Initialization in semidefinite programming via a self-dual skew-symmetric embedding

E. de Klerk*, C. Roos, T. Terlaky

Faculty of Technical Mathematics and Informatics, Delft University of Technology, P.O. Box 5031, 2600 GA Delft, Netherlands

Abstract

The formulation of interior point algorithms for semidefinite programming has become an active research area, following the success of the methods for large-scale linear programming. Many interior point methods for linear programming have now been extended to the more general semidefinite case, but the initialization problem remained unsolved.

In this paper we show that the initialization strategy of embedding the problem in a self-dual skew-symmetric problem can also be extended to the semidefinite case. This method also provides a solution for the initialization of quadratic programs and it is applicable to more general convex problems with conic formulation. © 1997 Elsevier Science B.V.

Keywords: Semidefinite programming; Duality; Self-dual embedding; Central path; Initialization

1. Introduction

The extension of interior point algorithms from linear programming (LP) to semidefinite programming (SDP) [1] has received much attention recently, as is clear from the number of recent papers dealing with the subject (e.g. [14, 10, 12]). The question of initialization has not been satisfactorily solved up to now, and 'big-M' methods are often employed in practice to obtain feasible starting points. In the LP case an elegant solution for the initialization problem is to embed the original problem in a skew-symmetric self-dual problem which has a known interior feasible solution [16, 7]. The solution of the embedding problem then yields the optimal solution to the original problem, or detects infeasibility or unboundedness. In this paper we show how to extend this idea to semidefinite programming.

We give a self-dual, skew-symmetric embedding with known interior feasible solution on its central path, and show that a so-called *maximally complementary* optimal solution of this problem yields information about the solution of the original problem. In particular, exactly one of the following three cases occurs:

(I) an optimal solution with zero duality gap for the original problem is obtained;

(II) a recession direction for either the primal and/or dual problem is obtained;

(III) a certificate is obtained that no optimal solution pair with zero duality gap exists and that neither the primal nor the dual problem has a recession direction. This can only happen if one or both of the primal and dual SDP problems fail to satisfy the Slater regularity condition [14].

---

* Corresponding author. E-mail: e.deklerk@twi.tudelft.nl.

Since our embedding problem has an initial interior feasible solution on its central path any interior point method can be applied to it.

It is known [14] that convex quadratic programming (QP) is a special case of SDP. Thus, the initialization via self-dual embedding for QP problems is covered as well. Via this reformulation and embedding the problems arising from the nonlinearity of the objective as mentioned in [15] are avoided. A homogeneous embedding of monotone nonlinear complementarity problems is discussed in [3]. An infeasible start algorithm is proposed there to solve the homogeneous model. The homogeneous embedding idea was first analysed for SDP by Potra and Sheng [10].

Contrary to [3, 10] our embedding problem has a nonempty interior and a perfectly centered initial starting point.

The paper is organized as follows. In Section 2 the primal–dual SDP pair is defined in the symmetric form. First, the orthogonality properties of the primal and dual solutions are proved. Further, after defining the central path of the SDP problem it will be proved that any convergent subsequence on the central path converges to a maximally complementary solution. In Section 3 the self-dual embedding model is presented. This problem has a trivial initial solution on its central path. We show that any limit point of the central path of the embedding problem either provides an optimal primal–dual pair with zero duality gap, or shows that such solutions do not exist.

2. Orthogonality, maximal complementarity in SDP

The semidefinite programming problem will be considered in the symmetric form. Thus our primal problem (P) is

$$\min_{X,z} \quad \mathrm{Tr}(CX)$$

$$\text{s.t.} \quad \mathrm{Tr}(A_i X) - z_i = b_i, \quad i = 1,\dots,m,$$

$$X \succcurlyeq 0,$$

$$z \geqslant 0,$$

with the associated dual problem (D)

$$\max_{S,y} \quad b^{\mathrm{T}} y$$

$$\text{s.t.} \quad \sum_{i=1}^{m} y_i A_i + S = C,$$

$$S \succcurlyeq 0,$$

$$y \geqslant 0,$$

where C and the A_i's are symmetric $n \times n$ matrices, $b \in \mathbb{R}^m$, and $X \succcurlyeq 0$ means X is positive semidefinite. The solutions (X,z) and (y,S) will be referred to as feasible solutions as they satisfy the primal and dual constraints, respectively. Observe that the duality gap for (P) and (D) is given by

$$\mathrm{Tr}(CX) - b^{\mathrm{T}} y = \mathrm{Tr}\left(\left(\sum_{i=1}^{m} y_i A_i + S \right) X \right)$$

$$- \sum_{i=1}^{m} y_i (\mathrm{Tr}(A_i X) - z_i)$$

$$= \mathrm{Tr}(SX) + y^{\mathrm{T}} z.$$

A basic property of LP problems known as the *orthogonality property* easily extends to SDP.

