Bull. Aust. Math. Soc. **79** (2009), 367–376
doi:10.1017/S0004972708001214

PREDICTOR–CORRECTOR SMOOTHING NEWTON METHOD FOR SOLVING SEMIDEFINITE PROGRAMMING

CAIYING WU ✉ and GUOQING CHEN

(Received 21 April 2008)

Abstract

There has been much interest recently in smoothing methods for solving semidefinite programming (SDP). In this paper, based on the equivalent transformation for the optimality conditions of SDP, we present a predictor–corrector smoothing Newton algorithm for SDP. Issues such as the existence of Newton directions, boundedness of iterates, global convergence, and local superlinear convergence of our algorithm are studied under suitable assumptions.

2000 *Mathematics subject classification*: primary 90C22.

Keywords and phrases: semidefinite complementarity, semidefinite programming, Fischer–Burmeister function, superlinear convergence.

1. Introduction

In this paper, we consider the following semidefinite programming (SDP) problem

$$\begin{aligned} &\min C \bullet X \\ (P) \quad &\text{such that } A_i \bullet X \geq b_i, \quad i = 1, 2, \ldots, m, \\ &X \succeq 0, \end{aligned} \tag{1}$$

and its dual

$$\begin{aligned} &\max b^T y \\ (D) \quad &\text{such that } \sum_{i=1}^{m} y_i A_i \preceq C, \\ &y \geq 0, \end{aligned} \tag{2}$$

where $C, A_i \in S^n$, $X \in S^n_+$, $y \in \mathbb{R}^m_+$, $b \in \mathbb{R}^m$. Here, we use S^n to denote the set of all $n \times n$ symmetric matrices and S^n_+ the set of all $n \times n$ symmetric positive semidefinite

Project supported by the Teaching and Research Award Program for the Outstanding Young Teachers in Higher Education Institutes of Ministry of Education, People's Republic of China.

matrices. By $A \bullet B = \langle A, B \rangle$ we denote the trace of $A^T B$. Here $\mathbb{R}^m_+$ denotes the set $\{y \in \mathbb{R}^m \mid y_i \geq 0, i = 1, 2, \ldots, m\}$. If $A_i \bullet X = b_i$, $i = 1, 2, \ldots, m$ and $y \in \mathbb{R}^m$, the (P) and (D) are the standard SDP. The standard SDP arise in a wide variety of applications from control theory to combinatorial optimization as well as structural computational complexity theory [15, 16].

There are many varied algorithms for solving standard SDP, and it is convenient to divide the methods into three groups according to their methodology. The first group is the primal-dual path-following interior point methods which have polynomial complexity (see, for example, [9, 7, 1, 4, 17, 8]). In each iteration of a primal-dual interior point algorithm, most of the computational work is devoted to the computation of a search direction by solving a linear system of equations exactly. The second group is similar to the first, but instead of solving for the Newton direction exactly at each iteration, inexact directions are used; see [2, 12, 13] for details. These two groups are interior point methods which first have to symmetrize the central path conditions of the standard SDP in order to guarantee that they obtain symmetric search directions. The final group is the smoothing Newton method [6]. The main idea of this method is that they reformulate the optimality conditions or central path conditions as a nonlinear equation. This reformulation system does not contain any explicitly inequality constraints such as $X \succeq 0$, $Z \succeq 0$ or $X \succ 0$, $Z \succ 0$, and Newton methods applied to this system automatically generate a symmetric search direction without any further transformations (unlike interior point methods).

In this paper, we first reformulate the optimality conditions for (P) and (D) as a semidefinite complementarity problem (SDCP) and then, by using Fischer–Burmeister function, we obtain a mapping H from $S^{n+m}_B \times S^{n+m}_B \times \mathbb{R}_+$ onto itself. Obviously, (P) and (D) are approximated by the smoothing equation $H(T) = 0$. Based on the algorithm in [6] for the standard SDP, we present a predictor–corrector smoothing Newton method. In our method, we solve the smoothing equation at each iteration and refine the approximation by reducing the smoothing parameter μ to zero. Moreover, we investigate the boundedness of the neighborhood $\mathcal{N}(\beta)$. Under suitable assumptions, we can establish both global convergence and local superlinear convergence of our algorithm. On the other hand, the method discussed here generates symmetric directions without any further transformations.

