A NEW PREDICTOR-CORRECTOR ALGORITHM FOR SDP WITH POLYNOMIAL CONVERGENCE[1]

Feixiang Chen

Yuming Feng

School of Mathematics and Statistics
Chongqing Three Gorges University
Wanzhou, Chongqing, 404000
P.R. China
e-mail: cfx2002@126.com

Abstract. We establish the polynomiality of primal-dual interior-point algorithms for SDP based on the direction of the M-Z family of search directions. We show that the polynomial iteration-complexity bounds of the well known algorithms for linear programming, namely, the predictor-corrector algorithm, carry over to the context of SDP.

Keywords: interior-point algorithm; polynomial complexity; path-following methods; semidefinite programming problems.

1. Introduction

Several authors have discussed generalizations of interior-point algorithms for linear programming (LP) to the context of semidefinite programming (SDP). The landmark work in this direction is due to Nesterov and Nemirovskii [1], where a general approach for using interior-point algorithms for solving convex programs is proposed, based on the notion of self-concordant functions. They show that the problem of minimizing a linear function over a convex set can be solved in 'polynomial time' as long as a self-concordant barrier function for the convex set is known. On the other hand, Alizadeh [2] extends Ye's projective potential reduction algorithm for LP to SDP and argues that many known interior point algorithms for LP can also be transformed into algorithms for SDP in a mechanical way. Since then many authors have proposed interior-point algorithms for solving the SDP, including Alizadeh, Haeberly and Overton [3], Kojima, Shida [4]and Shindoh, Kojima, and Hara [5], Monteiro [6], [7], Monteiro and Zhang [8], [9], and Zhang [10]. Most of these more recent works are concentrated on primal-dual algorithms.

[1]Research is supported by Natural Science Foundation of Chongqing Municipal Education Commission (No.KJ091104).

Notation and terminology

The set of all symmetric $n \times n$ matrices is denoted by S^n. For $Q \in S^n$, $Q \succeq 0$ means Q is positive semidefinite and $Q \succ 0$ means Q is positive definite, respectively. The trace of a matrix $Q \in R^{n\times n}$ is denoted by Tr $Q \equiv \sum_{i=1}^{n} Q_{ii}$. The inner product between P and Q in $R^{m\times n}$ is defined as $P \bullet Q \equiv$ Tr $P^T Q$. The Euclidean norm and its associated operator norm are both denoted by $\| \circ \|$; hence, $\|Q\| \equiv \max_{\|u\|=1} \|Qu\|$ for any $Q \in R^{n\times n}$. The Frobenius norm of $Q \in R^{n\times n}$ is $\|Q\|_F \equiv (Q \bullet Q)^{1/2}$. We frequently use the inequalities S^n_+ and S^n_{++} denote the set of all matrices in S^n which are positive semidefinite and positive definite, respectively. It is known that for each $V \in S^n_+$, there exists a unique $U \in S^n_+$, such that $U^2 = V$. The matrix U is called the square root of V and is denoted by $V^{1/2}$.

2. The SDP problem and preliminary discussion

In this section, we describe the SDP in symmetric matrices considered in this paper, state our assumptions, and derive the Newton direction for the central path equation. We also give some existence results for this Newton direction and state a generic path-following algorithm based on it.

Given $C \in S^n$ and $(A_i, b_i) \in S^n \times R$ for $i = 1, ..., m$, a primal-dual pair of SDP problems is defined as

$$(2.1) \qquad (P) \qquad \min\{C \bullet X : A_i \bullet X = b_i, i = 1, ..., m, X \succeq 0\},$$

$$(2.2) \qquad (D) \qquad \max\left\{ b^T y : \sum_{i=1}^{m} y_i A_i + S = C, S \succeq 0 \right\},$$

where $b \equiv (b_1, ..., b_m)^T$.

