Journal of Computational Mathematics
Vol.27, No.6, 2009, 787–801.

http://www.global-sci.org/jcm
doi:10.4208/jcm.2009.09-m2674

SOLVING A CLASS OF INVERSE QP PROBLEMS BY A SMOOTHING NEWTON METHOD*

Xiantao Xiao and Liwei Zhang
Department of Applied Mathematics, Dalian University of Technology, Dalian 116024, China
Email: xiaoxiantao82@yahoo.com.cn, lwzhang@dlut.edu.cn

Abstract

We consider an inverse quadratic programming (IQP) problem in which the parameters in the objective function of a given quadratic programming (QP) problem are adjusted as little as possible so that a known feasible solution becomes the optimal one. This problem can be formulated as a minimization problem with a positive semidefinite cone constraint and its dual (denoted IQD(A,b)) is a semismoothly differentiable (SC^1) convex programming problem with fewer variables than the original one. In this paper a smoothing Newton method is used for getting a Karush-Kuhn-Tucker point of IQD(A,b). The proposed method needs to solve only one linear system per iteration and achieves quadratic convergence. Numerical experiments are reported to show that the smoothing Newton method is effective for solving this class of inverse quadratic programming problems.

Mathematics subject classification: 90C20, 90C25, 90C90.
Key words: Fischer-Burmeister function, Smoothing Newton method, Inverse optimization, Quadratic programming, Convergence rate.

1. Introduction

For solving an optimization problem, we usually assume that the parameters, associated with decision variables in the objective function or in the constraint set, are known and we need to find an optimal solution to the problem. However, in the practice there are many instances in which we only know some estimates for parameter values, but we may have certain optimal solutions from experience, observations or experiments. An inverse optimization problem is to find values of parameters which make the known solutions optimal and which differ from the given estimates as little as possible.

Burton and Toint (1992) [3] first investigated an inverse shortest paths problem, since then there are many important contributions to inverse optimization and a large number of inverse combinatorial optimization problems have been studied, see the survey paper by Heuberger [6] and the references [1,2,4], etc. For continuous optimization, Zhang and Liu [14,15] first studied inverse linear programming, Iyengar and Kang [7] discussed inverse conic programming models and their applications in portfolio optimization. And recently, Zhang and Zhang [16] studied the rate of convergence of the augmented Lagrangian method for a type of inverse quadratic programming (IQP) problems. The quadratic programming problem, considered in [16], is of the form

$$\mathrm{QP}(G,c,A,b) \qquad \begin{aligned} \min \quad & f(x) := \frac{1}{2}x^T G x + c^T x \\ \text{s.t.} \quad & x \in \Omega_P := \{x' \in \mathbb{R}^n \mid Ax' \geq b\}, \end{aligned} \tag{1.1}$$

* Received October 29, 2007 / Revised version received March 24, 2008 / Accepted February 20, 2009 /

where $G \in \mathbb{R}^{n\times n}$ is a symmetric matrix, $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Let

$$A := (a_1, \dots, a_m)^T, \quad a_i \in \mathbb{R}^n,\, i = 1, \dots, m,$$

$\mathbb{S}^n$ denote the space of $n \times n$ symmetric matrices, and SOL(P) be the set of optimal solutions to a problem (P).

Given a feasible point $x^0 \in \Omega_P$, which should be the optimal solution to Problem (1.1) and a pair $(G^0, c^0) \in \mathcal{S}^n \times \mathbb{R}^n$ which is an estimate of (G, c). The inverse quadratic programming considered in this paper is to find a pair $(G, c) \in \mathbb{S}^n \times \mathbb{R}^n$ to solve

$$\text{IQP}(A,b) \qquad \begin{array}{ll} \min & \frac{1}{2}\|(G,c) - (G^0,c^0)\|^2 \\ \text{s.t.} & x^0 \in \text{SOL}(\text{QP}(G,c,A,b)), \\ & (G,c) \in \mathbb{S}^n_+ \times \mathbb{R}^n, \end{array} \tag{1.2}$$

where $\mathbb{S}^n_+$ is the cone of positively semi-definite symmetric matrices in $\mathbb{S}^n$ and $\|\cdot\|$ is defined by

$$\|(G', c')\| := \sqrt{\text{Tr}(G'^T G') + c'^T c'} \;\text{ for }\; (G', c') \in \mathbb{R}^{n\times n} \times \mathbb{R}^n.$$

Problem (1.2) is a cone-constrained optimization problem with a quadratic objective function. The scale of this problem will be quite large when n is a large number as the number of its decision variables is $n + n(n+1)/2$.

Without loss of generality, we assume that the first p constraints in Ω_P are active at x^0, or equivalently

$$I(x^0) := \{j : a_j^T x^0 = b_j, j = 1, \dots, m\} = \{1, \dots, p\}.$$

If $G \in \mathbb{S}^n_+$, then $x^0 \in \text{SOL}(\text{QP}(G, c, A, b))$ if and only if there exists $u \in \mathbb{R}^p$ such that

$$c + Gx^0 - \sum_{i=1}^{p} u_i a_i = 0,\; u_i \geq 0, i = 1, \dots, p.$$

Let $A_0 := (a_1, \dots, a_p)^T \in \mathbb{R}^{p\times n}$ and the $j-$th column of A_0 be $A_j \in \mathbb{R}^p$. Then $A_0 := (A_1, \dots, A_n)$ and the problem (1.2) can be equivalently expressed as follows

$$\begin{array}{ll} \min & \frac{1}{2}\|(G,c) - (G^0,c^0)\|^2 \\ \text{s.t.} & c + Gx^0 - A_0^T u = 0, \\ & (G,c,u) \in \mathbb{S}^n_+ \times \mathbb{R}^n \times \mathbb{R}^p_+. \end{array} \tag{1.3}$$

As the dimension of the above problem is $n(n+1)/2+n+p$, quite big when n is large, it would be helpful to consider its dual. It follows from [16] that the dual problem can be written as

$$\text{IQD}(A,b) \qquad \begin{array}{ll} \max & \upsilon_0(z) \\ \text{s.t.} & A_0 z \leq 0, \end{array} \tag{1.4}$$

where

$$\upsilon_0(z) = -\frac{1}{2}\|z\|^2 + c^{0T} z - \frac{1}{2}\|\Pi_{\mathcal{S}^n_+}(\bar{G}(z))\|_F^2 + \frac{1}{2}\|G^0\|_F^2, \tag{1.5}$$

and

$$\bar{G}(z) = G^0 - \mathcal{B}z, \quad \mathcal{B}z := \frac{zx^{0T} + x^0 z^T}{2}.$$

Obviously, $\mathcal{B}:\mathbb{R}^n\to\mathbb{S}^n$ is a linear operator and its adjoint $\mathcal{B}^*:\mathbb{S}^n\to\mathbb{R}^n$ is given by $\mathcal{B}^*G = Gx^0$. Problem (1.4) has a dimension n, much smaller than that of problem (1.3) when n is large. Based on [16, Lemma 3.3], if z^* is the unique solution to IQD(A,b), then $(G^*,c^*) = (\Pi_{\mathbb{S}^n_+}(\bar{G}(z^*)), c^0 - z^*)$ solves the original problem (1.2).