Throughout this paper, we use the following notation. We denote by S^n_{++} the set of symmetric positive definite matrices of dimension $n \times n$. We write $A \succeq 0$ and $A \succ 0$ to indicate that $A \in S^n_+$ and $A \in S^n_{++}$, respectively. If $A \succeq 0$, we denote by $A^{1/2}$ the unique positive semidefinite square root of A. For any $r_1, \ldots, r_n \in \mathbb{R}$, we denote by $\mathrm{diag}[r_i]$ the $n \times n$ diagonal matrix with diagonal entries $r_1, \ldots, r_n$. We denote for any $X \in \mathbb{R}^{n \times n}$ with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, $\min[X] = \min_i \lambda_i$. Here R_+, R_{++} denote the nonnegative and positive reals. The related symbol '$\circ$' is used for the composition of two mappings. In our analysis, $\|\cdot\|$ denotes the 2-norm for a vector and the Frobenius norm for a matrix. Define

$$S^{n+m}_B = \left\{ A \in S^{n+m} \;\middle|\; \overline{A} \in S^n,\, a \in \mathbb{R}^m,\, A = \begin{bmatrix} \overline{A} & 0 \\ 0 & \mathrm{diag}[a_i] \end{bmatrix} \right\}.$$

https://doi.org/10.1017/S0004972708001214 Published online by Cambridge University Press

We endow the vector space $\mathbb{R}^{n\times n} \times \mathbb{R}^m \times \mathbb{R}^{n\times n}$ with the norm

$$|||(X, y, S)||| = \sqrt{\|X\|^2 + \|y\|^2 + \|S\|^2}.$$

We use the same symbol for the norm

$$|||(X, y, S, \mu)||| = \sqrt{\|X\|^2 + \|y\|^2 + \|S\|^2 + \mu^2}$$

in the vector space $\mathbb{R}^{n\times n} \times \mathbb{R}^m \times \mathbb{R}^{n\times n} \times \mathbb{R}$.

2. Reformulation of the optimality conditions

We assume throughout this paper that there is a strictly feasible pair (X, y) for (P) and (D), that is, $A_i \bullet X > b_i$, $i = 1, 2, \ldots, m$, $X \succ 0$, $\sum_{i=1}^m y_i A_i \prec C$, $y > 0$. This assumption ensures that at least one primal-dual optimal pair exists. Under the strictly feasible assumption, the optimality conditions for the pair (P) and (D) problems are

$$\begin{gathered} A_i \bullet X \geq b_i, \quad i = 1, 2, \ldots, m, \quad X \succeq 0, \\ \sum_{i=1}^m y_i A_i \preceq C, \quad y \geq 0, \\ \sum_{i=1}^m (A_i \bullet X - b_i) y_i = 0, \quad X\Big(C - \sum_{i=1}^m y_i A_i\Big) = 0. \end{gathered} \tag{3}$$

It is easy to see that (3) is equivalent to

$$\begin{bmatrix} X & 0 \\ 0 & \mathrm{diag}[y_i] \end{bmatrix} \succeq 0, \quad \begin{bmatrix} C - \sum_{i=1}^m y_i A_i & 0 \\ 0 & \mathrm{diag}[A_i \bullet X - b_i] \end{bmatrix} \succeq 0 \tag{4}$$

and

$$\begin{bmatrix} X & 0 \\ 0 & \mathrm{diag}[y_i] \end{bmatrix} \begin{bmatrix} C - \sum_{i=1}^m y_i A_i & 0 \\ 0 & \mathrm{diag}[A_i \bullet X - b_i] \end{bmatrix} = 0. \tag{5}$$

Obviously, (4) and (5) are essentially a SDCP. A number of algorithms for solving SDCP are based on reformulating SDCP into a system of nonlinear equations. A very popular way to reformulate a SDCP as a nonlinear system consists of choosing a SDCP function, that is, a function $\varphi : S^n \times S^n \to S^n$ such that $\varphi(X, S) = 0$ if and only if $X \succeq 0$, $S \succeq 0$, $XS = 0$. One commonly used SDCP function is the Fischer–Burmeister function [14] $\varphi : S^n \times S^n \to S^n$ defined by

$$\varphi(X, S) = X + S - (X^2 + S^2)^{1/2}.$$

https://doi.org/10.1017/S0004972708001214 Published online by Cambridge University Press

Let $\mu > 0$ be any nonnegative number that will be viewed as an independent variable in this paper. Then define $\phi : S^n \times S^n \times \mathbb{R}_+ \to S^n$ by

$$\phi(X, S, \mu) = X + S - (X^2 + S^2 + 2\mu^2 I)^{1/2}. \tag{6}$$

This is the so-called smoothing Fischer–Burmeister function. It has been shown by Tseng that the mapping ϕ has the following property.

