A Potential Reduction Algorithm for an Extended SDP Problem

by

Jia-wang Nie and Ya-xiang Yuan

Report No.**ICM-99-011** March 1999

A Potential Reduction Algorithm for an Extended SDP Problem*

Jia-wang Nie and Ya-xiang Yuan
State Key Laboratory of Scientific and Engineering Computing,
Institute of Computational Mathematics and Scientific/Engineering Computing,
Chinese Academy of Sciences, P. O. Box 2719, Beijing 100080, China
Email-addresses: njw@lsec.cc.ac.cn, yyx@lsec.cc.ac.cn

March 12,1999

Abstract

An extended semi-definite programming, the SDP with an additional quadratic term in the objective function, is studied. Our generalization is similar to the generalization from linear programming to quadratic programming. Optimal conditions for this new class of problems are discussed. We present a potential reduction algorithm for solving QSDP problems. The convergence properties of this algorithms are also given.

key words: Qudratic semi-definite programming, potential reduction method, convergence.

1 Introduction

In recent years, semi-definite programming (SDP) has attracted much attention from researchers. Many interesting and important results on SDP have been obtained. The SDP has the following standard form:

$$\min \quad \langle C, X\rangle \tag{1.1}$$

$$s.t. \quad \langle A_i, X\rangle = b_i, \quad i = 1, \dots, m, \tag{1.2}$$

$$X \succeq 0, \tag{1.3}$$

where $C, A_i \in \mathcal{S}\Re^{n\times n}$, $\mathcal{S}\Re^{n\times n}$ being the set of all $n\times n$ real symmetric matrices. $X \succeq 0$ indicates that $X \in \mathcal{S}\Re^{n\times n}$ is positive semi-definite. $\langle A, B\rangle$ is the inner product of A and B in the space $\Re^{n\times n}$, namely

$$\langle A, B\rangle = tr(A^T B), \quad \forall\, A, B \in \Re^{n\times n}. \tag{1.4}$$

SDP can be viewed as a generalization of linear programming from vector spaces to matrix spaces. Even though there are nonlinear features in SDP problems since the set of positive semi-definite matrices is not polyhedral, SDP has exactly the same form as linear programming. Therefore, most interior point methods for linear programming can be extended to SDP. Indeed, most published works on SDP are about interior point methods, for example see [1], [2], [3], [8], [11], [12]. The SDP problem (1.1)-(1.3) is also called as SDPLP by [9]. Recently, there are some studies on extensions of the SDP. One extension is the semi-definite linear complementarity problems (SDPLCP). For more details on SDPLCP, please see [6], [7] and [9].

In this paper we consider a new extension of the SDP problems. It is well known that adding a quadratic term in the objective of a linear programming problem gives a quadratic programming problem.

*This work is partially supported by Chinese NNSF grants 19731010

It is natural for us to consider extending the SDP problem by adding a quadratic term in the objective function (1.1). In quadratic programming, the general quadratic term is $\frac{1}{2}x^T Hx$, where $x \in \Re^n$. However, in the matrix space $\Re^{n\times n}$, not all quadratic terms can be expressed in the form $\frac{1}{2}\langle HX, X\rangle$. Thus, we consider a quadratic form which can be written as

$$Q(X) = \frac{1}{2}\langle \varphi(X), X\rangle, \tag{1.5}$$

where $\varphi(X)$ has the form:

$$\varphi(X) = \sum_{i=1}^{\ell} H_i X W_i. \tag{1.6}$$

Here $H_i, W_i (i = 1, \dots, \ell)$ are matrices in $\mathcal{S}\Re^{n\times n}$. Throughout this paper, we assume that all H_i and W_i are positive semi-definite. Consequently, we have:

$$\langle \varphi(X), Y\rangle = \langle X, \varphi(Y)\rangle, \quad \forall\, X, Y \in \Re^{n\times n}, \tag{1.7}$$

and

$$\langle \varphi(X), X\rangle \geq 0, \quad \forall\, X \in \Re^{n\times n}. \tag{1.8}$$

Therefore, we derive the extended problem as follows.

$$(QSDP) \qquad \begin{aligned} \min \quad & q(X) = \langle C, X\rangle + \frac{1}{2}\langle \varphi(X), X\rangle && (1.9)\\ s.t. \quad & \langle A_i, X\rangle = b_i, \quad i = 1, \dots, m, && (1.10)\\ & X \succeq 0, && (1.11) \end{aligned}$$

where $\varphi(X)$ has the form (1.6). The dual of the above problem is

$$(QSDD) \qquad \begin{aligned} \max \quad & d(X, y) = b^T y - \frac{1}{2}\langle \varphi(X), X\rangle && (1.12)\\ s.t. \quad & \sum_{i=1}^{m} y_i A_i + S = C + \varphi(X), && (1.13)\\ & X, S \succeq 0, && (1.14) \end{aligned}$$

where $y = (y_1, ..., y_m)^T \in \Re^m$. The primal feasible region and dual feasible region are

$$\mathcal{F}_p = \{X \in \mathcal{S}\Re^{n\times n} : \langle A_i, X\rangle = b_i, \quad i = 1, \dots, m, X \succeq 0\} \tag{1.15}$$

and

$$\mathcal{F}_d = \{(X, y, S) \in \mathcal{S}\Re^{n\times n} \times R^m \times \mathcal{S}\Re^{n\times n} : C + \varphi(X) = \sum_{i=1}^{m} y_i A_i + S, X, S \succeq 0\} \tag{1.16}$$

respectively. In order to guarantee the existence of initial points, we make the following assumption:

Assumption 1.1 *(Slater regularity condition): There exist $X \succ 0$, $S \succ 0$ and $y \in \Re^m$ such that $X \in \mathcal{F}_p$ and $(X, y, S) \in \mathcal{F}_d$.*

For any pair $(X, y, S) \in \mathcal{F}_d$, it follows that

$$\begin{aligned} q(X) \quad - \quad & d(X, y)\\ = \quad & \langle C, X\rangle + \frac{1}{2}\langle \varphi(X), X\rangle - b^T y + \frac{1}{2}\langle \varphi(X), X\rangle\\ = \quad & \langle \varphi(X) + C - \sum_{i=1}^{m} y_i A_i, X\rangle\\ = \quad & \langle S, X\rangle \geq 0. \end{aligned} \tag{1.17}$$

Therefore, if we can find a pair $(X^\star, y^\star, S^\star) \in \mathcal{F}_d$ such that $\langle X^\star, S^\star\rangle = 0$, problems (QSDP) and (QSDD) are solved simultaneously.