In this paper, we consider the smoothing Newton method, developed by [13], for getting a Karush-Kuhn-Tucker point of IQD(A,b).

Throughout this paper the following notations will be used. We write $X\succ 0$ and $X\succeq 0$ if X is positive definite and positive semidefinite, respectively, we denote the symmetric square root of X by $X^{\frac{1}{2}}$ and its trace by Tr(X). Let $|X| := (X^2)^{\frac{1}{2}}$ and $\Pi_{S^n_+}(X) := (X+|X|)/2$ for any $X\in\mathbb{S}^n$. For matrices $X,Y\in\mathbb{S}^n$, the Frobenius inner product is defined as

$$X\bullet Y := \mathrm{Tr}(\mathrm{X}^{\mathrm{T}}\mathrm{Y}),$$

and the Frobenius norm of $X\in\mathbb{S}^n$ is

$$\|X\|_F := (X\bullet X)^{1/2}.$$

The Hadamard product of X and Y is denoted by $X\circ Y$, namely $(X\circ Y)_{ij} := X_{ij}Y_{ij}$. Let I be the identity matrix of appropriate dimension.

This paper is organized as follows. In Section 2, we give some results from nonsmooth analysis which shall be used in our convergence analysis. In Section 3, we describe the smoothing Newton method for IQD(A,b) and prove the global convergence and the quadratic convergence rate. Numerical results implemented by the smoothing Newton method are given in Section 4.

2. Preliminary

In this section, we recall some results on semi-smooth mappings and properties of some smoothing functions, which will be used in the sequel. Let X and Y be two finite dimensional real vector spaces. Let $\mathcal{O}$ be an open set in X and $\Psi:\mathcal{O}\subseteq X\to Y$ be a locally Lipschitz continuous function on the open set $\mathcal{O}$. By Rademacher's theorem, Ψ is almost everywhere Fréchet-differentiable in $\mathcal{O}$. We denote by $\mathcal{D}_\Psi$ the set of the point where Ψ is Fréchet-differentiable in $\mathcal{O}$. Then, the Bouligand-subdifferential of Ψ at $x\in\mathcal{O}$, denoted by $\partial_B\Psi(x)$, is

$$\partial_B\Psi(x) := \left\{\lim_{k\to\infty}\mathcal{J}\Psi(x^k)\,|\,x^k\in\mathcal{D}_\Psi, x^k\to x\right\},$$

where $\mathcal{J}\Psi(x^k)$ denotes the Jacobian of Ψ at x^k. Clarke's generalized Jacobian of Ψ at x is the convex hull of $\partial_B\Psi(x)$, i.e.,

$$\partial\Psi(x) = \mathrm{conv}\,\{\partial_B\Psi(x)\}\ .$$

The following concept of semismoothness was first introduced by Mifflin [8] for functionals and was extended by Qi and Sun [10] to vector valued functions.

Definition 2.1. *Let $\Psi:\mathcal{O}\subseteq X\to Y$ be a locally Lipschitz continuous function on the open set $\mathcal{O}$. We say that Ψ is semismooth at a point $x\in\mathcal{O}$ if*

(i) *Ψ is directionally differentiable at x; and*

(ii) *for any $\Delta x\in X$ and $V\in\partial\Psi(x+\Delta x)$ with $\Delta x\to 0$,*

$$\Psi(x+\Delta x)-\Psi(x)-V(\Delta x) = o(\|\Delta x\|)\,. \tag{2.1}$$

Furthermore, Ψ is said to be strongly semismooth at $x \in \mathcal{O}$ if Ψ is semismooth at x and for any $\Delta x \in X$ and $V \in \partial\Psi(x+\Delta x)$ with $\Delta x \to 0$,

$$\Psi(x+\Delta x) - \Psi(x) - V(\Delta x) = O(\|\Delta x\|^2). \tag{2.2}$$

The following lemma on the Bouligand-subdifferential of composite functions is proved in [12, Lemma 2.1] .

Lemma 2.1. *Let $F : X \to Y$ be a continuously differentiable function on an open neighborhood $\mathcal{O}$ of $\bar{x} \in X$ and $\Psi : \mathcal{O}_Y \subseteq Y \to X'$ be a locally Lipschitz continuous function on an open set $\mathcal{O}_Y$ containing $\bar{y} := F(\bar{x})$, where X' is a finite-dimensional real vector space. Suppose Ψ is directionally differentiable at every point in $\mathcal{O}_Y$, then*

$$\partial_B(\Psi \circ F)(\bar{x}) \subseteq \partial_B\Psi(\bar{y})\mathcal{J}F(\bar{x})\,. \tag{2.3}$$

Moreover, if the Jacobian mapping $\mathcal{J}F(\bar{x}) : X \to Y$ is onto, then the above inclusion becomes an equality, namely

$$\partial_B(\Psi \circ F)(\bar{x}) = \partial_B\Psi(\bar{y})\mathcal{J}F(\bar{x})\,. \tag{2.4}$$

For $\varepsilon \in \mathbb{R}$ and $X \in \mathbb{S}^n$, the *square smoothing function* $\Phi : \mathbb{R} \times \mathbb{S}^n \to \mathbb{S}^n$, see [13], is defined by

$$\Phi(\varepsilon, X) := (\varepsilon^2 I + X^2)^{1/2}, \ \ \forall\ (\varepsilon, X) \in \mathbb{R} \times \mathbb{S}^n.$$

Then, Φ is continuously differentiable at (ε, X) unless $\varepsilon = 0$ and for any $Y \in \mathbb{S}^n$,

$$\Phi(\varepsilon, X) \to |Y|, \quad \text{as} \ \ (\varepsilon, \mathrm{X}) \to (0, \mathrm{Y}).$$

For any $X \in \mathbb{S}^n$, let L_X be the Lyapunov operator:

$$L_X(Y) := XY + YX, \quad \forall\ Y \in \mathbb{S}^n,$$

with L_X^{-1} being its inverse (if it exists at all), i.e., for any $Y \in \mathbb{S}^n$, $L_X^{-1}(Y)$ is the unique $Z \in \mathbb{S}^n$ satisfying $XZ + ZX = Y$. The following result is proved in [13, Lemma 2.3, Theorem 2.5 and Proposition 3.1].

