that Assumptions 1 and 2 are satisfied and that the set of optimal solutions to the primal and dual problems are all X, y, S such that $y = 0$ and X and S are positive semidefinite matrices on the form

$$X = \begin{pmatrix} 1-x_{22} & x_{12} & 0 & 0 \\ x_{12} & x_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad S = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Note that the dual optimal solution is unique while the primal one is not. The analytic center problem (1.8) for the set of optimal solutions to this example is

$$\begin{array}{ll} \underset{X_{11}\in\mathcal{S}^2_{++}}{\text{maximize}} & \ln\det X_{11} \\ \text{subject to} & x_{11} + x_{22} = 1, \end{array}$$

with solution $X_{11} = \text{diag}(0.5, 0.5)$. However, the limit of $X_{11}(\mu)$ as $\mu \to 0^+$ is $\bar{X}_{11} = \text{diag}(0.4, 0.6)$; see Halická, de Klerk and Roos [9]. Since $Q = I$, we use X_{11} instead of U as variable matrix. In Section 3, we return to this example and illustrate the main result of this paper as well as the claim by Goldfarb and Scheinberg [6, Thm. 4.3].

2. Derivation of a well-behaved limit problem

To derive a problem that characterizes the limit point of the central path, we need to reformulate problem (1.3). This is done in three steps. First, we eliminate the term $\text{trace}(CX)$ from the objective. Then, we make a rotation to bring $\bar{X}$ and $\bar{S}$ to the form of (1.6) and fix the blocks that go to zero at their optimal values. Finally, we notice that the problem obtained has the same first-order optimality conditions as another convex problem which has better limiting properties. These properties allow us to let μ go to zero and obtain a well-defined limit problem in X_{11} only.

The first-order conditions for the problem

$$\begin{array}{lll} \underset{X\in\mathcal{S}^n_{++}}{\text{minimize}} & -\ln\det X & \\ \text{subject to} & \operatorname{trace}(A^{(i)}X)=b_i, & i=1,\dots,m, \\ & \operatorname{trace}(CX)=\operatorname{trace}(CX(\mu)), & \end{array} \tag{2.1}$$

state the existence of a vector $\pi\in I\!R^m$ and a scalar $\gamma\in I\!R$ such that

$$\begin{array}{rcll} -X^{-1} & = & \sum_{i=1}^m \pi_i A^{(i)} - \gamma C, & \\ \operatorname{trace}(A^{(i)}X) & = & b_i, & i=1,\dots,m, \\ \operatorname{trace}(CX) & = & \operatorname{trace}(CX(\mu)). & \end{array} \tag{2.2}$$

Since this problem and (1.3) both are convex problems and their optimality conditions (1.4) and (2.2) are equivalent (let $\gamma=1/\mu$ and $\pi=\lambda/\mu$), problems (1.3) and (2.1) have the same optimal solution.

In order to reformulate the objective of (2.1), we factorize Q^TXQ as

$$Q^TXQ=\begin{pmatrix} I & 0 \\ X_{b1}X_{11}^{-1} & I \end{pmatrix}\begin{pmatrix} X_{11} & 0 \\ 0 & X_{bb}-X_{b1}X_{11}^{-1}X_{1b} \end{pmatrix}\begin{pmatrix} I & X_{11}^{-1}X_{1b} \\ 0 & I \end{pmatrix}, \tag{2.3}$$

where

$$X_{b1}=\begin{pmatrix} X_{21} \\ X_{31} \end{pmatrix}, \quad X_{1b}=X_{b1}^T \quad \text{and} \quad X_{bb}=\begin{pmatrix} X_{22} & X_{23} \\ X_{32} & X_{33} \end{pmatrix}. \tag{2.4}$$

The factorization can be viewed as a triangular transformation as opposed to the orthonormal transformation used by Goldfarb and Scheinberg [6] which allows them to eliminate the extra term we get the objective of (2.6) below. The advantage of our transformation is that it allows us to formulate a problem whose feasible region is guaranteed to be well behaved as a function of μ as $\mu\to 0^+$. The orthonormality of Q in conjunction with (2.3) gives

$$\ln\det X=\ln\det(Q^TXQ)=\ln\det X_{11}+\ln\det(X_{bb}-X_{b1}X_{11}^{-1}X_{1b}). \tag{2.5}$$

