www.ElitesJournal.ir

مجله نخبگان علوم و مهندسی

Journal of Science and Engineering Elites

جلد ۴- شماره ۶- سال ۱۳۹۸

A New Primal-Dual Interior-Point Method for Semidefinite Optimization Based on a New Wide Neighborhood with Infinity-Norm

Afsaneh nasrollahi[1*], Behrouz kheirfam[2]

1- Department of Applied Mathematics, Azarbaijan Shahid Madani University, Tabriz, Iran
2- Department of Applied Mathematics, Azarbaijan Shahid Madani University, Tabriz, Iran

*a.nasrollahi@azaruniv.ac.ir

Received: January 2020 Accepted: February 2020

Abstract

In this paper, we present a new primal-dual interior-point algorithm based on a new large neighbourhood $N_\infty(\tau,\beta)$ for semidefinite optimization. This large neighbourhood is based on the infinity norm. It is larger than the $N(\tau,\beta)$ large neighborhood of the central path, which is popular wide neighborhood. We demonstraite the convergence of the proposed algorithm and show that the algorithm has $O(n^{1+\frac{2}{p}}\log\varepsilon^{-1})$ iteration complexity bound for the Nesterov-Todd direction.

Keywords: Semidefinite optimization, wide neighborhood, interior-point method, iteration complexity bound.

1- Introduction

During the last two decades, major developments in convex programming were focusing on semidefinite optimization (SDO), which optimizes a linear function with the matrix variable over the intersection of the cone of positive semidefinite matrices with an affine space.
The interior-point methods (IPMs) are one of the efficient methods for solving SDO. SDO has several applications in such as combinatorial optimization, system and control theory and eigenvalue optimization problems.The IPMs can be classified according to the length of the step that the algorithms take. In this way, we can talk about large-update algorithms which work in a wider neighbourhood of the central path and short-update ones determine the new iterates that are in a smaller neighbourhood. It is well known that wide neighborhood IPMs perform better in practice than small neighborhood IPMs. However, in theory, the iteration bound for small neighborhood IPMs is better than that proved for wide neighborhood IPMs. To bridge this gap, Ai and Zhang in [1] proposed a new wide neighborhood of the central path. They showed that their algorithm has the same complexity as the best known small neighborhood IPMs.

Recently, Asadi et al. in [2] proposed a new wide neighberhod using infnity norm for symmetric cone Cartesian $P_*(\kappa)$ -HLCP which is larger than Ai and Zhang's wide neighborhood. In this paper, we generalize this method for SDO problems and propose a new feasible IPM based on a new wide nighborhood.

2- SDO and preliminary discussions

we consider the following primal SDO problem

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

wher $C, X \in \mathbb{S}^n, b \in \Re^m$ and $A_i \in \mathbb{S}^n (i = 1,\ldots,m)$. Moreover, we assume that the matrices A_i are linearly independent. The dual problem of (P) is defined by

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

where $y \in \Re^m$ and $S \in \mathbb{S}^n$. We denote the primal-dual feasible set by

$$\mathcal{F} := \{(X, y, S) : A_i \bullet X = b_i, i = 1,\ldots,m,\ \ X \succeq 0,$$
$$\sum_{i=1}^{m} y_i A_i + S = C,\ \ S \succeq 0\},$$

and the set of primal-dual feasible interior points is denoted by

$$\mathcal{F}^0 := \{(X, y, S) : A_i \bullet X = b_i, i = 1,\ldots,m, X \succ 0,$$
$$\sum_{i=1}^{m} y_i A_i + S = C, S \succ 0\}.$$

Under the assumption that $\mathcal{F}^0$ is non-empty and the matrices $A_i, (i = 1,\ldots,m)$ are linearly independent, (X, y, S) an optimal solution of (P) and (D) if and only if

$$\begin{aligned} &A_i \bullet X = b_i, i = 1,\ldots,m,\\ &\sum_{i=1}^{m} y_i A_i + S = C, \qquad (1)\\ &XS = 0, \end{aligned}$$

The central path consists of points (X^μ, y^μ, S^μ) satisfying the perturbed system

$$\begin{aligned} &A_i \bullet X = b_i, i = 1,\ldots,m,\\ &\sum_{i=1}^{m} y_i A_i + S = C, \qquad (2)\\ &XS = \tau\mu I, \end{aligned}$$

where $\tau \in (0,1)$, $\mu > 0$ and I is the identity matrix. Nesterov and Nemirovsky [3] proved that there is a unique solution (X^μ, y^μ, S^μ) to the central path Eqs. (2) for any $\mu > 0$. Moreover, the limit point of (X^μ, y^μ, S^μ) as μ goes to zero is a primal-dual optimal solution of the corresponding SDO problem.