Let $L_R[X] = RX + XR$ denote the corresponding Lyapunov operator, where $R \in S^n_{++}$. Then the positive definiteness of R guarantees that the Lyapunov equation $L_R[X] = H$ has a unique solution within the set of symmetric matrices for every $H \in S^n$ (see [5]). Hence, we can define the inverse L_R^{-1} of L_R, that is, $L_R^{-1}[H]$ denotes the unique element X satisfying $RX + XR = H$.

LEMMA 1. [6] *Let $X, S \in S^n$ be two given matrices and any $\mu, \tau \in \mathbb{R}_{++}$. Then ϕ is continuously differentiable in (X, S, μ) with*

$$\nabla\phi(X, S, \mu)(U, V, \tau) = U + V - L_R^{-1}[XU + UX + SV + VS + 4\tau\mu I],$$

where $R = (X^2 + S^2 + 2\mu^2 I)^{1/2}$.

For any $W \in S_B^{n+m}$, define

$$G(W) = \begin{bmatrix} C - \sum_{i=1}^{m} w_i A_i & 0 \\ 0 & \operatorname{diag}[A_i \bullet \overline{W} - b_i] \end{bmatrix}. \tag{7}$$

DEFINITION 2. [14] Suppose that $F : S^n \to S^n$ is a matrix function. Then F is monotone if

$$\langle F(X) - F(X'), X - X' \rangle \geq 0, \quad \forall X, X' \in S^n.$$

LEMMA 3. *Let $W \in S_B^{n+m}$ and $G(W)$ be defined in (7). Then G is monotone.*

PROOF. By the definition of G, for any $W^1, W^2 \in S_B^{n+m}$,

$$G(W^1) - G(W^2) = \begin{bmatrix} -\sum_{i=1}^{m} (w_i^1 - w_i^2) A_i & 0 \\ 0 & \operatorname{diag}[A_i \bullet (\overline{W^1} - \overline{W^2})] \end{bmatrix}. \tag{8}$$

In addition,

$$W^1 - W^2 = \begin{bmatrix} \overline{W^1} - \overline{W^2} & 0 \\ 0 & \operatorname{diag}[w_i^1 - w_i^2] \end{bmatrix}.$$

Hence, $\langle G(W^1) - G(W^2), W^1 - W^2 \rangle = 0$. The conclusion follows. □

https://doi.org/10.1017/S0004972708001214 Published online by Cambridge University Press

To this end, we define the mapping $H : S_B^{n+m} \times S_B^{n+m} \times \mathbb{R}_+ \to S_B^{n+m} \times S_B^{n+m} \times \mathbb{R}_+$ by

$$H(W, D, \mu) = \begin{pmatrix} \phi(W, D, \mu) \\ G(W) - D \\ \mu \end{pmatrix}. \tag{9}$$

Therefore, the system $H(W, D, \mu) = 0$ is equivalent to the optimality conditions (3) themselves.

We use the notation $Y = (W, D) \in S_B^{n+m} \times S_B^{n+m}$, $T = (Y, \mu)$. As with interior point methods, a convergence analysis requires the iterates to lie in a neighborhood of the 'path' defined by $H(Y, \mu) = 0$. We use the following choice of neighborhood, based on [6]:

$$\mathcal{N}(\beta) = \{T \in S_B^{n+m} \times S_B^{n+m} \times \mathbb{R}_{++} : |||H(T)||| \leq \beta\mu\}, \tag{10}$$

where $\beta > 0$ is a constant.