Lemma 2.1 (Orthogonality). *Let* (X,z,S,y) *and* (X^0, z^0, S^0, y^0) *be two pairs of feasible solutions. The following orthogonality relation holds:*

$$\mathrm{Tr}((X - X^0)(S - S^0)) + (z - z^0)^{\mathrm{T}}(y - y^0) = 0.$$

Proof. The proof uses only the feasibility conditions of (P) and (D).

$$\mathrm{Tr}((X - X^0)(S - S^0)) + (z - z^0)^{\mathrm{T}}(y - y^0)$$

$$= \mathrm{Tr}\left((X - X^0) \left(\sum_{i=1}^{m} (y_i^0 - y_i) A_i \right) \right)$$

$$+ \sum_{i=1}^{m} \mathrm{Tr}(A_i (X - X^0)(y_i - y_i^0)) = 0. \qquad \square$$

It is well known that the pair of problems (P) and (D) has optimal solutions (X^*, S^*, z^*, y^*) with zero duality gap ($\mathrm{Tr}(CX^*) - b^{\mathrm{T}} y^* = 0$) if both the primal

and dual problems satisfy the Slater regularity condition [4], which will be assumed in the remainder of this section.

The optimality conditions for (P) and (D) are

$$\text{Tr}(A_i X) - z_i = b_i, \quad i = 1, \ldots, m, \quad z \geqslant 0, \ X \succcurlyeq 0,$$

$$\sum_{i=1}^{m} y_i A_i + S = C, \quad y \geqslant 0, \ S \succcurlyeq 0, \tag{1}$$

$$XS = 0,$$

$$y_i z_i = 0, \quad i = 1, \ldots, m.$$

Solutions (X, z) and (S, y) satisfying the last two equality constraints are called *complementary*. Since X and S are symmetric positive-semidefinite matrices the complementarity of X and S ($XS = 0$) is equivalent to $\text{Tr}(XS) = 0$. Complementary feasible solutions are optimal. If the complementarity conditions are relaxed to

$$\text{Tr}(A_i X) - z_i = b_i, \quad i = 1, \ldots, m,$$

$$\sum_{i=1}^{m} y_i A_i + S = C,$$

$$XS = \mu I,$$

$$y_i z_i = \mu, \quad i = 1, \ldots, m,$$

$$X, S \succcurlyeq 0,$$

$$z, y \geqslant 0,$$

with some $\mu > 0$, then this system has a unique solution $X(\mu)$, $z(\mu)$, $S(\mu)$, $y(\mu)$ if the matrices A_i are linearly independent. This solution can be seen as the parametric representation of a smooth curve (the *central path*) in terms of the parameter μ.

We show that any limit point of the central path in the optimal set is a so-called maximally complementary solution, defined as follows.

Definition 2.1. An optimal solution (X^*, S^*, z^*, y^*) is called a maximally complementary solution if z^* and y^* have the maximal number of strictly positive components, and if X^* and S^* have maximal rank. [1]

---

[1] Results pertaining to bounds on the rank of optimal solutions may be found in [8, 9], and on nondegeneracy and strict complementarity properties of optimal solutions in [2].

A maximally complementary solution can be obtained by a central path following algorithm – we prove below that all accumulation points of the central path in the optimal set are maximally complementary solutions. This result is analogous to the result that the central path of an LP problem converges to a strictly complementary solution, the so-called analytic center of the optimal set [6, 7]. [2] Recall that in the LP case a strictly complementary solution always exists [7]. This is not necessarily the case for nonlinear problems, not even for convex quadratic problems.

Theorem 2.1 (Maximal complementarity). *Consider any sequence $\{\mu_t\} \to 0$ with $\mu_t > 0$, $t = 1, \ldots$ Any convergent subsequence of $(X(\mu_t), z(\mu_t), S(\mu_t), y(\mu_t))$ on the central path converges to a maximally complementary solution as $\mu_t \to 0$.*

Proof. We will prove the theorem in two steps. First, we prove that a limit point (z, y) of $(z(\mu_t), y(\mu_t))$ exists. In a second step we show that any such limit point is maximally complementary. The existence of the limit points is proved by showing the following lemma.

Lemma 2.2. *Given $\bar{\mu} > 0$, the set*

$$\{(X(\mu), z(\mu), S(\mu), y(\mu)) : 0 \leqslant \mu \leqslant \bar{\mu}\}$$

is bounded.