Lemma 2.2. *For $(\varepsilon, X) \in \mathbb{R} \times \mathbb{S}^n$, assume there exist an orthogonal matrix P and a matrix $\Lambda = \mathrm{diag}(\mu_1, \cdots, \mu_\mathrm{n})$ of eigenvalues of X such that $X = P\Lambda P^T$, the following statements hold.*

(1) *If $\varepsilon^2 I + X^2$ is nonsingular, then Φ is continuously differentiable at (ε, X), where $\mathcal{J}\Phi(\varepsilon, X)$ satisfies the following equations*

$$\mathcal{J}\Phi(\varepsilon, X)(\tau, H) = L^{-1}_{\Phi(\varepsilon,X)}(L_X(H) + 2\varepsilon\tau I), \quad \forall\ (\tau, H) \in \mathbb{R} \times \mathbb{S}^n,$$

and for $i, j = 1, \cdots, n$,

$$(P^T\mathcal{J}\Phi(\varepsilon, X)(\tau, H)P)_{ij} = \begin{cases} \dfrac{(P^T HP)_{ij}(\mu_i + \mu_j)}{(\varepsilon^2 + \mu_i^2)^{1/2} + (\varepsilon^2 + \mu_j^2)^{1/2}}, & \text{if } i \neq j, \\[2ex] \dfrac{\mu_i(P^T HP)_{ii} + \varepsilon\tau}{(\varepsilon^2 + \mu_i^2)^{1/2}}, & \text{otherwise}. \end{cases} \tag{2.5}$$

(2) *Φ is strongly semismooth at $(0, X)$.*

(3) *For* $(0, H) \in \mathbb{R} \times \mathbb{S}^n$ *and* $V \in \partial_B \Phi(0, X)$*, it holds that*

$$V(0, H) = P(\Omega \circ P^T H P)P^T,$$

where $\Omega \in \mathbb{S}^n$ *has entries*

$$\Omega_{ij} = \begin{cases} t \in [-1, 1], & \text{if } \mu_i = \mu_j = 0, \\ \dfrac{\mu_i + \mu_j}{|\mu_i| + |\mu_j|}, & \text{otherwise.} \end{cases}$$

3. Smoothing Newton Method

This section focuses on the convergence analysis of the smoothing Newton method for getting a Karush-Kuhn-Tucker point of IQD(A, b). It is easy to see that v_0 is continuously differentiable with

$$\nabla v_0(z) = z - c^0 - \mathcal{B}^* \Pi_{S^n_+}(\bar{G}(z)).$$

The Lagrangian of the problem IQD(A, b) is

$$L(z, \lambda) := v_0(z) + \lambda^T A_0 z$$

and its gradient is

$$\nabla_z L(z, \lambda) = z - c^0 - \mathcal{B}^* \Pi_{S^n_+}(\bar{G}(z)) + A_0^T \lambda.$$

The Karush-Kuhn-Tucker optimality conditions for the problem IQD(A, b) are

$$\nabla_z L(z, \lambda) = z - c^0 - \mathcal{B}^* \Pi_{S^n_+}(\bar{G}(z)) + A_0^T \lambda = 0,$$

$$\lambda_i \geq 0,\ \lambda_i a_i^T z = 0,\ a_i^T z \leq 0,\ i = 1, \cdots, p,$$

which can be equivalently reformulated as

$$\nabla_z L(z, \lambda) = z - c^0 - \mathcal{B}^* (\bar{G}(z) + |\bar{G}(z)|)/2 + A_0^T \lambda = 0, \tag{3.1a}$$
$$\sqrt{\lambda_i^2 + (a_i^T z)^2} - \lambda_i + a_i^T z = 0,\ i = 1, \cdots, \tag{3.1b}$$

Inspired by [13], we define $F : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^{n+p}$ as follows

$$F(\varepsilon, z, \lambda) = \begin{bmatrix} z - c^0 - \mathcal{B}^* \bar{G}(z)/2 - \mathcal{B}^* (\bar{G}(z)^2 + \varepsilon^2 I)^{1/2}/2 + A_0^T \lambda \\ \sqrt{\lambda_1^2 + (a_1^T z)^2 + 4\varepsilon^2} - \lambda_1 + a_1^T z \\ \sqrt{\lambda_2^2 + (a_2^T z)^2 + 4\varepsilon^2} - \lambda_2 + a_2^T z \\ \vdots \\ \sqrt{\lambda_p^2 + (a_p^T z)^2 + 4\varepsilon^2} - \lambda_p + a_p^T z \end{bmatrix}. \tag{3.2}$$

Then $F(\varepsilon, z, \lambda)$ is continuously differentiable at $(\varepsilon, z, \lambda)$ with $\varepsilon \neq 0$ and strongly semismooth everywhere. Noting that $v_0(z)$ is strongly convex, we have that any solution to

$$E(\varepsilon, z, \lambda) = \begin{bmatrix} \varepsilon \\ F(\varepsilon, z, \lambda) \end{bmatrix} = 0 \tag{3.3}$$

is the optimal solution to the problem IQD(A, b). From [16, Lemma 3.3], there exists a unique solution to the problem IQD(A, b). Hence, (3.3) has a unique solution which is the solution to the problem IQD(A, b).

The *smoothing Newton method* is based on solving $E(\varepsilon, z, \lambda) = 0$ and uses the merit function $\phi(Z) := \|E(\varepsilon, z, \lambda)\|^2$ for the line search, where $Z = (\varepsilon, z, \lambda)$. Let $\bar{\varepsilon} > 0$ and $\eta \in (0, 1)$ be such that $\eta\bar{\varepsilon} < 1$. Define an auxiliary point $\bar{Z}$ by $\bar{Z} := (\bar{\varepsilon}, 0, 0) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^p$ and $\theta : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}_+$ by $\theta(Z) := \eta \min\{1, \phi(Z)\}$. The smoothing Newton method, proposed by [9, 13], can be described as follows:

Algorithm 3.1.

Step 1 *Select constants* $\delta \in (0, 1)$ *and* $\sigma \in (0, 1/2)$. *Let* $\varepsilon^0 := \bar{\varepsilon}$, $(z^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}^p$ *be an arbitrary point. Then let the initial point* $Z^0 = (\varepsilon^0, z^0, \lambda^0)$ *and* $k := 0$.

Step 2 *If* $E(Z^k) = 0$, *then stop; otherwise, let* $\theta_k := \theta(Z^k)$.