3. Invertibility of gradient of the mapping H and the boundedness of the neighborhood $\mathcal{N}(\beta)$

In this section, we have the following results showing that Lemma 3 is sufficient for $\nabla H(T)$ to be invertible for all $T \in S_B^{n+m} \times S_B^{n+m} \times \mathbb{R}_{++}$ and for the neighborhood $\mathcal{N}(\beta)$ to be bounded.

LEMMA 4. [14] *For any $A \in S_+^n$, $B \in S^n$, if $A^2 - B^2 \in S_+^n$, then $A - B \in S_+^n$.*

LEMMA 5. *For any $T \in S_B^{n+m} \times S_B^{n+m} \times \mathbb{R}_{++}$, $\nabla H(T)$ is invertible.*

PROOF. Let $T \in S_B^{n+m} \times S_B^{n+m} \times \mathbb{R}_{++}$ be fixed. Since $\nabla H(T)$ is linear mapping from the finite-dimensional vector space $S_B^{n+m} \times S_B^{n+m} \times \mathbb{R}$ onto itself, we only have to verify that this mapping is one-to-one. So, it suffices to show that the system $\nabla H(T)(\Delta T) = (0, 0, 0)$ or, equivalently, the system

$$\nabla\phi(W, D, \mu)(\Delta W, \Delta D, \Delta\mu) = 0, \tag{11}$$

$$M_1 \Delta W - \Delta D = 0, \tag{12}$$

$$\Delta\mu = 0 \tag{13}$$

has $\Delta T = 0$ as its only solution, where $M_1 = \nabla G(W)$. Set $R = (W^2 + D^2 + 2\mu^2 I)^{1/2}$. Then from Lemma 1, (11) and (13), we obtain

$$\Delta W + \Delta D - L_R^{-1}[W\Delta W + \Delta W W + D\Delta D + \Delta D D] = 0.$$

Applying L_R to both sides of the equation and rearranging terms yields

$$L_{R-W}[\Delta W] + L_{R-D}[\Delta D] = 0.$$

https://doi.org/10.1017/S0004972708001214 Published online by Cambridge University Press

Lemma 4 and its proof imply $R - D \succ 0$. Thus, the inverse L_{R-D}^{-1} exists, and we obtain

$$L_{R-D}^{-1} \circ L_{R-W}[\Delta W] + \Delta D = 0. \tag{14}$$

Taking the scalar product with ΔW yields

$$L_{R-D}^{-1} \circ L_{R-W}[\Delta W] \bullet \Delta W + \Delta D \bullet \Delta W = 0. \tag{15}$$

On the other hand, from Lemma 3, $\langle \Delta W, \Delta D \rangle = 0$. Using the fact that $R - D \succ 0$, $R - W \succ 0$, it follows from [6, Lemma 4.2(d)] that the operator $L_{R-D}^{-1} \circ L_{R-W}$ is strongly monotone. Therefore, (15) immediately gives $\Delta W = 0$. Then $\Delta D = 0$ by (14). The proof is complete. □

Similarly to the proof of [3, Lemma 8], we obtain the boundedness of the neighborhood $\mathcal{N}(\beta)$ as follows.

LEMMA 6. *For any* $\beta \in \mathbb{R}_{++}$ *and* $\mu_0 \in \mathbb{R}_{++}$, *if* $2\beta\mu_0 < \min[G(\widetilde{W})]$, $2\beta\mu_0 < \min[\widetilde{W}]$ *for some* $\widetilde{W} \in S_B^{n+m}$ *and* $0 < \mu \le \mu_0$, *then the set* $\mathcal{N}(\beta)$ *is bounded.*

4. The algorithm and its convergence analysis

In the following, we formally describe our method, parameterized by $\beta > 1$ and σ, $\alpha_1, \alpha_2 \in (0, 1)$. This method is a modification of a method studied in [6] for SDP.

STEP 0. Choose $T^0 \in S_B^{n+m} \times S_B^{n+m} \times \mathbb{R}_{++}$ with $D^0 = G(W^0)$. Choose $\beta > 1$ with $|||H(T^0)||| \le \beta\mu_0$, and set $k = 0$. Choose $\sigma, \alpha_1, \alpha_2 \in (0, 1)$ and $\varepsilon \ge 0$.

STEP 1. If $|||H(T^k)||| \le \varepsilon$, stop.

