International Journal of Innovative Computing, Information and Control
Volume **12**, Number **1**, February **2016**

ICIC International ©2016 ISSN 1349-4198
pp. 203–214

SOLVING A CLASS OF QUADRATIC PROGRAMS WITH LINEAR COMPLEMENTARITY CONSTRAINTS VIA A MAJORIZED PENALTY APPROACH

Shiyun Wang, Yong-Jin Liu and Li Wang

School of Science
Shenyang Aerospace University
No. 37, Daoyi South Avenue, Shenbei New Area, Shenyang 110136, P. R. China
wsy0902@163.com

Received April 2015; revised August 2015

Abstract. *This paper aims to solve a class of quadratic programming with linear complementarity constraints (QPLCCs). We transit it to an equivalent quadratic programming with nonsmooth equation constraints, and partially penalize the problem by setting the nonsmooth equation constraints as the penalty term. And then, we apply the majorization approach to solve the penalty form. We prove that this partially penalty method is exact. At last, by solving a sequence of convex semismooth quadratic optimization problems with linear constraints, the QPLCC is solved and the convergence analysis is obtained. Numerical results are displayed at the ending of this paper.*

Keywords: Quadratic program, Linear complementarity constraints, Majorization method, Penalty method

1. **Introduction.** Mathematical programming with complementarity constraints (MPCC) is a special case of mathematical programming with equilibrium constraints (MPEC), which plays a very important role in many fields such as inverse problem, engineering design, economic equilibria, transplantation science, multilevel game, and mathematical programming itself (for more applications of MPCC or MPEC, one can refer to [1, 8, 12, 30], etc.). As a special case of MPCCs, quadratic programming with linear complementarity constraints has many applications in data estimation and some inverse linear complementarity problems [19, 20]. In this paper, we consider a class of quadratic programming with linear complementarity constraints which can be specified as

$$\begin{array}{ll} \min_{x\in\mathcal{X}} & f(x) := \frac{1}{2}x^TGx + c^Tx \\ \text{s.t.} & 0 \le Ax + a \perp Bx + b \ge 0, \end{array} \tag{1}$$

where $c \in \Re^n$, $a \in \Re^m$, $b \in \Re^m$, $G \in S^n_+$, $A \in \Re^{m\times n}$, $B \in \Re^{m\times n}$ are given, $\mathcal{X}$ is a convex set in $\Re^n$, such as $\mathcal{X} = \{x : h(x) \le 0, g(x) = 0\}$. Many important classes of QPLCCs can be cast into Problem (1), see [7, 13, 21].

During the last three decades, tremendous progress has been achieved on algorithm to MPCCs, and these methods can be used to solve Problem (1). For example, interior-point methods (see, e.g., [14, 15]), matrix splitting methods (see, e.g., [16]) and active-set methods (see, e.g., [17, 18]), can be applied to solve MPCCs. Meanwhile, there are some excellent results with Problem (1): Fukushima et al. [7] presented a sequential quadratic programming algorithm for a mathematical program with a special form of linear complementarity constraints by using FB-function; Chen and Ye [13] considered a class of quadratic programs with linear complementary constraints by investigating the stationary conditions and proposed a Newton-like method; recently, Bai et al. [21]

proposed a two-stage approach to solve a more general situation of QPLCCs based on the Benders decomposition. Although so many methods can be used to solve Problem (1), they are not directed towards Problem (1). Majorization method, which was first introduced by de Leeuw [2, 3] and Heiser [4], is an efficient approach in solving many optimization problems [5, 10]. Majorized penalty approach has been successfully applied in calibrating rank constrained correlation matrix problem [9] and inverse linear second-order cone programming problems [29]. Due to its efficiency in solving the optimization problems, we will apply majorized penalty to solve Problem (1).

In our study, we transit the linear complementarity constraints in (1) to nonsmooth equations constraints like that in [13]. Then, we consider the penalized version of the problem by regarding the nonsmooth equation constraint as the penalty term and then apply the majorization method to the penalized problem by solving a sequence of semismooth optimization problems. The exactness of the penalty and the convergence of the majorized penalty approach will be analyzed. There are two main advantages in our method. On one hand, when we apply penalty method, Problem (1) is equivalent to Problem (4) which demonstrates the exactness of the penalty (see Proposition 2.1). On the other hand, when we apply the majorization method to solving Problem (4), the penalty term brings small error (see the algorithm $MA(\rho)$ and Table 1). The numerical results demonstrate the efficiency of our approach.

For a matrix $Z \in R^{m\times n}$, Z_i $(1 \leq i \leq m)$ denotes the i-th column of Z, and, Z_i^T $(1 \leq i \leq m)$ denotes the i-th column of Z^T. For a vector $z \in R^m$, z_i denotes the i-th entry. $[1 : m]$ is a set including the integers $1, 2, \ldots, m$.

The rest of this paper is organized as follows. In Section 2, we transmit the Problem (1) to an equivalent form. We propose a majorized penalty approach for solving it and analyze the convergent properties. In Section 3, we report the computational experiments on quadric objective function with linear constraints to test the efficiency of our approach, especially, an inverse linear programming to be solved. The final conclusions are made in Section 4.

2. **The Majorized Penalty Approach.**

In this section, we propose the so called majorized penalty approach to solve Problem (1). For this purpose, we first consider the penalized version to Problem (1) in Subsection 2.1 and then apply majorized method to solving the penalized problem and analyzing the convergency in Subsection 2.2.