Proof. Let (X^0, z^0, S^0, y^0) be any strictly feasible primal–dual solution, and $(X(\mu), z(\mu), S(\mu), y(\mu))$ a central solution corresponding to some $\mu > 0$. By orthogonality (Lemma 2.1) one has

$$\text{Tr}((X(\mu) - X^0)(S(\mu) - S^0))$$

$$+ (z(\mu) - z^0)^{\text{T}}(y(\mu) - y^0) = 0. \tag{2}$$

The centrality conditions imply $\text{Tr}(X(\mu)S(\mu)) = n\mu$ and $z(\mu)^{\text{T}}y(\mu) = m\mu$, which simplifies (2) to

$$\text{Tr}(X(\mu)S^0) + \text{Tr}(X^0 S(\mu)) + (z^0)^{\text{T}}y(\mu) + z(\mu)^{\text{T}}y^0$$

$$= (m + n)\mu + \text{Tr}(X^0 S^0) + (z^0)^{\text{T}}y^0. \tag{3}$$

---

[2] Since the first submission of this paper, Goldfarb and Scheinberg [5] have proved that the central path always converges in the SDP case. The limit is the analytic center of the optimal set.

The left-hand side terms of the last equation are nonnegative by feasibility. For a given $\bar{\mu} > 0$, and defining $\kappa := \mathrm{Tr}(X^0 S^0) + (z^0)^\mathrm{T} y^0$ one therefore has

$$\|X(\mu)\| < \frac{(m+n)\bar{\mu} + \kappa}{\lambda_{\min}(S^0)}, \quad \forall \mu \leqslant \bar{\mu},$$

where $\lambda_{\min}(S^0)$ denotes the smallest eigenvalue of S^0 and

$$y_i(\mu) < \frac{(m+n)\bar{\mu} + \kappa}{z_i^0}, \forall i = 1, \ldots, m, \mu \leqslant \bar{\mu}$$

with similar bounds for $S(\mu)$ and $z(\mu)$. $\quad\square$

In the following two lemmas we show that the limit points are also maximally complementary.

Lemma 2.3. *Any limit point (z, y) of the series $(z(\mu_t), y(\mu_t))$ is maximally complementary. Furthermore, for any maximally complementary solution (X^*, S^*, z^*, y^*) the series*

$$\mathrm{Tr}(S^{-1}(\mu_t) S^*) \quad and \quad \mathrm{Tr}(X^* X^{-1}(\mu_t))$$

are uniformly bounded.

Proof. Let (X^*, S^*, z^*, y^*) be a maximally complementary solution, and $(X(\mu), z(\mu), S(\mu), y(\mu))$ a central solution corresponding to some $\mu > 0$. Similar to the previous lemma, the orthogonality Lemma 2.1 yields

$$\mathrm{Tr}((X(\mu) - X^*)(S(\mu) - S^*))$$
$$+ (z(\mu) - z^*)^\mathrm{T} (y(\mu) - y^*) = 0.$$

Using the centrality conditions as before, and the optimality conditions $\mathrm{Tr}(X^* S^*) = 0$ and $(z^*)^\mathrm{T} y^* = 0$, this simplifies to

$$\mathrm{Tr}(X(\mu) S^*) + \mathrm{Tr}(X^* S(\mu))$$
$$+ (z^*)^\mathrm{T} y(\mu) + z(\mu)^\mathrm{T} y^* = (m+n)\mu.$$

Using the centrality conditions once more, we have

$$\mathrm{Tr}(S^{-1}(\mu) S^*) + \mathrm{Tr}(X^* X^{-1}(\mu))$$
$$+ \sum_{\substack{j=1 \\ z_j^* > 0}}^{m} \frac{z_j^*}{z_j(\mu)} + \sum_{\substack{j=1 \\ y_j^* > 0}}^{m} \frac{y_j^*}{y_j(\mu)} = m + n. \qquad (4)$$

By feasibility each term in (4) is nonnegative and consequently bounded from above by $m + n$. This proves the second part of the lemma. Furthermore, for the sequence μ_t, one has $\lim_{\mu_t \to 0} z_j(\mu_t) > 0$ if $z_j^* > 0$ and $\lim_{\mu_t \to 0} y_j(\mu_t) > 0$ if $y_j^* > 0$. $\quad\square$

To prove Theorem 2.1 we still need to show that limit points of the sequences $X(\mu_t)$ and $S(\mu_t)$ are also maximally complementary, i.e. if $X = \lim_{\mu_t \to 0} X(\mu_t)$ and $S = \lim_{\mu_t \to 0} S(\mu_t)$ then X and X^* have the same rank and, further, S and S^* have the same rank.

Lemma 2.4. *The matrices $X := \lim_{\mu_t \to 0} X(\mu_t)$ and $S := \lim_{\mu_t \to 0} S(\mu_t)$ are maximally complementary.*