This paper is organized as follows. In the next section, some optimal conditions for QSDP are presented. In section 3, we extend the potential reduction method using the NT direction for solving QSDP. Section 4 is devoted to the problem how to solve the linear system which is encountered in every iteration.

2 Optimal conditions

In this section, we study the optimal conditions for QSDP. Our analyses are mainly based on the studies on the following nonlinear system:

$$\begin{aligned} \langle A_i, X\rangle &= b_i,\ i=1,\dots,m, && (2.1)\\ \sum_i^m y_iA_i + S &= C+\varphi(X), && (2.2)\\ XS &= \mu I, \qquad X, S \succ 0, && (2.3) \end{aligned}$$

where $\mu > 0$ is a parameter. The solution of system (2.1)-(2.3) is called the central path for the QSDP. It is proved in the following that (2.1)-(2.3) has a unique solution for any given $\mu > 0$.

We consider the subproblem:

$$\begin{aligned} \min \quad & f^\mu(X) = q(X) - \mu \log\det(X) && (2.4)\\ s.t. \quad & \langle A_i, X\rangle = b_i, \quad i=1,\dots,m, && (2.5)\\ & X \succ 0. && (2.6) \end{aligned}$$

The first-order necessary conditions for the above subproblem are

$$\begin{aligned} C+\varphi(X) - \mu X^{-1} &= \sum_{i=1}^m y_iA_i, && (2.7)\\ \langle A_i, X\rangle &= b_i, \quad i=1,\dots,m, && (2.8)\\ X &\succ 0. && (2.9) \end{aligned}$$

Let $S = C+\varphi(X) - \sum_{i=1}^m y_iA_i$, then (2.7)-(2.9) are equivalent to (2.1)-(2.3). Therefore, (2.1)-(2.3) is also the first order conditions for problem (2.4)-(2.6). Thus, we only need to prove that (2.4)-(2.6) has a unique solution. Direct calculations show that

$$\nabla f^\mu(X) = C+\varphi(X) - \mu X^{-1}, \tag{2.10}$$

and

$$\nabla^2 f^\mu(X)Y = \sum_{i=1}^{\ell} H_iYW_i + \mu X^{-1}YX. \tag{2.11}$$

Recalling the assumption that $H_i, W_i \succeq 0$ and $\mu > 0$, we have that

$$\langle \nabla^2 f^{(\mu)}(X)Y, Y\rangle > 0, \quad \forall\, 0 \neq Y \in \Re^{n\times n}. \tag{2.12}$$

(2.11) and (2.12) imply that $f^\mu(X)$ is strictly convex in the feasible region.

Because of the Slater regularity condition, there exists a pair $(\bar X, \bar y, \bar S) \in \mathcal{F}_d^\circ$ with $\bar X \in \mathcal{F}_p^\circ$, where $\mathcal{F}_p^\circ$ and $\mathcal{F}_d^\circ$ are the sets of the interior points of $\mathcal{F}_p$ and $\mathcal{F}_d$ respectively. Now we prove that the level set

$$\Omega_\mu = \{X \in \mathcal{S}\Re^{n\times n} : f^\mu(X) \le f^\mu(\bar X)\} \tag{2.13}$$

is compact, which guarantees the existence of the solution. For every $X \in \Omega_\mu$,

$$q(X) - \mu \log(\det(X)) \leq f^{(\mu)}(\bar{X}). \tag{2.14}$$

Noting that $\sum_i^m \bar{y}_i A_i + \bar{S} = C + \varphi(\bar{X})$, we have

$$-\langle \varphi(\bar{X}), X\rangle + \frac{1}{2}\langle \varphi(X), X\rangle + \langle \bar{S}, X\rangle - \mu \log\det(X) \leq f^{\mu}(\bar{X}) - b^T \bar{y}. \tag{2.15}$$

The definition of $\varphi(X)$ implies that

$$\begin{aligned} \frac{1}{2}\langle \varphi(X), X\rangle - \langle \varphi(\bar{X}), X\rangle &= \frac{1}{2}\langle \varphi(X-\bar{X}), X-\bar{X}\rangle - \frac{1}{2}\langle \varphi(\bar{X}), \bar{X}\rangle \\ &\geq -\frac{1}{2}\langle \varphi(\bar{X}), \bar{X}\rangle. \end{aligned} \tag{2.16}$$

The fact that $\bar{S} \succ 0$ gives the following inequality

$$\langle \bar{S}, X\rangle \geq \lambda_{min}(\bar{S}) tr(X). \tag{2.17}$$

Now it follows from (2.15)-(2.17) that

$$\lambda_{min}(\bar{S}) tr(X) - \frac{1}{2}\langle \varphi(\bar{X}), \bar{X}\rangle - \mu \log(\det(X)) \leq f^{(\mu)}(\bar{X}) - b^T \bar{y}. \tag{2.18}$$

The above inequality implies that $tr(X)$ is bounded, which guarantees $\|X\|$ is also bounded. Thus, it is easy to see that (2.4)-(2.6) has a unique solution for each $\mu > 0$. Consequently (2.1)-(2.3) has a unique solution, which can be denoted by $(X(\mu), y(\mu), S(\mu))$.