The set of interior feasible solutions of (1) and (2) are

$$F^0(P) \equiv \{X \in S : X : A_i \bullet X = b_i,\ i = 1, ..., m,\ X \succ 0\}$$
$$F^0(D) \equiv \{(S, y) \in S \times R^m : \textstyle\sum_{i=1}^{m} y_i A_i + S = C,\ S \succ 0\}$$

respectively. Throughout this paper, we assume that $F^0(P) \times F^0(D) \neq \phi$ and that the matrices $A_i, i = 1, \ldots, m$, are linearly independent. Under the first assumption, it is well known that both (1) and (2) have optimal solutions X^* and (S^*, y^*) such that $C \bullet X^* = b^T y^*$, i.e., the optimal values of (1) and (2) coincide. The last condition, called the strong duality, can be alternatively expressed as $X^* \bullet S^* = 0$ or $X^* S^* = 0$. Hence, the set of primal and dual optimal solutions consists of all the solutions $(X, S, y) \in S^n_+ \times S^n_+ \times R^m$ to the following optimality system:

$$(2.3) \qquad XS = 0,$$

$$\sum_{i=1}^{m} y_i A_i + S = C, \tag{2.4}$$

$$A_i \bullet X = b_i, i = 1, \ldots, m. \tag{2.5}$$

It is known that for every $\mu > 0$, the perturbed system

$$XS = \mu I, \tag{2.6}$$

$$\sum_{i=1}^{m} y_i A_i + S = C, \tag{2.7}$$

$$A_i \bullet X = b_i, i = 1, \ldots, m. \tag{2.8}$$

has a unique solution, denoted by (X_μ, S_μ), for every $\mu > 0$, and that the limit $\lim\limits_{\mu \to 0}$ exists and is a solution of (1). The set $\{(X_\mu, S_\mu) : \mu > 0\}$ is called the central path associated with (1) and plays a fundamental role in the development of interior point algorithms for solving SDP. Using the square root $X^{1/2}$, (6) can also be alternatively expressed in the following symmetric form:

$$X^{1/2} S X^{1/2} = \mu I.$$

Path following algorithms for solving (1) are based on the idea of approximately tracing the central path. Application of Newton method for computing the solution of (2) with $\mu = \hat{\mu}$ leads to the Newton search direction $(\widehat{\Delta X}, \widehat{\Delta S})$ which solves the linear system

$$X\widehat{\Delta S} + \widehat{\Delta X} S = \hat{\mu} I - XS, \qquad (X + \widehat{\Delta X}, S + \widehat{\Delta S}) \in S_+^n \times S_+^n. \tag{2.9}$$

This system does not always have a solution. To overcome this bottleneck, if we adapt the M-Z search directions to the monotone SDP, we can describe it as a solution of the system of equations:

$$\begin{aligned} X^{-1/2}(X\Delta S + \Delta X S)X^{1/2} + X^{1/2}(\Delta S X + S \Delta X)X^{-1/2} \\ = 2(\hat{\mu} I - X^{1/2} S X^{1/2}). \end{aligned} \tag{2.10}$$

$$A_i \bullet \Delta X = 0, i = 1, \ldots, m. \tag{2.11}$$

$$\sum_{i=1}^{m} \Delta y_i A_i + \Delta S = 0, \tag{2.12}$$

It was shown in [11] that system (10), (11), (12) has a unique solution. The symmetric component $(\Delta X, \Delta S)$ of this solution is then used as a search direction to generate the next point.

Theorem 2.1 *System* (10), (11), (12) *has a unique solution.*

Lemma 2.1 *Let* $X \in F^0(P)$ *and* $(S, y) \in F^0(D)$ *be given and suppose that* $(\Delta X, \Delta S)$ *is a solution of system* (10), (11), (12) *with* $\hat{\mu} = \sigma\mu$, *then the following statements holds:*

(1) $\Delta S \bullet \Delta X = 0,$

(2) $(X + \alpha\Delta X) \bullet (S + \alpha\Delta S) = (1 - \alpha + \sigma\alpha)(X \bullet S),\ \forall \alpha \in R.$

Proof. Using (11) and (12), we obtain

$$\Delta S \bullet \Delta X = -\left(\sum_{i=1}^{m} \Delta y_i A_i\right) \bullet \Delta X = \sum_{i=1}^{m} \Delta y_i (A_i \Delta X) = 0.$$

and hence (1) follows. In view of (10) , we have

$$\begin{aligned} 2Tr(\sigma\mu I - XS) &= Tr[X\Delta S + \Delta XS] + Tr[\Delta SX + S\Delta X] \\ &= Tr[X\Delta S + \Delta XS + \Delta SX + S\Delta X] \\ &= 2Tr[X\Delta S + S\Delta X] \\ &= 2[X \bullet \Delta S + S \bullet \Delta X] \end{aligned}$$