Using (2.5) and letting all variables except those in the (1,1) block be equal to their optimal values allows us to rewrite problem (2.1) as

$$\begin{array}{lll} \underset{U\in\mathcal{S}^p_{++}}{\text{minimize}} & -\ln\det U-\ln\det(X_{bb}(\mu)-X_{b1}(\mu)U^{-1}X_{1b}(\mu)) & \\ \text{subject to} & \operatorname{trace}(A_{11}^{(i)}U)=\bar b_i(\mu), & i=1,\dots,m, \\ & (\operatorname{trace}(C_{11}U)=c(\mu)), & \end{array} \tag{2.6}$$

where

$$\begin{array}{rcll} \bar b_i(\mu) & = & \operatorname{trace}(A_{11}^{(i)}X_{11}(\mu)), & i=1,\dots,m, \\ c(\mu) & = & \operatorname{trace}(C_{11}X_{11}(\mu)). & \end{array} \tag{2.7}$$

To emphasize that the problem is in the rotated coordinates, we use U as the variable matrix. The reason for putting the last constraint in parentheses is that it is redundant. To see this, note that if y is an arbitrary vector which is optimal to the dual, then $\sum_i y_i A_{11}^{(i)} = C_{11}$. The last constraint is therefore a linear combination of the other constraints. Hence, it is redundant and will be omitted from hereon.

For notational convenience, we introduce the matrix function $M(\mu)$ in the following definition. Further, the limit of convergent sequences of values of this function as μ tends to zero is a key ingredient in the characterization of the limit point of the central path.

Definition 2.1. *For positive μ, let*

$$M(\mu) = X_{1b}(\mu)(X_{bb}(\mu) - X_{b1}(\mu)X_{11}(\mu)^{-1}X_{1b}(\mu))^{-1}X_{b1}(\mu),$$

where $X_{b1}(\mu)$, $X_{1b}(\mu)$ and $X_{bb}(\mu)$ are given by (2.4).

An alternative formulation of $M(\mu)$, which is used in proofs below, is given by the following lemma.

Lemma 2.1. *The function $M(\mu)$, given by Definition 2.1, can alternatively be written as*

$$M(\mu) = X_{1b}(\mu)S_{bb}(\mu)X_{b1}(\mu)/\mu, \quad \text{where} \quad S_{bb}(\mu) = \begin{pmatrix} S_{22}(\mu) & S_{23}(\mu) \\ S_{32}(\mu) & S_{33}(\mu) \end{pmatrix}.$$

Proof. The proof is based on the factorization (2.3) and the relation $X(\mu)S(\mu) = \mu I$. For simplicity, we suppress the dependency of X and S on μ in the proof. Let

$$L(\mu) = \begin{pmatrix} I & 0 \\ X_{b1}X_{11}^{-1} & I \end{pmatrix} \quad \text{and} \quad D(\mu) = \begin{pmatrix} X_{11} & 0 \\ 0 & X_{bb} - X_{b1}X_{11}^{-1}X_{1b} \end{pmatrix}.$$

It is straightforward to verify that $Q^TXQ = L(\mu)D(\mu)L(\mu)^T$ and, based on $XS = \mu I$, that $Q^TSQ = \mu L(\mu)^{-T}D(\mu)^{-1}L(\mu)^{-1}$. The inverses of $L(\mu)$ and $D(\mu)$ are

$$L(\mu)^{-1} = \begin{pmatrix} I & 0 \\ -X_{b1}X_{11}^{-1} & I \end{pmatrix} \quad \text{and} \quad D(\mu)^{-1} = \begin{pmatrix} X_{11}^{-1} & 0 \\ 0 & (X_{bb} - X_{b1}X_{11}^{-1}X_{1b})^{-1} \end{pmatrix}.$$

An explicit formation of S_{bb} then shows that $S_{bb}/\mu = (X_{bb} - X_{b1}X_{11}^{-1}X_{1b}^T)^{-1}$. Insertion in the definition of $M(\mu)$ gives the desired result. ∎

The following lemma shows that the optimal solution to (2.6) satisfies a set of conditions closely related to first-order optimality conditions for (2.6). In the subsequent lemma, these conditions are shown to be the first-order conditions for another convex problem.

Lemma 2.2. *Let Assumptions 1 and 2 hold and adopt notations (1.7) and (2.4). Let $M(\mu)$ be given by Definition 2.1. Then, $X_{11}(\mu)$ solves*

$$\begin{aligned} -U^{-1} + U^{-1}M(\mu)U^{-1} &= \textstyle\sum_{i=1}^m \pi_i(\mu)A_{11}^{(i)}, \\ \operatorname{trace}(A_{11}^{(i)}U) &= \bar{b}_i(\mu), \qquad i = 1,\ldots,m, \\ U &\in \mathcal{S}_{++}^p, \end{aligned} \tag{2.8}$$

where $\pi(\mu)$ is the set of multipliers to (2.6) and $\bar{b}(\mu)$ is given by (2.7).