Proof. We prove here that the rank of X is the same as the rank of X^*. The proof for S proceeds analogously. Recall from Lemma 2.3 that $\mathrm{Tr}(X(\mu_t)^{-1} X^*) \leqslant m + n$ for all $\mu_t > 0$. Let the rank of X^* be k, its eigenvalues be denoted by $(\lambda_1^*, \ldots, \lambda_k^*)$, and a set of its orthonormal eigenvectors be given as $(e_1^*, \ldots, e_k^*)$. Then $X^* = P^* \Lambda^* P^{*\mathrm{T}}$ where $P^* = (e_1^*, \ldots, e_k^*)$ is an $n \times k$ orthonormal and $\Lambda^* = \mathrm{diag}(\lambda_1^*, \ldots, \lambda_k^*)$ is a $k \times k$ diagonal matrix. Then we have

$$\mathrm{Tr}(X(\mu_t)^{-1} X^*) = \mathrm{Tr}(X(\mu_t)^{-1} P^* \Lambda^* P^{*\mathrm{T}})$$
$$= \mathrm{Tr}(P^{*\mathrm{T}} X(\mu_t)^{-1} P^* \Lambda^*) \leqslant m + n. \qquad (5)$$

Using this bound we derive

$$\mathrm{Tr}(P^{*\mathrm{T}} X(\mu_t)^{-1} P^*) = \mathrm{Tr}(P^{*\mathrm{T}} X(\mu_t)^{-1} P^* \Lambda^* \Lambda^{*-1})$$
$$\leqslant \frac{1}{\lambda_{\min}(\Lambda^*)} \mathrm{Tr}(P^{*\mathrm{T}} X(\mu_t)^{-1} P^* \Lambda^*) \leqslant \frac{m+n}{\lambda_{\min}(\Lambda^*)}.$$

For further reference we introduce the notation $K := (m + n)/\lambda_{\min}(\Lambda^*)$. Since $X(\mu_t)$ is symmetric positive definite, its orthonormal eigenvalue decomposition for each μ_t can be given as

$$X(\mu_t) = Q(\mu_t)^\mathrm{T} L(\mu_t) Q(\mu_t),$$

where $L(\mu_t)$ is an $n \times n$ positive diagonal matrix and $Q(\mu_t)^\mathrm{T} = Q(\mu_t)^{-1}$ is an $n \times n$ orthonormal matrix. Then $X(\mu_t)^{-1} = Q(\mu_t)^\mathrm{T} L(\mu_t)^{-1} Q(\mu_t)$ and further $B(\mu_t) =$

$Q(\mu_t)P^*$ is also an orthonormal matrix. Using this we have

$$\mathrm{Tr}(P^{*\mathrm{T}}X(\mu_t)^{-1}P^*)$$

$$= \mathrm{Tr}(P^{*\mathrm{T}}Q(\mu_t)^{\mathrm{T}}L(\mu_t)^{-1}Q(\mu_t)P^*)$$

$$= \mathrm{Tr}(B(\mu_t)^{\mathrm{T}}L(\mu_t)^{-1}B(\mu_t))$$

$$= \mathrm{Tr}(L(\mu_t)^{-1}B(\mu_t)B(\mu_t)^{\mathrm{T}})$$

$$= \sum_{i=1}^{n} \frac{B(\mu_t)_i B(\mu_t)_i^{\mathrm{T}}}{L(\mu_t)_i} \leqslant K, \qquad (6)$$

where $B(\mu_t)_i$ denotes the ith row of the matrix $B(\mu_t)$.

Introducing the notation $\eta(\mu_t)_i = B(\mu_t)_i B(\mu_t)_i^{\mathrm{T}}$, we have that $0 \leqslant \eta(\mu_t)_i \leqslant 1$ for all i and

$$\sum_{i=1}^{n} \eta(\mu_t)_i = \mathrm{Tr}(B(\mu_t)^{\mathrm{T}}B(\mu_t))$$

$$= \mathrm{Tr}(B(\mu_t)B(\mu_t)^{\mathrm{T}}) = k.$$

These last two relations imply that at least k of the n $\eta(\mu_t)_i$s are larger than or equal to $1/(n-k+1)$. We can choose an appropriate subsequence (indicated again by subscript t for the sake of simplicity) where these coordinates are fixed. Denote the set of these indices by I. Then we have

$$\eta(\mu_t)_i \geqslant \frac{1}{n-k+1}, \quad i \in I$$

and by (6)

$$L(\mu_t)_i \geqslant \frac{1}{(n-k+1)K}, \quad i \in I.$$

Using the notation $L = \lim_{\mu_t \to 0} L(\mu_t)$ we conclude that the diagonal matrix L has at least k nonzero diagonal elements. Since $Q(\mu_t)$ is orthonormal, for an appropriate subsequence (still indicated by subscript t) one has $Q = \lim_{\mu_t \to 0} Q(\mu_t)$ orthonormal, thus

$$X = \lim_{\mu_t \to 0} X(\mu_t) = \lim_{\mu_t \to 0} Q(\mu_t)^{\mathrm{T}}L(\mu_t)Q(\mu_t) = Q^{\mathrm{T}}LQ$$

has at least rank k. By noting that X^* has maximal rank among the optimal solutions, one has $\mathrm{rank}(X) \leqslant \mathrm{rank}(X^*)$ and thus $\mathrm{rank}(X) = k$. This completes the proof. $\quad\square$

To decide about the duality state of the problems (P) and (D) we need the following definition.