Step 3 *Compute* $\Delta Z^k := (\Delta\varepsilon^k, \Delta z^k, \Delta\lambda^k) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^p$ *by*

$$E(Z^k) + \mathcal{J}E(Z^k)(\Delta Z^k) = \theta_k \bar{Z}. \tag{3.4}$$

Step 4 *Let* l_k *be the smallest nonnegative integer* l *satisfying*

$$\phi(Z^k + \delta^l \Delta Z^k) \leq (1 - 2\sigma(1 - \eta\bar{\varepsilon})\delta^l)\phi(Z^k). \tag{3.5}$$

Define $Z^{k+1} = Z^k + \delta^{l_k} \Delta Z^k$,

Step 5 $k := k + 1$, *go to Step* 2.

Below we give the global convergence results on the above algorithm. The following propositions will be used in the proof of the forthcoming theorem.

Proposition 3.1. *For any* $(\varepsilon, z, \lambda) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^p$ *with* $\varepsilon \neq 0$, $\mathcal{J}E(\varepsilon, z, \lambda)$ *is nonsingular.*

Proof. By Lemma 2.2, we know that $\mathcal{J}E(\varepsilon, z, \lambda)$ exists. Suppose there exists $(\bar{\varepsilon}, \bar{z}, \bar{\lambda}) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^p$ such that

$$\mathcal{J}E(\varepsilon, z, \lambda)(\bar{\varepsilon}, \bar{z}, \bar{\lambda}) = 0,$$

i.e.,

$$\begin{bmatrix} \bar{\varepsilon} \\ \bar{z} + \frac{1}{2}\mathcal{B}^*\mathcal{B}\bar{z} - \frac{1}{2}\mathcal{B}^*\mathcal{J}\Phi(\varepsilon, \bar{G}(z))(\bar{\varepsilon}, -\mathcal{B}\bar{z}) + A_0^T\bar{\lambda} \\ \dfrac{\lambda_1\bar{\lambda}_1 + (a_1^Tz)(a_1^T\bar{z}) + 4\varepsilon\bar{\varepsilon}}{\sqrt{\lambda_1^2 + (a_1^Tz)^2 + 4\varepsilon^2}} - \bar{\lambda}_1 + a_1^T\bar{z} \\ \vdots \\ \dfrac{\lambda_p\bar{\lambda}_p + (a_p^Tz)(a_p^T\bar{z}) + 4\varepsilon\bar{\varepsilon}}{\sqrt{\lambda_p^2 + (a_p^Tz)^2 + 4\varepsilon^2}} - \bar{\lambda}_p + a_p^T\bar{z} \end{bmatrix} = 0, \tag{3.6}$$

which implies that $\bar{\varepsilon} = 0$ and

$$\frac{\lambda_i \bar{\lambda}_i + (a_i^T z)(a_i^T \bar{z})}{\sqrt{\lambda_i^2 + (a_i^T z)^2 + 4\varepsilon^2}} - \bar{\lambda}_i + a_i^T \bar{z} = 0, \ i = 1, \cdots, p,$$

i.e.,

$$\left(\frac{a_i^T z}{\sqrt{\lambda_i^2 + (a_i^T z)^2 + 4\varepsilon^2}} + 1 \right) a_i^T \bar{z} = \left(1 - \frac{\lambda_i}{\sqrt{\lambda_i^2 + (a_i^T z)^2 + 4\varepsilon^2}} \right) \bar{\lambda}_i, \quad i = 1, \cdots, p.$$

By denoting

$$r_i := \frac{a_i^T z}{\sqrt{\lambda_i^2 + (a_i^T z)^2 + 4\varepsilon^2}} + 1, \quad q_i := 1 - \frac{\lambda_i}{\sqrt{\lambda_i^2 + (a_i^T z)^2 + 4\varepsilon^2}}, \ \ i = 1, \cdots, p,$$

we have

$$r_i a_i^T \bar{z} = q_i \bar{\lambda}_i, \quad i = 1, \cdots, p.$$

Since $\varepsilon \neq 0$ implies $r_i > 0, q_i > 0, i = 1, \cdots, p$, we have

$$\bar{\lambda}_i = \frac{r_i}{q_i} a_i^T \bar{z}, \quad i = 1, \cdots, p.$$

Denote $C := \mathrm{diag}\,(r_1/q_1, r_2/q_2, \cdots, r_p/q_p)$. Then $C \in \mathbb{S}^p_+$ and $\bar{\lambda} = C A_0 \bar{z}$. From (3.6), we know that

$$\bar{z} + \frac{1}{2}\mathcal{B}^* \mathcal{B} \bar{z} - \frac{1}{2}\mathcal{B}^* \mathcal{J}\Phi(\varepsilon, \bar{G}(z))(0, -\mathcal{B}\bar{z}) + A_0^T \bar{\lambda} = 0,$$

i.e.,

$$\frac{1}{2}\mathcal{B}^* \mathcal{J}\Phi(\varepsilon, \bar{G}(z))(0, -\mathcal{B}\bar{z}) - \frac{1}{2}\mathcal{B}^* \mathcal{B} \bar{z} = (I + A_0^T C A_0)\bar{z},$$

which implies

$$\frac{1}{2}\bar{z}^T \mathcal{B}^* \mathcal{J}\Phi(\varepsilon, \bar{G}(z))(0, -\mathcal{B}\bar{z}) - \frac{1}{2}\bar{z}^T \mathcal{B}^* \mathcal{B} \bar{z} = \bar{z}^T (I + A_0^T C A_0)\bar{z}. \tag{3.7}$$

Suppose $\bar{z} \neq 0$, then together with the fact that $(I + A_0^T C A_0) \succ 0$, we have

$$\bar{z}^T (I + A_0^T C A_0)\bar{z} > 0.$$

Since $\bar{G}(z) \in \mathbb{S}^n$, there exist an orthogonal matrix P and a diagonal matrix

$$\Lambda = \mathrm{diag}(\mu_1, \cdots, \mu_\mathrm{n})$$

of eigenvalues of $\bar{G}(z)$ such that $\bar{G}(z) = P \Lambda P^T$. Hence

$$P^T \mathcal{J}\Phi(\varepsilon, \bar{G}(z))(0, -\mathcal{B}\bar{z})P = L^{-1}_{\Phi(\varepsilon,\Lambda)}(L_\Lambda(-P^T \mathcal{B}\bar{z}P)).$$

By a direct calculation, for $i, j = 1, \cdots, n$,

$$(P^T \mathcal{J}\Phi(\varepsilon, \bar{G}(z))(0, -\mathcal{B}\bar{z})P)_{ij} = \frac{(-P^T \mathcal{B}\bar{z}P)_{ij}(\mu_i + \mu_j)}{(\varepsilon^2 + \mu_i^2)^{1/2} + (\varepsilon^2 + \mu_j^2)^{1/2}}.$$

Let $\alpha := \max\{(\mu_i + \mu_j)((\varepsilon^2 + \mu_i^2)^{1/2} + (\varepsilon^2 + \mu_j^2)^{1/2})^{-1}, i, j = 1 \cdots, n\}$. Obviously $\alpha \leq 1$.