Proof. From the derivation of (2.6), we know that $X_{11}(\mu)$ is the unique optimal solution to that problem. Convexity together with the first-order optimality conditions of (2.6) imply that $U = X_{11}(\mu)$ solves

$$\begin{aligned} U^{-1}X_{1b}(\mu)(X_{bb}(\mu) - X_{b1}(\mu)U^{-1}X_{1b}(\mu))^{-1}X_{b1}(\mu)U^{-1} - U^{-1} &= \sum_{i=1}^{m} \pi_i(\mu)A_{11}^{(i)}, \\ \operatorname{trace}(A_{11}^{(i)}U) &= \bar{b}_i(\mu), \quad i = 1,\dots,m, \\ U &\in \mathcal{S}_{++}^{p}. \end{aligned}$$

Therefore, $U = X_{11}(\mu)$ solves the system even if the third U^{-1} of the first line is replaced $X_{11}(\mu)^{-1}$. Identification with $M(\mu)$ then completes the proof. ∎

We are now in a position to state a convex problem whose optimal solution is $X_{11}(\mu)$ and whose feasible region is well behaved as μ tends to zero.

Lemma 2.3. *Let Assumptions 1 and 2 hold and adopt notation (1.7). Let $M(\mu)$ be given by Definition 2.1 and let $\bar{b}$ be given by (2.7). Further, let $X(\mu)$ denote the unique solution to (1.3) for positive μ. Then, for every positive μ, $X_{11}(\mu)$ is the unique optimal solution to the problem*

$$\begin{aligned} &\underset{U\in\mathcal{S}_{++}^{p}}{\text{minimize}} && -\ln\det U + \operatorname{trace}(M(\mu)U^{-1}) \\ &\text{subject to} && \operatorname{trace}(A_{11}^{(i)}U) = \bar{b}_i(\mu), \qquad i = 1,\dots,m. \end{aligned} \tag{2.9}$$

Proof. The idea of the proof is the following. First, we prove strict convexity of the objective of (2.9). Since the feasible set is convex it then follows that any point satisfying the first-order conditions must be a unique minimizer. Comparing the first-order conditions with (2.8) shows that $X_{11}(\mu)$ is the unique minimizer.

Let $E^{kl} = e_k e_l^T + e_l e_k^T$ for $k \neq l$ and $E^{kk} = e_k e_k^T$ and write U as

$$U = \sum_{1\le k\le l\le p} u^{kl}E^{kl},$$

with u^{kl} being component (k, l) of U. Denote the second term of the objective of (2.9) by $h_\mu(U)$. Then,

$$\frac{\partial h_\mu}{\partial u^{kl}} = -\operatorname{trace}(M(\mu)U^{-1}E^{kl}U^{-1}) \quad \text{and}$$

$$\frac{\partial^2 h_\mu}{\partial u^{kl}u^{st}} = \operatorname{trace}(M(\mu)U^{-1}E^{st}U^{-1}E^{kl}U^{-1}) + \operatorname{trace}(M(\mu)U^{-1}E^{kl}U^{-1}E^{st}U^{-1}).$$

Since trace is a linear operation it follows that the Hessian of $h_\mu(U)$ is positive semidefinite if and only if

$$2\operatorname{trace}(M(\mu)U^{-1}PU^{-1}PU^{-1}) \ge 0 \tag{2.10}$$

for all symmetric $p \times p$ matrices P. Positive semidefiniteness of $M(\mu)$ follows, e.g., from Lemma 2.1 upon noticing that $S_{33}(\mu)$ is positive definite and μ is positive. In conjunction with the positive definiteness of U, (2.10) is equivalent to

$$\operatorname{trace}((M(\mu)^{1/2}U^{-1}PU^{-1/2})(M(\mu)^{1/2}U^{-1}PU^{-1/2})^T) \geq 0. \tag{2.11}$$