Lemma 2.1 *Assume that $A_1, \dots, A_m$ are linearly independent and Slater regularity condition holds. For any given $\bar{\mu} > 0$, the central path $(X(\mu), y(\mu), S(\mu))\,(0 < \mu < \bar{\mu})$ is bounded.*

Proof Let $(X(\bar{\mu}), y(\bar{\mu}), S(\bar{\mu}))$ be the interior point, and

$$C + \varphi(\bar{X}) = \sum_i^m \bar{y}_i A_i + \bar{S}. \tag{2.19}$$

By (2.1), it can be seen that

$$C + \varphi(X(\mu)) = \sum_{i=1}^m \bar{y}(\mu)_i A_i + S(\mu). \tag{2.20}$$

Combining the above two equalities, we obtain

$$tr((X(\mu) - \bar{X})(S(\mu) - \bar{S})) = \langle \varphi(X(\mu) - \bar{X}), X(\mu) - \bar{X}\rangle \geq 0. \tag{2.21}$$

From (2.3), $X(\mu)S(\mu) = \mu I$, which gives

$$tr(X(\mu)S(\mu)) = n\mu. \tag{2.22}$$

So, we can easily prove the inequality

$$tr(X(\mu)\bar{S}) + tr(\bar{X}S(\mu)) \leq n\bar{\mu} + tr(\bar{X}\bar{S}). \tag{2.23}$$

Using the notation $L = n\bar{\mu} + tr(\bar{X}\bar{S})$, we can see that, from $\bar{S} \succ 0$ and $\bar{X} \succ 0$,

$$\lambda_{min}(\bar{S}) tr(X(\mu)) + \lambda_{min}(\bar{X}) tr(S(\mu)) \leq L. \tag{2.24}$$

Since $S(\mu) \succ 0$ and $X(\mu) \succ 0$, the above relation implies

$$tr(X(\mu)) \leq (\lambda_{min}(\bar{S}))^{-1} L \tag{2.25}$$

and

$$tr(S(\mu)) \leq (\lambda_{min}(\bar{X}))^{-1} L. \tag{2.26}$$

From the fact that $\|X\| \leq tr(X)(if X \succ 0)$, we have

$$\|X(\mu)\| \leq (\lambda_{min}(\bar{S}))^{-1} L \tag{2.27}$$

and

$$\|S(\mu)\| \leq (\lambda_{min}(\bar{X}))^{-1} L. \tag{2.28}$$

The above two relations indicate that $X(\mu)$ and $S(\mu)$ are bounded. From the relation

$$C + \varphi(X(\mu)) = \sum_{i=1}^{m} y(\mu)_i A_i + S(\mu) \tag{2.29}$$

and the assumption that $A_1, \ldots, A_m$ are linearly independent, it is easy to see that $y(\mu) = (y(\mu)_1, \ldots, y(\mu)_m)^T$ is also bounded. **QED**

In particular, we define $\mu_k = \frac{1}{k}(k = 1, 2, \ldots)$. It follows from the above lemma that the sequence $(X(\mu_k), y(\mu_k), S(\mu_k))$ has a convergent subsequence $(X(\mu_{n_k}), y(\mu_{n_k}), S(\mu_{n_k}))$. Assume that

$$\lim_{k \to \infty} (X(\mu_{n_k}), y(\mu_{n_k}), S(\mu_{n_k})) = (X^\star, y^\star, S^\star). \tag{2.30}$$

Therefore, the limit point $(X^\star, y^\star, S^\star)$ satisfies the following relation

$$\langle A_i, x^\star \rangle = b_i, \; i = 1, \ldots, m, \tag{2.31}$$

$$\textstyle\sum_{i=1}^{m} y_i^\star A_i + S^\star = C + \varphi(x^\star), \tag{2.32}$$

$$X^\star S^\star = 0, \tag{2.33}$$

$$X^\star \succeq 0, S^\star \succeq 0. \tag{2.34}$$

Therefore we know that the limit point $X^\star$ is the solution.

Theorem 2.2 *If $A_1, \ldots, A_m$ are linearly independent, and under the Slater regularity condition, (1.9)-(1.11) has solutions. Furthermore, $\varphi \geq 0, X^\star \in \mathcal{F}_d$ is an solution to the problem (1.9)-(1.11) if and only if there exist $y^\star \in \Re^m$ and $S^\star \in \mathcal{S}\Re^{n \times n}$ such that $S^\star \succeq 0$,*

$$(X^\star, y^\star, S^\star) \in \mathcal{F}_d; \tag{2.35}$$

and

$$X^\star S^\star = 0. \tag{2.36}$$

Proof If $X^\star$ is a solution, it follows from first order stationary condition and the complementarity condition that there exist $y^\star \in \Re^m$ and $S^\star(\succeq 0) \in \mathcal{S}\Re^{n \times n}$ such that (2.35) and (2.36) hold. On the other hand, if (2.35)-(2.36) are satisfied, it can be easily seen that for any $X \in \mathcal{F}_d$,

$$\begin{aligned} q(X) - q(X^*) &= \langle C + \varphi(X^*), X - X^* \rangle + \frac{1}{2} \langle \varphi(X - X^*), X - X^* \rangle \\ &\geq \langle C + \varphi(X^*), X - X^* \rangle \quad (\varphi \geq 0) \\ &= \langle C + \varphi(X^*) - \sum_{i=1}^{m} y_i^\star A_i, X - X^* \rangle \\ &= \langle S^*, X - X^* \rangle \\ &= \langle S^*, X \rangle \geq 0. \qquad (X, S^* \succeq 0) \end{aligned} \tag{2.37}$$

Thus, X^* is an optimal solution. **QED**

3 The potential reduction algorithm

In this section, we give a potential reduction interior point method that uses the Nesterov-Todd direction to solve the problem (1.9)-(1.11). Under the Slater regularity condition and the assumption that H_i and W_i are positive definite, our algorithm is a polynomial-time algorithm.