2.1. **Exact penalty method.**

Denote $\Pi_{\Re^m_+}(\cdot)$ the metric projector over $\Re^m_+$. For $u, v \in \Re^m_+$, it is obvious that

$$u_i - \max\{u_i - v_i, 0\} \geq 0$$

and then

$$u \perp v \quad \Leftrightarrow \quad u = \Pi_{\Re^m_+}(u - v) \Leftrightarrow \textstyle\sum_{i=1}^m u_i - \sum_{i=1}^m \max\{u_i - v_i, 0\} = 0. \tag{2}$$

Defining

$$p(x) := \sum_{i=1}^m \left(x^T A_i^T + a_i\right) - \sum_{i=1}^m \max\left\{x^T(A_i^T - B_i^T) + (a_i - b_i), 0\right\},$$

then, Problem (1) can be written as a slightly different form as follows:

$$\begin{aligned} \min\nolimits_{x\in\mathcal{X}} \quad & f(x) \\ \text{s.t.} \quad & Ax + a \geq 0, \\ & Bx + b \geq 0, \\ & p(x) = 0. \end{aligned} \tag{3}$$

The constraints of nonsmooth equation in (3) appeared in reaction diffusion problems [22], thin stretched membranes partially covered with water [23] and other fields (for more applications of the constraints of nonsmooth equation as in (3), one can see [13] and its references).

For simplicity, we denote

$$\Omega = \{x \in \mathcal{X} : Ax + a \geq 0, Bx + b \geq 0\},$$

and

$$\Omega_0 = \{x \in \mathcal{X} : Ax + a \geq 0, Bx + b \geq 0, p(x) = 0\}.$$

And we define

$$\theta_\rho(x) : = f(x) + \rho p(x).$$

Then $p(x) \geq 0$ for $x \in \Omega$, and Ω_0 is the feasible set of Problem (3), for a given penalty parameter $\rho > 0$, the partially penalized form of Problem (3) by setting the nonsmooth constraint as the penalty term can be expressed as the following:

$$\min\nolimits_{x \in \Omega} \quad \theta_\rho(x) \tag{4}$$

We will show that Problem (4) is an exact l_1 partial penalty of Problem (3). For this purpose, we let

$$U := \{x \in \Re^n : p(x) = 0\}$$

and, for any $x \in \Omega$, define the Lagrange function of Problem (3)

$$L(x, \lambda) := f(x) + p(x)\lambda, \quad x \in \Omega.$$

Let

$$x^* \in \inf_{x \in \Omega} \sup_{\lambda \in R} L(x, \lambda)$$

and

$$\lambda^* \in \sup_{\lambda \in R} \inf_{x \in \Omega} L(x, \lambda)$$

We say that (x^*, λ^*) is a saddle point of the Lagrange function if

$$L(x^*, \lambda) \leq L(x^*, \lambda^*) \leq L(x, \lambda^*).$$

Proposition 2.1. *Suppose that (x^*, λ^*) is the saddle point of Problem (3). Denoting*

$$\begin{aligned} \rho^* &: = |\lambda^*|, \\ X^* &: = \operatorname{argmin}_{x \in \Omega \cap U} f(x), \\ Z(\rho) &: = \operatorname{argmin}_{x \in \Omega} \theta_\rho(x), \end{aligned}$$

then $X^* = Z(\rho)$ for $\rho > \rho^*$.

Proof: Since (x^*, λ^*) is the saddle point of Problem (3), then we have $p(x^*) = 0$ and $x^* \in X^*$. For any $x \in \Omega$,

$$\theta_\rho(x^*) = f(x^*) = L(x^*, \lambda^*) \leq L(x, \lambda^*).$$

Thus, it holds that

$$\begin{aligned} \theta_\rho(x^*) &\leq f(x) + \lambda^* p(x) \\ &\leq f(x) + |\lambda^*||p(x)| \\ &= f(x) + |\lambda^*| p(x) \\ &= f(x) + \rho^* p(x) \\ &\leq f(x) + \rho p(x) \\ &= \theta_\rho(x) \end{aligned} \tag{5}$$

which implies $X^* \subseteq Z(\rho)$, $\rho > \rho^*$.

On the other hand, we will show that $Z(\rho) \subseteq X^*$ when $\rho > \rho^*$. Otherwise, if there exist $\rho_1 > \rho^*$ and $x_1 \in \Omega$ such that $x_1 \in Z(\rho_1)$ and $x_1 \notin X^*$, recalling that (x^*, λ^*) is the saddle point of Problem (3), we have $x^* \in X^*$ (Theorem 1.6.1 in [32]) and then

$$\begin{aligned} f(x^*) &= \theta_{\rho_1}(x^*) \\ &= \theta_{\rho_1}(x_1) \text{ (because } x^* \in Z(\rho_1), x_1 \in Z(\rho_1)) \\ &= f(x_1) + \rho_1 p(x_1). \end{aligned}$$

Case 1. $x_1 \in U$. In this case, we have $p(x_1) = 0$ which means $f(x^*) = f(x_1)$, and then $x_1 \in X^*$. It is a contradiction.

Case 2. $x_1 \notin U$. In this case, we have $p(x_1) > 0$. Letting $\rho \in [\rho^*, \rho_1]$, then we obtain

$$\begin{aligned} \theta_\rho(x^*) &= f(x^*) + \rho p(x^*) \\ &= f(x^*) + \rho_1 p(x^*) \\ &\geq f(x_1) + \rho_1 p(x_1) \text{ (because } x_1 \in Z(\rho_1)) \\ &> f(x_1) + \rho p(x_1) \\ &= \theta_\rho(x_1) \end{aligned}$$

which contradicts with (5).