Since the trace of a positive semidefinite matrices is nonnegative, (2.11) holds and convexity of $\operatorname{trace}(M(\mu)U^{-1})$ follows. Strict convexity of the objective of (2.9) follows since the first term of the objective of (2.9) strictly convex; see, e.g., Monteiro and Todd [13, p. 273]. Together with the convexity of the feasible region, it follows that any point satisfying the first-order conditions for (2.9) is the unique optimal solution. However, the first-order conditions for (2.9) state that there exists a vector $\pi \in I\!R^m$ such that

$$\begin{aligned} -U^{-1} + U^{-1}M(\mu)U^{-1} &= \textstyle\sum_{i=1}^m \pi_i A_{11}^{(i)}, \\ \operatorname{trace}(A_{11}^{(i)}U) &= \bar{b}_i(\mu), \quad i = 1, \dots, m, \\ U &\in \mathcal{S}_{++}^p, \end{aligned} \tag{2.12}$$

A comparison with (2.8) shows that $U = X_{11}(\mu)$ and $\pi = \pi(\mu)$ satisfy (2.12) and hence, $X_{11}(\mu)$ is the unique optimal solution to (2.9). This completes the proof. ∎

The aim is of course to let μ tend to zero in (2.9). To obtain a well-defined limit problem, we need the following lemma which states boundedness of $M(\mu)$ as μ tends to zero. The proof is based on the alternative formulation of $M(\mu)$ given by Lemma 2.1.

Lemma 2.4. *Let Assumptions 1 hold and 2 and adopt the notation given by (1.7) and (2.4). Then, $M(\mu)$, given by Definition 2.1, is bounded as $\mu \to 0^+$.*

Proof. As previously noted, the two assumptions imply that the central path $(X(\mu), y(\mu), S(\mu))$ exists for all positive μ. Since this lemma only concerns the behavior along the central path, we may suppress the dependency on μ in the proof. An explicit formation of the expression for M given by Lemma 2.1 in terms of the partitioning (1.6) gives

$$M = \frac{1}{\mu}(X_{12}S_{22}X_{21} + X_{13}S_{32}X_{21} + X_{12}S_{23}X_{31} + X_{13}S_{33}X_{31}). \tag{2.13}$$

The proof is done by proving that each term in the parentheses is $O(\mu)$. First, note that the diagonal elements of X_{11} and S_{33} are all $\Theta(1)$ per definition. By feasibility, $\operatorname{trace}((X - \bar{X})(S - \bar{S})) = 0$ holds. Further, $\bar{X}$ and $\bar{S}$ are optimal, $Q^T\bar{X}Q$ and $Q^T\bar{S}Q$ are diagonal and X and S satisfy (1.5c). This, along with the facts that the trace is unchanged by rotations and that $QQ^T = I$, implies that

$$\begin{aligned} 0 &= \operatorname{trace}(XS) - \operatorname{trace}(Q^T\bar{X}QQ^TSQ) - \operatorname{trace}(Q^TXQQ^T\bar{S}Q) \\ &= n\mu - \sum_{i=1}^p \bar{x}_{ii}s_{ii} - \sum_{i=q+1}^n \bar{s}_{ii}x_{ii}. \end{aligned}$$

Hence,

$$\sum_{i=1}^{p} \bar{x}_{ii} s_{ii} + \sum_{i=q+1}^{n} \bar{s}_{ii} x_{ii} = O(\mu).$$

Since every term in the two sums is nonnegative and the present $\bar{x}_{ii}$ and $\bar{s}_{ii}$ are positive it follows that $x_{ii} = O(\mu)$ for $i = q+1, \dots, n$ and $s_{ii} = O(\mu)$ for $i = 1, \dots, p$. By positive definiteness of X_{33} and S_{11} for $\mu > 0$, it follows that $X_{33} = O(\mu)$ and $S_{11} = O(\mu)$. Note that the proof that X_{33} and S_{11} are $O(\mu)$ essentially follows Luo, Sturm and Zhang [12, Lemma 3.2].

Each component x_{ij} such that $1 \leq i \leq p$ and $q+1 \leq j \leq n$, i.e., a component of X_{13}, must be $O(\sqrt{\mu})$. To see this, consider the submatrix

$$\begin{pmatrix} x_{ii} & x_{ij} \\ x_{ij} & x_{jj} \end{pmatrix},$$

which by positive definiteness of X must be positive definite for all $\mu > 0$. Therefore, $x_{ij}^2 < x_{ii} x_{jj} = O(1)O(\mu)$ and X_{13} is $O(\sqrt{\mu})$. Similarly, $S_{13} = O(\sqrt{\mu})$.