At the beginning of an iteration, the current iterate $(X, y, S) \in \mathcal{F}_d^\circ$, $\mu = \langle X, S\rangle/n$. The potential functions we use are the primal

$$\mathcal{P}_{n+\rho}(X, S) = (n + \rho)\log\langle X, S\rangle - \log\det X, \tag{3.1}$$

and the primal-dual

$$\Psi_{n+\rho}(X, S) = \mathcal{P}_{n+\rho}(X, S) - \log\det S, \tag{3.2}$$

where $\rho > 0$ is a parameter. Similar to the Nesterov-Todd direction for SDP, our search direction (dX, dy, dS) is computed by solving the following system

$$\begin{aligned} D^{-1}dXD^{-1} + dX &= R & (3.3)\\ \langle A_i, dX\rangle &= 0, \qquad i = 1, \ldots, m, & (3.4)\\ \varphi(dX) - \sum_{i=1}^{m}(dy)_i A_i - dS &= 0, & (3.5)\end{aligned}$$

to generate the new point pair $(X^+, y^+, S^+) \in \mathcal{F}_d^\circ$, where D is the so-called scaling matrix:

$$D = X^{\frac{1}{2}}(X^{\frac{1}{2}}SX^{\frac{1}{2}})^{-\frac{1}{2}}X^{\frac{1}{2}} \tag{3.6}$$

and $R = \gamma\mu X^{-1} - S$. It should be noted that here we do not have the relation $\langle dX, dS\rangle = 0$ as in the SDP case. Instead, because φ is positive semi-definite, it follows from (3.5) that

$$\langle dX, dS\rangle = \langle dX, \varphi(dX)\rangle \geq 0. \tag{3.7}$$

By using the transformations

$$\left.\begin{array}{l} dX' = D^{-\frac{1}{2}}dXD^{-\frac{1}{2}}\\ dS' = D^{\frac{1}{2}}dSD^{\frac{1}{2}}\\ A_i' = D^{\frac{1}{2}}A_iD^{\frac{1}{2}}\\ R' = D^{\frac{1}{2}}RD^{\frac{1}{2}}\\ H' = D^{\frac{1}{2}}HD^{\frac{1}{2}}\\ W' = D^{\frac{1}{2}}WD^{\frac{1}{2}}\\ \bar{\varphi}(dX') = \sum_{i=1}^{\ell} H_i' dX' W_i', \end{array}\right\} \tag{3.8}$$

we can rewrite (3.3)-(3.5) into the following reduced system

$$dX' + dS' = R' \tag{3.9}$$

$$\langle A_i', dX'\rangle = 0, \quad i = 1, \ldots, m, \tag{3.10}$$

$$\bar{\varphi}(dX') - \sum_{i=1}^{m}(dy)_i A_i' - dS' = 0. \tag{3.11}$$

Denote

$$V^{\frac{1}{2}} = D^{-\frac{1}{2}}XD^{-\frac{1}{2}} = D^{\frac{1}{2}}SD^{\frac{1}{2}} \succ 0, \tag{3.12}$$

then relation

$$\langle X, S\rangle = \langle I, V\rangle \tag{3.13}$$

holds. In our next section, we will discuss the calculations of the scaling matrix D in (3.6) and the NT direction $(dX^{'}, dy, dS^{'})$ in (3.9)-(3.11). Once when $(dX^{'}, dy, dS^{'})$ are computed, we let

$$\begin{aligned} X^+ &= X + \theta dX, \\ y^+ &= y + \theta dy, \\ S^+ &= S + \theta dS, \end{aligned}$$

where $\theta > 0$ is a step-length such that $(X^+, y^+, S^+) \in \mathcal{F}_d^\circ$ and $\Psi_{n+\rho}(X^+, S^+) < \Psi_{n+\rho}(X, S)$. In order to ensure the above expected property, we need to analyze the reduction in the potential function. Our analysis is based on several famous inequalities.

Lemma 3.1 *Let $X \in \mathcal{S}\Re^{n\times n}, \|X\|_\infty < 1$, then the following inequality:*

$$tr(X) \geq \log\det(I + X) \geq tr(X) - \frac{1}{2}\frac{\|X\|^2}{1 - \|X\|_\infty} \tag{3.14}$$

holds. The norm $\|X\|_\infty$ denotes the maximum absolute of X's eigenvalues and $\|X\| = (tr(X^TX))^{\frac{1}{2}}$.

A proof for the above lemma can be found in Ye(1998).

Lemma 3.2 *Let the pair (dX, dy, dS) be the solution to system (3.3)-(3.5) with the parameter $\gamma = n/(n+\rho), \mu = \langle X, S\rangle/n$. If the step-length is*

$$\theta = \frac{\alpha}{\|V^{-\frac{1}{2}}\|_\infty \|\frac{\langle I,V\rangle}{n+\rho}V^{-\frac{1}{2}} - V^{\frac{1}{2}}\|}, \tag{3.15}$$

where $\alpha \in (0, 1)$ is a constant, then we have

$$(X^+, y^+, S^+) \in \mathcal{F}_d^\circ, \tag{3.16}$$

and

$$\Psi_{n+\rho}(X^+, S^+) - \Psi_{n+\rho}(X, S) \leq -\alpha\frac{\|V^{-\frac{1}{2}} - \frac{n+\rho}{\langle I,V\rangle}V^{\frac{1}{2}}\|}{\|V^{-\frac{1}{2}}\|_\infty} + \frac{\alpha^2}{2(1-\alpha)} + \frac{1}{2}(1 + \rho/n)\alpha^2. \tag{3.17}$$