Although $X, S \in \mathbb{S}^n$, the product XS is generally not in $\mathbb{S}^n$. Thus, system (2) is not a square system when X and S are restricted to $\mathbb{S}^n$ .To remedy, Zhang [5] introduced a general symmetrization scheme:

$$H_P(M) = \frac{1}{2}[PMP^{-1} + (PMP^{-1})^T], \forall M \in \Re^{n\times n},$$

where $P \in \Re^{n\times n}$ is some non-singular matrix. Zhang [5] also observed that if P is non-singular, then

$$H_P(M) = \tau\mu I \Leftrightarrow M = \tau\mu I.$$

Consequently, for any given non-singular matrix P, system (2) is equivalent to

$$\begin{aligned} &A_i \bullet X = b_i, i = 1,\dots,m, \\ &\sum_{i=1}^{m} y_i A_i + S = C, \\ &H_P(XS) = \tau\mu I, \end{aligned} \tag{3}$$

Applying Newton's method to the system (3) at (X, y, S) leads us to the linear system

$$\begin{aligned} &A_i \bullet \Delta X = 0, i = 1,\dots,m, \\ &\sum_{i=1}^{m} A_i \Delta y_i + \Delta S = 0, \\ &H_P(X\Delta S + \Delta XS) = \tau\mu I - H_P(XS) \end{aligned} \tag{4}$$

Different choices of the matrix P leads to the different search directions. Let us choose $P := W^{\frac{1}{2}}$ proposed by Nesterov and Todd [4] where

$$W = S^{\frac{1}{2}}(S^{\frac{1}{2}}XS^{\frac{1}{2}})^{-\frac{1}{2}}S^{\frac{1}{2}} = X^{-\frac{1}{2}}(X^{\frac{1}{2}}SX^{\frac{1}{2}})^{\frac{1}{2}}X^{-\frac{1}{2}}. \tag{5}$$

Clearly, the NT scaling matrix $P = W^{\frac{1}{2}}$ satisfies $PXSP^{-1} \in \mathrm{S}^n$. Let us introduce the following scaled quantities:

$$\begin{aligned} &\hat{X} = PXP, \quad \Delta\hat{X} = P\Delta XP, \\ &\hat{S} = P^{-1}SP^{-1}, \ \Delta\hat{S} = P^{-1}\Delta SP^{-1}, \ \hat{A}_i = P^{-1}A_iP^{-1}. \end{aligned} \tag{6}$$

Here, by using (6), we scale the search direction systems (4) as follows:

$$\begin{aligned} &\hat{A}_i \bullet \Delta\hat{X} = 0, i = 1,\dots,m, \\ &\sum_{i=1}^{m} \hat{A}_i (\Delta y)_i + \Delta\hat{S} = 0, \\ &H(\hat{X}\Delta\hat{S} + \Delta\hat{X}\hat{S}) = \tau\mu I - H(\hat{X}\hat{S}), \end{aligned} \tag{7}$$

where $H(\cdot) := H_I(\cdot)$ is the plain symmetrization operator.
In this paper, we introduce a new neighborhood for the central path similar to [2 as follows:

$$\mathrm{N}_\infty(\tau,\beta) = \{(X,y,S) \in \mathrm{F}^0 : \left\| (\tau\mu I - X^{\frac{1}{2}}SX^{\frac{1}{2}})^+ \right\|_\infty \le \beta\tau\mu\},$$

where $\beta, \tau \in (0,1)$ are given constants. The above defined neighborhood is a wide neighborhood since one can easily verify that:

$$\mathrm{N}(\tau,\beta) \subseteq \mathrm{N}_\infty(\tau,\beta),$$

where $\mathrm{N}(\tau,\beta) = \{(X,y,S) \in \mathrm{F}^0 : \left\| (\tau\mu I - X^{\frac{1}{2}}SX^{\frac{1}{2}})^+ \right\|_F \le \beta\tau\mu\}$ (8)

