

# A Modified Alternating Direction Method for Convex Quadratically Constrained Quadratic Semidefinite Programs

Jie Sun\* and Su Zhang<sup>†</sup>

July 14, 2010

## Abstract

We propose a modified alternate direction method for solving convex quadratically constrained quadratic semidefinite optimization problems. The method is a first-order method, therefore requires much less computational effort per iteration than the second-order approaches such as the interior point methods or the smoothing Newton methods. In fact, only a single inexact metric projection onto the positive semidefinite cone is required at each iteration. We prove global convergence and provide numerical evidence to show the effectiveness of this method.

**Key Words:** Alternating direction method, Conic programming, Quadratic semidefinite optimization

---

\*Corresponding author. Department of Decision Sciences and Risk Management Institute, National University of Singapore, Singapore 117245. Email:jsun@nus.edu.sg, Fax: +65-6779-2621.

<sup>†</sup>Department of Decision Sciences, National University of Singapore and Institute of Modern Management, Business School, Nankai University, PRC. Email:zhangsu@nankai.edu.cn.

## 1. Introduction

Let  $\mathbb{S}^n$  be the space of all  $n \times n$  real symmetric matrices. We consider a special class of nonlinear optimization problems defined over  $\mathbb{S}^n$ , called *Convex Quadratically Constrained Quadratic Semidefinite Programs* (CQCQSDPs), as follows.

$$\begin{aligned} \min \quad & q_0(X) \equiv \frac{1}{2} \langle X, Q_0(X) \rangle + \langle B_0, X \rangle \\ \text{s.t.} \quad & q_i(X) \equiv \frac{1}{2} \langle X, Q_i(X) \rangle + \langle B_i, X \rangle + c_i \leq 0, \quad i = 1, \dots, m \\ & X \succeq 0, \end{aligned} \tag{1.1}$$

where  $Q_i : \mathbb{S}^n \rightarrow \mathbb{S}^n, i = 0, 1, \dots, m$ , is a self-adjoint positive semidefinite linear operator;  $X, B_i \in \mathbb{S}^n$ , and  $c_i \in \mathbb{R}$  is a scalar. In addition,  $\langle \cdot, \cdot \rangle$  denotes the Frobenius inner product between square matrices, i.e.,  $\langle U, V \rangle = \text{Trace}(U^T V)$ . We denote by  $\mathbb{S}_+^n$  and  $\mathbb{S}_{++}^n$  the cones of positive semidefinite matrices and positive definite matrices, respectively. By writing  $X \succeq 0$  ( $X \succ 0$ , respectively) we mean that  $X \in \mathbb{S}_+^n$  ( $X \in \mathbb{S}_{++}^n$ , respectively).

Problem (1.1) is a convex programming problem in the symmetric matrix space. In a sense, it is also the most basic nonlinear semidefinite optimization problem. It has a number of important applications in engineering and management. For example, in order to find a positive semidefinite matrix that best approximates the solution to the matrix equation system

$$\langle A_i, X \rangle = b_i, \quad \forall i = 1, \dots, m,$$

we need to solve the matrix least-square problem

$$\min \quad \sum_{i=1}^m (\langle A_i, X \rangle - b_i)^2 \quad \text{s.t.} \quad X \succeq 0, \tag{1.2}$$

which is in the form of Problem (1.1). Another example is the following nearest covariance matrix problem in finance,

$$\begin{aligned} \min \quad & \|X - C\|^2 \\ \text{s.t.} \quad & \langle A_i, X \rangle = b_i, \quad i = 1, \dots, p, \\ & \langle A_i, X \rangle \leq b_i, \quad i = p + 1, \dots, m, \\ & X \text{ is a covariance matrix.} \end{aligned}$$

Since all covariance matrices are positive semidefinite. The resulting problem then becomes a special case of (1.1).

In [1], Beck studied quadratic matrix programming of order  $r$  which may not be convex. He constructed a special semidefinite relaxation and its dual and showed that under some mild conditions strong duality holds for the relaxed problem with at most  $r$  constraints. However, Beck's model does not include the semidefinite cone constraint. Therefore, it is essentially a vector optimization model, rather than a semidefinite optimization problem like (1.1).

A special case of (1.1) is the convex quadratic SDP (CQSDP) problem where the quadratic term only appears in the objective and the constraints are linear, together with the semidefinite cone constraint. In [15], a theoretical primal-dual potential reduction algorithm was proposed for solving CQSDP problems by Nie and Yuan. The authors suggested to use the conjugate gradient

method to compute an approximate search direction. Subsequent works include Qi and Sun [17] and Toh [20]. Qi and Sun used a Lagrangian dual approach. Toh introduced an inexact primal-dual path-following method with three classes of pre-conditioners for the augmented equation for fast convergence under suitable nondegeneracy assumptions. Unfortunately, all these new methods cannot be readily extended to solve (1.1) because the KKT conditions of (1.1) can no longer be written as a linear system as in CQSDP.

In two recent papers, Malick [13] and Boyd and Xiao [2], respectively applied classical quasi-Newton methods (in particular, the BFGS method) and the projected gradient method to the dual problem of CQSDP. More recently, Gao and Sun [8] designed an inexact smoothing Newton method to solve a reformulated semismooth system with two level metric projection operators and showed high efficiency of the proposed method in numerical experiments. In her thesis [23], Zhao proposed a semismooth Newton-CG augmented Lagrangian method for large scale CQSDPs. Again, all these algorithms are not applicable to CQCQSDPs because of the existence of quadratic constraints.

In the field of general nonlinear semidefinite optimization, we note the smoothing Newton algorithm of Sun, Sun, and Qi [18] and the augmented Lagrangian algorithm of Sun, Sun, and Zhang [19]. Although both algorithms could be used to solve (1.1), they are not specifically designed for CQCQSDP. Therefore, they may not be able to take advantage of the specific structure of the problem. It is also noted that the smoothing Newton method is a second-order algorithm while the proposed algorithm in this paper is a first-order method. Therefore, in each iteration, the smoothing Newton method requires much more computations than our algorithm. Similar comparison can be also made between the augmented Lagrangian method and our algorithm. It should be noted that the successive linearization method of Li and Sun [12] is also a first-order method, but it requires to solve a quadratic sub-problem at each iteration. Besides, it is targeted for general nonlinear semidefinite programs.

The general advantage of the first-order algorithms is twofold. First, this type of methods are relatively simple to implement, thus they are useful in finding an approximate solution of the problems, which may become the “first phase” of a hybrid second-order algorithm (e.g., a Newton method). Secondly, first-order methods usually require much less computation per iteration, therefore might be suitable for relatively large problems.

This paper is focused on analysis and test of an alternating direction method (ADM) for CQCQSDP. The ADM has been an effective first-order approach for solving large optimization problems with vector variables. It was probably first considered by Gabay [6] and Gabay and Mercier [7]. Further studies of such methods can be found, for instance, in [3, 5, 9]. The inexact versions of the ADM were proposed by Eckstein and Bertsekas [5] and Chen and Teboulle [3], respectively. He et al. [9] generalized the framework and proposed a new inexact ADM method with flexible conditions for structured monotone variational inequalities. All of the work above, however, was devoted to vector optimization problems. It appears to be new to apply the idea of ADM to develop a method for solving quadratic semidefinite programming problems.

