NUMERICAL ALGRBRA,
CONTROL AND OPTIMIZATION
Volume **1**, Number **1**, March **2011**

doi:10.3934//naco.2011.1.1

pp. **1–14**

ON METHODS FOR SOLVING NONLINEAR SEMIDEFINITE OPTIMIZATION PROBLEMS

Jie Sun

School of Business
National University of Singapore, 119245, Singapore

Abstract. The nonlinear semidefinite optimization problem arises from applications in system control, structural design, financial management, and other fields. However, much work is yet to be done to effectively solve this problem. We introduce some new theoretical and algorithmic development in this field. In particular, we discuss first and second-order algorithms that appear to be promising, which include the alternating direction method, the augmented Lagrangian method, and the smoothing Newton method. Convergence theorems are presented and preliminary numerical results are reported.

1. **Introduction.** The purpose of this paper is to review some recent algorithmic developments in solving nonlinear semidefinite optimization problems (also called nonlinear semidefinite *programs*). As a natural extension of (linear) semidefinite programs (SDP), nonlinear semidefinite programs (NSDP) arise in various application fields such as system control and financial engineering [1, 12, 20, 21, 30] and has caught much attention on its theory and algorithms [3, 5, 7, 8, 9, 10, 13, 23, 24, 25, 31, 32, 33, 34, 35, 36, 37, 46, 48, 49, 51, 54, 55, 57, 62, 63]. While the mathematical formats of NSDP may be different in various applications, we find it is convenient to start with the following *general model*

$$\begin{array}{ll} \min_{x\in X} & f(x) \\ \text{s.t.} & h(x) = 0, \\ & g(x) \in K, \end{array} \tag{1}$$

where $f : X \mapsto \Re$, $h : X \mapsto \Re^m$ and $g : X \mapsto Y$ are twice continuously differentiable functions, X and Y are two Hilbert spaces equipped with respective inner products and the respective induced norms, and K is a symmetric (homogeneous, self-dual) cone in Y. In addition, one of the spaces X and Y must be the space $\mathcal{S}^n$ of real $n \times n$ symmetric matrices. Here and below, unless otherwise stated, the vector norm is the Euclidean norm and the matrix norm is the Frobenius norm, namely $\langle x, y\rangle \equiv \text{Trace}\,(x^T y)$, where "$T$" stands for the transpose.

2000 *Mathematics Subject Classification.* Primary: 90C22; Secondary: 90C25.

Key words and phrases. Alternating direction method, augmented Lagrangian method, semismooth Newton method.

This paper is dedicated to Professor Masao Fukushima, a great friend and a distinguished mathematician, in celebration of his 60th birthday. The presentation contains some joint work of the author with Xiongda Chen, Defeng Sun, Liwei Zhang, Su Zhang, and others. The author would like to express his appreciation to his collaborators. The author is also responsible to all errors in this paper.

Example 1.1. (Pair of semidefinite programming) Let $x \in \mathcal{S}^n$ and $y \in \Re^m$. The primal and dual SDPs are defined as

$$(\mathrm{P}) \begin{cases} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b, x \succeq 0 \end{cases} \qquad (\mathrm{D}) \begin{cases} \max & \langle b, y\rangle \\ \text{s.t.} & A^*y \preceq c, \end{cases}$$

where $c \in \mathcal{S}^n, b \in \Re^m$ and A^* is the adjoint operator of the linear operator A. It is readily seen that both problems are in the form of (1).

Notice that in problem (P) the constraint $x \succeq 0$ or equivalently, $x \in S^n_+$, is a nonlinear convex constraint; whereas in problem (D), the constraint $A^*y \preceq c$ is also a nonlinear convex constraint. Hence (P) and (D) are essentially a pair of convex programming problems. Strong duality and efficient algorithms exist under more restrictive conditions than linear programming. This is also the case for problem (1) when $f(x), h(x)$ are linear and $g(x) \in K$ defines a convex set. However, problem (1) could be substantially more difficult if one of the functions $f(x), h(x)$, and $g(x)$, or one of the constraints in $g(x) \in K$ is not convex.

Example 1.2. (Convex semidefinite feasibility problems) Let $X = \mathcal{S}^n_+, K = \Re^m_-$, $f(x) = 0, h(x) = 0$, and let $g_i(x)$ be convex. Then problem (1) is a convex semidefinite feasibility problem.

The feasibility problem is important, not only because it arises in theory, for example, the feasibility of the Karush-Kuhn-Tucker (KKT) system is sometimes equivalent to the solvability to both primal and dual NSDP problems, but also because it may come directly from practice. For instance, arising from Lyapunov stability analysis of systems under uncertainty [2], it is desired to know whether the following system about matrix y is feasible

$$\begin{cases} \lambda y - (x_i^T y + y^T x_i) \in \mathcal{S}^m_+ \quad i = 1, ..., k \\ y - I \in \mathcal{S}^m_+ \\ y \in \mathcal{S}^m_+ \end{cases}$$

for given $\lambda \in \Re$ and $I, x_i \in \mathcal{S}^m,\ i = 1, ..., k$.

Similar to the case of nonlinear programming, associated to problem (1), there is a KKT system. Whenever we consider solving the KKT system, it is reduced to a problem of complementarity. A typical case is as follows.

Example 1.3. (The nonlinear matrix complementarity problem (NMCP))

$$\text{Find } x, y \text{ such that } \ x \in \mathcal{S}^n_+,\ y = F(x) \in \mathcal{S}^n_+, \text{ and } \ \langle x, y\rangle = 0, \tag{2}$$

where F is a given continuously differentiable mapping (operator). Problem (2) includes the nonlinear complementarity problem (NCP) as a special case.

In solving NCP, we often convert the NCP to a nonsmooth equation by using the so-called NCP functions and then solve the latter by a semismooth Newton approach (details later). For example, solving the problem

$$x, y \in \Re^n_+,\ x^T y = 0$$

is equivalent to solving the nonsmooth equation

$$\min[x, y] = 0; \qquad \text{(The min equation)}$$

it is also equivalent to solving

$$x + y - \sqrt{x^2 + y^2} = 0, \qquad \text{(The Fischer} - \text{Burmeister (FB) equation)}$$

where $\min(u)$, u^2, and $\sqrt{u}$ are understood coordinatewise.