3- The Algorithm

We assume that the algorithm begins with an iterate $(X, y, S) \in \mathrm{N}_\infty(\tau, \beta)$.We compute the search directions $(\Delta\hat{X}, \Delta y, \Delta\hat{S})$ by solving the following system:

$$\hat{A}_i \bullet \Delta\hat{X} = 0, i = 1, \ldots, m,$$

$$\sum_{i=1}^{m} \hat{A}_i (\Delta y)_i + \Delta\hat{S} = 0,$$

$$H(\hat{X}\Delta\hat{S} + \Delta\hat{X}) = (\tau\mu I - H(\hat{X}\hat{S}))^+ + \sqrt[p]{n}(\tau\mu I - H(\hat{X}\hat{S}))^-.$$

Let the new iterate be given by

$$(\hat{X}(\alpha), y(\alpha), \hat{S}(\alpha)) = (\hat{X}, y, \hat{S}) + \alpha(\Delta\hat{X}, \Delta y, \Delta\hat{S}),$$

where $0 \le \alpha \le 1$. It follows from the first two equations of (8) that $\Delta\hat{X} \bullet \Delta\hat{S} = 0$. Moreover, the third equation in (8) can be expressed in terms of the Kronecker product as follows:

$$\hat{\mathrm{E}} vec(\Delta\hat{X}) + \hat{\mathrm{F}} vec(\Delta\hat{S}) = vec((\tau\mu I - \hat{X}\hat{S})^+) + \sqrt[p]{n} vec((\tau\mu I - \hat{X}\hat{S})^-), \quad (9)$$

where

$$\hat{\mathrm{E}} := \frac{1}{2}(\hat{S} \otimes I + I \otimes \hat{S}, \; \hat{\mathrm{F}} := \frac{1}{2}(\hat{X} \otimes I + I \otimes \hat{X}).$$

Also, using (8), we have

$$H(\hat{X}(\alpha)\hat{S}(\alpha)) = \chi(\alpha) + \alpha^2 H(\Delta\hat{X}\Delta\hat{S}),$$

Where

$$\chi(\alpha) = \hat{X}\hat{S} + \alpha H(\hat{X}\Delta\hat{S} + \hat{S}\Delta\hat{X})$$
$$= \hat{X}\hat{S} + \alpha(\tau\mu I - H(\hat{X}\hat{S}))^+ + \alpha\sqrt[p]{n}(\tau\mu I - H(\hat{X}\hat{S}))^-. \quad (10)$$

The last equality is due to the third equation of (8). In the sequel, we will denote the eigenvalues of the matrix XS as $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$, and since the matrices $XS, SX, X^{\frac{1}{2}}SX^{\frac{1}{2}}, S^{\frac{1}{2}}XS^{\frac{1}{2}}, \hat{X}\hat{S}, \hat{S}\hat{X}$ are similar, they all have the same eigenvalues. Since $(\hat{X}, y, \hat{S}) \in \mathrm{F}^0$, we have

$$\lambda_i(\chi(\alpha)) = \lambda_i + \alpha(\tau\mu - \lambda_i), \quad \forall i \in \mathrm{I}^+ \quad (11)$$

and

$$\lambda_i(\chi(\alpha)) = \lambda_i + \sqrt[p]{n}\alpha(\tau\mu - \lambda_i), \quad \forall i \in \mathrm{I}^-$$

Where $\mathrm{I}^+ = \{i \mid \tau\mu - \lambda_i > 0\}$ and $\mathrm{I}^- = \{i \mid \tau\mu - \lambda_i < 0\}$.