Using the fact that $Tr(XS) = X \bullet S = n\mu$, we obtain

$$\begin{aligned} (X + \alpha\Delta X) \bullet (S + \alpha\Delta S) &= X \bullet S + \alpha(X \bullet \Delta S + S \bullet \Delta X) + \alpha^2 \Delta S \bullet \Delta X \\ &= X \bullet S + \alpha Tr(\sigma\mu I - XS) \\ &= X \bullet S + \alpha(\sigma n\mu - X \bullet S) \\ &= (1 - \alpha + \sigma\alpha)(X \bullet S) \end{aligned}$$

for every $\alpha \in R$. Hence, (2) holds. ■

Lemma 2.2 *For all* $Q \in R^{n\times n}$, *the following relations hold:*

$$\sum_{i=1}^{n} |\lambda_i(A)|^2 \le \|A\|_F^2 = \|A^T\|_F^2;$$

Lemma 2.3 *Suppose that* $W \in R^{n\times n}$ *is a nonsingular matrix, then for any* $E \in S^n$, *the following relations hold:*

$$\|E\|_F \le \frac{1}{2}\|WEW^{-1} + (WEW^{-1})^T\|_F. \tag{2.13}$$

Lemma 2.4 *Suppose that* $A_1, A_2 \in R^{n\times n}$, *then the following relations hold:*

$$\|A_1A_2\|_F \le \|A_1\|A_2\|_F.$$

For a nonsingular matrix $P \in R^{n\times n}$, consider the following operator $H_P : R^{n\times n} \to S^n$ defined as

$$H_P(M) \equiv \frac{1}{2}[PMP^{-1} + (PMP^{-1})^T, \quad \forall M \in R^{n\times n}.$$

The operator H_P has been used by Zhang [10] to characterize the central path of SDP problems.

Lemma 2.5 *Let* $(X, S) \in S^n_{++}$ *be such that* $\|X^{1/2}SX^{1/2} - \mu I\| \leq \mu\gamma$ *for some* $\gamma \in [0, 1)$ *and* $\mu > 0$*. Suppose that* $(\Delta X, \Delta S) \in S^{n\times n} \times S^{n\times n}$ *is a solution of* (4) *for* $\mathcal{W} \in R^{n\times n}$*, where* $\mathcal{W} = \sigma\mu I - X^{1/2}SX^{1/2}$*. Let* $\delta_x = \mu\|X^{-1/2}\Delta X X^{-1/2}\|_F$ *and* $\delta_s = \|X^{1/2}\Delta S X^{1/2}\|_F$*. Then,*

$$\delta_x\delta_s \leq \frac{1}{2}(\delta_x^2 + \delta_s^2) \leq \frac{\|\mathcal{W}\|_F^2}{2(1-\gamma)^2}.$$

Proof. We let $\mathcal{W} = H_{X^{-1/2}}[X\Delta S + \Delta X S]$. Using (4) and simple algebraic manipulation, we can obtain

$$\begin{aligned}\mathcal{W} &= X^{1/2}\Delta S X^{1/2} + \mu X^{-1/2}\Delta X X^{-1/2} + \frac{1}{2}X^{-1/2}\Delta X X^{-1/2}(X^{1/2}SX^{1/2} - \mu I)\\ &\quad + \frac{1}{2}(X^{1/2}SX^{1/2} - \mu I)X^{-1/2}\Delta X X^{-1/2},\end{aligned}$$

from which it follows that

$$\begin{aligned}\|\mathcal{W}\|_F &\geq \|X^{1/2}\Delta S X^{1/2} + \mu X^{-1/2}\Delta X X^{-1/2}\|_F\\ &\quad -\|X^{1/2}SX^{1/2} - \mu I\|\|X^{-1/2}\Delta X X^{-1/2}\|_F\\ &\geq (\|X^{1/2}\Delta S X^{1/2}\|_F^2 + \mu^2\|X^{-1/2}\Delta X X^{-1/2}\|_F^2)^{1/2} - \gamma\mu\delta_x/\mu\\ &\geq \sqrt{\delta_x^2 + \delta_s^2} - \gamma\delta_x\\ &\geq (1-\gamma)\sqrt{\delta_x^2 + \delta_s^2},\end{aligned}$$

where the second inequality follows from the assumption that $\|X^{1/2}SX^{1/2} - \mu I\| \leq \mu\gamma$ and the fact that $(X^{-1/2}\Delta X X^{-1/2}) \bullet (X^{1/2}\Delta S X^{1/2}) = \Delta X \bullet \Delta S \geq 0$, due to the monotonicity of $\mathcal{F}$. The result now follows trivially from the last inequality. ∎

Lemma 2.6 *With the notations above, we have*

$$\|H_{X^{-1/2}}[\Delta X \Delta S]\| \leq \frac{nK^2\mu}{2(1-\gamma)^2}.$$

where $K = max[|-\gamma + \sigma + 1|], |\gamma + \sigma - 1|] \leq 2$.