Coming to the realm of symmetric matrices, the first question is how we define the functions $\min(x), x^2$, and $\sqrt{x}$ for $x \in S^n$. A natural way to do it is to use the spectral function approach. Let $x \in \mathcal{S}^n$, let $\{\lambda_i(x)\}_{i=1}^n$ be the eigenvalues of x which are sorted in nonincreasing order, i.e.

$$\lambda_1(x) \geq \lambda_2(x) \geq \cdots \geq \lambda_k(x) \geq \cdots \geq \lambda_n(x), \quad d = \text{ diag}\,[\lambda_1, ..., \lambda_n],$$

where "diag" stands for the diagonal matrix. Let $x = qdq^T$ be an orthogonal decomposition of x. Then we define

$$\begin{aligned}
&\max[x, 0] \equiv q\,\text{diag}\,[\max(\lambda_1, 0), ..., \max(\lambda_n, 0)]q^T,\\
&\min[x, y] \equiv x - \max[x - y, 0],\\
&x^2 \equiv qd^2q, \text{ and } \sqrt{x} \equiv q\,\text{diag}\,[\sqrt{\lambda_1}, ..., \sqrt{\lambda_n}]q^T.
\end{aligned}$$

It can be shown that such defined x^2 is identical to the usual matrix multiplication $x \cdot x$.

The next question is whether the matrix min function and the matrix FB function, defined above, are NCP functions for the NMCP. The answer is affirmative. The proofs can be found in Tseng [58, Lemma 2.1 and Lemma 6.1]. Therefore, we have

$$\begin{aligned}
x, y = F(x) \in \mathcal{S}^n_+, \langle x, y\rangle = 0 \quad &\Longleftrightarrow \quad y = F(x), \min[x, y] = 0\\
&\Longleftrightarrow \quad y = F(x), x + y - \sqrt{x^2 + y^2} = 0.
\end{aligned}$$

These equivalence relationships are the basis for the so-called smoothing Newton methods. One of them is introduced in Section 4.

2. A first-order method: The alternating direction method.

The alternating direction method (ADM) is one of the effective first-order approaches for solving NSDP. It was developed for convex programming, see for example [11, 14, 17, 18]. Actually ADM can be seen as the block Gauss-Seidel variants of the augmented Lagrangian approach. The ADM requires much less computational effort per iteration than the second-order approaches such as the interior point methods or the smoothing Newton methods. To avoid solving expensive nonlinear variational problems on semidefinite cone at each iteration, we modify it by using both a prediction phase and a correction phase. In the proposed prediction-correction ADM, the main computational load of each iteration is only several metric projections onto semidefinite cone or onto a convex set defined by a single constaint.

We consider general convex nonlinear semidefinite programs (CNSDPs) defined in $\mathcal{S}^n$ as follows.

$$\min\ q_0(x) \quad \text{s.t.}\ \ q_i(x) \leq 0,\ i = 1, \cdots, m,\ x \succeq 0, \tag{CNSDP}$$

where $q_i : \mathcal{S}^n \to \Re,\ i = 0, 1, \cdots, m$, is a convex continuously differentiable function, and $x \in \mathcal{S}^n$.

By introducing auxiliary variables

$$y_i = x \text{ and } \Omega_i = \{y_i : q_i\,(y_i) \leq 0\},\ i = 1, \cdots, m,$$

we rewrite (CNSDP) equivalently as

$$\begin{array}{ll} \min & q_0(x) \\ \text{s.t.} & x = y_i,\ y_i \in \Omega_i,\ i = 1, \cdots, m \\ & x \succeq 0. \end{array} \tag{3}$$

The Lagrange dual of problem (3) is

$$\max_{\lambda_i} \min_{x \succeq 0, y_i \in \Omega_i} q_0(x) - \sum_{i=1}^{m} \langle \lambda_i, x - y_i \rangle .$$

Notice that the Lagrange multipliers $\lambda_i,\ i = 1, \cdots, m$, are symmetric matrices. It is well known that under mild constraint qualifications (e.g., Slater's condition), strong duality holds and hence, x^* is a solution of (3) if and only if there exists $\lambda_i^* \in \mathcal{S}^n,\ i = 1, \cdots, m$, such that $(x^*, y_i^*, \lambda_i^*)$ satisfies

$$\begin{array}{rcl} \left\langle x - x^*, \nabla q_0\left(x^*\right) - \sum_{i=1}^{m} \lambda_i^* \right\rangle & \geq & 0,\ \forall\ x \in \mathcal{S}_+^n \\ \langle y_i - y_i^*, \lambda_i^* \rangle & \geq & 0,\ \forall\ y_i \in \Omega_i,\ i = 1, \cdots, m \\ x^* - y_i^* & = & 0,\ \forall\ i = 1, \cdots, m \end{array} \tag{4}$$

Problem (4) is a variational inequality problem with a special structure. The variables (x, y_i, λ_i) are symmetric matrices, the underlying sets $\mathcal{S}_+^n$ and Ω_i are convex. For convenience, we state the basic assumption to guarantee that Problem (3) is solvable.

Assumption 2.1. The solution set $(x^*, y_i^*, \lambda_i^*)$ of KKT system (4) is nonempty.

In a nutshell, the original ADM, applied to (4), can be separated into three steps.

Algorithm 2.2. The Original Alternating Direction Method for CNSDP

Do at each iteration until a stopping criterion is met

Step 1. $\left(x^k, y_i^k, \lambda_i^k\right) \to \left(x^{k+1}, y_i^k, \lambda_i^k\right)$, where

$$\left\langle x - x^{k+1}, \nabla q_0\left(x^{k+1}\right) - \sum_{i=1}^{m} \left(\lambda_i^k - \beta_i \left(x^{k+1} - y_i^k\right)\right) \right\rangle \geq 0,\ \forall x \succeq 0$$

$\beta_i,\ i = 1, \cdots, m$, are certain positive scalars.