In [22], Yu applied an exact ADM for SDP. His main idea is to reformulate the primal-dual optimality conditions of SDP as a projection equation. However, there is a significant difficulty in applying Yu’s idea to CQCQSDP because one has to solve a nonlinear variational problem over the semidefinite cone at each iteration, which is very expensive. In this paper we propose a specially designed ADM for CQCQSDP. The new method has the advantage of being simple

and cheap in computation. At each iteration only one projection onto the semidefinite cone is necessary. Furthermore, this projection is allowed to be inexact. Under some mild conditions, we prove global convergence of this method.

After the first draft of this paper has been submitted, a referee kindly pointed us to two recent papers by Malick et al. [14] and Wen et al. [21], respectively. Both papers use the idea of quadratic regularity (e.g. augmented Lagrangian) combined with first-order numerical techniques (e.g., ADM) in the inner loop of the algorithms, therefore bear some similarity to our proposed algorithm. However, there are some major differences. First, the methods of Malick et al. and Wen et al. are aimed at (linear) SDPs while our algorithm is aimed at CQCQSDPs; Second, both algorithms in [14, 21] solve the dual form of the problem and take advantage of the efficient solvability of the dual problem. For example, an explicit formula is given to compute the optimal point of the dual augmented Lagrangian of SDPs in the paper of Wen et al. This advantage may not exist if the same idea is applied to CQCQSDPs. It should be also noted that our proposed method directly solves Problem (1.1). Hence it is a primal approach.

The paper is organized as follows. In Section 2 we reformulate Problem (1.1) as a variational inequality problem and present a specialization of the ADM, called the “original ADM”. We then simplify the original ADM into a “modified ADM”, which takes into account the special structure of the reformulated variational inequality problem. In Section 3 we prove the convergence of the modified ADM. Section 4 is devoted to the inexact case and its convergence. We introduce the testing problems and report our numerical results in Section 5, and make some concluding remarks in Section 6.

## 2. The Algorithms

Recall that  $q_i(X) \equiv \frac{1}{2} \langle X, Q_i(X) \rangle + \langle B_i, X \rangle + c_i \leq 0$ . By introducing artificial constraints

$$Y_i = X \quad \text{and} \quad \Omega_i = \{Y_i : q_i(Y_i) \leq 0\}, \quad \forall i = 1, \dots, m, \quad (2.1)$$

we may rewrite (1.1) equivalently as

$$\begin{aligned} \min \quad & q_0(X) \\ \text{s.t.} \quad & X = Y_i, \quad Y_i \in \Omega_i, \quad i = 1, \dots, m \\ & X \succeq 0. \end{aligned} \quad (2.2)$$

The Lagrange dual of Problem (2.2) is

$$\max_{\lambda_i} \min_{X \succeq 0, Y_i \in \Omega_i} L(X, Y_1, \dots, Y_m, \lambda_1, \dots, \lambda_m) \equiv q_0(X) - \sum_{i=1}^m \langle \lambda_i, X - Y_i \rangle.$$

Notice that the Lagrange multipliers  $\lambda_i$ ,  $i = 1, \dots, m$ , are symmetric matrices. It is well known that under mild constraint qualifications (e.g., Slater’ condition), strong duality holds and hence,  $X^*$  is an optimal solution of (2.2) if and only if there exists  $\lambda_i^* \in \mathbb{S}^n$ ,  $i = 1, \dots, m$  such that  $(X^*, Y_i^*, \lambda_i^*)$  satisfies the following KKT system in variational inequality form

$$\left\{ \begin{array}{l} \left\langle X - X^*, Q_0(X^*) + B_0 - \sum_{i=1}^m \lambda_i^* \right\rangle \geq 0, \quad \forall X \in \mathbb{S}_+^n \\ \langle Y_i - Y_i^*, \lambda_i^* \rangle \geq 0, \quad \forall Y_i \in \Omega_i, \quad i = 1, \dots, m \\ X^* = Y_i^*, \quad i = 1, \dots, m. \end{array} \right. \quad (2.3)$$

Problem (2.3) is a variational inequality problem with a special structure. The variables  $(X, Y_i, \lambda_i)$  are symmetric matrices, the underlying sets  $\mathbb{S}_+^n$  and  $\Omega_i$  are convex and, in general, non-polyhedral. The alternating direction method, when applied to Problem (2.3), can be separated into three steps. The motivation of these steps is to solve the variational inequalities in (2.3) consecutively and iteratively, using the most recent updates of the variables. In this sense, the ADM can be seen as the block Gauss-Seidel variants of the augmented Lagrangian approach.

**Algorithm 2.1. The Original Alternating Direction Method for CQCQSDP**

*Step 1.*  $(X^k, Y_i^k, \lambda_i^k) \rightarrow (X^{k+1}, Y_i^k, \lambda_i^k)$ , where  $\beta_i$  is certain positive scalar and  $X^{k+1}$  satisfies

$$\left\langle X - X^{k+1}, Q_0(X^{k+1}) + B_0 - \sum_{i=1}^m (\lambda_i^k - \beta_i (X^{k+1} - Y_i^k)) \right\rangle \geq 0, \quad \forall X \succeq 0. \quad (2.4)$$

*Step 2.*  $(X^{k+1}, Y_i^k, \lambda_i^k) \rightarrow (X^{k+1}, Y_i^{k+1}, \lambda_i^k)$ ,  $i = 1, \dots, m$ , where  $Y_i^{k+1}$  satisfies

$$\left\langle Y_i - Y_i^{k+1}, \lambda_i^k - \beta_i (X^{k+1} - Y_i^{k+1}) \right\rangle \geq 0, \quad \forall Y_i \in \Omega_i. \quad (2.5)$$

*Step 3.*  $(X^{k+1}, Y_i^{k+1}, \lambda_i^k) \rightarrow (X^{k+1}, Y_i^{k+1}, \lambda_i^{k+1})$ ,  $i = 1, \dots, m$ , where

$$\lambda_i^{k+1} = \lambda_i^k - \beta_i (X^{k+1} - Y_i^{k+1}). \quad (2.6)$$

At Step 1 and Step 2 we should solve variational inequalities. In the following, we will convert them to simple projection operations. For this purpose, all we need is the following fact from convex analysis.

**Lemma 2.2.** ([11] Theorem 2.3) *Let  $\Omega$  be a closed convex set in a Hilbert space and let  $P_\Omega(x)$  be the projection of  $x$  onto  $\Omega$ . Then*

$$\langle z - y, y - x \rangle \geq 0, \quad \forall z \in \Omega \iff y = P_\Omega(x). \quad (2.7)$$

Taking  $x = Y_i^{k+1} - \alpha_i (\lambda_i^k - \beta_i (X^{k+1} - Y_i^{k+1}))$ ,  $y = Y_i^{k+1}$ , and  $z = Y_i$  in (2.7), we see that (2.5) is equivalent to the following nonlinear equation

$$Y_i^{k+1} = P_{\Omega_i} \left[ Y_i^{k+1} - \alpha_i (\lambda_i^k - \beta_i (X^{k+1} - Y_i^{k+1})) \right],$$

where  $\alpha_i$  can be any positive number. Thus by choosing  $\alpha_i = \frac{1}{\beta_i}$ , the right hand side item  $Y_i^{k+1}$  is cancelled. Therefore, in order to solve (2.5), we only have to compute

$$Y_i^{k+1} = P_{\Omega_i} \left[ X^{k+1} - \frac{1}{\beta_i} \lambda_i^k \right]. \quad (2.8)$$

However, this trick does not work for (2.4) due to the existence of the term  $Q_0(X^{k+1})$ . We therefore suggest the following approximate approach. For a certain constant  $\gamma$ , let

$$R(X^k, X^{k+1}) \equiv Q_0(X^{k+1}) - Q_0(X^k) - \gamma (X^{k+1} - X^k)$$

be the residual (the difference) between  $Q_0(X^{k+1})$  and its linearization at  $X^k$ .