Proof From the transformations (3.8), it can be seen that

$$\begin{aligned} R^{'} &= \gamma\mu D^{\frac{1}{2}}X^{-1}D^{\frac{1}{2}} - D^{\frac{1}{2}}SD^{\frac{1}{2}} \\ &= \gamma\mu V^{-\frac{1}{2}} - V^{\frac{1}{2}}, \end{aligned}$$

and

$$\gamma\mu = \langle I, V\rangle/(n + \rho). \tag{3.18}$$

So we have

$$\begin{aligned} D^{-\frac{1}{2}}X^+D^{-\frac{1}{2}} &= V^{\frac{1}{2}} + \theta dX^{'} \\ D^{\frac{1}{2}}S^+D^{\frac{1}{2}} &= V^{\frac{1}{2}} + \theta dS^{'}. \end{aligned} \tag{3.19}$$

Because $dX^{'} + dS^{'} = R^{'}$ and $\langle dX^{'}, dS^{'}\rangle \geq 0$, it holds that

$$\begin{aligned} \|dX^{'}\|^2 &= \langle dX^{'}, dX^{'}\rangle \\ &\leq \langle dX^{'}, R^{'}\rangle \\ &\leq \|dX^{'}\|\|R^{'}\|, \end{aligned}$$

which implies that

$$\|dX'\| \le \|R'\|. \tag{3.20}$$

Similarly, $\|dS'\| \le \|R'\|$. From the choice of θ in (3.15), it is easy to show that

$$X^+, S^+ \succ 0. \tag{3.21}$$

Now we consider the reductions in the potential functions.

$$\begin{aligned}(n+\rho)\log\langle X^+, S^+\rangle &- (n+\rho)\log\langle X, S\rangle = (n+\rho)[\log\langle V^{\frac{1}{2}} + \theta dX', V^{\frac{1}{2}} + \theta dS'\rangle - \log\langle I, V\rangle] \\ &= (n+\rho)\log(1 + \frac{\theta\langle V^{\frac{1}{2}}, R'\rangle + \theta^2\langle dX', dS'\rangle}{\langle I, V\rangle}) \\ &\le \frac{(n+\rho)\theta}{\langle I, V\rangle}(\langle V^{\frac{1}{2}}, R'\rangle + \theta\langle dX', dS'\rangle). \end{aligned} \tag{3.22}$$

We also estimate the reduction in $\log\det(X)$

$$\begin{aligned}\log\det(X^+) &- \log\det(X) = \log\det(I + \theta X^{-1}dX) \\ &\ge \theta tr(X^{-1}dX) - \frac{1}{2}\frac{\|\theta X^{-1}dX\|^2}{1 - \|\theta X^{-1}dX\|_\infty}. \end{aligned} \tag{3.23}$$

Noting that θ is given by (3.15), we have

$$\|\theta V^{-\frac{1}{2}} dX'\|_\infty \le \alpha, \tag{3.24}$$

so

$$\log\det(X^+) - \log\det(X) \ge \theta\langle V^{-\frac{1}{2}}, dX'\rangle - \frac{1}{2}\frac{\|\theta V^{-\frac{1}{2}} dX'\|^2}{1-\alpha}. \tag{3.25}$$

Similarly, we also have the result on S:

$$\log\det(S^+) - \log\det(S) \ge \theta\langle V^{-\frac{1}{2}}, dS'\rangle - \frac{1}{2}\frac{\|\theta V^{-\frac{1}{2}} dS'\|^2}{1-\alpha}. \tag{3.26}$$

Using the above inequalities, we give the estimation of the reduction in $\Psi_{n+\rho}(X, S)$.

$$\begin{aligned}\Psi_{n+\rho}(X^+, S^+) &- \Psi_{n+\rho}(X, S) \le \frac{(n+\rho)}{\langle I, V\rangle}(\theta\langle V^{\frac{1}{2}}, R'\rangle \\ &+ \theta^2\langle dX', dS'\rangle) - \theta\langle V^{-\frac{1}{2}}, dS' + dX'\rangle \\ &+ \frac{1}{2(1-\alpha)}(\|\theta V^{-\frac{1}{2}} dS'\|^2 + \|\theta V^{-\frac{1}{2}} dX'\|^2) \\ &= -\theta\frac{\langle I, V\rangle}{(n+\rho)}\|V^{-\frac{1}{2}} - \frac{n+\rho}{\langle I, V\rangle}V^{\frac{1}{2}}\|^2 \\ &+ \frac{\theta^2(n+\rho)}{\langle I, V\rangle}\langle dX', dS'\rangle \\ &+ \frac{1}{2(1-\alpha)}(\|\theta V^{-\frac{1}{2}} dS'\|^2 + \|\theta V^{-\frac{1}{2}} dX'\|^2). \end{aligned} \tag{3.27}$$

It follows from (3.9) that

$$dX' + dS' = R'. \tag{3.28}$$

The relation $\langle dX', dS'\rangle \ge 0$, the above equation and (3.15) imply that

$$\|\theta V^{-\frac{1}{2}} dS'\|^2 + \|\theta V^{-\frac{1}{2}} dX'\|^2 \le \alpha^2. \tag{3.29}$$

Furthermore, it is obvious to see that

$$\langle dX^{'}, dS^{'} \rangle \leq \frac{1}{2}\|dX^{'} + dS^{'}\|^2 = \frac{1}{2}\|R^{'}\|^2. \tag{3.30}$$