Step 2. $\left(x^{k+1}, y_i^k, \lambda_i^k\right) \to \left(x^{k+1}, y_i^{k+1}, \lambda_i^k\right)$, $i = 1, \cdots, m$, where

$$\left\langle y_i - y_i^{k+1}, \lambda_i^k - \beta_i \left(x^{k+1} - y_i^{k+1}\right) \right\rangle \geq 0,\ \forall y_i \in \Omega_i$$

Step 3. $\left(x^{k+1}, y_i^{k+1}, \lambda_i^k\right) \to \left(x^{k+1}, y_i^{k+1}, \lambda_i^{k+1}\right)$, $i = 1, \cdots, m$, where

$$\lambda_i^{k+1} = \lambda_i^k - \beta_i \left(x^{k+1} - y_i^{k+1}\right)$$

Arbitrary x^0, y_i^0 and λ_i^0 are chosen as the starting point. The alternating direction method reaches optimality by taking alternating steps in the primal and dual space. Primal feasibility, dual feasibility, and complementary slackness are not maintained; instead, all are satisfied as the algorithm finds a fixed point of the recursions in Steps 1-3.

Algorithm 2.2 is not easy to implement since Steps 1 and 2 require to solve certain (probably nonlinear) variational inequality problems. A prediction-correction

approach is therefore used, which is stated as follows, where the notation $\Pi_{\Omega}(x)$ represents the projection of x onto the set Ω.

Algorithm 2.3. The Prediction-Correction Alternating Direction Method for CNSDP
Do at each iteration until a stopping criterion is met
Prediction Phase:

Step 1. $\left(x^k, y_i^k, \lambda_i^k\right) \to \left(\widetilde{x}^k, y_i^k, \lambda_i^k\right)$, where

$$\widetilde{x}^k = \Pi_{\mathcal{S}_+^n}\left[x^k - \alpha\left(\nabla q_0\left(x^k\right) - \sum_{i=1}^{m}\left(\lambda_i^k - \beta_i\left(x^k - y_i^k\right)\right)\right)\right]$$

Step 2. $\left(\widetilde{x}^k, y_i^k, \lambda_i^k\right) \to \left(\widetilde{x}^k, \widetilde{y}_i^k, \lambda_i^k\right)$, $i = 1, \cdots, m$, where

$$\widetilde{y}_i^k = \Pi_{\Omega_i}\left[\widetilde{x}^k - \frac{1}{\beta_i}\lambda_i^k\right]$$

Step 3. $\left(\widetilde{x}^k, \widetilde{y}_i^k, \lambda_i^k\right) \to \left(\widetilde{x}^k, \widetilde{y}_i^k, \widetilde{\lambda}_i^k\right)$, $i = 1, \cdots, m$, where

$$\widetilde{\lambda}_i^k = \lambda_i^k - \beta_i\left(\widetilde{x}^k - \widetilde{y}_i^k\right)$$

Correction Phase:

Step 4. $\left(\widetilde{x}^k, \widetilde{y}_i^k, \widetilde{\lambda}_i^k\right) \to \left(x^{k+1}, y_i^{k+1}, \lambda_i^{k+1}\right)$, where

$$\begin{aligned} x^{k+1} &= \Pi_{\mathcal{S}_+^n}\left[x^k - \gamma^k R\left(x^k, \widetilde{x}^k\right)\right] \\ y_i^{k+1} &= \Pi_{\Omega_i}\left[y_i^k - \gamma^k\left(y_i^k - \widetilde{y}_i^k\right)\right], \ i = 1, \cdots, m \\ \lambda_i^{k+1} &= \lambda_i^k - \gamma^k\left(\lambda_i^k - \widetilde{\lambda}_i^k\right), \ i = 1, \cdots, m \end{aligned}$$

The positive scalar γ^k is the step-length and $R(x)$ is a certain residual function.

The motivation behind Algorithm 2.3 are roughly as follows. By a well-known equivalence between a variational inequality with respect to a set Ω and the projection onto Ω (e.g., Theorem 2.3 of [27]), one can connect the fixed point problems about the variational inequalities in Algorithm 2.2 to the fixed point problems about projections in Algorithm 2.3. However, to guarantee convergence, a correction step is necessary after each prediction (i.e., projection) step in the "fixed projection" algorithm. This is much in the same spirit as what is done in the predictor-corrector interior point path following algorithms for nonlinear programming (e.g. see [56]).

We have the following convergence result for Algorithm 2.3.

Theorem 2.4. [63, Theorem 3.2] *The sequence $\left\{x^k\right\}$ generated by the prediction-correction alternating direction method converges to an optimal solution point x^* of (CNSDP).*

For other first-order methods, see Kanzow, Nagel, Kato, and Fukushima [26] and Noll and Apkarian [35] and references therein. It should be noted that ADM has been shown to be one of the fastest methods in solving linear SDP in a recent work of Wen, Goldfarb, and Yin [60].

A test result is shown in Table 1, where the problem is a nearest correlation matrix problem [20]. The problem is formulated as a convex quadratic SDP, where C+1e-3*E, C+1e-2*E, C+1e-1*E represent different trust regions of the target matrices, n is the size of these matrices, and (a),(b), and (c) mean different starting points.

		C+1e-3*E		C+1e-2*E		C+1e-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
	b)	20	0.9	21	0.9	28	1.2
	c)	21	1.0	20	0.9	24	1.1
500	a)	10	47.3	14	60.4	23	90.7
	b)	20	95.0	21	92.1	27	111.5
	c)	23	105.1	23	98.5	25	109.1
1000	a)	10	370.7	15	537.9	24	777.3
	b)	20	701.2	22	730.2	29	957.5
	c)	23	809.7	23	791.2	26	843.2
2000	a)	11	2972	14	3843	25	6321
	b)	20	5485	23	6377	31	7956
	c)	24	6362	24	6408	27	6823

TABLE 1. Numerical Results for Nearest Covariance Problems (Source: Table 1 of [55])

3. **The augmented Lagrangian method: A 1^+-order method.** Recall problem (1)

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & h(x) = 0, \\ & g(x) \in K, \\ & x \in X. \end{array}$$

The augmented Lagrangian function with the penalty parameter c for the problem above is defined as

$$L_c(x, \zeta, \xi) \equiv f(x) + \langle \zeta, h(x) \rangle + \frac{c}{2}\|h(x)\|^2 + \frac{1}{2c}\left[\|\Pi_{K^*}(\xi - cg(x))\|^2 - \|\xi\|^2\right], \quad (5)$$

where $(x, \zeta, \xi) \in X \times \Re^m \times Y$, and K^* is the positive dual cone of K: $K^* \equiv \{y : \langle x, y \rangle \geq 0 \ \forall \ x \in K\}$. The augmented Lagrangian method consists of two steps.