Now instead of computing

$$X^{k+1} = P_{\mathbb{S}_+^n} \left[ X^{k+1} - \alpha \left( Q_0 (X^{k+1}) + B_0 - \sum_{i=1}^m (\lambda_i^k - \beta_i (X^{k+1} - Y_i^k)) \right) \right],$$

we compute

$$\begin{aligned} X^{k+1} &= P_{\mathbb{S}_+^n} \left[ X^{k+1} - \alpha \left( Q_0 (X^{k+1}) + B_0 - \sum_{i=1}^m (\lambda_i^k - \beta_i (X^{k+1} - Y_i^k)) - R (X^k, X^{k+1}) \right) \right] \\ &= P_{\mathbb{S}_+^n} \left[ X^{k+1} - \alpha \left( \left( \sum_{i=1}^m \beta_i + \gamma \right) X^{k+1} + B_0 - \sum_{i=1}^m (\lambda_i^k + \beta_i Y_i^k) - \gamma X^k + Q_0 (X^k) \right) \right] \end{aligned} \quad (2.9)$$

We choose  $\gamma$  so that  $\gamma > \lambda_{\max}(Q_0)$ , where  $\lambda_{\max}(Q_0)$  is the largest eigenvalue of  $Q_0$ . The reason for this choice will be clear in the proof of Theorem 3.2 below. Setting

$$\alpha = \left( \sum_{i=1}^m \beta_i + \gamma \right)^{-1} \quad \text{and} \quad D_k = B_0 - \sum_{i=1}^m (\lambda_i^k + \beta_i Y_i^k) - \gamma X^k + Q_0 (X^k),$$

we obtain

$$X^{k+1} = P_{\mathbb{S}_+^n} [-\alpha D_k], \quad (2.10)$$

which will be used as an approximation to the solution of variational inequality (2.4). In summary, the modified alternating direction method is given as follows.

**Algorithm 2.3. The Modified Alternating Direction Method for CQCQSDP**

*Step 1.*  $(X^k, Y_i^k, \lambda_i^k) \rightarrow (X^{k+1}, Y_i^k, \lambda_i^k)$ , where  $\beta_i > 0$  and

$$X^{k+1} = P_{\mathbb{S}_+^n} \left[ - \left( \sum_{i=1}^m \beta_i + \gamma \right)^{-1} D_k \right]. \quad (2.11)$$

*Step 2.*  $(X^{k+1}, Y_i^k, \lambda_i^k) \rightarrow (X^{k+1}, Y_i^{k+1}, \lambda_i^k)$ ,  $i = 1, \dots, m$ , where

$$Y_i^{k+1} = P_{\Omega_i} \left[ X^{k+1} - \frac{1}{\beta_i} \lambda_i^k \right]. \quad (2.12)$$

*Step 3.*  $(X^{k+1}, Y_i^{k+1}, \lambda_i^k) \rightarrow (X^{k+1}, Y_i^{k+1}, \lambda_i^{k+1})$ ,  $i = 1, \dots, m$ , where

$$\lambda_i^{k+1} = \lambda_i^k - \beta_i (X^{k+1} - Y_i^{k+1}). \quad (2.13)$$

From Step 3, we could interpret  $\beta_i$  as the dual stepsize. In our computational test, it is set as one although its choice is flexible.

In order to solve (2.3) by this modified alternating direction method, we need to compute the metric projection of a matrix onto  $\Omega_i$  and  $\mathbb{S}_+^n$ . The projection onto  $\Omega_i$  can be computed in the same way as computing the Euclidean projection of a vector onto an ellipsoid in the real vector space. Therefore, the computation of this projection can be very fast, see, for example, [4] for the corresponding algorithms. The projection onto  $\mathbb{S}_+^n$  requires a full (or at least, positive) eigenvalue decomposition. To handle potentially large problems, we allow this computation to be performed inexactly according to the result in Section 4.

### 3. Convergence Results

**Proposition 3.1.** *The sequence  $\{X^k, Y_i^k, \lambda_i^k\}$  generated by the modified alternating direction method for CQCQSDP satisfies*

$$\sum_{i=1}^m \frac{1}{\beta_i} \langle \lambda_i^{k+1} - \lambda_i^*, \lambda_i^k - \lambda_i^{k+1} \rangle + \sum_{i=1}^m \beta_i \langle Y_i^{k+1} - Y_i^*, Y_i^k - Y_i^{k+1} \rangle + \langle X^{k+1} - X^*, R(X^k, X^{k+1}) \rangle \geq 0, \quad (3.1)$$

where  $(X^*, Y_i^*, \lambda_i^*)$  are defined as in (2.3).

**Proof.** Using (2.3) and (2.5), we have

$$\langle Y_i^{k+1} - Y_i^*, \lambda_i^* - \lambda_i^{k+1} \rangle \geq 0. \quad (3.2)$$

Similarly, we get

$$\langle Y_i^{k+1} - Y_i^k, \lambda_i^k - \lambda_i^{k+1} \rangle \geq 0. \quad (3.3)$$

Note that (2.9) can be written equivalently as

$$\left\langle X - X^{k+1}, Q_0(X^{k+1}) + B_0 - \sum_{i=1}^m \lambda_i^{k+1} - \sum_{i=1}^m \beta_i (Y_i^k - Y_i^{k+1}) - R(X^k, X^{k+1}) \right\rangle \geq 0, \quad \forall X \in \mathbb{S}_+^n,$$

in view of the relationship (2.7) and (2.13). Setting  $X = X^*$  in it, we obtain

$$\left\langle X^{k+1} - X^*, -Q_0(X^{k+1}) - B_0 + \sum_{i=1}^m \lambda_i^{k+1} + \sum_{i=1}^m \beta_i (Y_i^k - Y_i^{k+1}) + R(X^k, X^{k+1}) \right\rangle \geq 0. \quad (3.4)$$

Let  $X = X^{k+1}$  in the first inequality of (2.3). Then

$$\left\langle X^{k+1} - X^*, Q_0(X^*) + B_0 - \sum_{i=1}^m \lambda_i^* \right\rangle \geq 0. \quad (3.5)$$