The $(1,1)$, $(1,2)$ and $(1,3)$ blocks of the equality $Q^T X Q Q^T S Q = \mu I$ give

$$X_{11}S_{11} + X_{12}S_{21} + X_{13}S_{31} = \mu I, \tag{2.14a}$$
$$X_{11}S_{12} + X_{12}S_{22} + X_{13}S_{32} = 0, \tag{2.14b}$$
$$X_{11}S_{13} + X_{12}S_{23} + X_{13}S_{33} = 0. \tag{2.14c}$$

From the results above, the first and third term of (2.14c) are $O(\sqrt{\mu})$ and therefore $X_{12}S_{23} = O(\sqrt{\mu})$. The first and the third term as well as the right-hand side of (2.14a) are all $O(\mu)$, implying that $X_{12}S_{21} = O(\mu)$. Multiplying (2.14b) from the right by X_{21} gives

$$X_{11}S_{12}X_{21} + X_{12}S_{22}X_{21} + X_{13}S_{32}X_{21} = 0.$$

Since $X_{11} = O(1)$ and $S_{12}X_{21} = O(\mu)$, the first term is $O(\mu)$. Further, the third term is also $O(\mu)$ since $X_{13} = O(\sqrt{\mu})$ and $S_{32}X_{21} = O(\sqrt{\mu})$. Hence, the second term is also $O(\mu)$.

To summarize, we have proved that

$$\begin{aligned} &X_{11} = O(1), \quad X_{13} = O(\sqrt{\mu}), \quad X_{12}S_{21} = O(\mu), \qquad X_{12}S_{22}X_{21} = O(\mu), \\ &S_{33} = O(1), \quad S_{13} = O(\sqrt{\mu}) \quad X_{12}S_{23} = O(\sqrt{\mu}). \end{aligned}$$

Insertion into (2.13) shows that M is bounded as $\mu \to 0^+$. ∎

The boundedness of $M(\mu)$, given by the previous lemma, implies the existence of a convergent subsequence as μ tends to zero. The following theorem shows that a reasonable limit problem of (2.9) is obtained when $M(\mu)$ is replaced with the limit of any such subsequence and $\bar{b}(\mu)$ is replaced with its limit, b. The limit problem may be viewed as a "corrected" analytic center problem in that a convex term is added to the objective.

Theorem 2.1. *Let Assumptions 1 and 2 hold. Further, let $X(\mu)$ denote the unique solution to (1.3) for positive μ and let $\bar{X} = \lim_{\mu\to 0^+} X(\mu)$. Also, let $M(\mu)$ be given by Definition 2.1 and let $\{\mu_k\}_{k=1}^{\infty}$ be an arbitrary positive sequence converging to zero such that $\{M(\mu_k)\}$ converges. Denote the limit matrix of $\{M(\mu_k)\}$ by M_0. Finally, adopt the notation (1.7) so that $\bar{X} = Q_1\bar{X}_{11}Q_1^T$. Then $\bar{X}_{11}$ is the unique optimal solution to the convex problem*

$$\begin{array}{lll} \underset{U\in\mathcal{S}_{++}^p}{\text{minimize}} & -\ln\det U + \operatorname{trace}(M_0U^{-1}) & \\ \text{subject to} & \operatorname{trace}(A_{11}^{(i)}U) = b_i, & i = 1,\dots,m. \end{array} \tag{2.15}$$

Proof. By (1.6), $\bar{X} = Q_1\bar{X}_{11}Q_1^T$, where $\bar{X}_{11}$ is diagonal and positive definite. It remains to show that $\bar{X}_{11}$ is the unique optimal solution to the convex problem (2.15).

The existence of a positive sequence $\{\mu_k\}$ converging to zero such that the sequence $\{M(\mu_k)\}$ converges follows from Lemma 2.4. The feasible region of (2.15) is convex. Since $M(\mu) \in \mathcal{S}_+^p$ for each positive μ and $\mathcal{S}_+^p$ is closed, $M_0 \in \mathcal{S}_+^p$. The objective of (2.15) is strictly convex by the same argument as the objective of (2.9) is strictly convex; see Lemma 2.3.

Let $f_0(U)$ denote the objective of (2.15) and let $f_\mu(U)$ denote the objective of (2.9). Further, let the set of feasible points with respect to (2.15) and (2.9) be denoted by $\mathcal{F}_0$ and $\mathcal{F}_\mu$ respectively.