STEP 2 (Predictor step). Let $(\Delta Y^k, \Delta\mu_k) \in S_B^{n+m} \times S_B^{n+m} \times \mathbb{R}$ be a solution of the system

$$\nabla H(T^k) \begin{pmatrix} \Delta Y \\ \Delta\mu \end{pmatrix} = -H(T^k). \tag{16}$$

If $|||H(Y^k + \Delta Y^k, \mu_k)||| > \beta\mu_k$, then let $\widehat{Y^k} = Y^k$, $\widehat{\mu_k} = \mu_k$, $\eta_k = 1$; otherwise let $\eta_k = (\alpha_1)^s$ where s is the nonnegative number with

$$|||H(Y^k + \Delta Y^k, (\alpha_1)^r\mu_k)||| \le \beta(\alpha_1)^r\mu_k, \quad r = 0, 1, \ldots, s,$$
$$|||H(Y^k + \Delta Y^k, (\alpha_1)^{s+1}\mu_k)||| > \beta(\alpha_1)^{s+1}\mu_k,$$

and set

$$\widehat{\mu_k} = \eta_k\mu_k, \quad \widehat{Y^k} = \begin{cases} Y^k & s = 0, \\ Y^k + \Delta Y^k & s \ne 0. \end{cases}$$

STEP 3 (Corrector step). Let $(\widehat{\Delta Y^k}, \widehat{\Delta\mu_k})$ be a solution of

$$\nabla H(\widehat{T^k}) \begin{pmatrix} \widehat{\Delta Y} \\ \widehat{\Delta\mu} \end{pmatrix} = -H(\widehat{T^k}) + \begin{pmatrix} 0 \\ \frac{1}{\beta}|||H(\widehat{T^k})||| \end{pmatrix}. \tag{17}$$

https://doi.org/10.1017/S0004972708001214 Published online by Cambridge University Press

Let $\widehat{\eta_k}$ be the maximum of the numbers $1, \alpha_2, \alpha_2^2, \ldots,$ with

$$|||H(\widehat{Y^k} + \widehat{\eta_k}\widehat{\Delta Y^k}, \widehat{\mu_k} + \widehat{\eta_k}\Delta\widehat{\mu_k})||| \le \left[1 - \sigma\left(1 - \frac{1}{\beta}\right)\widehat{\eta_k}\right]|||H(\widehat{T^k})|||.$$

Set $Y^{k+1} = \widehat{Y^k} + \widehat{\eta_k}\widehat{\Delta Y^k}$, $\mu_{k+1} = \widehat{\mu_k} + \widehat{\eta_k}\Delta\widehat{\mu_k}$, $k = k + 1$ and go to Step 1.

REMARK. (i) It is obvious that all iterations $\{\widehat{T^k}\}$ and $\{T^k\}$ generated by the algorithm satisfy $\widehat{D^k} = G(\widehat{W^k})$, $D^k = G(W^k)$. Moreover, we shall see that the algorithm is well defined and the sequences $\{\widehat{T^k}\}, \{T^k\} \subset S_B^{n+m} \times S_B^{n+m} \times \mathbb{R}_+$. However, in interior point methods, the symmetric search directions are obtained by symmetrizing the central path conditions.

(ii) Since the matrices are not necessarily positive definite or positive semidefinite, the algorithm is a infeasible noninterior point method.

(iii) Similar to the algorithm in [6], the predictor step will be responsible for the local fast convergence of the algorithm, whereas the corrector step will be used in order to prove global convergence. Furthermore, the idea of Step 3 is derived from the one-step continuation methods for the nonlinear complementarity problem [11, 18, 19].

We now start to analyze the properties of the algorithm formally.

LEMMA 7. (i) *The algorithm is well defined.* (ii) *If infinite sequences* $\{T^k\}$ *and* $\{\widehat{T^k}\}$ *are generated by the algorithm, then*

$$T^k, \widehat{T^k} \in \mathcal{N}(\beta), \quad 0 < \mu_{k+1} \le \mu_k \quad \textit{for all } k \ge 0. \tag{18}$$

PROOF. (i) If $\mu_k > 0$, then it follows from Lemma 5 that $\nabla H(T^k)$ and $\nabla H(\widehat{T^k})$ are invertible. Hence, (16) and (17) are well defined at the kth iteration. For any $\alpha \in (0, 1]$, Define

$$r(\alpha) = H(\widehat{T^k} + \alpha\Delta\widehat{T^k}) - H(\widehat{T^k}) - \alpha\nabla H(\widehat{T^k})\Delta\widehat{T^k}. \tag{19}$$