Therefore,

$$\begin{aligned}
\frac{\theta^2(n+\rho)}{\langle I, V\rangle}\langle dX^{'}, dS^{'}\rangle &\leq \frac{\alpha^2}{\|V^{-\frac{1}{2}}\|_\infty^2 \|\frac{\langle I,V\rangle}{n+\rho}V^{-\frac{1}{2}} - V^{\frac{1}{2}}\|^2} \\
&= \frac{\alpha^2(n+\rho)}{\|V^{-\frac{1}{2}}\|_\infty^2 \langle I, V\rangle} \\
&= \frac{1}{2}(n+\rho)\alpha^2[(\frac{1}{\sqrt{\lambda_1}})^2(\lambda_1 + \ldots + \lambda_n)]^{-1} \\
&= \frac{1}{2}\alpha^2 \frac{n+\rho}{\frac{\lambda_1}{\lambda_1} + \ldots + \frac{\lambda_n}{\lambda_1}} \\
&\leq \frac{1}{2}(1+\rho/n)\alpha^2, \qquad (3.31)
\end{aligned}$$

where $0 < \lambda_1 \leq \ldots, \lambda_n$ are the eigenvalues of V. Now inequalities (3.27), (3.29) and (3.31) show that

$$\begin{aligned}
\Psi_{n+\rho}(X^+, S^+) - \Psi_{n+\rho}(X, S) &\leq -\alpha \frac{\|V^{-\frac{1}{2}} - \frac{n+\rho}{\langle I,V\rangle}V^{\frac{1}{2}}\|}{\|V^{-\frac{1}{2}}\|_\infty} \\
&+ \frac{\alpha^2}{2(1-\alpha)} + \frac{1}{2}(1+\rho/n)\alpha^2. \qquad (3.32)
\end{aligned}$$

which completes the proof. **QED**

The above lemma gives the bound for the reduction in the potential function $\Psi_{n+\rho}(X, S)$. The first order term on α in the right hand side of (3.17) depends on V. An estimate to this term is given by the following result, which was proved by Ye (1998).

Lemma 3.3 *Let $V \in \mathcal{S}\Re^{n\times n}, V \succ 0$, and $\rho \geq \sqrt{n}$, then the inequality*

$$\frac{\|V^{-\frac{1}{2}} - \frac{n+\rho}{\langle I,V\rangle}V^{\frac{1}{2}}\|}{\|V^{-\frac{1}{2}}\|_\infty} \geq \sqrt{\frac{3}{4}} \tag{3.33}$$

holds.

From (3.32) and (3.33), it follows that

$$\Psi_{n+\rho}(X^+, S^+) - \Psi_{n+\rho}(X, S) \leq g(\alpha), \tag{3.34}$$

where $g(\alpha) = -\frac{\sqrt{3}}{2}\alpha + \frac{\alpha^2}{2(1-\alpha)} + \frac{1}{2}(1+\rho/n)\alpha^2$. Because $g(0) = 0$ and $g'(0) = -\frac{\sqrt{3}}{2}$, there exists a number $\alpha^* > 0$ such that

$$g(\alpha) \leq -\frac{\sqrt{3}}{4}\alpha, \quad \forall\, 0 < \alpha \leq \alpha^*. \tag{3.35}$$

If $\alpha = \alpha^*$ in every iteration, we have

$$\Psi_{n+\rho}(X^+, S^+) - \Psi_{n+\rho}(X, S) \leq -\frac{\sqrt{3}}{4}\alpha^* < 0. \tag{3.36}$$

From the above analysis, we can construct our algorithm for QSDP (1.9)-(1.11) as follows.

Algorithm 3.4 *Given an initial central path point $(X^\circ, y^\circ, S^\circ) \in \mathcal{F}_d^\circ$ and $X^\circ \in \mathcal{F}_p^\circ$, set $\rho = \sqrt{n}$ and $k = 0$. Given $\epsilon > 0$. While $\langle X^k, S^k\rangle \geq \epsilon$ do*

Step 1 *Set* $(X, S) = (X^k, S^k), \gamma = n/(n+\rho)$.

Step 2 *Apply Algorithm 4.1 to solve system (3.3)-(3.5).*

Step 3 *Let* $(X^{k+1}, y^{k+1}, S^{k+1}) = (X^k, y^k, S^k) + \bar{\theta}(dX, dy, dS)$, *where* $\bar{\theta} = argmin_{\alpha \geq 0} \Psi_{n+\rho}(X + \theta dX, S + \theta dS)$.

Step 4 $k := k + 1$, *and go to Step 1.*

The following theorem gives an upper bound for the number of iterations of the above algorithm.

Theorem 3.5 *Let* $\rho = \sqrt{n}$, *Algorithm 3.4 terminates after at most* $O(\sqrt{n} \log(\langle X^\circ, S^\circ \rangle/\epsilon))$ *iterations, namely the stopping condition*

$$\langle X^k, S^k \rangle < \epsilon \tag{3.37}$$

holds for some $k = O(\sqrt{n} \log(\langle X^\circ, S^\circ \rangle/\epsilon))$.

Proof From Step 3 of the algorithm and (3.36), we have

$$\Psi_{n+\rho}(X^{k+1}, S^{k+1}) \leq \Psi_{n+\rho}(X^k, S^k) - \delta, \tag{3.38}$$

where $\delta = \frac{\sqrt{3}}{4}\alpha^* > 0$. The above inequality leads to

$$\Psi_{n+\rho}(X^k, S^k) \leq \Psi_{n+\rho}(X^\circ, S^\circ) - k\delta. \tag{3.39}$$

Using definitions (3.1) and (3.2), inequality (3.39), the following inequality (see Ye(1998))

$$n \log\langle X, S \rangle - \log\det(XS) \geq n \log n, \tag{3.40}$$

and the assumption that $(X^\circ, y^\circ, S^\circ)$ is on the central path, we can obtain that

$$\begin{aligned}
\rho \log\langle X^k, S^k \rangle &= \Psi_{n+\rho}(X^k, S^k) - n \log\langle X^k, S^k \rangle + \log(\det(X^k S^k)) && (3.41)\\
&\leq \Psi_{n+\rho}(X^\circ, S^\circ) - k\delta - n \log\langle X^k, S^k \rangle + \log(\det(X^k S^k))\\
&\leq \Psi_{n+\rho}(X^\circ, S^\circ) - k\delta - n \log n\\
&= \rho \log\langle X^\circ, S^\circ \rangle - k\delta + n \log\langle X^\circ, S^\circ \rangle - n \log(n) - \log(\det(X^\circ S^\circ))\\
&= \rho \log\langle X^\circ, S^\circ \rangle - k\delta. && (3.42)
\end{aligned}$$

Therefore,

$$\log\langle X^k, S^k \rangle \leq \log\langle X^\circ, S^\circ \rangle - \frac{\delta}{\rho} k, \tag{3.43}$$

which yields

$$\langle X^k, S^k \rangle \leq \langle X^\circ, S^\circ \rangle \exp\left\{-\frac{\delta}{\rho} k\right\}. \tag{3.44}$$

When the right hand of the above inequality is less than ϵ, the algorithm terminates.