Also, using (8), $\Delta\hat{X} \bullet \Delta\hat{S} = 0$ and $TrH_P(M) = Tr(M)$, we have

$$\mu(\alpha) = \frac{Tr(X(\alpha)S(\alpha))}{n} = \frac{Tr(H(\hat{X}(\alpha)\hat{S}(\alpha)))}{n}$$
$$= \mu + \frac{\alpha}{n} Tr(\tau\mu I - \hat{X}\hat{S})^+ + \frac{\alpha}{n^{(1-\frac{1}{p})}} Tr(\tau\mu I - \hat{X}\hat{S})^- \quad (12)$$

Algorithm 1: Feasible Interior point algorithm

Input: an accuracy parameter $\varepsilon > 0$, neighborhood parameters $0 < \tau \le \frac{1}{4}, 0 < \beta \le \frac{1}{2}$, and an initial point $(X^0, y^0, S^0) \in \mathcal{N}_\infty(\tau, \beta)$ with $X^0 \bullet S^0 = n\mu^0$

Output: a sequence of iterates $\{(X^k, y^k, S^k) : k = 0,1,2,\cdots\}$.

Step 0: Set $k := 0$
Step 1: If $X^k \bullet S^k \le \varepsilon$, then stop. Otherwise go to step 2.

Step 2: Compute the NT scaling $P^k = (W^k)^{\frac{1}{2}}$ by (5).
Step 3: Compute the Newton direction $(\Delta\hat{X}, \Delta y, \Delta\hat{S})$ from (8) .
Step 4: Find the largest step size $0 \le \alpha^k \le 1$, such that

$$(X(\alpha), y(\alpha), S(\alpha)) \in \mathcal{N}_\infty(\tau, \beta), \quad (13$$

Step 5: Let $(X^{k+1}, y^{k+1}, S^{k+1}) := (X(\alpha^k), y(\alpha^k), S(\alpha^k))$. Set $k := k+1$ and go to Step 1.

4- Complexity analysis

In this section, we present the convergence analysis of the algorithm and derive the $O(n^{1+\frac{2}{p}} \log \varepsilon^{-1})$ iteration bound for the algorithm. Firstly, we recall some useful results.

Lemma 1 Assume $U, V \in \mathcal{S}^n$, then we have

$$\|(U+V)^+\|_\infty \le \|U^+ + V^+\|_\infty \le \|U^+\|_\infty + \|V^+\|_\infty.$$

Lemma 2 Assume $U, V \in \mathcal{S}^n$, then we have

$$\|U\|_\infty \le \|U\|_F,$$

thus

$$\|< U, V >\|_\infty \le \|< U, V >\|_F \le \|U\|_F \|V\|_F.$$

Lemma 3 For $U \in \mathcal{S}^n$, we have $\|U\|_1 \le n\|U\|_\infty$.

Lemma 4 Suppose that $Q \in \mathcal{R}^{n\times n}$ is a non-singular matrix. Then, for any $M \in \mathcal{S}^n$, we have

$$\|M^+\|_\infty \le \frac{1}{2}\|[QMQ^{-1} + (QMQ^{-1})^T]^+\|_\infty.$$

Lemma 5 Let $(X, y, S) \in \mathcal{F}^0$. Then, we have

(i) $Tr([\tau\mu I - X^{\frac{1}{2}} S X^{\frac{1}{2}}]^-) \le -(1-\tau)n\mu,$

(ii) $Tr([\tau\mu I - X^{\frac{1}{2}} S X^{\frac{1}{2}}]^+) \le n\beta\tau\mu.$

Lemma 6 Let $\alpha \in (0,1]$, $\mu(\alpha) = \dfrac{Tr(X(\alpha)S(\alpha))}{n}$ and $(X, y, S) \in \mathcal{F}^0$. Then, we have

$\mu(\alpha) \geq (1-\alpha\sqrt[p]{n})\mu.$

Lemma 7 Let $(X,y,S)\in \mathrm{N}(\tau,\beta)$. Then $\mu(\alpha)$ is strictly monotonically decreasing in $\alpha\in(0,1]$.

Lemma 8 For any $u,v\in\Re^n$ and $G\in \mathrm{S}^n_{++}$, we have

$$\|u\|\ \|v\| \leq \sqrt{cond(G)}\left\|G^{-1/2}u\right\|\ \left\|G^{1/2}v\right\|$$
$$\leq \frac{1}{2}\sqrt{cond(G)}(\left\|G^{-1/2}u\right\|^2+\left\|G^{1/2}v\right\|^2),$$

where $cond(G)=\dfrac{\lambda_{max}(G)}{\lambda_{min}(G)}$.