Definition 2.2. We say that the primal problem (P) has a ray if there is a symmetric matrix $\bar{X} \succcurlyeq 0$ such that $\mathrm{Tr}(A_i\bar{X}) \geqslant 0, \forall i$ and $\mathrm{Tr}(C\bar{X}) < 0$. Analogously, the dual problem (D) has a ray if there is a vector $0 \leqslant \bar{y} \in \mathbb{R}^m$ such that $-\sum_{i=1}^{m} \bar{y}_i A_i \succcurlyeq 0$ and $b^{\mathrm{T}}\bar{y} > 0$.

If there is either a primal or dual ray, then no optimal solution to either (P) or (D) exists: If there is a dual ray $\bar{y}$ then (P) is infeasible, since by assuming primal feasibility one has $0 \geqslant \sum_{i=1}^{m} \mathrm{Tr}(A_iX)\bar{y}_i \geqslant b^{\mathrm{T}}\bar{y} > 0$ for any primal feasible X, which is a contradiction. If one has a primal ray $\bar{X}$ then (D) is infeasible since by assuming dual feasibility one has $0 \leqslant \sum_{i=1}^{m} \mathrm{Tr}(A_i\bar{X})y_i \leqslant \mathrm{Tr}(C\bar{X}) < 0$ for any dual feasible y, which is also a contradiction.

A maximally complementary solution to our embedding problem, presented in the next section, will enable us to construct optimal solutions with zero duality gap to (P) and (D) if they exist, or to detect rays, or to exclude these two possibilities.

3. Self-dual embedding

We now embed the primal–dual pair of problems (P) and (D) in a larger problem. We no longer require the Slater regularity condition for (P) or (D). The embedding is done in two steps: A homogeneous, self-dual SDP problem is constructed which is then extended to a problem with nonempty interior of the feasible region. The following model can be obtained by homogenizing the optimality conditions of (P) and (D).

Find a $\tau > 0$ such that

$$\mathrm{Tr}(A_iX) - \tau b_i - z_i = 0 \quad \forall i,$$

$$-\sum_{i=1}^{m} y_i A_i + \tau C - S = 0,$$

$$b^{\mathrm{T}}y - \mathrm{Tr}(CX) - \rho = 0, \qquad (7)$$

$$y \geqslant 0, \quad X \succcurlyeq 0, \quad \tau \geqslant 0,$$

$$z \geqslant 0, \quad S \succcurlyeq 0, \quad \rho \geqslant 0,$$

is feasible. Note that the model (7) always admits the trivial zero solution. Moreover, a solution $(X^*, S^*, z^*, y^*, \tau^*, \rho^*)$ to (7) with $\tau^* > 0$ yields an optimal solution to the original problem pair (P) and (D) (scale the variables X^*, S^*, z^* and y^* by $1/\tau^*$). This is true due to the last constraint which requires the duality gap for the original pair of problems (P) and (D) be nonpositive. As the gap is always nonnegative by weak duality, optimality of the scaled solutions is established. Clearly, if (P) and (D) have no optimal solutions with zero duality gap then the skew-symmetric problem has no solution with $\tau^* > 0$.

By the above arguments, the last constraint of problem (7) prohibits an interior solution: if $\tau > 0$ one cannot have $\rho > 0$ as well, due to weak duality. The formulation must therefore be extended. By introducing some new vectors, parameters and variables a self-dual model with known initial interior solution is obtained. The construction is analogous to that given in [7] for LP. In what follows e denotes the all one vector:

$$\min_{y, X, \tau, \theta, z, S, \rho, v} \theta\beta$$

$$\text{s.t.} \quad \mathrm{Tr}(A_i X) - \tau b_i + \theta \bar{b}_i - z_i = 0 \quad \forall i,$$

$$-\sum_{i=1}^{m} y_i A_i + \tau C - \theta \bar{C} - S = 0,$$

$$b^{\mathrm{T}} y - \mathrm{Tr}(CX) + \theta \alpha - \rho = 0, \qquad (8)$$

$$-\bar{b}^{\mathrm{T}} y + \mathrm{Tr}(\bar{C}X) - \tau \alpha - v = -\beta,$$

$$y \geqslant 0, \quad X \succcurlyeq 0, \quad \tau \geqslant 0, \quad \theta \geqslant 0,$$

$$z \geqslant 0, \quad S \succcurlyeq 0, \quad \rho \geqslant 0, \quad v \geqslant 0,$$

where

$$\bar{b}_i := b_i + 1 - \mathrm{Tr}(A_i), \quad i = 1, \ldots, m,$$

$$-\bar{C} := I + \sum_{i=1}^{m} A_i - C,$$

$$\alpha := 1 + \mathrm{Tr}(C) - b^{\mathrm{T}} e,$$

$$\beta := m + n + 2.$$

It is straightforward to verify that a feasible interior starting solution is given by $y^0 = z^0 = e$, $X^0 = S^0 = I$, and $\theta^0 = \rho^0 = \tau^0 = v^0 = 1$. It is also easy to check that the embedding problem is self-dual via Lagrangean duality. This implies that the duality gap equals $2\theta\beta$ and

therefore $\theta^* = 0$ at an optimal solution since the self-dual embedding problem satisfies the Slater condition. It also proves existence of the central path. Lemma 2.1 furthermore guarantees that any limit point of the central path is a maximally complementary optimal solution.