Note that $\bar{b}(\mu)$ is continuous and converges to b as $\mu \to 0^+$. The inner product is also continuous and, therefore, $\bar{X}_{11} \in \mathcal{F}_0$ since $\mu_k \to 0$ as $k \to \infty$. Further, $f_0(\bar{X}_{11})$ is finite and $f_0(U)$ is unbounded at the relative boundary of $\mathcal{F}_0$. Hence, any optimal solution to (2.15) must lie in $\mathcal{F}_0$. Let $\bar{W}$ be arbitrary in $\mathcal{F}_0$ and let $W_k = \bar{W} + X_{11}(\mu_k) - \bar{X}_{11}$. Then, there is a constant K such that $W_k \in \mathcal{F}_{\mu_k}$ for all $k \geq K$ since trace is a linear operation, $\mathcal{S}_{++}^p$ is an open set and $X_{11}(\mu_k) - \bar{X}_{11} \to 0$ as $k \to \infty$. By optimality of $X_{11}(\mu_k)$ with respect to (2.9) for $\mu = \mu_k$ it follows that $f_{\mu_k}(X_{11}(\mu_k)) \leq f_{\mu_k}(W_k)$ for all $k \geq K$. Furthermore, both $\ln\det U$ and $\operatorname{trace}(MU^{-1})$ are continuous on $\mathcal{S}_{++}^p$. Along with the following facts: trace is a linear operation, $M(\mu_k)$ and W_k converge to M_0 and $\bar{W}$ respectively and $\bar{W}, \bar{X}_{11}, X_{11}(\mu_k)$ and W_k all belong to $\mathcal{S}_{++}^p$, this implies that $f_0(\bar{X}_{11}) \leq f_0(\bar{W})$. Since $\bar{W}$ is arbitrary in $\mathcal{F}_0$, $\bar{X}_{11}$ solves (2.15). By the strict convexity of $f_0(U)$ and the convexity of $\mathcal{F}_0$, $\bar{X}_{11}$ is also the unique optimal solution. ∎

In the next section, we discuss the example in Section 1.2 in light of this theorem. Then, in Section 4, we show that in presence of strict complementarity, the result of Theorem 2.1 simplifies to the ordinary analytic center result since $M(\mu)$ in that case converges to zero.

3. Discussion based on the example in Section 1.2

Let us return to the example (1.9) and see what is the problem with the proof in Goldfarb and Scheinberg [6, Theorem 4.3]. Remember that the example is on a form for which $Q = I$ is admissible. As in Section 1.2, the ordinary interpretation of indices can therefore be used. Empirically, problem (10) in Goldfarb and

Scheinberg [6] applied to (1.9) gives that $X_{11}(\mu)$ is the solution to a problem on the form

$$\begin{aligned} \underset{X_{11}\in\mathcal{S}^2_{++}}{\text{maximize}} \quad & \ln\det X_{11} \\ \text{subject to} \quad & x_{11} + k_1(\mu)x_{22} = k_2(\mu), \\ & l_i(\mu)x_{22} = l_i(\mu)\gamma(\mu), \qquad i = 1,\dots,4, \end{aligned} \tag{3.1}$$

where $k_1(\mu)$ and $k_2(\mu)$ converges to one as μ goes to zero, while $l_i(\mu) \to 0$ for $i = 1,\dots,4$ and $\gamma(\mu) \to 3/5$. This means that for small positive μ, the problem is approximately

$$\begin{aligned} \underset{X_{11}\in\mathcal{S}^2_{++}}{\text{maximize}} \quad & \ln\det X_{11} \\ \text{subject to} \quad & x_{11} + x_{22} = 1, \\ & x_{22} = 3/5, \end{aligned}$$

which is not “close” to the analytic center problem

$$\begin{aligned} \underset{X_{11}\in\mathcal{S}^2_{++}}{\text{maximize}} \quad & \ln\det X_{11} \\ \text{subject to} \quad & x_{11} + x_{22} = 1. \end{aligned}$$

The reason for the discrepancy is clearly that $l_i(\mu) \to 0$. This means that for small μ values the last four constraints of (3.1) approximately state that $x_{22}(\mu) = 3/5$. However, these constraints vanish if we put $\mu = 0$ since $l_i(0)$ is equal to zero. This is what happens in the analytic center problem.

The discussion above demonstrates that it is essential to make sure that the feasible points of the limit problem are close to feasible points of the parameterized problem for parameter values that are small enough. This is not necessarily the case in problem (10) in Goldfarb and Scheinberg [6] as demonstrated by the example above. However, for problems (2.9) and (2.15) this is the case because the feasible region is described as the intersection of a constant cone and the solutions to a set of linear equations with a *constant* matrix and a right-hand side continuously parameterized by μ.