By (17), we obtain for any $\alpha \in (0, 1]$ that

$$\widehat{\mu_k} + \alpha\Delta\widehat{\mu_k} = (1-\alpha)\widehat{\mu_k} + \frac{\alpha}{\beta}|||H(\widehat{T^k})||| > 0.$$

It follows from (17) that

$$\begin{aligned} |||H(\widehat{T^k} + \alpha\Delta\widehat{T^k})||| &= |||H(\widehat{T^k}) + \alpha\nabla H(\widehat{T^k})\Delta\widehat{T^k} + r(\alpha)||| \\ &= \left|\left|\left|(1-\alpha)H(\widehat{T^k}) + \begin{pmatrix} 0 \\ \dfrac{\alpha}{\beta}|||H(\widehat{T^k})||| \end{pmatrix} + r(\alpha)\right|\right|\right| \\ &\le \left[1 - \alpha\left(1 - \frac{1}{\beta}\right)\right]|||H(\widehat{T^k})||| + |||r(\alpha)|||. \end{aligned} \tag{20}$$

https://doi.org/10.1017/S0004972708001214 Published online by Cambridge University Press

Therefore, for any $\mu_k > 0$, we have $\widehat{\mu_k} = \eta_k \mu_k > 0$, which implies that $H(\cdot)$ is continuously differentiable around $(\widehat{Y^k}, \widehat{\mu_k})$. Thus, (19) implies that $|||r(\alpha)||| = o(\alpha)$. We further obtain by (20) that there exists a constant $\overline{\alpha} \in (0, 1]$ such that

$$|||H(\widehat{T^k} + \alpha \Delta \widehat{T^k})||| \le \left[1 - \sigma \left(1 - \frac{1}{\beta}\right) \alpha\right] |||H(\widehat{T^k})||| \tag{21}$$

holds for any $\alpha \in (0, \overline{\alpha}]$. This shows that Step 3 is well defined at the kth iteration. Therefore, from $\mu_0 > 0$ and the above statements, we obtain that the algorithm is well defined and generates two infinite sequences $\{\widehat{T^k}\}$ and $\{T^k\}$ with $\mu_k > 0$ and $\widehat{\mu_k} > 0$ for all $k \ge 0$.

(ii) Next, we prove $T^k, \widehat{T^k} \in \mathcal{N}(\beta)$ for all $k \ge 0$ by induction on k. Obviously, $T^0, \widehat{T^0} \in \mathcal{N}(\beta)$. Suppose that $T^k, \widehat{T^k} \in \mathcal{N}(\beta)$, then, by the line search in Step 3,

$$\begin{aligned} \beta \mu_{k+1} - |||H(T^{k+1})||| &= \beta \left[(1 - \widehat{\eta_k}) \widehat{\mu_k} + \frac{\widehat{\eta_k}}{\beta} |||H(\widehat{T^k})|||\right] - |||H(T^{k+1})||| \\ &\ge (1 - \widehat{\eta_k}) |||H(\widehat{T^k})||| + \widehat{\eta_k} |||H(\widehat{T^k})||| - |||H(T^{k+1})||| \ge 0, \end{aligned} \tag{22}$$

which proves $T^{k+1} \in \mathcal{N}(\beta)$. Consequently, $\widehat{T^{k+1}} \in \mathcal{N}(\beta)$ by Step 2. Finally,

$$\mu_{k+1} \le (1 - \widehat{\eta_k}) \widehat{\mu_k} + \frac{\widehat{\eta_k}}{\beta} \beta \widehat{\mu_k} = \widehat{\mu_k} = \eta_k \mu_k \le \mu_k. \qquad \square$$

LEMMA 8. *If the sequence $\{T^k\}$ generated by the algorithm has an accumulation point, then the sequence $\{\mu_k\}$ converges to zero. In particular, every accumulation point of the sequence $\{T^k\}$ satisfies $H(\cdot) = 0$.*