2.2. **Majorization method and its convergency.** To apply majorization method for Problem (4), we first introduce some notations and symbols. Suppose that h is defined over $X \subset R^n$, we say that g on $X \times X$ is a **majorization function** of h if $h(x) \leq g(x, y)$, $\forall x, y \in X$ and $h(x) = g(x, x), \forall x \in X$. Letting

$$\theta_\rho^1(x) := \frac{1}{2} x^T G x + c^T x + \rho \langle e, Ax + a \rangle$$

and

$$\theta_\rho^2(x) := \rho \sum_{i=1}^m \max\left\{ x^T (A_i^T - B_i^T) + (a_i - b_i), 0 \right\},$$

then the objective function of Problem (4) can be rewritten as

$$\theta_\rho(x) = \theta_\rho^1(x) - \theta_\rho^2(x) \tag{6}$$

which yields $\theta_\rho(\cdot)$ is the difference of two convex quadratic functions $\theta_\rho^1(\cdot)$ and $\theta_\rho^2(\cdot)$.

By the convexity of the function $\theta_\rho^2(\cdot)$, we have

$$\theta_\rho^2(x) \geq \theta_\rho^2(s) + \langle \xi, x - s \rangle \quad \forall s \in \Re^n,$$

where ξ is the subgradient of $\theta_\rho^2(\cdot)$ at the point s, i.e., $\xi \in \partial \theta_\rho^2(s)$. For any $s \in \Omega$, define $\hat{\theta}_\rho(\cdot, s) : \Omega \to \Re$ as

$$\hat{\theta}_\rho(x, s) := \theta_\rho^1(x) - \theta_\rho^2(s) - \xi^T (x - s) \quad \forall x \in \Omega. \tag{7}$$

Then, we readily obtain that

$$\hat{\theta}_\rho(x, s) \geq \theta_\rho(x) \text{ and } \hat{\theta}_\rho(s, s) = \theta_\rho(s) \quad \forall x \in \Omega.$$

Then, the function $\hat{\theta}_\rho(\cdot, s)$ is a majorization function of $\theta_\rho(\cdot)$ at any $s \in \Omega$. Moreover, from (7), we know that for any $s \in \Omega$, $\hat{\theta}_\rho(\cdot, s)$ is convex quadratic in Ω.

For any $x \in \Omega$, let $N_\Omega(x)$ denote the normal cone of Ω at the point $x \in \Omega$, i.e.,

$$N_\Omega(x) := \{ d \in \Re^n : \langle d, s - x \rangle \leq 0, \forall s \in \Omega \}.$$

A point $x^* \in \Omega$ is said to be a stationary point of Problem (4) if

$$0 \in \nabla \theta_\rho^1(x^*) - \partial \theta_\rho^2(x^*) + N_\Omega(x^*), \tag{8}$$

where $\nabla\theta^1_\rho(x^*)$ is the gradient of θ^1_ρ at x^* and $\partial\theta^2_\rho(x^*)$ is the subgradient of θ^2_ρ at x^*.

Let

$$\begin{aligned}\alpha &:= \{i \in [1:m] : x^T(A_i^T - B_i^T) + (a_i - b_i) > 0\},\\ \beta &= \{i \in [1:m] : x^T(A_i^T - B_i^T) + (a_i - b_i) = 0\},\\ \gamma &:= \{i \in [1:m] : (A^T - B^T)_i x + (a_i - b_i) < 0\}.\end{aligned}$$

Then the differential properties of $\theta^2_\rho(\cdot)$ at x can be described as the following results.

Theorem 2.1. *$\theta^2_\rho(\cdot)$ is continuously differentiable at x if and only if $\beta = \emptyset$, and then we have $\partial\theta^2_\rho(x) = \{\nabla\theta^2_\rho(x) = \rho\sum_{i\in\alpha}(A_i^T - B_i^T)\}$, where $\nabla\theta^2_\rho(x)$ is the gradient of θ^2_ρ at x.*

Theorem 2.2. *Assuming that $\beta \neq \emptyset$, we have*

$$\partial\theta^2_\rho(x) = \left\{\xi : \xi = \rho\left(\sum_{i\in\alpha}\left(A_i^T - B_i^T\right) + \sum_{i\in\beta}\lambda_i\left(A_i^T - B_i^T\right)\right), \lambda_i \in [0,1]\right\}.$$

Proof: Letting $f_i(x) = \max\{x^T(A_i^T - B_i^T) + (a_i - b_i), 0\}$ then: (1) if $i \in \alpha$, we have $\nabla f_i(x) = (A_i^T - B_i^T)$; (2) if $i \in \gamma$, we have $\nabla f_i(x) = 0$; (3) if $i \in \beta$, we have $\partial f_i(x) = \lambda_i(A_i^T - B_i^T)$ where $\lambda_i \in [0,1]$. Then, we can get the following from Theorem 23.8 in [31]:

$$\partial\theta^2_\rho(x) = \rho\partial\left(\sum_i^m f_i(x)\right) = \rho\left(\sum_i^m \partial f_i(x)\right) = \rho\left(\sum_{i\in\alpha}(A_i^T - B_i^T) + \sum_{i\in\beta}\lambda_i(A_i^T - B_i^T)\right).$$

Now we are ready to state the majorization algorithm based on the majorization function $\hat\theta_\rho(\cdot, s)$ at any $s \in \Omega$ for solving Problem (4).