1. Determine x^k by minimizing $L_{c_k}(x, \zeta^k, \xi^k)$,
2. Compute (ζ^{k+1}, ξ^{k+1}) by

$$\begin{cases} \zeta^{k+1} = \zeta^k + c_k h(x^k), \\ \xi^{k+1} = \Pi_{K^*}(\xi^k - c_k g(x^k)), \end{cases}$$

and update c_{k+1} by

$$c_{k+1} := c_k \quad \text{or} \quad c_{k+1} := \kappa c_k \quad (\text{for some fixed } \kappa > 1).$$

3.1. **General discussions.** The classical analysis on the augmented Lagrangian method can be found in [19, 38, 39, 42, 43, 44, 45, 47]. Basically, it is proved in theory that this method is globally convergent for certain $c \geq \bar{c}$ with a linear

rate of convergence. However, our analysis [52] shows that the coefficient of the convergence rate is proportional to $\bar{c}^{-\frac{1}{2}}$, namely one has

$$\|(\zeta,\xi)^{k+1} - (\zeta,\xi)^*\| \le \bar{c}^{-\frac{1}{2}}\|(\zeta,\xi)^k - (\zeta,\xi)^*\|.$$

Therefore, as $\bar{c}$ becomes large, the convergence is nearly superlinear. Ignoring a bit of mathematical accuracy, we may call such algorithm a 1^+-order algorithm. This result could be sharpened for nonlinear programming (NP), and is true for both NP and NSDP.

The augmented Lagrangian is continuously differentiable at x with

$$\nabla_x L_c(x,\zeta,\xi) = \nabla f(x) + \nabla h(x)(\zeta + ch(x)) - \nabla g(x)\Pi_K(\xi - cg(x)). \tag{6}$$

Clearly, (6) is not differentiable due to the term $\Pi_K(\xi - cg(x))$. We denote $\partial_B(\nabla_x L_c)(x,\zeta,\xi)$ the B-differential of $\nabla_x L_c$ with respect to (x,ζ,ξ) (See Section 4.1 for the definition of B-differentials) and $\pi_x\,\partial_B(\nabla_x L_c)(x,\zeta,\xi)$ the projection of $\partial_B(\nabla_x L_c)(x,\zeta,\xi)$ onto the X-space.

Define

$$\mathcal{A}_c(\overline{\zeta},\overline{\xi},W) \equiv \nabla^2_{xx} L_0(\overline{x},\overline{\zeta},\overline{\xi}) + c\,\nabla h(\overline{x})\nabla h(\overline{x})^T + c\,\nabla g(x) W(\nabla g(\overline{x})^*).$$

Then

$$\pi_x\,\partial_B(\nabla_x L_c)(\overline{x},\overline{\zeta},\overline{\xi}) = \left\{\mathcal{A}_c(\overline{\zeta},\overline{\xi},W) \mid W \in \partial_B\Pi_K(\overline{\xi} - cg(\overline{x}))\right\}.$$

Assumption B1.*$(\overline{\zeta},\overline{\xi})$ is the unique Lagrange multiplier at $\overline{x}$ and there exist two positive numbers c_0 and $\underline{\eta}$ such that for any $c \ge c_0$ and $W \in \partial_B\Pi_K(\overline{\xi} - cg(\overline{x}))$,*

$$\left\langle d, \mathcal{A}_c(\overline{\zeta},\overline{\xi},W)d\right\rangle \ge \underline{\eta}\,\langle d,d\rangle \quad \forall\, d \in x\,.$$

Let $\vartheta_c : \Re^m \times Y \mapsto \Re$ be defined as

$$\vartheta_c(\zeta,\xi) \equiv \min_{x\in B_\varepsilon(\overline{x})} L_c(x,\zeta,\xi), \quad (\zeta,\xi) \in \Re^m \times Y\,.$$

Let $\overline{y} \equiv (\overline{\zeta},\overline{\xi})$ and $\triangle y \equiv (\triangle\zeta,\triangle\xi) \in \Re^m \times Y$. Define the set $\overline{\mathcal{V}}_c(\Delta y)$ as

$$\left\{ -\begin{pmatrix} \nabla h(\overline{x})^T \\ -W(\nabla g(\overline{x})^*)\end{pmatrix} \mathcal{A}_c(\overline{y},W)^{-1} \begin{pmatrix} \nabla h(\overline{x})^T \\ -W(\nabla g(\overline{x})^*)\end{pmatrix}^T (\triangle y) \right.$$
$$\left. + \begin{pmatrix} 0 \\ -c^{-1}\triangle\xi + c^{-1}W(\triangle\xi)\end{pmatrix} \mid \forall \triangle y \in \Re^m \times Y,\ W \in \partial_B\Pi_K(\overline{\xi} - cg(\overline{x}))\right\}.$$

Then one can show that

$$\partial_B(\nabla\vartheta_c)(\overline{y})(\Delta y) \subseteq \overline{\mathcal{V}}_c(\Delta y)\,.$$

Assumption B2.*There exist positive numbers $\overline{c} \ge c_0$, $\mu_0 > 0$, $\varrho_0 > 0$, and $\tau > 1$ such that for any $c \ge \overline{c}$ and $\Delta y \in \Re^m \times Y$,*

$$\|(x_c)'(\overline{y};\Delta y))\| \le \varrho_0\|\Delta y\|/c$$

and

$$\left\langle V(\Delta y) + c^{-1}\Delta y,\ \Delta y\right\rangle \in \mu_0\,[-1,1]\,\|\Delta y\|^2/c^\tau \quad \forall\, V(\Delta y) \in \overline{\mathcal{V}}_c(\Delta y)\,.$$

The main result on the rate of convergence of the augmented Lagrangian method for the constrained optimization problem (1) is as follows.