Numerically, $M(\mu)$ converges for the example (1.9) and the limit is

$$M_0 = \begin{pmatrix} 0 & 0 \\ 0 & 0.3 \end{pmatrix}.$$

Applying (2.15) to the example (1.9) then gives the problem

$$\begin{aligned} \underset{X_{11}\in\mathcal{S}^2_{++}}{\text{minimize}} \quad & -\ln(x_{11}x_{22} - x_{12}^2) + \frac{0.3x_{11}}{x_{11}x_{22} - x_{12}^2} \\ \text{subject to} \quad & x_{11} + x_{22} = 1. \end{aligned}$$

Eliminating x_{22} and setting the partial derivatives with respect to x_{11} and x_{12} to zero gives

$$\begin{aligned} -\frac{1 - 2x_{11}}{x_{11}(1 - x_{11}) - x_{12}^2} + \frac{0.3(x_{11}(1 - x_{11}) - x_{12}^2) - 0.3x_{11}(1 - 2x_{11})}{(x_{11}(1 - x_{11}) - x_{12}^2)^2} &= 0, \\ \frac{2x_{12}}{x_{11}(1 - x_{11}) - x_{12}^2} + \frac{0.6x_{11}x_{12}}{(x_{11}(1 - x_{11}) - x_{12}^2)^2} &= 0. \end{aligned}$$

It is straightforward to verify that $x_{11} = 0.4$ and $x_{12} = 0$, i.e., the limit of the trajectory, solve these equations. Because of convexity, this is the solution to the problem.

Note that since $\bar{X}_{11}$ is diagonal by construction, (2.15) can alternatively be posed as a convex nonlinear programming problem in the diagonal elements of U only. Therefore, if the eigenvectors corresponding to the p positive eigenvalues are known, these eigenvalues are given by the optimal solution to (2.15).

4. The strict complementarity case

In the strict complementarity case, the central path converges to the analytic center; see, e.g., Luo, Sturm and Zhang [12, Lemma 3.4] or de Klerk [4, Thm 2.3.2]. The following corollary to Theorem 2.1 shows that, in this case, problem (2.15) simplifies to the ordinary analytic center problem since $M(\mu)$ tends to zero as μ tends to zero.

Corollary 4.1. *Let Assumptions 1 and 2 hold and adopt notation (1.7). Assume that strict complementarity holds and let $X(\mu)$ be the unique solution to (1.3). Finally, let $\bar{X} = \lim_{\mu\to 0^+} X(\mu)$. Then, $\bar{X}_{11}$ is the unique solution to the convex problem*

$$\begin{array}{ll} \underset{U\in\mathcal{S}^p_{++}}{\text{minimize}} & -\ln\det U \\ \text{subject to} & \operatorname{trace}(A^{(i)}_{11}U) = b_i, \quad i = 1,\dots,m. \end{array} \tag{4.1}$$

Proof. First, note that strict complementarity implies that $p = q$, $X_{b1} = X_{31}$ and $X_{bb} = X_{33}$. The result is a consequence of Theorem 2.1 since $\lim_{\mu\to 0^+} M(\mu) = 0$. To see this, use the formulation of $M(\mu)$ given by Lemma 2.1. It implies that

$$\|M(\mu)\| \le \|X_{13}(\mu)\|\|S_{33}(\mu)\|\|X_{31}(\mu)\|/\mu.$$

We know that $\|S_{33}(\mu)\| = \Theta(1)$. From Luo, Sturm and Zhang [12, Lemma 3.3] we get $\|X_{13}(\mu)\| = o(\sqrt{\mu})$. Therefore, $\|M(\mu)\|$ converges to zero and the result follows from Theorem 2.1. ∎

It should be noted that apart from getting a problem with a simpler objective, less information is needed to formulate the problem defining the limit point of the central path in case strict complementarity holds. In that case, only the optimal partition is needed, while otherwise the asymptotic behavior of the central path is needed to determine the matrix M_0 needed in the limit problem.

5. Summary and discussion

In the present paper we have given a characterization of the limit point of the central path in linear semidefinite programming. This characterization consists of an optimization problem over the set of optimal solutions, just like the problem defining the analytic center of the set of optimal solutions. However, the problems differ in that our problem has an extra term in the objective which guarantees that the solution to the problem is the limit point of the central path. This holds without the strict complementarity assumption which, in all proofs we are aware of, is required to guarantee that the central path converges to the analytic center.