A Majorization Algorithm $[MA(\rho)]$

Step 0. Given a penalty parameter $\rho > 0$. Choose $x^0 \in \Omega$. Set $j := 0$.

Step 1. Find the optimal solution

$$x^{j+1} = \underset{x\in\Omega}{\operatorname{argmin}}\left\{\hat\theta^j_\rho(x) := \hat\theta_\rho(x, x^j)\right\},$$

where $\hat\theta_\rho(\cdot, x^j)$ is the majorization function of $\theta_\rho(\cdot)$ at $x^j \in \Omega$ defined as in (7) and $\xi^j \in \partial\theta^2_\rho(x^j)$. That is, x^{j+1} is the optimal solution of the following problem:

$$\begin{aligned}\min \quad & \hat\theta_\rho(x, x^j) = \theta^1_\rho(x) - \theta^2_\rho(x^j) - (\xi^j)^T(x - x^j)\\ \text{s.t.} \quad & x \in \Omega.\end{aligned} \tag{9}$$

Step 2. If $x^{j+1} = x^j$, then stop; otherwise, set $j := j+1$ and go to Step 1.

Theorem 2.3. *Let $\{x^j\}$ be the sequence generated by the $MA(\rho)$. Then, the following properties hold.*

- **(a):** *The sequence $\{\theta_\rho(x^j)\}$ is monotonically nonincreasing.*
- **(b):** *If $x^{j+1} = x^j$ for some integer $j \geq 0$, then x^{j+1} is a stationary point of Problem (4). Otherwise, the infinite sequence $\{\theta_\rho(x^j)\}$ satisfies*

$$\theta_\rho(x^{j+1}) - \theta_\rho(x^j) \leq -\frac{1}{2}\langle x^{j+1} - x^j, M(x^{j+1} - x^j)\rangle, \quad j = 1, 2, \ldots, \tag{10}$$

where the matrix M is given by

$$M = \begin{pmatrix} G & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}. \tag{11}$$

(c): *Moreover, suppose that the matrix G is positive definite, then the sequence $\theta_\rho\{x^j\}$ is convergent, the sequence $\{x^j\}$ is bounded and any accumulation point of $\{x^j\}$ is a stationary point of Problem (4).*

Proof: (a) For any $j \geq 0$, since x^{j+1} is the optimal solution to $\min_{x\in\Omega}\left\{\hat{\theta}^j_\rho(x)\right\}$ and $x^j \in \Omega$, we obtain that

$$\hat{\theta}^j_\rho(x^{j+1}) \leq \hat{\theta}^j_\rho(x^j),$$

which, combining with $\theta_\rho(x^{j+1}) \leq \hat{\theta}^j_\rho(x^{j+1})$ and $\theta_\rho(x^j) = \hat{\theta}^j_\rho(x^j)$, yields that $\{\theta_\rho(x^j)\}$ is a monotonically nonincreasing sequence.

(b) If $x^j = x^{j+1}$ for some integer $j \geq 0$. Note that x^{j+1} is the optimal solution to $\min_{x\in\Omega}\left\{\hat{\theta}^j_\rho(x)\right\}$, then we know that there exist $\xi^j \in \partial\theta^2_\rho(x^j)$ such that

$$0 \in \nabla\theta^1_\rho(x^{j+1}) - \xi^j + N_\Omega(x^{j+1}). \tag{12}$$

By $x^j = x^{j+1}$, one has $\xi^j \in \partial\theta^2_\rho(x^{j+1})$. Thus, (12) shows that

$$0 \in \nabla\theta^1_\rho(x^{j+1}) - \partial\theta^2_\rho(x^{j+1}) + N_\Omega(x^{j+1}),$$

which implies x^{j+1} is a stationary point of Problem (4).

If $x^j \neq x^{j+1}$ for any $j \geq 0$, for each $j \geq 0$, since x^{j+1} solves the problem of minimizing $\hat{\theta}^j_\rho(\cdot)$ over Ω, we have there existing $(D^{j+1})^T \in N_\Omega(x^{j+1})$ such that

$$\nabla\hat{\theta}^j_\rho(x^{j+1}) = (D^{j+1})^T. \tag{13}$$

Recalling that $\theta_\rho(x^{j+1}) \leq \hat{\theta}^j_\rho(x^{j+1})$, $\theta_\rho(x^j) = \hat{\theta}^j_\rho(x^j)$, the function $\hat{\theta}^j_\rho(\cdot)$ is quadratic and the definition of M_ρ in (11), we can easily derive that

$$\begin{aligned}
&\theta_\rho(x^{j+1}) - \theta_\rho(x^j)\\
\leq\ & \hat{\theta}^j_\rho(x^{j+1}) - \hat{\theta}^j_\rho(x^j)\\
=\ & \langle\nabla\hat{\theta}^j_\rho(x^j), x^{j+1} - x^j\rangle + \frac{1}{2}\langle x^{j+1} - x^j, M_\rho(x^{j+1} - x^j)\rangle\\
=\ & \langle\nabla\hat{\theta}^j_\rho(x^{j+1}) + M(x^j - x^{j+1}), x^{j+1} - x^j\rangle + \frac{1}{2}\langle x^{j+1} - x^j, M(x^{j+1} - x^j)\rangle\\
=\ & \langle D^{j+1} + M(x^j - x^{j+1}), x^{j+1} - x^j\rangle + \frac{1}{2}\langle x^{j+1} - x^j, M(x^{j+1} - x^j)\rangle\\
=\ & \langle D^{j+1}, x^{j+1} - x^j\rangle - \langle x^{j+1} - x^j, M(x^{j+1} - x^j)\rangle + \frac{1}{2}\langle x^{j+1} - x^j, M(x^{j+1} - x^j)\rangle\\
\leq\ & -\frac{1}{2}\langle x^{j+1} - x^j, M(x^{j+1} - x^j)\rangle.
\end{aligned} \tag{14}$$