Adding (3.4) and (3.5) together, we have

$$\begin{aligned} & \left\langle \sum_{i=1}^m (\lambda_i^{k+1} - \lambda_i^*), X^{k+1} - X^* \right\rangle + \left\langle \sum_{i=1}^m \beta_i (Y_i^k - Y_i^{k+1}), X^{k+1} - X^* \right\rangle \\ & + \left\langle X^{k+1} - X^*, R(X^k, X^{k+1}) \right\rangle \geq \left\langle X^{k+1} - X^*, Q_0(X^{k+1}) - Q_0(X^*) \right\rangle \geq 0. \end{aligned} \quad (3.6)$$

It follows from (3.2), (3.3), and (3.6) that

$$\begin{aligned} & \left\langle \sum_{i=1}^m (\lambda_i^{k+1} - \lambda_i^*), X^{k+1} - X^* \right\rangle + \left\langle \sum_{i=1}^m \beta_i (Y_i^k - Y_i^{k+1}), X^{k+1} - X^* \right\rangle \\ & + \left\langle X^{k+1} - X^*, R(X^k, X^{k+1}) \right\rangle + \sum_{i=1}^m \langle Y_i^{k+1} - Y_i^*, \lambda_i^* - \lambda_i^{k+1} \rangle + \sum_{i=1}^m \langle Y_i^{k+1} - Y_i^k, \lambda_i^k - \lambda_i^{k+1} \rangle \\ & = \sum_{i=1}^m \frac{1}{\beta_i} \langle \lambda_i^{k+1} - \lambda_i^*, \lambda_i^k - \lambda_i^{k+1} \rangle + \sum_{i=1}^m \beta_i \langle Y_i^{k+1} - Y_i^*, Y_i^k - Y_i^{k+1} \rangle \\ & + \left\langle X^{k+1} - X^*, R(X^k, X^{k+1}) \right\rangle \geq 0. \quad \square \end{aligned}$$

The following is the main result of this section.

**Theorem 3.2.** *The sequence  $\{X^k\}$  generated by the modified alternating direction method converges to a solution point  $X^*$  of system (2.3).*

**Proof.** We denote

$$W \equiv \begin{pmatrix} X \\ Y_i \\ \lambda_i \end{pmatrix} \quad \text{and} \quad G \equiv \begin{pmatrix} \gamma I - Q_0 & 0 & 0 \\ 0 & H & 0 \\ 0 & 0 & H^{-1} \end{pmatrix},$$

where  $H$  is a diagonal matrix with  $\beta_i I$  blocks on its diagonal. Clearly,  $G$  is positive definite. We therefore define the  $G$ -inner product of  $W$  and  $W'$  as

$$\langle W, W' \rangle_G \equiv \sum_{i=1}^m \frac{1}{\beta_i} \langle \lambda_i, \lambda'_i \rangle + \sum_{i=1}^m \beta_i \langle Y_i, Y'_i \rangle + \langle X, \gamma X' - Q_0(X') \rangle$$

and the associated  $G$ -norm as

$$\|W\|_G \equiv \left( \sum_{i=1}^m \frac{1}{\beta_i} \|\lambda_i\|^2 + \sum_{i=1}^m \beta_i \|Y_i\|^2 + \|X\|_{\gamma I - Q_0}^2 \right)^{\frac{1}{2}},$$

where  $\|X\|_{\gamma I - Q_0}^2 = \gamma \|X\|^2 - \langle X, Q_0(X) \rangle$  (This explains why we take  $\gamma > \lambda_{\max}(Q_0)$  in (2.9)). Observe that, by Lemma 2.2, solving the optimal condition (2.3) for Problem (2.2) is equivalent to finding a zero point of the residual function

$$\|e(W)\|_G \equiv \left\| \begin{array}{c} X - P_{\mathbb{S}_+^n} \left[ X - \alpha \left( Q_0(X) + B_0 - \sum_{i=1}^m \lambda_i \right) \right] \\ Y_i - P_{\Omega_i} [Y_i - \alpha_i \lambda_i] \\ \beta_i (X - Y_i) \end{array} \right\|_G.$$

Then we have from (2.13) and the first equation in (2.9) that

$$\begin{aligned} X^{k+1} &= P_{\mathbb{S}_+^n} \left[ X^{k+1} - \alpha \left( Q_0(X^{k+1}) + B_0 - \sum_{i=1}^m (\lambda_i^k - \beta_i (X^{k+1} - Y_i^k)) - R(X^k, X^{k+1}) \right) \right] \\ &= P_{\mathbb{S}_+^n} \left[ X^{k+1} - \alpha \left( Q_0(X^{k+1}) + B_0 - \sum_{i=1}^m \lambda_i^{k+1} - \sum_{i=1}^m \beta_i (Y_i^k - Y_i^{k+1}) - R(X^k, X^{k+1}) \right) \right], \end{aligned}$$

which, together with (2.6), (2.8), and the non-expansion property of the projection operator, implies that

$$\|e(W^{k+1})\|_G \leq \left\| \begin{array}{c} \alpha \left( \sum_{i=1}^m \beta_i (Y_i^{k+1} - Y_i^k) + R(X^k, X^{k+1}) \right) \\ 0 \\ \lambda_i^k - \lambda_i^{k+1} \end{array} \right\|_G \leq \delta \|W^k - W^{k+1}\|_G \quad (3.7)$$

where  $\delta$  is a positive constant depending on  $\alpha$ ,  $\beta_i$ ,  $\gamma$ , and the largest eigenvalue of  $Q_0$ .

Note that (3.1) can be written as

$$\langle W^{k+1} - W^*, W^k - W^{k+1} \rangle_G \geq 0,$$

which implies

$$\langle W^k - W^*, W^k - W^{k+1} \rangle_G \geq \|W^k - W^{k+1}\|_G^2.$$

Thus

$$\begin{aligned}
\|W^{k+1} - W^*\|_G^2 &= \|(W^k - W^*) - (W^k - W^{k+1})\|_G^2 \\
&= \|W^k - W^*\|_G^2 - 2\langle W^k - W^*, W^k - W^{k+1} \rangle_G + \|W^k - W^{k+1}\|_G^2 \\
&\leq \|W^k - W^*\|_G^2 - \|W^k - W^{k+1}\|_G^2 \\
&\leq \|W^k - W^*\|_G^2 - \frac{1}{\delta^2} \|e(W^{k+1})\|_G^2.
\end{aligned} \tag{3.8}$$

From the above inequality, we have

$$\|W^{k+1} - W^*\|_G^2 \leq \|W^k - W^*\|_G^2 \leq \dots \leq \|W^0 - W^*\|_G^2. \tag{3.9}$$

That is, the sequence  $\{W^k\}$  is bounded. Thus there exists at least one cluster point of  $\{W^k\}$ .