From (3.7), we have

$$
\begin{aligned}
0 &< \frac{1}{2}\bar{z}^T\mathcal{B}^*\mathcal{J}\Phi(\varepsilon,\bar{G}(z))(0,-\mathcal{B}\bar{z}) - \frac{1}{2}\bar{z}^T\mathcal{B}^*\mathcal{B}\bar{z} \\
&= \frac{1}{2}(\mathcal{B}\bar{z}) \bullet \mathcal{J}\Phi(\varepsilon,\bar{G}(z))(0,-\mathcal{B}\bar{z}) - \frac{1}{2}(\mathcal{B}\bar{z}) \bullet (\mathcal{B}\bar{z}) \\
&= \frac{1}{2}\mathrm{Tr}(\mathcal{B}\bar{z}\mathcal{J}\Phi(\varepsilon,\bar{G}(z))(0,-\mathcal{B}\bar{z})) - \frac{1}{2}\mathrm{Tr}(\mathcal{B}\bar{z}\mathcal{B}\bar{z}) \\
&= \frac{1}{2}\mathrm{Tr}(PP^T\mathcal{B}\bar{z}PP^T\mathcal{J}\Phi(\varepsilon,\bar{G}(z))(0,-\mathcal{B}\bar{z})PP^T) - \frac{1}{2}\mathrm{Tr}(\mathcal{B}\bar{z}\mathcal{B}\bar{z}) \\
&\le \frac{1}{2}(\alpha-1)\mathrm{Tr}(\mathcal{B}\bar{z}\mathcal{B}\bar{z}) \le 0.
\end{aligned}
$$

This shows $\bar{z}=0$, then $\bar{\lambda}=CA_0\bar{z}=0$. Thus $\mathcal{J}E(\varepsilon,z,\lambda)$ is nonsingular. □

Proposition 3.1 shows that Algorithm 3.1 is well defined. We state it formally in the following theorem.

Theorem 3.1. *The sequence $\{Z^k\}$ generated by Algorithm* 3.1 *converges to the solution of $E(Z)=0$.*

Proof. From (3.5), we know that

$$
\infty > \sum_{k=0}^{\infty}(\phi(Z^k)-\phi(Z^{k+1})) \ge \sum_{k=0}^{\infty} 2\sigma(1-\eta\bar{\varepsilon})\delta^l)\phi(Z^k)
$$

and $\phi(Z^k)$ is strict monotone decreasing, then we have $\{\phi(Z^k)\}$ converges to 0, which, together with the fact that $E(Z)=0$ has a unique solution, implies $\{Z^k\}$ converges to the solution of $E(Z)=0$. The proof is complete. □

The next proposition is about the nonsingularity of the B-differential of E at $(0,z,\lambda)\in\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^p$.

Proposition 3.2. *Suppose A_0 is of full row rank. Then any element $U\in\partial_B E(0,z,\lambda)$ is nonsingular.*

Proof. Let U be an element of $\partial_B E(0,z,\lambda)$. Assume that there exists $(\bar{\varepsilon},\bar{z},\bar{\lambda})\in\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^p$ such that $U(\bar{\varepsilon},\bar{z},\bar{\lambda})=0$. We have $\bar{\varepsilon}=0$ and

$$
\begin{bmatrix}
\bar{z}+\frac{1}{2}\mathcal{B}^*\mathcal{B}\bar{z}-\frac{1}{2}\mathcal{B}^*V(0,-\mathcal{B}\bar{z})+A_0^T\bar{\lambda} \\
\\
W_1(\bar{z},\bar{\lambda})-\bar{\lambda}_1+a_1^T\bar{z} \\
\vdots \\
W_p(\bar{z},\bar{\lambda})-\bar{\lambda}_p+a_p^T\bar{z}
\end{bmatrix} = 0, \tag{3.8}
$$

where $V\in\partial_B\Phi(0,\bar{G}(z))$ and $W_i\in\partial\sqrt{(a_i^Tz)^2+(\lambda_i)^2}, i=1,\cdots,p$.

Since $\bar{G}(z)\in\mathbb{S}^n$, there exist an orthogonal matrix P and a diagonal matrix

$$
\Lambda=\mathrm{diag}(\mu_1,\cdots,\mu_{\mathrm{n}})
$$

of eigenvalues of $\bar{G}(z)$ such that $\bar{G}(z)=P\Lambda P^T$. From Lemma 2.2, for $i,j=1,\cdots,n$,

$$
(P^TV(0,-\mathcal{B}\bar{z})P)_{ij}=\Omega_{ij}*(P^T(-\mathcal{B}\bar{z})P)_{ij},
$$

where the matrix $\Omega \in \mathbb{S}^n$ has entries

$$\Omega_{ij} = \begin{cases} t \in [-1,1], & \text{if } \mu_i = \mu_j = 0, \\ \dfrac{\mu_i + \mu_j}{|\mu_i| + |\mu_j|}, & \text{otherwise.} \end{cases}$$

Obviously $\Omega_{ij} \le 1$, for $i, j = 1, \cdots, n$.

Noting the fact that the Euclidean norm $f(x) = \sqrt{x^T x}$ is subdifferentiable at every $x \in \mathbb{R}^n$ and the set $\partial f(0)$ is the Euclidean unit ball, we have that for $i = 1, \cdots, p$

$$W_i(\bar{z}, \bar{\lambda}) - \bar{\lambda}_i + a_i^T \bar{z} = \begin{cases} w_{i1} a_i^T \bar{z} + w_{i2} \bar{\lambda}_i - \bar{\lambda}_i + a_i^T \bar{z} & \text{if } \lambda_i = a_i^T z = 0, \\ \dfrac{\lambda_i \bar{\lambda}_i + (a_i^T z)(a_i^T \bar{z})}{\sqrt{(\lambda_i)^2 + (a_i^T z)^2}} - \bar{\lambda}_i + a_i^T \bar{z} & \text{otherwise,} \end{cases} \tag{3.9}$$

where $(w_{i1}, w_{i2}) \in \{(\zeta_1, \zeta_2) \in \mathbb{R}^2 | \ \zeta_1^2 + \zeta_2^2 \le 1\}$. Hence, from (3.8), for $i = 1, \cdots, p$

$$\bar{\lambda}_i = \begin{cases} 0 & \text{if } \lambda_i = a_i^T z = 0 \ \text{ and } (w_{i1}, w_{i2}) = (-1, 0), \\ t_i a_i^T \bar{z} & \text{otherwise,} \end{cases} \tag{3.10}$$

where $t_i \ge 0$. Then we obtain

$$\bar{z}^T A_0^T \bar{\lambda} = \sum_{i=1}^{p} (a_i^T \bar{z}) \bar{\lambda}_i \ge 0.$$