(c) Since the matrix G is positive definite, then, $\hat{\theta}^j_\rho(\cdot)$ is strongly convex for all $j \geq 0$, which implies that x^{j+1} is bounded. Thus, the sequence $\{x^j\}$ is bounded. Assume that x^* is an accumulation point of $\{x^j\}$ and $\{x^{j_k}\}$ is a subsequence such that $\lim_{k\to\infty} x^{j_k} = x^*$. Since the sequence $\{\theta_\rho(x^j)\}$ is monotonically nonincreasing and has a lower bound, we know that the sequence $\{\theta_\rho(x^j)\}$ is convergent. Moreover, by (10), we know that

$$\begin{aligned}
\lim_{i\to+\infty}\frac{1}{2}\sum_{j=1}^{i}\langle x^{j+1} - x^j, M(x^{j+1} - x^j)\rangle &\leq \lim_{i\to+\infty}\sum_{j=1}^{i}\left(\theta_\rho(x^j) - \theta_\rho(x^{j+1})\right)\\
&= \theta_\rho(x^0) - \lim_{i\to\infty}\theta_\rho(x^{i+1})\\
&< \infty,
\end{aligned}$$

which, together with the positive semi-definiteness of M_ρ, implies that

$$\lim_{k\to\infty} x^{j_k+1} = \lim_{k\to\infty} x^{j_k} = x^*. \tag{15}$$

Noting that $x^{j_k+1} = \underset{x\in\Omega}{\operatorname{argmin}} \hat{\theta}_\rho^j(x)$, we have

$$0 = \nabla\theta_\rho^1(x^{j_k+1}) - \xi_{j_k} - D^{j_k+1}, \tag{16}$$

where

$$D^{j_k+1} \in N_\Omega(x^{j_k+1}), \tag{17}$$

and

$$\xi_{j_k} \in \partial\theta_\rho^2(x^{j_k}).$$

Noting that $\theta_\rho^2(\cdot)$ is proper convex, then $\{\xi_{j_k}\}$ is bounded [31]. By taking a subsequence if necessary, we assume that there exists $\xi^* \in \partial\theta_\rho^2(x^*)$ such that

$$\lim_{k\to+\infty} \xi_{j_k} = \xi^*.$$

From (16), we have $\{D^{j_k+1}\}$ having a limit denoted by D^*. To verify that x^* is a stationary point of Problem (4), we only need to show that $\xi^* \in \partial\theta_\rho^2(x^*)$ and $D^* \in N_\Omega(x^*)$. In fact, if $\xi^* \notin \partial\theta_\rho^2(x^*)$, then there exists $\tilde{x} \in \Omega$ such that

$$\theta_\rho^2(\tilde{x}) < \theta_\rho^2(x^*) + (\xi^*)^T(\tilde{x} - x^*).$$

Noting that $\xi_k \in \partial\theta_\rho^2(x^{j_k})$ for all $j_k \geq 0$, it holds that

$$\theta_\rho^2(\tilde{x}) \geq \theta_\rho^2(x^{j_k}) + (\xi_k)^T\left(\tilde{x} - x^{j_k}\right).$$

Taking the limit $k \to +\infty$ in the above formula, we obtain that

$$\theta_\rho^2(\tilde{x}) \geq \theta_\rho^2(x^*) + (\xi^*)^T(\tilde{x} - x^*),$$

which is a contradiction.

Now, we will show that $D^* \in N_\Omega(x^*)$. By $D^{j_k+1} \in N_\Omega(x^{j_k+1})$, for each $\tilde{x} \in \Omega$ we have

$$\langle D^{j_k+1}, \tilde{x} - x^{j_k+1}\rangle \leq 0$$

which implies

$$\langle D^*, \tilde{x} - x^*\rangle \leq 0.$$

Then x^* is a stationary point of Problem (4). This completes the proof.

2.3. **Implementation issues.** Associating the penalty method and the majorization method, we obtain the following so-called majorized penalty algorithm for solving Problem (3).

A Majorized Penalty Algorithm $[MPA]$

Step 0: Given the tolerance $\varepsilon > 0$, penalty parameter $\rho_1 > 0$ and a real number $\mu > 1$. Choose $x^0 \in \Omega$.

Step 1: For $k = 1, 2, \ldots$

Step 1.1 Starting with x^{k-1} as the initial point, apply the majorization algorithm $MA(\rho_k)$ to find x^k such that

$$x^k = \underset{z\in\Omega}{\operatorname{argmin}}\{\theta_{\rho_k}(x) := f(x) + \rho_k p(x)\}.$$

Step 1.2 If $p(x^k) < \varepsilon$, then stop; otherwise, goto Step 1.3.

Step 1.3 Update ρ_{k+1} by $\rho_{k+1} = \mu\rho_k$.

Remark 2.1. *By the classic convergence results of the penalty methods (cf.* [11, Chapter 12.1]*) and Theorem 2.3, we can obtain the convergence of the MPA.*

In this subsection, we shall address some practical issues in the implementation of the majorized penalty method for solving Problem (3).

(1) *The choice of the subdifferential* ξ_j. In order to make the calculation fast, we select the $\xi^j \in \partial\theta_\rho^2(x^j)$ satisfying that $\xi_j = \rho\sum_{i\in\alpha_j}(A_i^T - B_i^T)$.