PROOF. Since the sequence $\{\mu_k\}$ is monotonically decreasing and bounded from below by zero, it converges to a nonnegative number μ_∞. If $\mu_\infty = 0$, then by $|||H(T^k)||| \le \beta \mu_k$, we can show $|||H(T^\infty)||| = 0$. So assume that $\mu_\infty \ne 0$. Then the update rules in Step 2 of the algorithm give

$$\eta_k = 1, \quad \widehat{Y^k} = Y^k, \quad \widehat{\mu_k} = \mu_k \tag{23}$$

for all k sufficiently large; subsequently, if necessary, we assume without loss of generality that (23) holds for all $k \ge 0$. Then we obtain from the line search in Step 3 and Lemma 7 that

$$\mu_{k+1} \le |||H(T^{k+1})||| \le |||H(T^k)||| \le \beta \mu_k \le \beta \mu_0.$$

So $\lim_{k \to \infty} |||H(T^k)||| = |||H(T^\infty)||| \ne 0$. From $|||H(T^\infty)||| > 0$, we have that $\lim_{k \to \infty} \widehat{\eta_k} = 0$. Thus, the step size $\widetilde{\eta} = \widehat{\eta_k} / \alpha_2$ does not satisfy the line search criterion in Step 3 for any sufficiently large k, that is, the following inequality holds:

$$|||H(T^k + \widetilde{\eta} \Delta \widehat{T^k})||| > \left[1 - \sigma \left(1 - \frac{1}{\beta}\right) \widetilde{\eta}\right] |||H(T^k)|||$$

https://doi.org/10.1017/S0004972708001214 Published online by Cambridge University Press

for any sufficiently large k, which implies that

$$\frac{[|||H(T^k + \widetilde{\eta}\widehat{\Delta T^k})||| - |||H(T^k)|||]}{\widetilde{\eta}} > -\sigma\left(1 - \frac{1}{\beta}\right)|||H(T^k)|||.$$

From $\mu_\infty \neq 0$, we know that $H(T)$ is continuously differentiable at T^∞. Letting $k \to \infty$, then the above inequality gives

$$\frac{H(T^\infty)}{|||H(T^\infty)|||} \bullet \left[\nabla H(T^\infty)\begin{pmatrix}\Delta Y^\infty \\ \Delta\mu_\infty\end{pmatrix}\right] \geq -\sigma\left(1 - \frac{1}{\beta}\right)|||H(T^\infty)|||. \tag{24}$$

In addition, by taking the limit on (17), we obtain

$$\nabla H(T^\infty)\begin{pmatrix}\Delta Y^\infty \\ \Delta\mu_\infty\end{pmatrix} = -H(T^\infty) + \begin{pmatrix}0 \\ \frac{1}{\beta}|||H(T^\infty)|||\end{pmatrix}. \tag{25}$$

Combining (24) with (25)

$$-|||H(T^\infty)||| + \frac{|||H(T^\infty)|||}{\beta} \geq -|||H(T^\infty)||| + \frac{\mu_\infty}{\beta} \geq -\sigma\left(1 - \frac{1}{\beta}\right)|||H(T^\infty)|||. \tag{26}$$

This indicates that $-1 + 1/\beta + \sigma(1 - 1/\beta) \geq 0$, which contradicts the fact that $\beta > 1$ and $\sigma \in (0, 1)$. Thus, $\mu_\infty = 0$ and $H(T^\infty) = 0$ by the relation $|||H(T^k)||| \leq \beta\mu_k$. □

THEOREM 9. *Suppose that all of the conditions assumed in Lemma 6 are satisfied. If $\{T^k\}$ is the sequence generated by the algorithm, then its accumulation points satisfy $H(\cdot) = 0$ and $\lim_{k\to\infty} \mu_k = 0$.*

PROOF. This proof is obvious by Lemmas 6 and 8. □

THEOREM 10. *Suppose that all of the conditions assumed in Lemma 6 are satisfied and that T^* is an accumulation point of the sequence $\{T^k\}$ generated by the algorithm. Let the constant β satisfy the inequality $\sqrt{\beta^2 - 1} > \kappa\sqrt{n + m}$, where κ denotes the constant from [6, Lemma 3.1]. If all $V \in \partial H(T^*)$ are invertible, then $\mu_{k+1} = o(\mu_k)$, that is, the smoothing parameter converges locally superlinearly to zero.*