It also follows from (3.8) that

$$\sum_{k=0}^{\infty} \frac{1}{\delta^2} \|e(W^{k+1})\|_G^2 < +\infty$$

and thus

$$\lim_{k \rightarrow \infty} \|e(W^k)\|_G = 0.$$

Let  $\overline{W}$  be a cluster point of  $\{W^k\}$ , and let  $\{W^{k_j}\}$  be a corresponding subsequence converging to  $\overline{W}$ . We have

$$\|e(\overline{W})\|_G = \lim_{j \rightarrow \infty} \|e(W^{k_j})\|_G = 0,$$

which means that  $\overline{W}$  is a zero point of the residual function. Therefore  $\overline{W}$  satisfies (2.3). Setting  $W^* = \overline{W}$  in (3.9), we have

$$\|W^{k+1} - \overline{W}\|_G^2 \leq \|W^k - \overline{W}\|_G^2, \quad \forall k \geq 0.$$

Thus, the sequence  $\{W^k\}$  has a unique cluster point and

$$\lim_{k \rightarrow \infty} W^k = \overline{W}.$$

□

#### 4. The Inexact Case

The modified alternating direction method requires to compute the projection onto the semidefinite cone. However, there seems to be little justification of the effort to do it exactly. In fact, inspired by Pang [16] and He et al. [9], we next develop an inexact version of the modified alternating direction method.

##### Algorithm 4.1. The Inexact Version of the Modified Alternating Direction Method

For a given nonnegative sequence  $\{\eta_k\}$  satisfying  $\sum_{k=0}^{\infty} \eta_k^2 < +\infty$ , we solve, in each iteration, (2.11) inexactly as follows

$$X^{k+1} = P_{\mathbb{S}_+^n} \left[ X^{k+1} - \alpha \left( Q_0(X^{k+1}) + B_0 - \sum_{i=1}^m (\lambda_i^k - \beta_i(X^{k+1} - Y_i^k)) - R(X^k, X^{k+1}) + \Theta_k(X^{k+1}) \right) \right], \tag{4.1}$$

where  $\Theta_k$  is an operator such that

$$\left\| \Theta_k \left( X^{k+1} \right) \right\|_{(\gamma I - Q_0)^{-1}} \leq \eta_k \left\| e \left( W^{k+1} \right) \right\|_G. \quad (4.2)$$

We next show that the sequence  $\{X^k, Y_i^k, \lambda_i^k\}$  generated by the inexact version will satisfy a weak contractive property.

**Proposition 4.2.** *Let  $\{X^k, Y_i^k, \lambda_i^k\}$  be the sequence generated by the inexact version of the modified alternating direction method. Then there is a  $k_0 \geq 0$  such that*

$$\left\| W^{k+1} - W^* \right\|_G^2 \leq \left( 1 + 4(\delta')^2 \eta_k^2 \right) \left\| W^k - W^* \right\|_G^2 - \frac{1}{6(\delta')^2} \left\| e \left( W^{k+1} \right) \right\|_G^2, \quad \forall k \geq k_0, \quad (4.3)$$

where  $\delta'$  is a positive constant depending on  $\alpha$ ,  $\beta_i$ ,  $\gamma$ , and the eigenvalues of  $Q_0$ .

**Proof.** Using the relationship of (2.7) and (2.13) in the inexact method, we have

$$\begin{aligned} & \left\langle X - X^{k+1}, Q_0 \left( X^{k+1} \right) + B_0 - \sum_{i=1}^m \lambda_i^{k+1} - \sum_{i=1}^m \beta_i \left( Y_i^k - Y_i^{k+1} \right) \right. \\ & \left. - R \left( X^k, X^{k+1} \right) + \Theta_k \left( X^{k+1} \right) \right\rangle \geq 0, \quad \forall X \in \mathbb{S}_+^n. \end{aligned}$$

Similarly to the proof of Proposition 3.1, by adding inequalities (3.2), (3.3), and (3.5) to the inequality above, we obtain

$$\begin{aligned} & \sum_{i=1}^m \frac{1}{\beta_i} \left\langle \lambda_i^{k+1} - \lambda_i^*, \lambda_i^k - \lambda_i^{k+1} \right\rangle + \sum_{i=1}^m \beta_i \left\langle Y_i^{k+1} - Y_i^*, Y_i^k - Y_i^{k+1} \right\rangle \\ & + \left\langle X^{k+1} - X^*, R \left( X^k, X^{k+1} \right) \right\rangle + \left\langle X^{k+1} - X^*, -\Theta_k \left( X^{k+1} \right) \right\rangle \geq 0. \end{aligned}$$

According to the definitions of  $\langle \cdot, \cdot \rangle_G$  and  $\| \cdot \|_G$ , the inequality above can be written as

$$\left\langle W^{k+1} - W^*, W^k - W^{k+1} \right\rangle_G + \left\langle X^{k+1} - X^*, -\Theta_k \left( X^{k+1} \right) \right\rangle \geq 0,$$

which implies

$$\left\langle W^k - W^*, W^k - W^{k+1} \right\rangle_G \geq \left\| W^k - W^{k+1} \right\|_G^2 + \left\langle X^{k+1} - X^*, \Theta_k \left( X^{k+1} \right) \right\rangle.$$

Thus,

$$\begin{aligned} \left\| W^{k+1} - W^* \right\|_G^2 &= \left\| \left( W^k - W^* \right) - \left( W^k - W^{k+1} \right) \right\|_G^2 \\ &= \left\| W^k - W^* \right\|_G^2 - 2 \left\langle W^k - W^*, W^k - W^{k+1} \right\rangle_G + \left\| W^k - W^{k+1} \right\|_G^2 \\ &\leq \left\| W^k - W^* \right\|_G^2 - \left\| W^k - W^{k+1} \right\|_G^2 - 2 \left\langle X^{k+1} - X^*, \Theta_k \left( X^{k+1} \right) \right\rangle \\ &= \left\| W^k - W^* \right\|_G^2 - \left\| W^k - W^{k+1} \right\|_G^2 - 2 \left\langle X^{k+1} - X^k, \Theta_k \left( X^{k+1} \right) \right\rangle \\ &\quad - 2 \left\langle X^k - X^*, \Theta_k \left( X^{k+1} \right) \right\rangle. \end{aligned} \quad (4.4)$$

Note that with (4.2), it holds

$$\begin{aligned} -2 \left\langle X^k - X^*, \Theta_k \left( X^{k+1} \right) \right\rangle &\leq 4(\delta')^2 \eta_k^2 \left\| X^k - X^* \right\|_{\gamma I - Q_0}^2 + \frac{1}{4(\delta')^2 \eta_k^2} \left\| \Theta_k \left( X^{k+1} \right) \right\|_{(\gamma I - Q_0)^{-1}}^2 \\ &\leq 4(\delta')^2 \eta_k^2 \left\| W^k - W^* \right\|_G^2 + \frac{1}{4(\delta')^2} \left\| e \left( W^{k+1} \right) \right\|_G^2. \end{aligned} \quad (4.5)$$

It follows from (4.2) and  $\sum_{k=0}^{\infty} \eta_k^2 < +\infty$  that there is a  $k_1 \geq 0$  such that for all  $k \geq k_1$

$$\begin{aligned} -2 \left\langle X^{k+1} - X^k, \Theta_k \left( X^{k+1} \right) \right\rangle &\leq 4(\delta')^2 \eta_k^2 \left\| X^{k+1} - X^k \right\|_{\gamma I - Q_0}^2 + \frac{1}{4(\delta')^2 \eta_k^2} \left\| \Theta_k \left( X^{k+1} \right) \right\|_{(\gamma I - Q_0)^{-1}}^2 \\ &\leq \frac{1}{4} \left\| W^{k+1} - W^k \right\|_G^2 + \frac{1}{4(\delta')^2} \left\| e \left( W^{k+1} \right) \right\|_G^2. \end{aligned} \quad (4.6)$$