From the first equation of (3.8), we have

$$\bar{z} + \frac{1}{2} \mathcal{B}^* \mathcal{B} \bar{z} - \frac{1}{2} \mathcal{B}^* V(0, -\mathcal{B} \bar{z}) + A_0^T \bar{\lambda} = 0,$$
$$\bar{z}^T \bar{z} + \frac{1}{2} \bar{z}^T \mathcal{B}^* \mathcal{B} \bar{z} - \frac{1}{2} \bar{z}^T \mathcal{B}^* V(0, -\mathcal{B} \bar{z}) + \bar{z}^T A_0^T \bar{\lambda} = 0.$$

By applying the proof of Proposition 3.1, we have $\bar{z} = 0$. Thus, $A_0^T \bar{\lambda} = 0$. As A_0 is of full row rank , we have $\bar{\lambda} = 0$. Hence, U is nonsingular. □

We now state the quadratic convergence of Algorithm 3.1 in the following theorem. Based on Proposition 3.1 and Proposition 3.2, its proof is a direct application of [13, theorem 4.2].

Theorem 3.2. *Suppose A_0 is of full row rank. Then the whole sequence $\{Z^k\}$ generated by Algorithm 3.1 converges to the solution Z^* of $E(Z) = 0$ with*

$$\|Z^{k+1} - Z^*\| = \mathcal{O}(\|Z^k - Z^*\|^2)$$

and

$$\varepsilon^{k+1} = \mathcal{O}((\varepsilon^k)^2).$$

Proof. From Proposition 3.1 and Proposition 3.2, as A_0 is of full row rank, then for all Z^k sufficiently close to Z^*, $U \in \partial_B E(Z^k)$ is nonsingular. Then compute ΔZ^k by (3.4), we have $\Delta Z^k = U^{-1}[-E(Z^k) + \theta_k \bar{Z}]$. Hence, as E is strongly semismooth at Z^*, from the definition of strongly semismooth, for all Z^k sufficiently close to Z^*, we have

$$\begin{aligned}\|Z^k + \Delta Z^k - Z^*\| &= \|Z^k + U^{-1}[-E(Z^k) + \theta_k \bar{Z}] - Z^*\| \\ &= \mathcal{O}(\|E(Z^k) - E(Z^*) - U(Z^k - Z^*)\| + \theta_k\|\bar{\varepsilon}\|) \\ &= \mathcal{O}(\|Z^k - Z^*\|^2) + \mathcal{O}(\phi(Z^k)) \\ &= \mathcal{O}(\|Z^k - Z^*\|^2) + \mathcal{O}((E(Z^k) - E(Z^*))^2) \\ &= \mathcal{O}(\|Z^k - Z^*\|^2).\end{aligned}$$

Noting that for all Z^k sufficiently close to Z^*,

$$\begin{aligned}\phi(Z^k + \Delta Z^k) &= \mathcal{O}(\|Z^k + \Delta Z^k - Z^*\|^2) \\ &= \mathcal{O}(\|Z^k - Z^*\|^4) = \mathcal{O}(\phi(Z^k)^2),\end{aligned}$$

we have $Z^{k+1} = Z^k + \Delta Z^k$. Then we can easily get

$$\|Z^{k+1} - Z^*\| = \mathcal{O}(\|Z^k - Z^*\|^2).$$

Next, from the definition of θ_k and the fact that $Z^k \to Z^*$ as $k \to \infty$, for all k sufficiently large, we have

$$\theta_k = \eta\phi(Z^k).$$

Also, for all k sufficiently large, because $Z^{k+1} = Z^k + \Delta Z^k$, we have

$$\varepsilon^{k+1} = \varepsilon^k + \Delta\varepsilon^k = \theta_k\bar{\varepsilon} = \eta\bar{\varepsilon}\phi(Z^k).$$

Hence, for all k sufficiently large,

$$\lim_{k\to\infty} \frac{\varepsilon^{k+1}}{(\varepsilon^k)^2} = \lim_{k\to\infty} \frac{\phi(Z^k)}{\eta\bar{\varepsilon}\phi(Z^{k-1})^2} = \mathcal{O}(1).$$

This proves $\varepsilon^{k+1} = \mathcal{O}((\varepsilon^k)^2)$. The proof is complete. □

4. Numerical Experiments

In this section, we report our numerical experiments of Algorithm 3.1 carried out in Matlab 7.1 running on a PC Intel Pentium IV of 2.80 GHz CPU. After a few preliminary tests, the following choices were made and used in all the test examples.

- In Step 3 of Algorithm 3.1, as $\mathcal{J}E(Z^k)$ is nonsymmetric and its explicit form is complicate, we use CGS method (*conjugate gradient square method*) [11] to solve (3.4).
- In Step 2 of Algorithm 3.1, instead of $E(Z^k) = 0$, the stopping criterion in the experiments we adopted is $Tol. := \phi(Z^k) < 10^{-5}$.
- We set other parameters in the algorithm as $\eta = 0.5, \sigma = 0.3, \delta = 0.5$.

We test the following class of examples:

Example 4.1. *Consider the following simple quadratic program,*

$$\begin{aligned} \min \quad & f(x) := \frac{1}{2}x^T G x + c^T x \\ \text{s.t.} \quad & Ax \geq b, \quad x \geq 0, \end{aligned} \tag{4.1a}$$

where

$$G = \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix}, \quad c = \begin{bmatrix} -2 \\ -6 \end{bmatrix}, \quad A = \begin{bmatrix} -0.5 & -0.5 \\ 1 & -2 \end{bmatrix}, \quad b = \begin{bmatrix} -1 \\ -2 \end{bmatrix}. \tag{4.1b}$$

The optimal solution of (4.1a) *is* $x^* = (0.8, 1.2)$. *Let* $x^0 = (0, 0)$ *and* (G^0, c^0) *be an estimate of (G,c), we consider the following inverse quadratic programming problem,*

$$\begin{aligned} \min \quad & \frac{1}{2}\|(G, c) - (G^0, c^0)\|^2 \\ \text{s.t.} \quad & x^0 \in SOL(4.1a), \\ & (G, c) \in \mathbb{S}^2_+ \times \mathbb{R}^2. \end{aligned} \tag{4.2}$$