Theorem 3.1. [52, Theorem 4.1] *Suppose that K is a self-dual cone and that $\Pi_K(\cdot)$ is semismooth everywhere. Let Assumptions B1 and B2 be satisfied. Let c_0, $\underline{\eta}$, $\overline{c}$, μ_0, ϱ_0, and τ be the positive numbers defined in these assumptions. Define*

$$\varrho_1 \equiv 2\varrho_0 \quad \text{and} \quad \varrho_2 \equiv 4\mu_0.$$

Then for any $c \geq \overline{c}$, there exist two positive numbers ε and δ (both depending on c) such that for any $(\zeta, \xi) \in B_\delta(\overline{\zeta}, \overline{\xi})$, the problem

$$\min \; L_c(x, \zeta, \xi) \quad \text{s.t. } x \in B_\varepsilon(\overline{x})$$

has a unique solution denoted $x_c(\zeta, \xi)$. The function $x_c(\cdot, \cdot)$ is locally Lipschitz continuous on $B_\delta(\overline{\zeta}, \overline{\xi})$ and is semismooth at any point in $B_\delta(\overline{\zeta}, \overline{\xi})$, and for any $(\zeta, \xi) \in B_\delta(\overline{\zeta}, \overline{\xi})$, we have

$$\|x_c(\zeta, \xi) - \overline{x}\| \leq \varrho_1 \|(\zeta, \xi) - (\overline{\zeta}, \overline{\xi})\|/c$$

and

$$\|(\zeta_c(\zeta, \xi), \xi_c(\zeta, \xi)) - (\overline{\zeta}, \overline{\xi})\| \leq \varrho_2 \|(\zeta, \xi) - (\overline{\zeta}, \overline{\xi})\|/c^{\tau-1},$$

where

$$\zeta_c(\zeta, \xi) \equiv \zeta + ch(x_c(\zeta, \xi)) \quad \text{and} \quad \xi_c(\zeta, \xi) \equiv \Pi_K(\xi - cg(x_c(\zeta, \xi))).$$

Under Assumptions B1 and B2, Theorem 3.1 shows that if for all k sufficiently large with $c_k \equiv c$ larger than a threshold and if (x^k, ζ^k, ξ^k) is sufficiently close to $(\overline{x}, \overline{\zeta}, \overline{\xi})$, then the augmented Lagrangian method can be regarded locally as the gradient ascent method applied to the dual problem

$$\max \; \vartheta_c(\zeta, \xi) \quad \text{s.t. } (\zeta, \xi) \in \Re^m \times Y$$

with a constant step-length c, i.e., for all k sufficiently large

$$\begin{pmatrix} \zeta^{k+1} \\ \xi^{k+1} \end{pmatrix} = \begin{pmatrix} \zeta^k \\ \xi^k \end{pmatrix} + c\nabla\vartheta_c(\zeta^k, \xi^k).$$

We make an intuitive explanation on why this method performs "near superlinearly" in practice. By Assumption B2, $-c^{-1}I$ is a good estimate to elements in $\partial\nabla\vartheta_c(\zeta^k, \xi^k)$ for all (ζ^k, ξ^k) sufficiently close to $(\overline{\zeta}, \overline{\xi})$ as every element in $\partial\nabla\vartheta_c(\overline{\zeta}, \overline{\xi})$ is in the form of $-c^{-1}I + O(c^{-\tau})$, where I is the identity operator in $\Re^m \times Y$. Hence the gradient direction is "close" to the Newton direction. Since $\nabla\vartheta_c(\cdot, \cdot)$ is semismooth at $(\overline{\zeta}, \overline{\xi})$, the fast local convergence of the augmented Lagrangian method comes no surprise for those who are familiar with the theory of superlinear convergence of the generalized Newton method for semismooth equations.

The near-superlinear rate of convergence has made the augmented Lagrangian method very attractive, not only in theory, but also in practice. A recent paper of Zhao, Sun, and Toh [64] addresses the semidefinite programming problem. They tested hundreds of problems from a standard data base. Their conclusion is "Numerical experiment on a variety of large scale SDPs with matrix dimensions up to 1,600 and number of constraints up to 1,283,258 shows that the proposed method is very efficient". We note that SDP is a linear model while our proposed research is on nonlinear models. However, as reported in some recent works [16, 40, 59], the performance of the augmented Lagrange method (combined with a semismooth Newton procedure) is impressive in solving some important classes of NSDP problems as well.

For numerical results, we refer the reader to [16, 40, 59, 64]. In particular, hundreds of SDPs from various applications are tested and compared with other methods of SDP in [64].

4. A second-order method: The squared smoothing Newton method.

4.1. Semismooth equations and Newton's method.

Newton's Method

$$x^{k+1} = x^k - [F'(x^k)]^{-1}F(x^k)$$

has been used for solving $F(x) = 0$ if F is continuously differentiable (smooth).

If F is only Lipschitzian, say $F(x) = \max\{x^2, x\}$, then we may use generalized Jacobian $V_k \in \partial F(x^k)$ and

$$x^{k+1} = x^k - V_k^{-1}F(x^k) \tag{7}$$

to solve $F(x) = 0$. See [41] for details.

For a locally Lipschitz $F : \Re^n \to \Re^n$, the *B-subdifferential* at x is the set $\partial_B F(x)$ of all $n \times n$ matrices

$$V = \lim_{x^k \to x} JF(x^k),$$

where $\{x^k\}$ is a sequence of F(Fréchet)-differentiable points of F converging to x. The *C(Clarke)-subdifferential* [6] is defined as

$$\partial F(x) \equiv \text{conv } \partial_B F(x).$$

Does the nonsmooth Newton method (7) converge and how fast does it converge? Simple computation shows that

$$\begin{aligned}
\|x^{k+1} - x^*\| &= \|x^k - V_k^{-1}F(x^k) - x^*\| \\
&= \|V_k^{-1}[F(x^k) - F(x^*) - V_k(x^k - x^*)]\| \\
&= O(\|F(x^k) - F(x^*) - V_k(x^k - x^*)\|) \\
&\overset{?}{=} o(\|x^k - x^*\|).
\end{aligned}$$

Thus, the key property of F that guarantees superlinear convergence of method (7) is

$$F(x^k) - F(x^*) - V_k(x^k - x^*) = o(\|x^k - x^*\|) \ \forall \ V_k \in \partial F(x^k), x^k \to x^*,$$

which we shall use as the definition of *semismoothness* of F at x^*. If we replace $o(\|x^k - x^*\|)$ by $O(\|x^k - x^*\|^2)$, then we have *strong semismoothness*. In applying semismooth Newton methods (7), strong semismoothness leads to quadratic convergence.

Good news: Many functions are semismooth such as

- Convex and concave functions, smooth functions;
- "Piecewise smooth" functions, popular NCP-functions;
- Eigenvalues as functions of symmetric matrices;
- Spectral matrix functions defined by eigenvalues and a semismooth scalar function f; i.e.

$$F(x) = p\,\text{diag}\,(f(\lambda_1), ..., f(\lambda_n))p^T, \quad \text{where } x = p\,\text{diag}\,(\lambda_1, ..., \lambda_n)p^T;$$

- Combination and composition of the above.