Similarly to the proof of (3.7) and using (4.2), we have

$$\begin{aligned} \left\| e \left( W^{k+1} \right) \right\|_G &\leq \left\| \begin{array}{c} \alpha \left( \sum_{i=1}^m \beta_i \left( Y_i^{k+1} - Y_i^k \right) + R \left( X^k, X^{k+1} \right) - \Theta_k \left( X^{k+1} \right) \right) \\ 0 \\ \lambda_i^k - \lambda_i^{k+1} \end{array} \right\|_G \\ &\leq \delta' \left( \left\| W^k - W^{k+1} \right\|_G + \left\| \Theta_k \left( X^{k+1} \right) \right\|_{(\gamma I - Q_0)^{-1}} \right) \\ &\leq \delta' \left( \left\| W^k - W^{k+1} \right\|_G + \eta_k \left\| e \left( W^{k+1} \right) \right\|_G \right), \end{aligned}$$

where  $\delta'$  is a positive constant depending on  $\alpha$ ,  $\beta_i$ ,  $\gamma$ , and the eigenvalues of  $Q_0$ . Because  $\sum_{k=0}^{\infty} \eta_k^2 < +\infty$ , there exists a  $k_2 \geq 0$  such that for all  $k \geq k_2$

$$\frac{8}{9} \left\| e \left( W^{k+1} \right) \right\|_G \leq \delta' \left\| W^k - W^{k+1} \right\|_G. \quad (4.7)$$

Substituting (4.5) and (4.6) into inequality (4.4) and using (4.7), we obtain that for all  $k \geq k_0 \equiv \max(k_1, k_2)$

$$\begin{aligned} \left\| W^{k+1} - W^* \right\|_G^2 &\leq \left( 1 + 4(\delta')^2 \eta_k^2 \right) \left\| W^k - W^* \right\|_G^2 - \frac{3}{4} \left\| W^k - W^{k+1} \right\|_G^2 + \frac{1}{2(\delta')^2} \left\| e \left( W^{k+1} \right) \right\|_G^2 \\ &\leq \left( 1 + 4(\delta')^2 \eta_k^2 \right) \left\| W^k - W^* \right\|_G^2 - \frac{2}{3(\delta')^2} \left\| e \left( W^{k+1} \right) \right\|_G^2 + \frac{1}{2(\delta')^2} \left\| e \left( W^{k+1} \right) \right\|_G^2 \\ &= \left( 1 + 4(\delta')^2 \eta_k^2 \right) \left\| W^k - W^* \right\|_G^2 - \frac{1}{6(\delta')^2} \left\| e \left( W^{k+1} \right) \right\|_G^2. \end{aligned}$$

□

The following theorem guarantees the convergence of the inexact method.

**Theorem 4.3.** *The sequence  $\{X^k\}$  generated by the inexact version of the modified alternating direction method for CQCQSDP converges to a solution point  $X^*$ .*

**Proof.** Note that under the assumption  $\sum_{k=0}^{\infty} \eta_k^2 < +\infty$  the product  $\prod_{k=0}^{\infty} \left( 1 + 4(\delta')^2 \eta_k^2 \right)$  is bounded. We denote

$$C_s \equiv 4(\delta')^2 \sum_{k=0}^{\infty} \eta_k^2 \quad \text{and} \quad C_p \equiv \prod_{k=0}^{\infty} \left( 1 + 4(\delta')^2 \eta_k^2 \right).$$

Let  $\widehat{W}$  be an optimal solution of the problem. From Proposition 4.2, we have for all  $k \geq k_0$

$$\begin{aligned} \|W^k - \widehat{W}\|_G^2 &\leq \left( \prod_{l=k_0}^k (1 + 4(\delta')^2 \eta_l^2) \right) \|W^{k_0} - \widehat{W}\|_G^2 \\ &\leq C_p \|W^{k_0} - \widehat{W}\|_G^2. \end{aligned}$$

Therefore, there exists a constant  $\tau > 0$  such that

$$\|W^k - \widehat{W}\|_G^2 \leq \tau, \quad \forall k \geq 0. \quad (4.8)$$

It follows that the sequence  $\{W^k\}$  is bounded and therefore it has at least a cluster point. From (4.3) and (4.8), we get

$$\begin{aligned} \frac{1}{6(\delta')^2} \sum_{k=k_0}^{\infty} \|e(W^{k+1})\|_G^2 &\leq \|W^{k_0} - \widehat{W}\|_G^2 + 4(\delta')^2 \sum_{k=k_0}^{\infty} \left( \eta_k^2 \|W^k - \widehat{W}\|_G^2 \right) \\ &\leq (1 + C_s) \tau. \end{aligned}$$

It follows that

$$\lim_{k \rightarrow \infty} \|e(W^k)\|_G = 0.$$

Let  $W^*$  be a cluster point of  $\{W^k\}$ . Suppose the subsequence  $\{W^{k_j}\}$  converges to  $W^*$ . Then

$$\|e(W^*)\|_G = \lim_{j \rightarrow \infty} \|e(W^{k_j})\|_G = 0$$

and  $W^*$  is a solution. Since  $W^*$  is a solution, for all  $k \geq k_0$ ,  $l \geq 0$ , we have

$$\|W^{k+l} - W^*\|_G^2 \leq C_p \|W^k - W^*\|_G^2. \quad (4.9)$$

It follows from (4.9) that the sequence  $\{W^k\}$  has a unique cluster point and

$$\lim_{k \rightarrow \infty} W^k = W^*. \quad \square$$

## 5. Numerical Results

### 5.1. The Nearest Correlation Matrix Problem and Its Extensions

Higham [10] introduced the following nearest correlation matrix problem. For arbitrary symmetric matrix  $C$ , one solves the optimization problem

$$\min \frac{1}{2} \|X - C\|^2 \quad \text{s.t. } X \in \mathbb{S}_+^n, \quad X_{ii} = 1, \quad i = 1, \dots, n. \quad (5.1)$$

In [10], Higham used the modified alternating projections method to compute the solution for (5.1). Later, Qi and Sun [17] proposed a quadratically convergent Newton method for solving it.

Recently, Gao and Sun [8] extended this model to allow general linear constraints such as bounds on the matrix entries. We further extend their model by adding some quadratic constraints. For instance, when we want to compute a sample covariance matrix, one natural

question comes out: How should we treat data collected in different historical periods? In fact, we face the tradeoff between long-term data and short-term data. By using long-term data the obtained sample covariance matrix might be more stable, but less updated information has been caught; by using short-term data we can focus on current situation, but the sample covariance matrix could contain a lot of errors because of a smaller data size. It thus makes sense to combine two kinds of approaches to achieve better estimation for covariance matrix. We propose a new model for robust estimation of covariance matrix as follows.

$$\begin{aligned}
\min \quad & \frac{1}{2} \|X - C\|^2 \\
\text{s.t.} \quad & \frac{1}{2} \|X - C'\|^2 \leq \varepsilon \\
& \langle A_i, X \rangle = b_i, \quad i = 1, \dots, p \\
& \langle A_i, X \rangle \geq b_i, \quad i = p + 1, \dots, m \\
& X \in \mathbb{S}_+^n,
\end{aligned} \tag{5.2}$$

where  $C, C'$  are the sample covariance matrices from short-term data and long-term data respectively and  $\varepsilon$  is a positive constant to control the size of trust region from the long-term stable estimation. Note that  $C$  can be just a given symmetric matrix if some observations are missing or wrong. Basically, Problem (5.2) is to find the nearest covariance matrix from short-term sample estimation within the trust region from the long-term sample estimation. Furthermore, additional linear equality and inequality constraints can be included. The optimal solution of (5.2) can be desirable because it will not be too far away from the long-term stable estimation while at the same time it can contain current information as much as possible through minimizing the distance with the short-term estimation. By doing so, the estimation error can be also systematically reduced.