We can now use a maximally complementary solution of the embedding problem (8) to obtain information about the original problem pair (P) and (D). In particular, we will distinguish between the three possibilities as discussed in the Introduction, namely

(I)   A primal–dual optimal pair (X^*, S^*, z^*, y^*) is obtained with zero duality gap $\mathrm{Tr}(CX^*) - b^{\mathrm{T}} y^* = 0$;
(II) A primal and/or dual ray is detected;
(III) A certificate is obtained that no optimal pair with zero duality gap exists, and that neither (P) nor (D) has a ray.

Theorem 3.1. *Let* $(X^*, S^*, z^*, y^*, \tau^*, \theta^*, S^*, \rho^*, v^*)$ *be a maximally complementary solution to the self-dual embedding problem. Then*

(i)   *if* $\tau^* > 0$ *then case* (I) *holds*;
(ii) *if* $\tau^* = 0$ *and* $\rho^* > 0$ *then case* (II) *holds*;
(iii) *if* $\tau^* = \rho^* = 0$ *then case* (III) *holds*.

Proof. We first consider the two possibilities $\tau^* = 0$ and $\tau^* > 0$.

If $\tau^* > 0$, then $(X^*, z^*)/\tau^*$ and $(y^*, S^*)/\tau^*$ are maximally complementary and optimal for (P) and (D), respectively, i.e. case (I) holds.

If $\tau^* = 0$ then we have $\tau = 0$ in any optimal solution of the embedding problem. This implies that we cannot have a pair of optimal solutions for (P) and (D) with duality gap zero, because if such a pair exists we can construct an optimal solution of the embedding problem with $\tau = 1$. If $\tau^* = 0$ it also follows that $\mathrm{Tr}(A_i X^*) \geqslant 0$ for all i and $\sum_{i=1}^{m} y_i^* A_i \preccurlyeq 0$. We further distinguish between two sub-cases: $\rho^* > 0$ and $\rho^* = 0$.

If $\rho^* > 0$ then $b^{\mathrm{T}} y^* - \mathrm{Tr}(CX^*) > 0$, i.e. $b^{\mathrm{T}} y^* > 0$ and/or $\mathrm{Tr}(CX^*) < 0$. In other words, there are primal and/or dual rays and case (II) applies. If $b^{\mathrm{T}} y^* > 0$ then y^* is a dual ray. In this case (P) is infeasible, and if (D) is feasible it is unbounded. If $\mathrm{Tr}(CX^*) < 0$ then we have a primal ray. In this case (D) is infeasible, and if (P) is feasible it is unbounded. If both $b^{\mathrm{T}} y^* > 0$ and $\mathrm{Tr}(CX^*) < 0$ then we have both a primal and a dual ray and in this case both (P) and (D) are infeasible.

Conversely, we must show that if there exists a primal and/or dual ray, then any maximally complementary solution of the embedding problem must have $\rho^* > 0$ and $\tau^* = 0$. Given a primal ray $\bar{X}$, one can construct an optimal solution to the embedding by setting $X^* = \kappa\bar{X}$, where $\kappa > 0$ is a constant to be specified later, and further setting $\tau^* = 0$, $y^* = 0$ and $\theta^* = 0$ (which guarantees optimality), to obtain

$$\rho^* = -\kappa\operatorname{Tr}(C\bar{X}) > 0,$$

$$z_i^* = \kappa\operatorname{Tr}(A_i\bar{X}) \geqslant 0, \quad i = 1,\ldots,m,$$

$$S^* = 0,$$

$$v^* = 2n + 2 + \kappa\operatorname{Tr}\left(C\bar{X} - \bar{X} - \sum_{i=1}^{m} A_i\bar{X}\right).$$

The first three equations show that ρ^*, z^* and S^* are feasible. It remains to prove that v^* is nonnegative. This is ensured by choosing κ sufficiently small to guarantee

$$2n + 2 \geqslant -\kappa\operatorname{Tr}\left(C\bar{X} - \bar{X} - \sum_{i=1}^{m} A_i\bar{X}\right)$$

which in turn ensures $v^* \geqslant 0$. The proof for a dual ray proceeds analogously.