Lemma 9 For the NT search directions, $cond(G)=1$.

Lemma 10 If $(X,y,S)\in \mathrm{N}(\tau,\beta), \tau\leq\frac{1}{4}$ and $\beta\leq\frac{1}{2}$, then

(i) $\left\|(\hat{E}\hat{F})^{-\frac{1}{2}}vec(\tau\mu I-\hat{X}\hat{S})^-\right\|^2 \leq (1-\frac{\tau}{2})n\mu,$

(ii) $\left\|(\hat{E}\hat{F})^{-\frac{1}{2}}vec(\tau\mu I-\hat{X}\hat{S})^+\right\|^2 \leq \dfrac{n\beta^2\tau\mu}{1-\beta}.$

Lemma 11 Let $(\hat{X},y,\hat{S})\in \mathrm{N}(\tau,\beta), G=\hat{E}^{-1}\hat{F}$, then

$$\left\|H(\Delta\hat{X}\Delta\hat{S})\right\|_F \leq \left\|\Delta\hat{X}\right\|_F\left\|\Delta\hat{S}\right\|_F \leq \frac{1}{2}n^{\frac{2}{p}+1}\mu(\beta\tau+1-\frac{\tau}{2})$$

Lemma 12 Let $\alpha\in(0,1]$, $\mu(\alpha)=\dfrac{Tr(X(\alpha)S(\alpha))}{n}$ and $(X,y,S)\in \mathrm{F}^0$. Then, we have

$\mu(\alpha)\leq(1-\gamma\alpha)\mu.$

where $\gamma=1-\tau$

Lemma 13 Let $\mu(\alpha)>0$ and $(\hat{X},y,\hat{S})\in \mathrm{N}(\tau,\beta)$, then , we have

$$\left\|[\tau\mu(\alpha)I-\chi(\alpha)]^+\right\|_\infty \leq (1-\alpha)\beta\tau\mu(\alpha).$$

Lemma 15 Let $\beta\leq\frac{1}{2},\tau\leq\frac{1}{4}$ and $(\hat{X},y,\hat{S})\in \mathrm{N}_\infty(\tau,\beta)$. Then for all $0\leq\alpha\leq\dfrac{\beta\tau}{n^{1+\frac{2}{p}}}$, we have $(X(\alpha),y(\alpha),S(\alpha))\in \mathrm{N}_\infty(\tau,\beta)$.

5- Polynomial complexity

The next theorem gives an iteration-complexity bound for the algorithm .

Theorem 16 Suppose that $\tau\leq\frac{1}{4}$ and $\beta\leq\frac{1}{2}$ are fixed for all iterations. Then Algorithm 1 will terminate in $\mathrm{O}(n^{1+\frac{2}{p}}\log\varepsilon^{-1})$ iterations with a solution X^k,y^k,S^k for which

$X^k \bullet S^k \le \varepsilon (X^0 \bullet S^0)$.

6- Conclusions

In this paper, we proposed a new primal-dual interior-point algorithm for SDO problems acting in a new large neighborhood. We analyzed this algorithm and we showed its globally convergence using the symmetrization of the NT-directions. We showed that the algorithm has $O(n^{1+\frac{2}{p}} \log \varepsilon^{-1})$ iteration complexity.

7- References

1. Ai, W., Zhang, S., ”An $O(\sqrt{n}L)$ iteration primal-dual path-following method, based on wide neighborhoods and large updates for monotone LCP” , SAIM J.Optim., Vol.16, No.2 2005, pp.400-417.

2. Asadi, S., Mansouri, H., Lesaja, G., Zangiabadi, M., “A long-step interior-point algorithm for symmetric cone Cartesian $P_*(\kappa)$ -HLCP”, J. Optimization, DOI: 10.1080/02331934.2018.1512604

3. Nesterov, Y.E. and Nemirovski, A.S., Interior point methods in convex programming: theory and applications SIAM, Philadelphia, PA. 1994.

4. Nesterov, Y.E., Todd, M.J., “Primal-dual interior-point methods for self-scaled cones”, SIAM J. Optim., Vol.8, No.1 1998, pp.324-364.

5. Zhang, Y., “On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming”, SIAM J. Optim., Vol.8, No.1 1998, pp.365-386.