We have, for brevity, focused on results for the primal part of the central path. For the dual part, analogous results can be established in a similar fashion to the results concerning the primal part. Further, the results of course hold for problem classes which are special cases of semidefinite programming, in particular for classes where strict complementarity is not guaranteed to hold, e.g. second order cone programming.

An open question is whether $M(\mu)$ can be proven to converge as $\mu \to 0^+$. This would simplify the theorem of the paper since M_0 then could be defined as the limit of $M(\mu)$ as $\mu \to 0^+$ instead of the limit of some convergent subsequence $\{M(\mu_k)\}$.

References

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

[2] F. Alizadeh, J.-P. A. Haeberly, and M. L. Overton. Complementarity and nondegeneracy in semidefinite programming. *Math. Program.*, 77:111–128, 1997.

[3] F. Alizadeh, J.-P. A. Haeberly, and M. L. Overton. Primal-dual interior-point methods for semidefinite programming: Convergence rates, stability and numerical results. *SIAM J. Optim.*, 8:746–768, 1998.

[4] E. de Klerk. *Interior Methods for Semidefinite Programming.* PhD thesis, Technische Universiteit Delft, 1997.

[5] E. de Klerk, J. Peng, C. Roos, and T. Terlaky. A scaled Gauss–Newton primal-dual search direction for semidefinite optimization. *SIAM J. Optim.*, 11:870–888, 2001.

[6] D. Goldfarb and K. Scheinberg. Interior point trajectories in semidefinite programming. *SIAM J. Optim.*, 8:871–886, 1998.

[7] G. H. Golub and C. F. Van Loan. *Matrix Computations.* The Johns Hopkins University Press, Baltimore, Maryland, third edition, 1996. ISBN 0-8018-5414-8.

[8] L. M. Graña Drummond and Y. Peterzil. The central path in smooth convex semidefinite programs. *Optimization.* to appear.

[9] M. Halická, E. de Klerk, and C. Roos. On the convergence of the central path in semidefinite optimization. *SIAM J. Optim.*, 12:1090–1099, 2002.

[10] C. Helmberg, F. Rendl, R. J. Vanderbei, and H. Wolkowicz. An interior-point method for semidefinite programming. *SIAM J. Optim.*, 6:342–361, 1996.

[11] M. Kojima, S. Shindoh, and S. Hara. Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices. *SIAM J. Optim.*, 7:86–125, 1997.

[12] Z. Luo, J. F. Sturm, and S. Zhang. Superlinear convergence of a symmetric primal-dual path following algorithm for semidefinite programming. *SIAM J. Optim.*, 8:59–81, 1998.

[13] R. Monteiro and M. Todd. Path-following methods. In H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors, *Handbook of Semidefinite Programming*, pages 267–306. Kluwer Academic Publisher, 2000.

[14] R. D. C. Monteiro. Primal-dual path following methods for semidefinite programming. *SIAM J. Optim.*, 7:663–678, 1997.

[15] R. D. C. Monteiro and T. Tsuchiya. Polynomial convergence of a new family of primal-dual algorithms for semidefinite programming. *SIAM J. Optim.*, 9:551–577, 1999.

[16] R. D. C. Monteiro and Y. Zhang. A unified analysis for a class of long-step primal-dual path-following interior-point algorithms for semidefinite programming. *Math. Prog.*, 81:281–299, 1998.

[17] Y. Nesterov and A. Nemirovskii. *Interior-point polynomial algorithms in convex programming.* SIAM, Philadelphia, 1994. ISBN 0-89871-319-6.

[18] F. A. Potra and R. Sheng. A superlinearly convergent primal-dual infeasible-interior-point algorithm for semidefinite programming. *SIAM J. Optim.*, 8:1007–1028, 1998.

[19] M. V. Ramana. An exact duality theory for semidefinite programming and its complexity implications. *Math. Program.*, 77:129–162, 1997.

[20] R. T. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, New Jersey, 1970. ISBN 0-691-08069-0.

[21] M. Todd. Semidefinite optimization. In A. Iserles, editor, *Acta Numerica 2001*, pages 515–560. Cambridge University Press, 2001. ISBN 0-521-8031-8.

[22] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Rev.*, 38:49–95, 1996.

[23] H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors. *Handbook of semidefinite programming.* Kluwer Academic Publishers, Boston, MA, 2000.

[24] Y. Zhang. On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming. *SIAM J. Optim.*, 8:365–386, 1998.