Now we consider the number k such that

$$\langle X^\circ, S^\circ \rangle \exp\left\{-\frac{\delta}{\rho} k\right\} \geq \epsilon. \tag{3.45}$$

The above inequality and the fact that $\rho = \sqrt{n}$ imply that

$$\begin{aligned}
k &\leq \frac{\rho}{\delta} \log(\langle X^\circ, S^\circ \rangle/\epsilon)\\
&= \frac{\rho}{\delta} \log(\langle X^\circ, S^\circ \rangle/\epsilon)\\
&= O(\sqrt{n} \log(\langle X^\circ, S^\circ \rangle/\epsilon)). && (3.46)
\end{aligned}$$

Hence, the theorem is true. **QED**

4 The calculations of the search direction

In our algorithm given in the previous section, the search direction (dX, dy, dS) is computed by solving system (3.3)-(3.5). Using transformations (3.8), we only need to solve system (3.9)-(3.11). Substituting (3.9) into (3.11), we rewrite the system as:

$$\bar{\varphi}(dX^{'}) - \sum_{i=1}^{m}(dy)_i A_i^{'} + dX^{'} = R^{'}, \tag{4.1}$$

$$\langle A_i^{'}, dX^{'} \rangle = 0, i = 1, \ldots, m. \tag{4.2}$$

Using the notation (Golub and van Loan (1996))

$$vec(X) = (X(:,1)^T, \ldots, X(:,n)^T), \tag{4.3}$$

we can reformulate system (4.1)-(4.2) as

$$\begin{aligned} (Q+I)vec(dX^{'}) - Ady &= vec(R^{'}) & (4.4)\\ -A^T vec(dX^{'}) &= 0, & (4.5) \end{aligned}$$

where $A = (a_1, \ldots, a_m), a_i = vec(A_i^{'})$, and

$$Q = \begin{pmatrix} \sum_{i=1}^{\ell} W_i^{'}(1,1)H_i^{'} & \ldots & \sum_{i=1}^{\ell} W_i^{'}(1,n)H_i^{'} \\ \vdots & \ldots & \vdots \\ \sum_{i=1}^{\ell} W_i^{'}(n,1)H_i^{'} & \ldots & \sum_{i=1}^{\ell} W_i^{'}(n,n)H_i^{'} \end{pmatrix} \tag{4.6}$$

is a block matrix in $\Re^{n^2 \times n^2}$. (4.4)-(4.5) can be rewritten in matrix form:

$$\begin{pmatrix} Q+\mathrm{I} & -A \\ -A^T & O \end{pmatrix} \begin{pmatrix} vec(dX^{'}) \\ dy \end{pmatrix} = \begin{pmatrix} vec(R^{'}) \\ o \end{pmatrix}. \tag{4.7}$$

Under our assumption that $H_i, W_i \succeq 0$, it can be seen that $Q \succeq 0$. The linearly independence of $\{A_i\}$ implies that A is of full column rank. Therefore the matrix

$$\begin{pmatrix} Q+I & -A \\ -A^T & O \end{pmatrix} \tag{4.8}$$

is nonsingular. Thus, the vectors $vec(dX^{'})$ and dy can be solved from (4.7) uniquely. Here, one would ask whether $dX^{'} = vec^{1}(vec(dX^{'}))$ is symmetric. We claim that the answer is yes. Since (4.3) has a unique solution, we only need to prove that the equivalent form (3.3) has a symmetric solution. Let $\mathrm{A}_{\mathrm{m}+1}^{'}, \ldots, \mathrm{A}_{\frac{1}{2}\mathrm{n}(\mathrm{n}+1)}^{'}$ be the orthonormal basis of the null space of $\mathrm{A}_1^{'}, \ldots, \mathrm{A}_{\mathrm{m}}^{'}$ in $\mathcal{S}\Re^{n \times n}$. (4.2) implies that $dX^{'}$ can be expressed in the form

$$dX^{'} = \sum_{k=m+1}^{\frac{1}{2}n(n+1)} \nu_k \mathrm{A}_{\mathrm{k}}^{'}. \tag{4.9}$$

No matter what $\nu = (\nu_{m+1}, \ldots, \nu_{\frac{1}{2}n(n+1)})^T$ is, $dX^{'}$ is symmetric. Because system (4.1) corresponds to the subproblem :

$$\begin{aligned} \min \quad & \frac{1}{2}\langle \bar{\varphi}(dX^{'}), dX^{'} \rangle + \frac{1}{2}\langle dX^{'}, dX^{'} \rangle - \langle R^{'}, dX^{'} \rangle \\ s.t. \quad & \langle A_i^{'}, dX^{'} \rangle = 0, i = 1, \ldots, m. \end{aligned}$$