(2) *The choice of the initial point* $x^0 \in \Omega$. We take the solution of $\min_{x\in\Omega} f(x)$ as the initial point with the Matlab command "quadprog".

(3) *The solver for the subproblems.* In the algorithm $MA(\rho)$, we need to solve the following quadratic optimization problem with linear constraints:

$$\begin{array}{ll} \min_{x\in\mathcal{X}} & \frac{1}{2}x^TGx + c^Tx + \rho\langle e, Ax+a\rangle - (\xi^j)^T(x-x^j) \\ \text{s.t.} & Ax+a \geq 0, \quad Bx+b \geq 0. \end{array} \tag{18}$$

In the implementation, we consider $\mathcal{X} = R^n$ and use "quadprog" to solve (18).

(4) *The stopping criterion.* In the implementation, we terminate the algorithm $MA(\rho)$ for a given penalty parameter ρ, if

$$\|x^{j+1} - x^j\| \leq 10^{-6} \quad \text{or} \quad \frac{|\sqrt{\theta_\rho(x^j)} - \sqrt{\theta_\rho(x^{j-1})}|}{\max\{1, \sqrt{\theta_\rho(x^j)}\}} \leq 10^{-6},$$

and terminate the outer part of the penalized majorization algorithm if

$$|p(x^j)| \leq 10^{-8}.$$

3. Numerical Results.

In this section, we report our numerical experiments conducted for demonstrating the performance of our algorithm.

Many functions can be selected to be penalty term. For instance, if we denote $u = Ax+a$ and $v = Bx+b$, we can set $p(u,v)$ by:

Method 1: $p(u,v) := \sum_{i=1}^m u_i - \sum_{i=1}^m \max\{u_i - v_i, 0\}$;
Method 2: $p(u,v) := \langle u, v\rangle$;
Method 3: $p(u,v) := \|\sum_{i=1}^m u_i - \sum_{i=1}^m \max\{u_i - v_i, 0\}\|_p (p \neq 1)$.

In this paper, we select $p(u,v)$ as Method 1 (see (4)). There mainly exist three reasons. One is the exactness of Method 1. The second is that, when we apply majorization method in Subsection 2.2, we can discover that, the error only exist at the break points (it is just the same with the original problem at most points). The third is that: in Method 2, A^TB is not necessarily positive definite and then it may bring difficulty in solving the parital penalized problem; moreover, the quadratic objective function will be majored by a linear function, and then it may bring a larger error for the approximation according to the analysis of the following majorization method; as a general case of Method 1, Method 3 is obviously more complicated to deal with than Method 1.

If we set $P(x) = \langle Ax+a, Bx+b\rangle$ as the penalty term (denoted Method 2), with the consideration that $\langle Ax+a, Bx+b\rangle = \frac{1}{4}\left(\|(Ax+a)+(Bx+b)\|^2 - \|(Ax+a)-(Bx+b)\|^2\right)$, the corresponding penalty function, majorization function and the objecctive function Θ_ρ, $\tilde{\Theta}_\rho$, and $\tilde{\Theta}_\rho^j$ can be written as

$$\begin{aligned} \Theta_\rho(x) &= \frac{1}{2}x^TGx + c^Tx + \frac{\rho}{4}\|(A+B)x + (a+b)\|^2 - \frac{\rho}{4}\|(A-B)x + (a-b)\|^2, \\ \Theta_\rho^1(x) &= \frac{1}{2}x^TGx + c^Tx + \frac{\rho}{4}\|(A+B)x + (a+b)\|^2, \\ \Theta_\rho^2(x) &= \frac{\rho}{4}\|(A-B)x + (a-b)\|^2, \end{aligned}$$

$$\begin{aligned}\tilde{\Theta}_\rho(x,s) &= \frac{1}{2}x^TGx + c^Tx + \frac{\rho}{4}\|(A+B)x+(a+b)\|^2 - \frac{\rho}{4}\|(A-B)s+(a-b)\|^2 \\ &\quad -\frac{\rho}{2}((A-B)s+(a-b))^T(A-B)(x-s), \\ \tilde{\Theta}^j_\rho(x) &= \frac{1}{2}x^TGx + c^Tx + \frac{\rho}{4}\|(A+B)x+(a+b)\|^2 - \frac{\rho}{4}\|(A-B)x^j+(a-b)\|^2 \\ &\quad -\frac{\rho}{2}((A-B)x^j+(a-b))^T(A-B)(x-x^j).\end{aligned}$$

In Method 1, a piecewise linear function $\theta^2_\rho(x)$ is approximated as a linear function with the majorization function. In Method 2, a quadratic function $\Theta^2_\rho(x)$ is approximated as a linear function with the majorization function. The figures (Figure 1 and Figure 2) are to show the error of the Method 1 and Method 2. When we set $m = n = 1$, $X = R$, $A = 2$, $B = 1$, $a = b = 0$, $\rho = 1$ and $s = 0$, then we have $\theta^2_\rho(x) = \max\{x, 0\}$. By applying the