## 5.2. Computational Experiment and Results

The codes were written in MATLAB (version 6.5) and the computations were performed on a 1.86 GHz Intel Core 2 PC with 3GB of RAM. We consider the following testing examples.

**Example 1.** CQSDPs arising from the nearest correlation matrix problem (5.1). The matrix  $C$  is generated from the MATLAB segment:  $x = 10^{\wedge}[-4 : 4/(n - 1) : 0]$ ;  $C = \text{gallery}('randcorr', n * x / \text{sum}(x))$ . For the test purpose, we perturb  $C$  to

$$C = C + 10^{-3} * E, \text{ or } C = C + 10^{-2} * E, \text{ or } C = C + 10^{-1} * E,$$

where  $E$  is a randomly generated symmetric matrix with entries in  $[-1, 1]$ . The MATLAB code for generating  $E$  is:  $E = \text{rand}(n)$ ;  $E = (E + E')/2$ ; for  $i = 1 : n$ ;  $E(i, i) = 1$ ; end. Note that we make the perturbation larger than  $10^{-4} * E$  considered in [10]. To observe the robustness of our algorithm, we use three sets of starting point:

- a)  $(X^0, Y^0, \lambda^0) = (C, C, 0)$ ;
- b)  $(X^0, Y^0, \lambda^0) = (I_n, I_n, 0)$ ;
- c)  $X^0 = \text{rand}(n)$ ;  $X^0 = [X^0 + (X^0)']/2$ ; for  $i = 1 : n$ ;  $X^0(i, i) = 1$ ; end;  $Y^0 = X^0$ ;  $\lambda^0 = 0$ .

We test for  $n = 100, 500, 1000, 2000$ , respectively.

**Example 2.** CQCQSDPs without linear constraints arising from the extended nearest correlation problem (5.2). The matrix  $C$  is generated from the MATLAB segment:  $x = 10^{\wedge}[-4 :$

$4/(n-1) : 0]$ ;  $C = \text{gallery}(\text{randcorr}', n * x / \text{sum}(x))$ . For the test purpose, we perturb  $C$  in the following four situations:

- $C' = C + 10^{-1} * E'$ ;  $C = C + 10^{-1} * E$ ;
- $C' = C + 10^{-2} * E'$ ;  $C = C + 10^{-2} * E$ ;
- $C' = C + 10^{-2} * E'$ ;  $C = C + 10^{-1} * E$ ;
- $C' = C + 10^{-1} * E'$ ;  $C = C + 10^{-2} * E$ ;

where  $E$  and  $E'$  are two random symmetric matrices generated as in Example 1. We use  $(X^0, Y^0, \lambda^0) = (C, C, 0)$  as the starting point. We take  $\epsilon = r * \text{norm}(C - C')$  in (5.2) with  $r = 0 : 0.2 : 1.2$  to consider the effect of trust region size for  $n = 100$ . Then using  $r = 0.8$ , we test for  $n = 100, 500, 1000$ , and  $2000$ , respectively.

**Example 3.** The same as Example 2 but the diagonal entries of variable matrix are additionally required to be ones, i.e., it is a correlation matrix. We use  $(X^0, Y_1^0, Y_2^0, \lambda_1^0, \lambda_2^0) = (C, C, C, 0, 0)$  as the starting point and set  $\epsilon = 0.8 * \text{norm}(C - C')$ .

We simply use  $\gamma = 1$  and  $\beta = 1$  in the modified ADMs. The convergence was checked at the end of each iteration using the condition,

$$\frac{\max\{\|X^k - X^{k-1}\|_\infty, \|Y^k - Y^{k-1}\|_\infty, \|\lambda^k - \lambda^{k-1}\|_\infty\}}{\max\{\|X^1 - X^0\|_\infty, \|Y^1 - Y^0\|_\infty, \|\lambda^1 - \lambda^0\|_\infty\}} \leq 10^{-6}.$$

We also set the maximum number of iterations to 500.

The main computational cost at each iteration is matrix eigenvalue decomposition. The performance results of our modified ADMs are reported in Tables 1 – 4. The columns corresponding to “No. It” give the iteration numbers and the columns corresponding to “CPU Sec.” give the CPU time in seconds. The “\*” entries in Table 2 indicate the algorithm could not find the optimal solution within 500 iterations, which actually means that Problem (5.2) is infeasible in practice.

From the numerical results reported in Table 1, we can see for problems of all sizes, the algorithm obtains the solutions mostly in less than 30 iterations and with reasonable accuracy  $10^{-6}$ . These results are comparative with those in [8, 20]. Actually the CPU time by our proposed algorithm is between the results reported in [8] and the results reported in [20]. Furthermore, we see that the modified ADM is quite robust for solving the nearest correlation problem (5.1) because the number of iterations is little affected by the choices of starting point.

For cases with quadratic constraints, to which the algorithms in [8, 20] cannot apply, the numerical results reported in Tables 2 – 4 are also promising. If  $r$  is too small, the problem might be infeasible; while if  $r$  is too large, the trust region constraint becomes redundant. These are all verified by the numerical results reported in Table 2. It seems  $r = 0.8$  is a suitable parameter regardless of different choices of  $C$  and  $C'$ . Using this  $r$  for problem’s size  $n = 100, 500, 1000, 2000$ , the numerical results reported in Tables 3 – 4 show that our algorithm is effective to solve CQCQSDPs with or without linear constraints.