Proof. Follows immediately the assumption that $(X, S) \in N(\mu, \gamma)$ and Lemma 2.3, we can obtain

$$\begin{aligned}\|H_{X^{-1/2}}[\Delta X\Delta S]\| &\leq \|X^{-1/2}\Delta X\Delta S X^{-1/2}]\| \leq \|X^{-1/2}\Delta X\Delta S X^{-1/2}]\|_F\\ &\leq \|X^{-1/2}\Delta X X^{-1/2}\|_F\|X^{1/2}\Delta S X^{1/2}\|_F\\ &\leq \frac{\|\sigma\mu I - X^{1/2}SX^{1/2}\|_F^2}{2(1-\gamma)^2\mu} \leq \frac{n\|\sigma\mu I - X^{1/2}SX^{1/2}\|^2}{2(1-\gamma)^2\mu}\\ &\leq \frac{nK^2\mu}{2(1-\gamma)^2}.\end{aligned}$$

Hence, the relation holds. ∎

3. The predictor-corrector algorithm

In this section, we give the polynomial convergence analysis of a predictor-corrector algorithm for SDP which is a director extension of the LP predictor-corrector algorithm studied by Mizuno, Todd and Ye.

The algorithm considered in this subsection is as follows.

Algorithm-I

Choose constants $0 < \tau < 1/2$ satisfying the conditions of Theorem 3.1 below. let $\epsilon \in (0,1)$ and $(X^0, S^0) \in F^0(P) \times F^0(D)$ and $\mu_0 = X^0 \bullet S^0/n$ be such that $(X^0, S^0) \in N_F(\mu_0, \tau)$ and set $k = 0$.

Repeat until $\mu_k \leq \varepsilon\mu_0$, do

(1) Compute the solution $(\Delta X_P^k, \Delta S_P^k)$ of system (10), (11), (12) with $(X, S) = (X^k, S^k)$ and $\hat{\mu} = 0$;

(2) Let $\alpha_k \equiv max\{\alpha \in [0,1] : (X^k(\alpha^{'}), S^k(\alpha^{'})) \in N_F((1-\alpha^{'})\mu_k, 2\tau)$, $\forall \alpha^{'} \in [0, \alpha]\}$, where $X^k(\alpha) = X^k + \alpha\Delta X_P^k$, $S^k(\alpha) = S^k + \alpha\Delta S_P^k$;

(3) Let $(\widehat{X}^k, \widehat{S}^k) \equiv (X^k, S^k) + \alpha_k(\Delta X_P^k, \Delta S_P^k)$ and $\mu_{k+1} = (1-\alpha_k)\mu_k$;

(4) Compute the solution $(\Delta X_C^k, \Delta S_C^k)$ of system (10), (11), (12) with $(X, S) = (\widehat{X}^k, \widehat{S}^k)$ and $\hat{\mu} = \mu_{k+1}$;

(5) Set $(X^{k+1}, S^{k+1}) \equiv (\widehat{X}^k, \widehat{S}^k) + (\Delta X_C^k, \Delta S_C^k)$;

(6) Increment k by 1.

End

The proof of the next lemma is straightforward and, therefore, we omit the details.

Lemma 3.7 *With the notations above, the following relations hold:*

(1) $H_{X^{-1/2}}(X(\alpha)Z(\alpha)) = (1-\alpha)H_{X^{-1/2}}(XZ) + \alpha\gamma\mu I + \alpha^2 H_{X^{-1/2}}(\Delta X \Delta Z)$;

(2) $\mu(\alpha) = (1-\alpha)\mu + \gamma\alpha\mu$;

(3) $H_{X^{-1/2}}(X(\alpha)Z(\alpha)) - \mu(\alpha)I = (1-\alpha)[H_{X^{-1/2}}(XZ) - \mu I] + \alpha\gamma\mu I + \alpha^2 H_{X^{-1/2}}(\Delta X \Delta Z)$.

By Lemma 3.1, we can obtain that the improvement of the objective value depends on the size of α, so we wish to bound α from below.