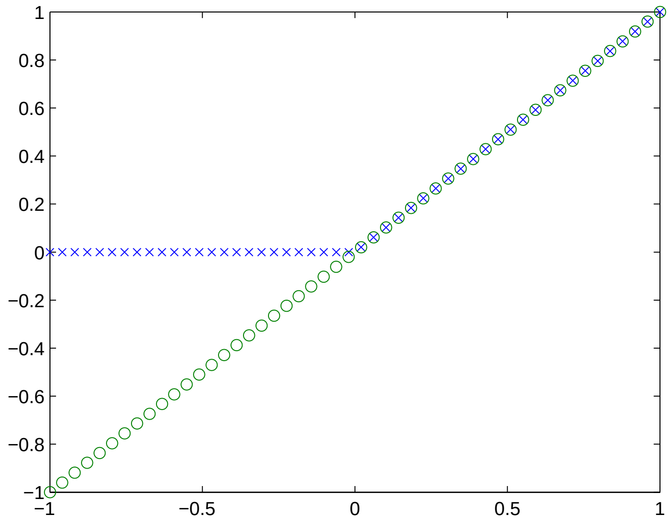

FIGURE 1. The error with Method 1 ($\xi = 1$)

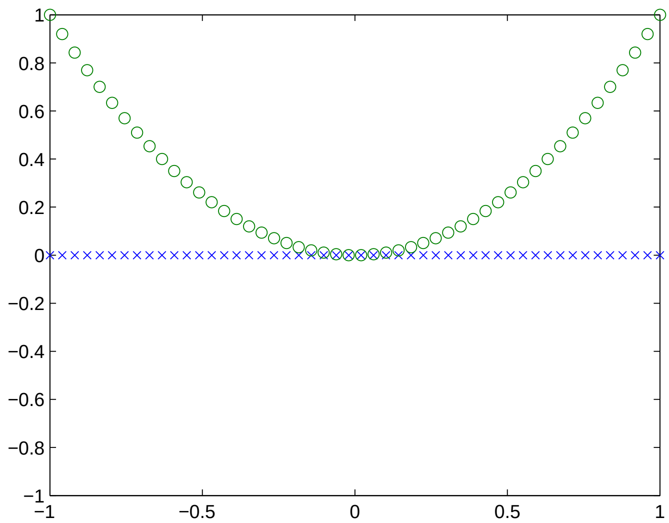

FIGURE 2. The error with Method 2

majorization method in Subsection 2.2, we know $\theta_\rho^2(x)$ is approximated as the function ξx, $\xi \in [0,1]$ (see Figure 1). At this time, $\Theta_\rho^2(x) = x^2$, and it is approximated as 0 (see Figure 2).

To show the efficiency of our penalty term, we select the following examples are to illustrate the difference between Method 1 and Method 2. The algorithm is implemented in `MATLAB` language. All numerical experiments are performed on a Laptop of Intel Core 2 Duo CPU 2.8GHz with 4GB RAM memory, running Windows 7 and `MATLAB` 8.3 (`R2014a`).

Example 3.1. *Consider the following optimization problem*

$$\begin{array}{ll} \min_{x\in\Re} & \frac{1}{2}x^2 + x \\ \text{s.t.} & 0 \le (x+1) \perp (x+3) \ge 0. \end{array} \tag{19}$$

Obviously, we know the optimal solution is $x^* = -1$. Now we apply the penalty majorization approach to solve (19).

Example 3.2. *Let $x \in \Re^4$. Now we consider the Problem 2 in [7] by removing the box constraint. To do this, we re-describe the problem in the formation (1):*

$$a = (-36, -25)^T,\ b = (0,0)^T,\ c = (-30, -30, -15, -15)^T;$$

$$G = \begin{pmatrix} 2 & 2 & 1 & 1 \\ 2 & 2 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix};\ A = \begin{pmatrix} 8/3 & 2 & 2 & 8/3 \\ 2 & 5/4 & 5/4 & 2 \end{pmatrix};\ B = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Let "infeas" denote the infeasibility of Problem (4) at the final iteration of the algorithm, i.e.,

$$\text{infeas} := p(x^*),$$

where x^* is the finally iterative value when the majorized penalty algorithm terminates. The numerical results are shown in the following tables.

Table 1. Example 3.1

	x^*	infeas	time
Method 1	-1	0	0.018
Method 2	-1	5.8091e-12	0.04

Table 2. Example 3.2

	x^*	infeas	time
Method 1	$(5.95, 6.17, 1.83, 1.55)^T$	-7.1054e-15	0.17
Method 2	$(5.95, 6.17, 1.83, 1.55)^T$	6.5743e-11	0.5

4. **Conclusions.** This paper has studied the inverse linear second order cone programming problem in which the parameters in both the objective function and the constraint set need to be adjusted. This inverse problem can be formulated as a linear second order cone complementarity constrained optimization problem. To solve the formulated problem, we propose the majorized penalty method. The main idea of the majorized penalty method is to first consider a sequence of penalized problem and then to apply the majorization method to the penalized problems. Numerical results conducted for randomly generated inverse LSOCP problems demonstrate that our approach is quite effective. Future research work related to this paper is in progress studying the majorized penalty algorithm for solving other inverse linear/quadratic conic optimization problems.

Acknowledgment. The first author's research is supported by the general project of science and technology, the Education Department of Liaoning Province (Grant No. L20130-80) and the Doctor startup fund of Shenyang Aerospace University (Grant No. 13YB14). The second author's research is supported by the National Natural Science Foundation of China (Grant No. 11371255) and the Program for Liaoning Excellent Talents in University. The third author's research is supported by the National Natural Science Foundation of China (Grant No. 11326187), and Doctor startup fund of Shenyang Aerospace University (Grant No. 13YB15).

REFERENCES

[1] N. H. Chieu and G. M. Lee, Constraint qualifications for mathematical programs with equilibrium constraints and their local preservation property, *Journal of Optimization Theory and Applications*, vol.163, no.3, pp.755-776, 2014.