4.2. **Semismooth NCP equations and complementarity problems.** Consider the nonlinear matrix complementarity problem (NMCP)

$$x \succeq 0,\ F(x) \succeq 0,\ x^T F(x) = 0. \quad (F : \mathcal{S}^n \to \mathcal{S}^n \text{ smooth})$$

First transform it to a semismooth equation via NCP-functions, e.g. $\min\{x, F(x)\} = 0$ or, equivalently

$$x \text{ solves NMCP} \Leftrightarrow x - [x - F(x)]_+ = 0$$

where $[\cdot]_+$ is the Euclidean projector onto $\mathcal{S}^n_+$.

Another widely used NCP-function is the Fisher-Burmeister function

$$x + y - (x^2 + y^2)^{1/2}, \text{ where } y = F(x).$$

It has been shown [49, 53] that both functions and their smoothing counterparts

$$G(\varepsilon, x) = \begin{cases} x + F(x) - \sqrt{x^2 + F(x)^2 + \varepsilon^2 I} \\ x + F(x) - \sqrt{(x - F(x))^2 + \varepsilon^2 I} \end{cases}$$

are strongly semismooth in (ε, x).

4.3. **A squared smoothing Newton method for NMCP.** For $\varepsilon \in \Re$ and $x \in \mathcal{S}^n$, let

$$G(\varepsilon, x) := x + F(x) - \sqrt{(x - F(x))^2 + \varepsilon^2 I}\,.$$

The *squared smoothing Newton method*, in particular, solves the auxiliary equation

$$E(\varepsilon, x) \equiv \begin{bmatrix} \varepsilon \\ G(\varepsilon, x) \end{bmatrix} = 0$$

and uses the merit function $\phi(z) \equiv \varepsilon^2 + \|G(z)\|^2$ for the line search, where $z \equiv (\varepsilon, x)$.

Let $\bar{\varepsilon} \in \Re_{++}$ and $\eta \in (0, 1)$ be such that $\eta\bar{\varepsilon} < 1$. Define an auxiliary point $\bar{z}$ by

$$\bar{z} := (\bar{\varepsilon}, 0) \in \Re \times \mathcal{S}^n$$

and $\theta : \Re \times \mathcal{S}^n \mapsto \Re_+$ by

$$\theta(z) := \eta \min\{1, \phi(z)\}\,.$$

Let

$$\mathcal{N} := \{\, z = (\varepsilon, x) \in \Re \times \mathcal{S}^n \,:\, \varepsilon \geq \theta(z)\bar{\varepsilon}\,\}\,.$$

Algorithm 4.1.

Step 0. Select constants $\delta \in (0, 1)$ and $\sigma \in (0, 1/2)$. Let $\varepsilon^0 = \bar{\varepsilon}$, $x^0 \in \mathcal{S}^n$ be an arbitrary point and $k := 0$.

Step 1. If $E(z^k) = 0$, then stop. Otherwise, let $\theta_k := \theta(z^k)$.

Step 2. Compute $\Delta z^k := (\Delta\varepsilon^k, \Delta x^k) \in \Re \times \mathcal{S}^n$ by

$$E(z^k) + \nabla E(z^k)(\Delta z^k) = \theta_k \bar{z}\,.$$

Step 3. Let l_k be the smallest nonnegative integer l satisfying

$$\phi(z^k + \delta^l \Delta z^k) \leq [\,1 - 2\sigma(1 - \eta\bar{\varepsilon})\delta^l\,]\phi(z^k)\,.$$

Define $z^{k+1} := z^k + \delta^{l_k}\Delta z^k$.

Step 4. Replace k by $k + 1$ and go to Step 1.

		Smoothing		SDPT3	
m	N	iter	cpu(s)	iter	cpu(s)
50	100	8.9	0.312	10.1	0.206
100	200	9.0	1.094	10.0	0.502
150	300	9.2	2.958	11.0	1.485
200	400	9.0	6.409	10.0	2.853
250	500	8.9	13.070	11.0	5.665
300	600	9.1	19.260	10.0	8.131
350	700	8.9	33.107	11.0	13.823
400	800	8.8	44.847	11.0	39.879

TABLE 2. Random SOCP with random initial points. (Source: Table 6.5 of [4])

Theorem 4.2. [53, Theorem 6.3] *Assume that*

(i) for every $k \geq 0$, *if* $\varepsilon^k \in \Re_{++}$ *and* $z^k \in \mathcal{N}$, $\nabla E(z^k)$ *is nonsingular; and*

(ii) for any accumulation point $z^* = (\varepsilon^*, x^*)$ *of* $\{z^k\}$ *if* $\varepsilon^* > 0$ *and* $z^* \in \mathcal{N}$, *then* $\nabla E(z^*)$ *is nonsingular.*

Then an infinite sequence $\{z^k\} \subset \mathcal{N}$ *is generated by Algorithm 4.1 and each accumulation point* z^* *of* $\{z^k\}$ *is a solution of* $E(z) = 0$. *Moreover, if* E *is strongly semismooth at* z^* *and if all* $V \in \partial_B E(z^*)$ *are nonsingular, then the whole sequence* $\{z^k\}$ *converges to* z^*,

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

and

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

Chen and Tseng [5] tested a number of mid-size SDPs by using a smoothing Newton method. The results show that their method is about 2-4 times slower in CPU time than the interior point method while the iteration counts are similar.

We tested a smoothing Newton method for second-order cone programs (SOCP) [4]. The test problems are randomly generated with sizes N from 100 to 800. The random problems of each size are generated 10 times, and thus we have totally 80 random problems. Table 2 contains the result of the smoothing method compared with SDPT3, a well-known software package for cone optimization problems. The iteration number and CPU time in the table are averages in solving the 10 random problems.

5. **Conclusions.** This paper surveys some algorithms and numerical results in solving nonlinear semidefinite optimization problems. While the alternative direction method requires the problem to be convex, all other methods are aimed at general nonlinear semidefinite programs. We introduced two first-order approaches (the alternating direction method and the augmented Lagrangian method), in which the major computation in an iteration involves only projections onto the semidefinite cone. We showed that the second-order method (the squared smoothing Newton method) has higher convergence rate, but needs to solve a Newton equation per iteration.