Theorem 3.1 *With the notations above, we let*

$$\hat{\alpha} = \max\{\alpha \in (0,1], (X(\alpha), Z(\alpha)) \in N(2\eta)\},$$

then

$$\hat{\alpha} \geq \frac{2}{1 + \sqrt{1 + 16\|H_{X^{-1/2}}(\Delta X \Delta Z)/\mu\|_F}}.$$

Proof. Using Lemma 2.4, we have the following inequality:

$$\begin{aligned}
&\|H_{X^{-1/2}}(X(\alpha)Z(\alpha)) - \mu(\alpha)I\|_F \\
&\quad = \|(1-\alpha)(H_{X^{-1/2}}(XZ) - \mu I) + \alpha^2 H_{X^{-1/2}}(\Delta X \Delta Z)\|_F \\
&\quad \leq \|(1-\alpha)H_{X^{-1/2}}(XZ - \mu I)\|_F + \alpha^2 \|H_{X^{-1/2}}(\Delta X \Delta Z)\|_F \\
&\quad \leq (1-\alpha)\eta\mu + \alpha^2 \|H_{X^{-1/2}}(\Delta X \Delta Z)\|_F.
\end{aligned}$$

We see that for

$$0 \leq \hat{\alpha} \leq \frac{2}{1 + \sqrt{1 + 16\|H_{X^{-1/2}} \Delta X \Delta Z/\mu\|_F}};$$

$$\begin{aligned}
\|H_{X^{-1/2}}(X(\alpha)Z(\alpha)) - \mu(\alpha)I\|_F &\leq (1-\alpha)\eta\mu + \alpha^2\|H_{X^{-1/2}}(\Delta X \Delta Z)\|_F \\
&\leq 2\eta(1-\alpha).
\end{aligned}$$

This because the quadratic term in θ:

$$\|H_{X^{-1/2}}(\Delta X \Delta Z)/\mu\|_F \alpha^2 + \eta\alpha - \eta \leq 0$$

for α between zero and the root

$$\frac{2}{1 + \sqrt{1 + 16\|H_{X^{-1/2}}(\Delta X \Delta Z)/\mu\|_F}}.$$

Thus, $\|H_{X^{-1/2}}(X(\theta)Z(\theta)) - \mu(\alpha)I\|_F \leq 2\eta(1-\alpha)\mu = 2\eta\mu(\alpha)$, then we complete the proof. ∎

Theorem 3.2 *Algorithm-I terminates in at most* $\mathcal{O}(\sqrt{n}\log\varepsilon^{-1})$ *iterations.*

Proof. The proof follows immediately from Theorem 3.2 and Lemma 2.7 and a simple induction argument. ∎

References

[1] Nesterod, Y.E. and Nemirovskii, A.S., *Interior Point methods in Convex Programming: Theory and Applications*, SIAM, Philadelphia, 1994.

[2] Alizadeh, F., *Interior point methods in semidefinite programming with applications to combinatorial optimization*, SIAM J. Optim. 5(1995), 13-51.

[3] Alizadeh, F., Haeberly, J.P.A., Overton, M.L., *Primal-dual interior-point methods for semidefinite programming: convergence rates, stability and numerical results.* SIAM Journal on Optimization, 1998, 746-768.

[4] Kojima, M., Shida, M., Shindoh, S., *Local convergence of predictor-corrector infeasible-interior-point algorithms for SDPs and SDLCPs*, Mathematical Programming, 1998, 129-160.

[5] Kojima, M., Shindoh, S., Hara, S., *Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices*. SIAM Journal on Optimization , vol.7, no. 1 (1997), 86-125.

[6] Monteiro, R.D.C., *Primal-Dual Path-Following Algorithms for Semidefinite Programming*, SIAM Journal on Optimization, vol. 7, (3) (1997), 663-678.

[7] Monteiro, R.D.C., *Polynomial convergence of primal-dual algorithms for semidefinite programming based on the Monteiro and Zhang family of directions*, SIAM Journal on Optimization, vol. 8 (1998), 797-812.

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

[9] Monteiro, R.D.C. and Zhang, Y., *Polynomial convergence of a new family of primal-dual algorithms for semidefinite programming,* SIAM Journal on Optimization, vol. 9, no. 3 (1999), 551-577.

[10] Zhang, Yin, *On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming,* SIAM Journal on Optimization, 1998, 365-386.

Accepted: 27.07.2009