[2] J. De Leeuw, Applications of convex analysis to multidimensional scaling, in *Recent Developments in Statistics*, J. R. Barra, F. Brodeau, G. Romier and B. van Cutsem (eds.), Amsterdam, The Netherlands, 1977.

[3] J. De Leeuw, Convergence of the majorization method for multidimensional scaling, *Journal of Classification*, vol.5, pp.163-180, 1988.

[4] J. De Leeuw and W. J. Heiser, Convergence of correction matrix algorithms for multidimensional scaling, in *Geometric Representations of Relational Data*, J. C. Lingoes, I. Borg and E. E. C. I. Roskam (eds.), Mathesis Press, 1977.

[5] J. De Leeuw, *A Decomposition Method for Weighted Least Squares Low-Rank Approximation of Symmetric Matrices*, Department of Statistics, UCLA, 2006.

[6] L. C. Evans and R. F. Gariepy, Measure theory and tine properties of functions, *Stud. Adv. Math.*, CRC, 1992.

[7] M. Fukushima, Z.-Q. Luo and J.-S. Pang, Globally convergent sequential quadratic programming algorithm for mathematical problems with linear complementarity constraints, *Computational Optimization and Applications*, vol.10, pp.5-34, 1998.

[8] L. Guo, G.-H. Lin and J. J. Ye, Solving mathematical programs with equilibrium constraints, *Journal of Optimization Theory and Applications*, doi:10.1007/s10957-014-0699-z, 2015.

[9] Y. Gao and D. F. Sun, A majorized penalty approach for calibrating rank constrained correlation matrix problems, *Technical Report*, National University of Singapore, 2010.

[10] H. A. L. Kiers, Setting up alternating least squares and iterative majorization algorithm for solving various matrix optimization problems, *Computational Statistics & Data Analysis*, vol.41, pp.157-170, 2002.

[11] D. G. Luenberger, *Linear and Nonlinear Programming*, 2nd Edition, Kluwer Academic Pulishers, 1984.

[12] Z.-Q. Luo, J.-S. Pang and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, 1996.

[13] X. Chen and J. J. Ye, A class of quadratic programs with linear complementarity constraints, *Set-Value Anal.*, vol.17, pp.113-133, 2009.

[14] Z.-Q. Luo, Convergence analysis of primal-dual interior point algorithms for convex quadratic programs, *Recent Trends in Optimization Theory and Applications*, pp.255-270, 1995.

[15] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM, Philadelphia, PA, 1994.
[16] Y. Y. Lin and J. S. Pang, Iterative methods for large quadratic programms: A survey, *SIAM J. Control Optim.*, vol.25, pp.383-411, 1987.
[17] J. J. Moré and S. J. Wright, *Optimization Software Guide*, SIAM, Philadelphia, PA, 1993.
[18] J. S. Pang, Methods for quadratic programmming: A survey, *Computers Chem. Engineering*, vol.7, pp.583-594, 1983.
[19] J. Hu, J. E. Mitchell, J. S. Pang, K. P. Bennett and G. Kunapuli, On the global solution of linear programs with linear complementarity constraints, *SIAMJ. Optim.*, vol.19, pp.445-471, 2008.
[20] S. Scholtes, Active set methods for inverse linear complementarity problems, *Research Papers in Management Studies*, University of Cambridge, No.28, 1999.
[21] L. J. Bai, J. E. Mitchell and J. S. Pang, On convex quadratic programs with linear complementarity constraints, *Comput. Optim. Appl.*, vol.54, pp.517-554, 2013.
[22] A. K. Aziz, A. B. Stephens and M. Suri, Nemerical methods for reaction-diffusion problems with non-differentiablekinetics, *Numer. Math.*, vol.51, pp.1-11, 1988.
[23] F. Kikuchi, K. Nakazato and T. Ushijima, Finite element approximation of a nonlinear eigenvalue problem related to MHD equilibra, *Japan. J. Appl. Math.*, vol.1, pp.369-404, 1984.
[24] J. Rappaz, Approximation of a nondifferentiable nonlinear problem related to MHD equilibria, *Numer. Math.*, vol.45, pp.117-133, 1984.
[25] X. M. Hu and D. Ralph, A note on sensitivity of value functions of mathematical programs with complementarity constraints, *Math. Program., Ser. A*, vol.93, pp.265-279, 2002.
[26] H. Scheel and S. Scholtes, Mathematical programs with complementarity constraints: Stationary, optimality and sensitivity, *Mathematics of Operations Research*, vol.25, no.1, pp.1-22, 2000.
[27] J. J. Ye, Optimal conditions for optimization problems with complementarity constraints, *SIAM. J. OPTIM.*, vol.9, no.2, pp.374-387, 1999.
[28] C. Ding, D. F. Sun and J. J. Ye, First order optimality conditions for mathematical programs with semidefinite cone complementarity constraints, *Technical Report*, 2010.
[29] S. Y. Wang, Y.-J. Liu and Y. Jiang, A majorized penalty approach to inverse linear second order cone programming problems, *Journal of Industrial and Management Optimization*, vol.10, no.3, pp.965-976, 2014.
[30] J. V. Outrata, On the numerical solution of a class of stachelberg problems, *Zeitschrigt for Operations Research*, vol.4, pp.255-278, 1990.
[31] R. T. Rockfellar, *Convex Analysis*, Princeton University Press, Prince, 1970.
[32] Y. G. Evtushenko, *Numerical Optimization Yechniques*, Optimization Software, Inc., 1985.