As described in the Section Introduction, A_0 (the 2×2 identity matrix) can be simply derived and problem (4.2) can be equivalently expressed as

$$\begin{aligned} \min \quad & \frac{1}{2}\|(G, c) - (G^0, c^0)\|^2 \\ \text{s.t.} \quad & c + Gx^0 - A_0^T u = 0, \\ & (G, c, u) \in \mathbb{S}^2_+ \times \mathbb{R}^2 \times \mathbb{R}^2_+. \end{aligned} \tag{4.3}$$

Since x^0 is an zero vector and A_0 is an identity matrix, we can easily know that the solution (G^*, c^*) to problem (4.2) is always $(\Pi_{\mathbb{S}^2_+}(G^0), \Pi_{\mathbb{R}^2_+}(c^0))$ and $u^* = c^*$. Table 1 presents the results of our experiments on Example 4.1, in which z^* denotes the optimal solution to the dual problem of (4.2), (G^*, c^*) denotes the optimal solution to problem (4.2).

From Table 1, it can be verified that every pair of (G^*, c^*) actually equals to

$$(\Pi_{\mathbb{S}^2_+}(G^0), \Pi_{\mathbb{R}^2_+}(c^0)),$$

which means that the smoothing Newton method indeed gives the correct solutions. Noticing that u^* is made up of the Lagrangian multipliers with respect to the active constraints and $u^* = c^*$, from Table 1 we can see that the strict complementary condition is not required in this method.

Table 1: Numerical results of Example 4.1

G^0	c^0	z^*	G^*	c^*
$\begin{bmatrix} 2.5 & -2.8 \\ -2.8 & 4.5 \end{bmatrix}$	$\begin{bmatrix} -2.5 \\ -6.5 \end{bmatrix}$	$\begin{bmatrix} -2.50 \\ -6.50 \end{bmatrix}$	$\begin{bmatrix} 2.50 & -2.80 \\ -2.80 & 4.50 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$
$\begin{bmatrix} 1 & -2 \\ -2 & 2 \end{bmatrix}$	$\begin{bmatrix} 0.5 \\ -5.5 \end{bmatrix}$	$\begin{bmatrix} 0 \\ -5.50 \end{bmatrix}$	$\begin{bmatrix} 1.35 & -1.73 \\ -1.73 & 2.21 \end{bmatrix}$	$\begin{bmatrix} 0.50 \\ 0 \end{bmatrix}$
$\begin{bmatrix} 0 & -1 \\ -1 & 2 \end{bmatrix}$	$\begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0.35 & -0.85 \\ -0.85 & 2.06 \end{bmatrix}$	$\begin{bmatrix} 0.50 \\ 0.50 \end{bmatrix}$

Example 4.2. *We still consider problem* (4.1*a*)*, let*

$$x^0 = \begin{bmatrix} \frac{2}{3} \\ \frac{4}{3} \end{bmatrix}, \quad G^0 = \begin{bmatrix} 3 & -1 \\ -1 & 5 \end{bmatrix}, \quad c^0 = \begin{bmatrix} -1 \\ -5 \end{bmatrix}.$$

Then the inverse problem of (4.1*a*) *is formed as*

$$\begin{array}{ll} \min & \frac{1}{2}\|(G,c)-(G^0,c^0)\|^2 \\ \text{s.t.} & c+Gx^0-A_0^T u=0, \\ & (G,c,u)\in \mathbb{S}^2_+ \times \mathbb{R}^2 \times \mathbb{R}^2_+, \end{array} \tag{4.4}$$

where A_0 is of the form A given in (4.1b)*.*

Example 4.3. *The following problem named* HS76 *comes from CUTEr set of problems* [5]*:*

$$\begin{array}{ll} \min & f(x) := \frac{1}{2}x^T Gx + c^T x \\ \text{s.t.} & Ax \geq b, \end{array} \tag{4.5}$$

where

$$G = \begin{bmatrix} 2 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 2 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \quad c = \begin{bmatrix} -1 \\ -3 \\ 1 \\ -1 \end{bmatrix}, \quad A = \begin{bmatrix} -1 & -2 & -1 & -1 \\ -3 & -1 & -2 & 1 \\ 0 & 1 & 4 & 0 \end{bmatrix}, \quad b = \begin{bmatrix} -5 \\ -4 \\ 1.5 \end{bmatrix}.$$

Let

$$x^0 = \begin{bmatrix} 0 \\ 1.5 \\ 0 \\ 2 \end{bmatrix}, \quad G^0 = \begin{bmatrix} 3 & 0 & -1 & 0 \\ 0 & 2 & 0 & 0 \\ -1 & 0 & 3 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}, \quad c^0 = \begin{bmatrix} 0 \\ -2 \\ 2 \\ 0 \end{bmatrix}.$$

Then we consider the inverse problem of (4.3) *as follows*

$$\begin{array}{ll} \min & \frac{1}{2}\|(G,c)-(G^0,c^0)\|^2 \\ \text{s.t.} & c+Gx^0-A_0^T u=0, \\ & (G,c,u)\in \mathbb{S}^4_+ \times \mathbb{R}^4 \times \mathbb{R}^2_+, \end{array} \tag{4.6}$$

where $A_0 = \begin{bmatrix} -1 & -2 & -1 & -1 \\ 0 & 1 & 4 & 0 \end{bmatrix}$.

Example 4.4. *The following problem named* S268 *comes from CUTEr set of problems* [5]*:*

$$\begin{array}{ll} \min & f(x) := \frac{1}{2}x^T Gx + c^T x \\ \text{s.t.} & Ax \geq b, \end{array} \tag{4.7}$$

where

$$G = \begin{bmatrix} 20394 & -24908 & -2026 & 3896 & 658 \\ -24908 & 41818 & -3466 & -9828 & -372 \\ -2026 & -3466 & 3510 & 2178 & -348 \\ 3896 & -9828 & 2178 & 3030 & -44 \\ 658 & -372 & -348 & -44 & 54 \end{bmatrix}, \quad c = \begin{bmatrix} 18340 \\ -34198 \\ 4542 \\ 8672 \\ 86 \end{bmatrix},$$

$$A = \begin{bmatrix} -1 & -1 & -1 & -1 & -1 \\ 10 & 10 & -3 & 5 & 4 \\ -8 & 1 & -2 & -5 & 3 \\ 8 & -1 & 2 & 5 & -3 \\ -4 & -2 & 3 & -5 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} -5 \\ 20 \\ -40 \\ 11 \\ -30 \end{bmatrix}.$$