Finally, if a maximally complementary solution is obtained with $\tau^* = \rho^* = 0$, then we again have that all optimal solutions yield $\rho = \tau = 0$, i.e. cases (I) and (II) cannot occur. This completes the proof. $\square$

Remark 1. Item (iii) in Theorem 3.1 (where $\tau^* = \rho^* = 0$) covers different situations (which can only occur if one or both of (P) and (D) are not strictly feasible), including

1. Existence of optimal solutions to (P) and (D) with nonzero duality gap;
2. Existence of a zero duality gap for (P) and (D) in the sense that $\inf\operatorname{Tr}(CX) = \sup b^{\mathrm{T}}y$, but where no optimal solutions exists to both (P) and (D).

If $\tau^* = \rho^* = 0$, some additional information can be obtained by monitoring the sequence $\rho(\mu_t)/\tau(\mu_t)$ of centered iterates. The variables ρ and τ are complementary and along the central path of the embedding one has $\rho(\mu_t)\tau(\mu_t) = \mu_t$. On the other hand, $\beta\theta$ must be proportional to the dual-

ity gap $\theta(\mu_t) = (n + m + 2)\mu_t/\beta$. It follows that $\theta(\mu_t)/\tau(\mu_t) \to 0$ as $\mu_t \to 0$. This implies the sequences $(X(\mu_t), z(\mu_t))/\tau(\mu_t)$ and $(y(\mu_t), S(\mu_t))/\tau(\mu_t)$ are asymptotically feasible for (P) and (D), respectively. If one also has $\rho(\mu_t)/\tau(\mu_t) \to 0$ then there also holds

$$\operatorname{Tr}(CX(\mu_t)/\tau(\mu_t)) - b^{\mathrm{T}}y(\mu_t)/\tau(\mu_t) \to 0$$

as $\mu_t \to 0$. In this case, the sequences $X(\mu_t)/\tau(\mu_t)$ and $S(\mu_t)/\tau(\mu_t)$ cannot both converge, since then an optimal feasible pair with zero duality gap is obtained.

Infeasibility of (P) (resp. (D)) is only detected by the embedding approach if it corresponds to a ray in (D) (resp. (P)). This excludes pathological cases where, e.g. (P) is solvable but (D) is infeasible. A suggestion for future research is to extend the results of Todd and Ye [13] concerning approximate Farkas lemmas to the semidefinite cone. This may allow one to use the above to detect infeasibility in general.

It would also be interesting to study the self-dual embedding problem of the dual semidefinite programming formulation of Ramana [11], where a zero duality gap is always guaranteed for a feasible primal–dual problem pair without requiring the Slater condition. This may extend the analysis to include case 1 of Remark 1, should it occur. To do this is the subject of further research.

Remark 2. The self-dual embedding may be solved directly with any primal algorithm, yielding an infeasible primal–dual algorithm for the original problem. It is therefore unnecessary to use the primal–dual KKT conditions of the embedding problem.

Remark 3. The embedding strategy can be adapted to accomodate given, strictly feasible starting solutions. For example, if $X^0 \succ 0$ and $z^0 > 0$ are feasible for (P) and $\mu^0 > 0$ is given, one can construct $\tau^0 > 0$, $\theta^0 > 0$, $\rho^0 > 0$, $v^0 > 0$, $y^0 > 0$, and $S^0 \succ 0$ such that $\tau^0\rho^0 = \mu^0$, $\theta^0 v^0 = \mu^0$, $y_i^0 z_i^0 = \mu^0$ for all i and $X^0 S^0 = \mu^0 I$. The parameters for the embedding problem now become

$$\bar{b}_i := \frac{\tau^0 b_i + z_i^0 - \operatorname{Tr}(A_i X^0)}{\theta^0}, \quad i = 1,\ldots,m,$$

$$-\bar{C} := \frac{S^0 + \sum_{i=1}^{m} y_i^0 A_i - \tau^0 C}{\theta^0},$$

$$\alpha := \frac{\rho^0 + \text{Tr}(CX^0) - b^{\text{T}}y^0}{\theta^0},$$

$$\beta := v^0 + \tau^0\alpha - \text{Tr}(\bar{C}X^0) + \bar{b}^{\text{T}}y^0.$$

It is straightforward to verify that the self-dual problem obtained with the use of these vectors admits the above specified initial positive solution which is on its central path associated to the point with $\mu^0 > 0$. It is also easy to check by simple calculation that $\beta = (m + n + 2)\mu^0 > 0$. In this way, re-optimization can be done if some problem data change. Whether such a 'warm-start' strategy via the embedding can be made computationally efficient remains a question for future research.