PROOF. We regard the $n \times n$ symmetric matrix space as a special case of $\mathbb{R}^{n(n+1)/2}$. Hence, [10, Proposition 3.1] applies to matrix variables as well, namely, if T^k sufficiently close to T^*, then there exists a constant $\overline{\beta} > 0$ such that

$$|||\nabla H^{-1}(T^k)||| \leq \overline{\beta}. \tag{27}$$

Thus, $|||\Delta T^k||| = ||| - \nabla H^{-1}(T^k)H(T^k)||| \leq \overline{\beta}|||H(T^k)||| \leq \beta\overline{\beta}\mu_k$, that is, $|||\Delta T^k||| = O(\mu_k)$. The following proof is similar to [6, Lemma 5.6]. □

https://doi.org/10.1017/S0004972708001214 Published online by Cambridge University Press

References

[1] F. Alizadeh, J. P. A. Haeberly and M. L. Overton, 'Primal-dual interior-point methods for semidefinite programming: convergence rates, stability and numerical results', *SIAM J. Optim.* **8**(3) (1998), 746–768.

[2] S. Burer, 'Semidefinite programming in the space of partial positive semidefinite matrices', *SIAM J. Optim.* **14**(1) (2003), 139–172.

[3] X. Chen and P. Tseng, 'Noninterior continuation methods for solving semidefinite complementarity problems', *Math. Program., Ser.* A **95** (2003), 431–474.

[4] C. Helmberg, F. Rendl, R. J. Vanderbei and H. Wolkowicz, 'An interior-point method for semidefinite programming', *SIAM J. Optim.* **6**(2) (1996), 342–361.

[5] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, New York, 1985).

[6] C. Kanzow and C. Nagel, 'Semidefinite programs: new search directions, smoothing-type methods, and numerical results', *SIAM J. Optim.* **13**(1) (2002), 1–23.

[7] M. Kojima, S. Shindoh and S. Hara, 'Interior point methods for the monotone semidefinite linear complementarity problem in symmetric matrices', *SIAM J. Optim.* **7** (1997), 86–125.

[8] R. D. C. Monteiro, 'Primal-dual path following algorithms for semidefinite programming', *SIAM J. Optim.* **7** (1997), 663–678.

[9] R. D. C. Monteiro and Y. Zhang, 'A unified analysis for a class of path-following primal-dual interior-point algorithms for semidefinite programming', *Math. Program.* **81** (1998), 281–299.

[10] L. Q. Qi and J. Sun, 'A nonsmooth version of Newton's method', *Math. Program.* **58** (1993), 353–368.

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

[12] K. C. Toh, 'Solving large scale semidefinite programs via an iterative solver on the agumented systems', *SIAM J. Optim.* **14** (2004), 670–698.

[13] K. C. Toh and M. Kojima, 'Solving some large scale semidefinite programs via the conjugate residual method', *SIAM J. Optim.* **12** (2002), 669–691.

[14] P. Tseng, 'Merit functions for semidefinite complementarity problems', *Math. Program.* **83** (1998), 159–185.

[15] H. Wolkowicz, Bibliography on semidefinite programming, http://liinwww.ira.uka.de/bibliography/math/psd.html.

[16] H. Wolkowicz, R. Saigal and L. Vandenberghe, *Handbook of Semidefinite Programming* (Kluwer Academic, Boston, MA, 2000).

[17] Y. Zhang, 'On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming', *SIAM J. Optim.* **8**(2) (1998), 365–386.

[18] L. P. Zhang and Z. Y. Gao, 'Superlinearly/quadratic one-step smoothing Newton method for P_0-NCP without strict complementarity', *Math. Methods Oper. Res.* **56** (2002), 231–241.

[19] L. P. Zhang, J. Y. Han and Z. H. Huang, 'Superlinear/quadratic one-step smoothing Newton method for P_0-NCP', *Acta Math. Sin., Engl. Ser.* **21**(1) (2005), 117–128.

CAIYING WU, College of Mathematics Science, Inner Mongolia University, Hohhot 010021, People's Republic of China
e-mail: wucaiyingun@163.com

GUOQING CHEN, College of Mathematics Science, Inner Mongolia University, Hohhot 010021, People's Republic of China
e-mail: cgq@imu.edu.cn