## 6. Concluding Remarks

We propose a modified alternating direction method for solving convex quadratically constrained quadratic semidefinite problems. The advantage of the proposed method is that it does

**Table 1: Numerical results for Example 1**

Example 1		$C+10^{-3}*E$		$C+10^{-2}*E$		$C+10^{-1}*E$	
n=	case	No. It	CPU Sec.	No. It	CPU Sec.	No. It	CPU Sec.
100	a)	7	0.4	14	0.6	24	1.0
100	b)	20	0.9	21	0.9	28	1.2
100	c)	21	1.0	20	0.9	24	1.1
500	a)	10	47.3	14	60.4	23	90.7
500	b)	20	95.0	21	92.1	27	111.5
500	c)	23	105.1	23	98.5	25	109.1
1000	a)	10	370.7	15	537.9	24	777.3
1000	b)	20	701.2	22	730.2	29	957.5
1000	c)	23	809.7	23	791.2	26	843.2
2000	a)	11	2972	14	3843	25	6321
2000	b)	20	5485	23	6377	31	7956
2000	c)	24	6362	24	6408	27	6823

**Table 2: Numerical results for Example 2 with different trust region sizes**

Example 2 n=100	$C=C+10^{-1}*E$	$C=C+10^{-2}*E$	$C=C+10^{-1}*E$	$C=C+10^{-2}*E$
	$C'=C+10^{-1}*E'$	$C'=C+10^{-2}*E'$	$C'=C+10^{-2}*E'$	$C'=C+10^{-1}*E'$
r=	No. It	No. It	No. It	No. It
0	*	*	*	*
0.2	*	*	47	58
0.4	*	26	22	26
0.6	23	17	15	15
0.8	17	17	14	15
1.0	1	9	2	1
1.2	1	1	1	1

**Table 3: Numerical results for Example 2 with  $r = 0.8$** 

Example 2 r=0.8	$C=C+10^{-1}*E$		$C=C+10^{-2}*E$		$C=C+10^{-1}*E$		$C=C+10^{-2}*E$	
	$C'=C+10^{-1}*E'$		$C'=C+10^{-2}*E'$		$C'=C+10^{-2}*E'$		$C'=C+10^{-1}*E'$	
n=	No. It	CPU Sec.	No. It	CPU Sec.	No. It	CPU Sec.	No. It	CPU Sec.
100	17	0.9	16	0.9	15	0.9	16	0.9
500	16	84.2	16	80.8	13	65.2	17	84.7
1000	16	622.2	16	648.3	13	516.3	17	693.7
2000	15	4638	15	5047	13	4054	18	5726

**Table 4: Numerical results for Example 3**

Example 3 r=0.8	$C=C+10^{-1}*E$		$C=C+10^{-2}*E$		$C=C+10^{-1}*E$		$C=C+10^{-2}*E$	
	$C'=C+10^{-1}*E'$		$C'=C+10^{-2}*E'$		$C'=C+10^{-2}*E'$		$C'=C+10^{-1}*E'$	
n=	No. It	CPU Sec.	No. It	CPU Sec.	No. It	CPU Sec.	No. It	CPU Sec.
100	53	2.5	33	1.6	36	1.8	36	1.7
500	38	161.2	33	147.0	35	151.6	35	157.3
1000	37	1248	33	1169	36	1257	35	1254
2000	36	9571	33	9207	36	9665	36	10103

not require to solve sub-variational inequality problems over the semidefinite cone; instead, in each iteration it requires only one projection onto the semidefinite cone, plus  $m$  easy vector projections. The convergence of the method is analyzed and it is shown that as long as the KKT variational inequality system of the problem has an optimal solution, the method will produce a sequence that converges to a solution of the CQCQSDP. The method can be relaxed to allow inexact projection onto the semidefinite cone, while preserving the same convergence properties.

The proposed modified method does not require second-order information and it is easy to implement. In addition, for problems with a large number of quadratic constraints, the vector projections in Step 2 can be computed simultaneously by difference processors, either within a single computer or within a multi-grid computer network. These additional features appear to be attractive in solving large scale convex quadratically constrained quadratic semidefinite programs. Our numerical experiment shows that the method can effectively solve the covariance matrix problems of matrix size up to  $2000 \times 2000$  within reasonable time and accuracy by using a desktop computer.

## References

- [1] Beck, A. (2007). Quadratic matrix programming. *SIAM Journal on Optimization*, 17, 1224-1238.
- [2] Boyd, S. & Xiao, L.(2005). Least-squares covariance matrix adjustment. *SIAM Journal on Matrix Analysis and Applications* 27, 532-546.
- [3] Chen, G. & Teboulle, M. (1994). A proximal-based decomposition method for convex minimization problems. *Mathematical Programming*, 64, 81-101.
- [4] Dai, Y.H. (2006) Fast algorithms for projection on an ellipsoid. *SIAM Journal on Optimization*, 16, 986-1006.
- [5] Eckstein, J. & Bertsekas, D.P. (1992). On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming*, 55, 293-318.
- [6] Gabay, D. (1983). Applications of the method of multipliers to variational inequalities. In: M. Fortin and R. Glowinski (eds.), *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-value Problem* (pp. 299-331). Amsterdam, North-Holland.
- [7] Gabay, D. & Mercier, B. (1976). A dual algorithm for the solution of nonlinear variational problems via finite element approximations. *Computers and Mathematics with Applications*, 2, 17-40.
- [8] Gao, Y. & Sun, D. (2009). Calibrating least squares covariance matrix problems with equality and inequality constraints. *SIAM Journal on Matrix Analysis and Applications*, 31, 1432-1457.
- [9] He, B.S., Liao, L.Z., Han, D.R. & Yang, H. (2002). A new inexact alternating directions method for monotone variational inequalities. *Mathematical Programming*, 92, 103-118.

- [10] Higham, N.J. (2002). Computing the nearest correlation matrix - a problem from finance. *IMA Journal of Numerical Analysis*, 22, 329-343.
- [11] Kinderlehrer, D. & Stampacchia, G. (1980). *An Introduction to Variational Inequalities and Their Applications*. New York, Academic Press.
- [12] Li, C. & Sun, W. (2009). On filter-successive linearization methods for nonlinear semidefinite programming. *Science in China, Series A*, 52, 2341-2361.
- [13] Malick, J. (2004). A dual approach to semidefinite least-squares problems. *SIAM Journal on Matrix Analysis and Applications*, 26, 272-284.
- [14] Malick, J., Povh, J., Rendl, F. & Wiegale, A. (2009). Regularization methods for semidefinite programming. *SIAM Journal on Optimization*, 20, 336-356.
- [15] Nie, J.W. & Yuan, Y.X. (2001). A predictor- corrector algorithm for QSDP combining Dikin-type and Newton centering steps. *Annals of Operations Research*, 103, 115-133.
- [16] Pang, J.-S. (1986). Inexact methods for the nonlinear complementarity problem. *Mathematical Programming*, 36, 54-71.
- [17] Qi, H. & Sun, D. (2006). A quadratically convergent Newton method for computing the nearest correlation matrix. *SIAM Journal on Matrix Analysis and Applications*, 28, 360-385.
- [18] Sun, J., Sun, D. & Qi, L. (2004). A squared smoothing Newton method for nonsmooth matrix equations and its applications in semidefinite optimization problems. *SIAM Journal on Optimization*, 14, 783-806.
- [19] Sun, D., Sun, J. & Zhang, L. (2008). Rates of convergence of the augmented Lagrangian method for nonlinear semidefinite programming. *Mathematical Programming*, 114, 349-391.
- [20] Toh, K. (2007). An inexact primal-dual path following algorithm for convex quadratic SDP. *Mathematical Programming*, 112, 221-254.
- [21] Wen, Z., Goldfarb, D. & Yin, W. (2009). Alternating direction augmented Lagrangian methods for semidefinite programming. *Optimization Online*.
- [22] Yu, Z.S. (2006). Solving semidefinite programming problems via alternating direction methods. *Journal of Computational and Applied Mathematics*, 193, 437-445.
- [23] Zhao, X.Y. (2010). A semismooth Newton-CG augmented Lagrangian method for large scale linear and convex quadratic SDPs. *PhD Thesis*, National University of Singapore.