Remark 4. The results of the previous sections can be generalized to primal–dual convex problems in the conic formulation. Consider the primal problem as

$$\min_x \{c^{\text{T}}x \mid Ax - b \in \mathscr{C}_1, \ x \in \mathscr{C}_2\}$$

and its dual problem as

$$\max_y \{b^{\text{T}}y \mid -A^{\text{T}}y + c \in \mathscr{C}_2^*, \quad y \in \mathscr{C}_1^*\}$$

where $\mathscr{C}_1, \mathscr{C}_2$ are convex cones, $\mathscr{C}_1^*, \mathscr{C}_2^*$ are their dual cones, respectively, A is an $m \times n$ matrix, $b, y \in \mathbb{R}^m$ and $c, x \in \mathbb{R}^n$. These problems can be embedded in the skew–symmetric self-dual problem with nonempty interior as follows.

$$\min_{x,y,\tau,\theta} \quad \theta\beta$$

$$\text{s.t.} \quad Ax - \tau b + \theta\bar{b} \in \mathscr{C}_1,$$

$$-A^{\text{T}}y + \tau c - \theta\bar{c} \in \mathscr{C}_2^*,$$

$$b^{\text{T}}y - c^{\text{T}}x + \theta\alpha \geqslant 0,$$

$$-\bar{b}^{\text{T}}y + \bar{c}^{T}x - \tau\alpha \geqslant -\beta,$$

$$y \in \mathscr{C}_1^*, \quad x \in \mathscr{C}_2, \quad \tau \geqslant 0, \quad \theta \geqslant 0,$$

where

$$\bar{b} := b + e_m - Ae_n,$$

$$-\bar{c} := e_n + A^{\text{T}}e_n - c,$$

$$\alpha := 1 + c^{\text{T}}e_n - b^{\text{T}}e_m,$$

$$\beta := m + n + 2,$$

where e_m and e_n denote the all-one vectors in $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively.

One can show via Lagrangean duality that this embedding problem is self-dual with the all one solution as an initial interior feasible solution. Analogous results can be derived as for the positive semidefinite case – maximal complementarity of the scalar variables (which are in $\mathbb{R}_+$) can be proved the same way as in Lemma 2.3. This is sufficient to prove the validity of the above embedding. To define and prove maximal complementarity of the general conic variables is the subject of further research.

Acknowledgement

References

[1] F. Alizadeh, Combinatorial optimization with interior point methods and semi-definite matrices, Ph.D. Thesis, University of Minnesota, Minneapolis, USA, 1991.

[2] F. Alizadeh, J.-P.A. Haeberley, M.L. Overton, Complementarity and nondegeneracy in semidefinite programming, Working Paper, RUTCOR, Rutgers University, New Brunswick, NJ, 1995.

[3] E.D. Andersen, Y. Ye, On a homogeneous algorithm for the monotone complementarity problem, Tech. Report, Department of Management Sciences, University of Iowa, Iowa City, USA, 1995.

[4] M.S. Bazarraa, H.D. Sherali, C.M. Shetty, Nonlinear Programming: Theory and Algorithms, Wiley, New York, 1993.

[5] D. Goldfarb, K. Scheinberg, Interior point trajectories in semidefinite programming, Working Paper, Dept. of IEOR, Columbia University, New York, NY, 1996.

[6] O. Güler, Y. Ye, Convergence behavior of interior-point algorithms, Math. Programming 60 (1993) 215–228.

[7] B. Jansen, C. Roos, T. Terlaky, The theory of linear programming: skew symmetric self-dual problems and the central path, Optimization 29 (1994) 225–233.

[8] G. Pataki, On the rank of extreme matrices in semidefinite programming and the multiplicity of optimal eigenvalues, Tech. Report MSRR-604, Carnegie Mellon University, Pittsburgh, USA, Revised August 1995.

[9] G. Pataki, On the facial structure of cone-lp's and semi-definite programs, Tech. Report MSRR-595, Carnegie Mellon University, Pitsburgh, USA, 1995.

[10] F.A. Potra, R. Sheng, Homogeneous interior-point algorithms for semidefinite programming, Reports on Computational Mathematics 82, Department of Mathematics, The University of Iowa, Iowa City, USA, 1995.

[11] M. Ramana, An exact duality theory for semidefinite programming and its complexity implications, DIMACS Tech. Report 95-02, RUTCOR, Rutgers University, New Brunswick, NJ, USA, 1995.

[12] J.F. Sturm, S. Zhang, Symmetric primal–dual path following algorithms for semidefinite programming, Tech. Report 9554/A, Tinbergen Institute, Erasmus University Rotterdam, 1995.

[13] M.J. Todd, Y. Ye, Approximate Farkas lemmas and stopping rules for iterative infeasible point algorithms for linear programming, Tech. Report 1109, School of OR and Industrial Engineering, Cornell University, Ithaca, New York, 1994.

[14] L. Vanderberghe, S. Boyd, Semidefinite Programming, SIAM Rev. 38 (1996) 49–95.

[15] Y. Ye, A homogeneous and self-dual algorithm for LCP, Math. Programming 76 (1997) 211–222.

[16] Y. Ye, M.J. Todd, S. Mizuno, An $\mathcal{O}(\sqrt{n}L)$-iteration homogeneous and self-dual linear programming algorithm, Math. Oper. Res. 19 (1994) 53–67.