We note that there is a similarity between the nonlinear semidefinite optimization algorithms and the nonlinear programming algorithms. However, analysis of the former algorithms often involves special algebraic structure of the matrix space and the "real dimension" of the nonlinear semidefinite optimization problem is often quite large, e.g., a problem in $\mathcal{S}^n$ with $n = 200$ has $n(n+1)/2 = 20,100$ independent variables. Overall, much work is yet to be done, and, due to its importance in applications, the area of methodology for nonlinear semidefinite optimization problems is expected to grow and to attract intensive research in the foreseeable future.

Parallel to the work on NSDP, there have been lots of work on nonlinear second-order cone programming (NSOCP) problems since 1990s (e.g., [22]). It is not surprising that the algorithms for NSDP, NSOCP, and NP have many common characters. In fact, the n dimensional Lorenz space, where the second-order cone is defined, the space $\mathcal{S}^n$, and the space $\Re^n$ are three special cases of the so-called Euclid-Jordan space, in which a special binary operation can be defined and used in the analysis of algorithms. The interested reader may refer to the papers [15] and [31, 51] for some details.

Acknowledgments. This research is partially supported by a Provost's Chair Professorship at National University of Singapore. The author is thankful to Professor Kanzow of University of Würzburg for his invitation to the special issue of *Numerical Algebra Control and Optimization* and his very efficient editorial help.

REFERENCES

[1] A. Ben-Tal, F. Jarre, M. Kocvara, A. Nemirovski and J. Zowe, *Optimal design of trusses under a nonconvex global buckling constraint*, Optim. and Eng., **1** (2000), 189–213.
[2] S. Boyd, L. El Ghaoui, E. Feron and V. Balakrishnan. "Linear Matrix Inequalities in System and Control Theory," SIAM Studies in Applied Mathematics, Philadelphia, 1994.
[3] X. Chen, H. D. Qi and P. Tseng, *Analysis of nonsmooth symmetric-matrix functions with applications to semidefinite complementarity problems*, SIAM J. Optim., **13** (2003), 960–985.
[4] X. Chen, D. Sun and J. Sun, *Complementarity functions and numerical experiments on some smoothing Newton methods for second-order-cone complementarity problems,* Comput. Optim. Appl., **25** (2003), 39–56.
[5] X. Chen and P. Tseng, *Non-interior continuation methods for solving semidefinite complementarity problems*, Math. Program., **95** (2003), 431–473.
[6] F. H. Clarke, "Optimization and Nonsmooth Analysis," John Wiley and Sons, New York, 1983.
[7] R. Correa and C. H. Ramirez, *A global algorithm for nonlinear semidefinite programming*, SIAM J. Optim., **15** (2004), 303–318.
[8] M. Diehl, F. Jarre and C. H. Vogelbusch, *Loss of superlinear convergence for an SQP-type method with conic constraints*, SIAM J. Optim., **16** (2006), 1201–1210.
[9] M. Doljansky, *An interior proximal algorithm and the exponential multiplier method for semidefinite programming*, SIAM J. Optim., **9** (1999), 1–13.
[10] A. Forsgren, *Optimality conditions for nonconvex semidefinite programming*, Math. Program., **88** (2000), 105–128.
[11] J. Eckstein, "Splitting methods for monotone operators with applications to parallel optimization," PhD thesis, Massachusetts Institute of Technology, 1989.
[12] B. Fares, D. Noll and P. Apkarian, *Robust control via sequential semidefinite programming*, SIAM J. Contr. and Optim., **40** (2002), 1791–1820.
[13] M. L. Flegel and C. Kanzow, *Equivalence of two nondegeneracy conditions for semidefinite programs*, J. Optim. Theory Appl., **135** (2007), 381–397.
[14] M. Fukushima, *Application of the alternating directions method of multipliers to separable convex programming problems*, Comput. Optim. Appl., **1** (1992), 93–111.
[15] M. Fukushima, Z. Q. Luo and P. Tseng, *Smoothing functions for second-order cone complementarity problems*, SIAM J. Optim., **12** (2002), 436–460.