Let $x^0 = (1,1,1,1,1)^T$ *and*

$$G^0 = \begin{bmatrix} 20000 & -20000 & -2000 & 3000 & 600 \\ -20000 & 4000 & -3000 & -10000 & -300 \\ -2000 & -3000 & 3000 & 2000 & -300 \\ 3000 & -10000 & 2000 & 3000 & -40 \\ 600 & -300 & -300 & -40 & 50 \end{bmatrix}, \quad c^0 = \begin{bmatrix} 10000 \\ -30000 \\ 4000 \\ 8000 \\ 80 \end{bmatrix}.$$

Then we consider the inverse problem of (4.3) *as follows*

$$\begin{array}{ll} \min & \frac{1}{2}\|(G,c) - (G^0,c^0)\|^2 \\ \text{s.t.} & c + Gx^0 - A_0^T u = 0, \\ & (G,c,u) \in \mathbb{S}^5_+ \times \mathbb{R}^5 \times \mathbb{R}^2_+, \end{array} \tag{4.8}$$

where $A_0 = \begin{bmatrix} -1 & -1 & -1 & -1 & -1 \\ 8 & -1 & 2 & 5 & -3 \end{bmatrix}$.

Table 2: Numerical results of Examples 4.2, 4.3 and 4.4

Example 4.2		Example 4.3		Example 4.4	
k	$\phi(Z^k)$	k	$\phi(Z^k)$	k	$\phi(Z^k)$
0	5.85×10^1	0	3.75×10^2	0	7.26×10^8
1	6.22×10^{-1}	1	1.73×10^1	2	9.93×10^6
2	6.35×10^{-2}	2	3.93×10^{-1}	4	1.40×10^1
3	1.81×10^{-2}	3	1.71×10^{-2}	6	6.25×10^{-2}
4	3.70×10^{-4}	4	2.05×10^{-3}	7	2.44×10^{-4}
5	3.69×10^{-6}	5	2.89×10^{-6}	8	3.73×10^{-9}

Table 2 demonstrates the asymptotic convergence rate of Algorithm 3.1 on Examples 4.2, 4.3 and 4.4, where k denotes the kth iteration ($k = 0$ stands for the initial iteration) and $\phi(Z^k)$ denotes the value of the merit function $\phi(Z^k)$. As shown in Table 2, the convergence is stable and the quadratic rate is observable.

Example 4.5. *Let* G^0 *and* c^0 *be a random* $n \times n$ *symmetric matrix and a random* $n \times 1$ *vector, respectively.* A_0 *is a random* $p \times n$ *matrix. For convenience, we set the elements of* x^0 *all* 1. *We report our numerical results for* $n = 100, 200, 500, 1000$ *and* $p = n/10$.

When solving Example 4.5, the initial point (z^0, λ^0) is chosen to be the vector whose entries are all ones. Our numerical results are reported in Table 3, where Iter., Func., Res0. and Res∗. stand for, respectively, the number of iterations, the number of function evaluations and the residuals $\Phi(\cdot)$ at the starting point and the final iterate of implementation.

Table 3: Numerical results of Example 4.5

n	p	cputime	Iter.	Func.	Res0.	Res$*$.
20	2	0.4s	6	7	2.01×10^{4}	3.48×10^{-7}
50	5	3.1s	8	9	2.92×10^{5}	1.29×10^{-7}
100	10	24.7s	10	11	2.29×10^{6}	1.06×10^{-7}
200	20	2m48.5s	12	13	1.82×10^{7}	3.83×10^{-9}
500	50	55m41.4s	16	17	2.82×10^{8}	9.80×10^{-8}
1000	100	5h26m25.0s	13	16	2.26×10^{9}	3.97×10^{-6}

Based on Table 3, the largest numerical example we tested in this paper is $n = 1000, p = 100$. In this case there are roughly 500,000 unknowns in the primal problem. In consideration of the scale of problem solved,the smoothing Newton method is very effective. Since the vast majority of our computer cputime is spent on the preconditioner conjugate gradient square method for solving the linear system (3.4), it would save much computing time if a better preconditioner for (3.4) is found.

Acknowledgments. The authors are grateful to the anonymous referees for their helpful comments and suggestions on improving the quality of this paper. The research is supported by the National Natural Science Foundation of China under project No. 10771026 and by the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry of China.

References

[1] R.K. Ahuja, J.B. Orlin, Inverse optimization, *Oper. Res.*, **49** (2001), 771-783.

[2] R.E. Burkard, Y. Lin, J. Zhang, Weight reduction problems with certain bottleneck objectives, *Eur. J. Oper. Res.*, **153** (2004), 191-199.

[3] W.R. Burton, Ph.L. Toint, On an instance of the inverse shortest paths problem, *Math. Program.*, **53** (1992), 45-61.

[4] M. Cai, X. Yang, J. Zhang, The complexity analysis of the inverse center location problem, *J. Glob. Optim.*, **15** (1999), 213-218.

[5] A.R. Conn, N. Gould, Ph.L. Toint, Constrained and unconstrained testing environment. `http://hsl.rl.ac.uk/cuter-www/`.

[6] C. Heuberger, Inverse combinatorial optimization: a survey on problems, methods and results, *J. Comb. Optim.*, **8** (2004),329-361.

[7] G. Iyengar, W. Kang, Inverse conic programming and applications, *Oper. Res. Lett.*, **33** (2005), 319-330.

[8] R. Mifflin, Semismooth and semiconvex functions in constrained optimization, *SIAM J. Control. Optim.*, **15** (1997), 957-972.

[9] L. Qi, D. Sun, G. Zhou, A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities, *Math. Program.*, **87** (2000), 1-35.

[10] L. Qi, J. Sun, A nonsmooth version of Newton's method, *Math. Program.*, **58** (1993), 353-367.

[11] P. Sonneveld, CGS: A fast Lanczos-type solver for nonsymmetric linear systems, *SIAM J. Sci. Comput.*, **10** (1989), 36-52.

[12] D. Sun, The strong second order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications, *Math. Oper. Res.*, **31** (2006), 761-776.

[13] J. Sun, D. Sun, L. Qi, A Squared Smoothing Newton Method for Nonsmooth Matrix Equations and Its Applications in Semidefinite Optimization Problems, *SIAM J. Optimiz.*, **14** (2004), 783-806.

[14] J. Zhang, Z. Liu, Calculating some inverse linear programming problems, *J. Comput. Appl. Math.*, **72** (1996), 261-273.

[15] J. Zhang, Z. Liu, A further study on inverse linear programming problems, *J. Comput. Appl. Math.*, **106** (1999), 345-359.

[16] J. Zhang, L. Zhang, *An Augument Lagrangian Method for a Type of Inverse Quadratic Programming Problems*, Report, City University of Hong Kong, 2006.