It is equivalent to the following subproblem:

$$\min \quad \frac{1}{2}\nu^T M\nu + c^T\nu \tag{4.10}$$

$$s.t. \quad \nu \in \Re^{\frac{1}{2}n(n+1)-m}, \tag{4.11}$$

where $c_i = \langle R^{'}, A_i^{'}\rangle$, and

$$M(i,j) = \langle A_i^{'}, A_j^{'}\rangle + \sum_{k=1}^{\ell}\langle (H_k^{'} A_i^{'} W_k^{'}, A_j^{'}\rangle,$$

for $i, j = 1+m, \ldots, \frac{1}{2}n(n+1)$. The facts that $\varphi \succeq 0$ and $A_{m+1}, \ldots, \mathrm{A}^{'}_{\frac{1}{2}\mathrm{n(n+1)}}$ are linearly independent indicate that $\mathrm{M} \succ 0$. So subproblem (4.10) has a unique solution ν^*. Let

$$d\hat{X}' = \sum_{k=m+1}^{\frac{1}{2}n(n+1)} \nu_k^* A_k^{'}, \tag{4.12}$$

$\hat{dy}$ be the solution of the following linear system:

$$\begin{pmatrix} \langle A_1^{'}, A_1^{'}\rangle & \ldots & \langle A_1^{'}, A_m^{'}\rangle \\ \vdots & \ldots & \vdots \\ \langle A_1^{'}, A_1^{'}\rangle & \ldots & \langle A_m^{'}, A_m^{'}\rangle \end{pmatrix} \hat{dy} = \begin{pmatrix} \langle R^{'} - d\hat{X}' - \bar{\varphi}(dX^{'}), A_1^{'}\rangle \\ \vdots \\ \langle R^{'} - d\hat{X}' - \bar{\varphi}(dX^{'}), \mathrm{A}_{\mathrm{m}}^{'}\rangle \end{pmatrix}. \tag{4.13}$$

We can see that $(d\hat{X}', \hat{dy})$ is a solution to system (4.1)-(4.2) and $d\hat{X}'$ is symmetric.

In our algorithm, we also need to compute the scaling matrix $D = X^{\frac{1}{2}}(X^{\frac{1}{2}}SX^{\frac{1}{2}})^{-\frac{1}{2}}X^{\frac{1}{2}}$. We here adopt the technique introduced by Todd, Toh and Tütüncü[8]. Let the Cholesky factorizations of the matrices X and S be

$$X = L_1 L_1^T, \quad \mathrm{S} = \mathrm{L}_2\mathrm{L}_2^{\mathrm{T}}. \tag{4.14}$$

we can compute the SVD of $L_2^T L_1$:

$$L_2^T L_1 = U\Lambda V^T. \tag{4.15}$$

Define $J = \mathrm{L}_1^{-1}\mathrm{X}^{\frac{1}{2}}$, which is an orthogonal matrix. From the fact that

$$\begin{aligned} X^{\frac{1}{2}}SX^{\frac{1}{2}} &= J^T(L_1^T L_2)(L_2^T L_1)J \\ &= (J^T V)\Lambda^2(V^T J), \end{aligned}$$

it follows

$$(X^{\frac{1}{2}}SX^{\frac{1}{2}})^{-\frac{1}{2}} = (J^T V)\Lambda^{-1}(V^T J). \tag{4.16}$$

Thus, $D = L_1 V\Lambda^{-1}V^T L_1 = GG^T$, where $G = L_1 V\Lambda^{-\frac{1}{2}}$. Thus, the following algorithm can be used to solve system (3.3)-(3.5).

Algorithm 4.1

Step 1 *Compute* $X = L_1L_1^T$, $S = L_2L_2^T$, *and* $L_2^T L_1 = U\Lambda V^T$.

Step 2 $G = L_1 V\Lambda^{-\frac{1}{2}}, D = GG^T$

Step 3 *Solve system (4.7).*
Let $dX^{'} = vec^{-1}(vec(dX^{'}))$ *and* $dS^{'} = R^{'} - dX^{'}$.

Step 4 *Compute dX,dS from (3.8), Stop.*

In the above algorithm, Step 1 and Step 2 are used to compute the scaling matrix D. Step 3 and Step 4 are to solve system (3.3)-(3.5). The above algorithm is used in Step 2 of Algorithm 3.4 for computing the search direction.

5 Conclusion

An extended SDP with a quadratic term in the objective function is studied in this paper. This work can be viewed as a generalization of quadratic programming just as semidefinite programming being a generalization of linear programming. We have given an interior point method that uses the Nesterov-Todd search direction for the extended problem. The polynomial-time property of the algorithm is proved.

References

[1] F. Alizadeh, Interior point methods in semidefinite programming with application to combinatorial optimization, *SIAM J. Opt.* 5(1995) 13-51.

[2] F. Alizadeh, J.-P.A. Haeberly, M.L. Overton, Primal-dual interior point methods for semidefinite programming, Working Paper, RUTCOR Rutgers

[3] K. Fujisawa, M. Kojima and K. Nakata, Exploiting sparsity in primal-dual interior-point methods for semidefinite programming, *Math. Prog.* 79(1997) 235-253.

[4] G.H. Golub and C.F. Van Loan, *Matrix Computations, 3rd ed.* (Johns Hopskins University Press, Baltimore, 1996).

[5] E.de Klerk, *Interior Point Methods for SDP*, Ph.D thesis, Delft University of Technology, Delft, The Netherlands, 1997.

[6] M. Kojima, S. Shindoh and S. Hara, Interior-point methods for the monotone semidefinite linear complementarity problems, *SIAM J. Opt.* 7(1997) 86-125.

[7] H.D. Qi and X. Chen, On stationary points of merit functions for semi-definite complementarity problems, Report, ICMSEC, Chinese Academy of Sciences, 1997.

[8] M.J. Todd, K.C. Toh and R.H. Tütüncü, On the Nesterov-Todd Direction in semidefinite programming, Technical Report No. 1154, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, N.Y., 1996.

[9] P. Tseng, Search directions and convergence analysis of some infeasible path-following methods for the monotone semi-definite LCP, *Optimization Methods and Software*, 9(1998) 245-268.

[10] YinYu Ye, *Interior Point Algorithm Theory and Analysis*, (John Wiley and Sons, New York, 1998).

[11] Y. Zhang, On extending primal-dual interior-point algorithms from linear programming to semidefinite programming, to appear in *SIAM J. Opt.*

[12] Q. Zhao, S.E. Karische, F. Rendl and H. Wolkowicz, Semidefinite programming relaxations for the quadratic assignment problem, CORR Rept 95-27, Dept of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, Canada, 1996.