[16] Y. Gao and D. Sun, *Calibrating least squares covariance matrix problems with equality and inequality constraints*, SIAM J. Matrix Anal. Appl. **31** (2009), 1432–1457.
[17] B. S. He, L. Z. Liao, D. R. Han and H. Yang, *A new inexact alternating directions method for monotone variational inequalities*, Math. Program., **92** (2002), 103–118.
[18] B. S. He, L. Z. Liao and M. J. Qian, *Alternating projection based prediction-correction methods for structured variational inequalities*, J. Compu. Math., **24** (2002), 693–710.
[19] M. R. Hestenes, *Multiplier and gradient methods*, J. Optim. Theory Appl., **4** (1969), 303–320.
[20] N. J. Higham, *Computing a nearest symmetric positive semidefinite matrix*, Linear Algebra Appl., **103** (1998), 103–118.
[21] F. Jarre, *An interior method for nonconvex semidefinite programs*, Optim. and Eng., **1** (2000), 347–372.
[22] C. Kanzow, I. Ferenczi and M. Fukushima, *On the local convergence of semismooth newton methods for linear and nonlinear second-order cone programs without strict complementarity*, SIAM J. Optim., **20** (2009), 297–320.
[23] C. Kanzow and C. Nagel, *Semidefinite programs: New search directions, smoothing-type methods, and numerical results*, SIAM J. Optim., **13** (2002), 1–23.
[24] C. Kanzow and C. Nagel, *Some structural properties of a Newton-type method for semidefinite programs*, J. Optim. Theory Appl., **122** (2004), 219–226.
[25] C. Kanzow and C. Nagel, *Quadratic convergence of a nonsmooth newton-type method for semidefinite programs without strict complementarity*, SIAM J. Optim., **15** (2005), 654–672.
[26] C. Kanzow, C. Nagel, H. Kato and M. Fukushima, *Successive linearization methods for nonlinear semidefinite programs*, Comput. Optim. Appl., **31** (2005), 251–273.
[27] D. Kinderlehrer and G. Stampacchia, "An Introduction to Variational Inequalities and Their Applications," Academic Press, New York, 1980.
[28] M. Kocvara and M. Stingl, *PENNON - A Generalized augmented Lagrangian method for semidefinite programming,* In "High Performance Algorithms and Software for Nonlinear Optimization"(eds. G. Di Pillo and A. Murli), Kluwer Academic Publishers, Dordrecht, (2003), 297–315.
[29] M. Kocvara and M. Stingl, *PENNON: a code for convex nonlinear and semidefinite programming*, Optim. Meth. Soft., **18** (2003) 317–333.
[30] M. Kocvara and M. Stingl, *Solving nonconvex SDP problems of structural optimization with stability control*, Optim. Meth. Soft., **19** (2004), 595–609.
[31] L. Kong, J. Sun and N. Xiu, *A regularized smoothing Newton method for symmetric cone complementarity problems*, SIAM J. Optim., **19** (2008), 1028–1047.
[32] F. Leibfritz, *COMPl$_e$ib 1.1: Constraint matrix-optimization problem library - a collection of test examples for nonlinear semidefinite programs, control system design and related problems*, Technical Report, Department of Mathematics, University of Trier, Germany, 2005.
[33] F. Leibfritz and M. E. Mostafa, *An interior point constrained trust region method for a special class of nonlinear semidefinite programming problems*, SIAM J. Optim., **12** (2002), 1048–1074.
[34] C. Li and W. Sun, *On filter-successive linearization methods for nonlinear semidefinite programming*, China Sci. Ser. A, **52** (2009), 2341-2361 .
[35] D. Noll and P. Apkarian, *Spectral bundle methods for non-convex maximum eigenvalue functions: first-order methods*, Math. Program., **104** (2005), 701–727.
[36] D. Noll and P. Apkarian, *Spectral bundle methods for non-convex maximum eigenvalue functions: second-order methods*, Math. Program., **104** (2005), 729–747.
[37] J. S. Pang, D. Sun and J. Sun, *Semismooth homeomorphisms and strong stability of semidefinite and Lorentz complementarity problems*, Math. Oper. Res., **28** (2003), 39–63.
[38] T. Pennanen, *Local convergence of the proximal point algorithm and multiplier methods without monotonicity*, Math. Oper. Res., **27** (2002), 170–191.
[39] M. J. D. Powell, *A method for nonlinear constraints in minimization problems*, In "Optimization"(eds. R. Fletcher), Academic Press, New York, (1972), 283–298.
[40] H. Qi and D. Sun, *An augmented Lagrangian dual approach for the H-weighted nearest correlation matrix problem*, IMA J. Numer. Anal. (2011), to appear.
[41] L. Qi and J. Sun, *A nonsmooth version of Newton's method*, Math. Program., **58** (1993), 353–367.
[42] R. T. Rockafellar, *Augmented Lagrange multiplier functions and duality in nonconvex programming*, SIAM J. Control, 12 (1974) 268-285.

[43] R. T. Rockafellar, *Monotone operators and the proximal point algorithm*, SIAM J. Control Optim., **14** (1976), 877–898.
[44] R. T. Rockafellar, *Augmented Lagrangians and applications of the proximal point algorithm in convex programming*, Math. Oper. Res., **1** (1976), 97–116.
[45] R. T. Rockafellar, *Lagrange multipliers and optimality*, SIAM Review, **35** (1993), 183–238.
[46] A. Shapiro, *First and second order analysis of nonlinear semidefinite programs*, Math. Program., **77** (1997), 301–320.
[47] A. Shapiro and J. Sun, *Some properties of the augmented Lagrangian in cone constrained optimization*, Math. Oper. Res., **29** (2004), 479–491.
[48] 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.
[49] D. Sun and J. Sun, *Semismooth matrix valued functions*, Math. Oper. Res., **27** (2002), 150–169.
[50] D. Sun and J. Sun, *Strong semismoothness of the Fischer-Burmeister SDC and SOC complementarity functions*, Math. Program., **103** (2005), 575–582.
[51] D. Sun and J. Sun, *Löwner's operator and spectral functions in Euclidean Jordan algebras*, Math. Oper. Res., **33** (2008), 421–445.
[52] D. Sun, J. Sun and L. Zhang, *Rates of convergence of the augmented Lagrangian method for nonlinear semidefinite programming*, Math. Program., **114** (2008), 349–391.
[53] J. Sun, D. Sun and L. Qi, *A squared smoothing Newton method for nonsmooth matrix equations and its applications in semidefinite optimization problems*, SIAM J. Optim., **14** (2004), 783–806.
[54] J. Sun, L. Zhang and Y. Wu, *Properties of the augmented Lagrangian in nonlinear semidefinite optimization*, J. Optim. Theory Appl., **129** (2006), 437–456.
[55] J. Sun and S. Zhang, *A modified alternating direction method for convex quadratically constrained quadratic semidefinite programs*, Eur. J. Oper. Res., **207** (2010), 1210–1220.
[56] J. Sun, G. Zhao and J. Zhu, *A predictor-corrector algorithm for a class of nonlinear saddle point problems*, SIAM J. Contr. Optim., **35** (1997), 532–551.
[57] N. K. Tsing, M. K. H. Fan and E. I. Verriest, *On analyticity of functions involving eigenvalues*, Linear Algebra Appl., **207** (1994), 159–180.
[58] P. Tseng, *Merit functions for semidefinite complementarity problems*, Math. Program., **83** (1998), 159–185.
[59] C. Wang, D. Sun and K. C. Toh, *Solving log-determinant optimization problems by a Newton-CG proximal point algorithm*, SIAM J. Optim., **20** (2010), 2994–3013.
[60] Z. Wen, D. Goldfarb and W. Yin, *Alternating direction augmented Lagrangian methods for semidefinite programming,* Optimization Online, 2009.
[61] H. Yamashita, H. Yabe and K. Harada, *A primal-dual interior point method for nonlinear semidefinite programming*, Technical report, Tokyo University of Science, 2007.
[62] Z. S. Yu, *Solving semidefinite programming problems via alternating direction methods*, J. Compu. Appl. Math., **193** (2006), 437–445.
[63] S. Zhang, J. Ang and J. Sun, *An alternating direction method for solving convex nonlinear semidefinite programming problems*, Technical Report, School of Business, National University of Singapore, 2010.
[64] X. Zhao, D. Sun and K. Toh, *A Newton-CG augmented Lagrangian method for semidefinite programming*, SIAM J. Optim., **20** (2010), 1737–1765.

Received Nov. 9, 2010; revised Nov. 25, 2010.

E-mail address: jsun@nus.edu.sg