Operations Research Letters 44 (2016) 319–323

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

A primal–dual interior-point method for semidefinite optimization based on a class of trigonometric barrier functions

M. Reza Peyghami [a,b,*], S. Fathi-Hafshejani [c], S. Chen [d]

[a] Faculty of Mathematics, K.N. Toosi University of Tech., P.O. Box 16315-1618, Tehran, Iran

[b] Scientific Computations in Optimization and Systems Engineering (SCOPE), K.N. Toosi University of Tech., Tehran, Iran

[c] Faculty of Mathematics, Shiraz University of Tech., P.O. Box 71555-313, Shiraz, Iran

[d] Department of Mathematics and Statistics, York University, Toronto, Ontario, Canada

ARTICLE INFO

Article history:
Received 11 December 2015
Received in revised form
21 February 2016
Accepted 29 February 2016
Available online 5 March 2016

Keywords:
Proximity function
Semidefinite optimization
Interior-point methods
Large-update methods

ABSTRACT

A primal–dual interior-point method (IPM) based on a new class of proximity functions is proposed for solving Semidefinite Optimization (SDO) problems. The proposed functions are induced from the kernel functions with trigonometric barrier terms. We derive iteration complexity of large-update IPMs for SDO as $O\left(\sqrt{n}\log n\log\frac{n}{\epsilon}\right)$. This improves the result obtained in Li and Zhang (2015) for linear optimization and matches to the bound for the so-called self-regular kernel functions.

1. Introduction

Semidefinite Optimization (SDO) problems are convex optimization problems over the intersection of an affine set and the cone of positive semidefinite matrices. These problems have wide applications in the real world problems [17] and became an active area of research since invention of Interior Point Methods (IPMs) by Karmarkar [4] in 1984. Note that some IPMs designed for Linear Optimization (LO) have been successfully extended to SDO. The first work in this area was proposed in [8] using the self-concordant barrier functions which consist of the so-called logarithmic barrier function.

Peng et al. in [9] proposed a new variant of IPMs for solving LO and conic problems in which the logarithmic barrier function is replaced by the so-called Self-Regular (SR) functions. For a given accuracy $\epsilon > 0$ and problem size n, they derived the so far best known worst case complexity results for small- and large-update primal–dual IPMs as $O\left(\sqrt{n}\log\frac{n}{\epsilon}\right)$ and $O\left(\sqrt{n}\log n\log\frac{n}{\epsilon}\right)$, respectively. These results narrowed up the gap between small- and large-update methods to a factor of $\log n$, which was $\sqrt{n}$ based on the logarithmic barrier function. Since then, the main focus is on finding a kernel function for which the complexity of large-update methods is equal to (or even better than) $O\left(\sqrt{n}\log\frac{n}{\epsilon}\right)$, or show that such a kernel function does not exist. In this regard, many attempts for introducing non-SR kernel functions have been done, see e.g. [1,10,11,16]. The kernel function with trigonometric barrier term was first introduced in [3]. They derived the worst case complexity for large-update IPMs as $O\left(n^{\frac{3}{4}}\log\frac{n}{\epsilon}\right)$. Following this work, some other kernel functions with trigonometric barrier terms have been proposed in the literature, see e.g. [2,5,12,13]. Recently, Li and Zhang in [6] introduced a new trigonometric kernel function and obtained the complexity of IPMs for LO as $O\left(n^{\frac{2}{3}}\log\frac{n}{\epsilon}\right)$.

In this paper, we propose a new kernel function with trigonometric barrier term and analyze the worst case iteration complexity of large-update primal–dual IPMs for SDO problems. Using a simple analysis, we show that the iteration complexity is bounded above by $O\left(\sqrt{n}\log n\log\frac{n}{\epsilon}\right)$, which narrows up the result obtained in [6] to a factor of $O\left(n^{\frac{1}{6}}\log n\right)$. Moreover, the obtained complexity result coincides to the best known results by self-regular kernel functions.

The paper is organized as follows: In Section 2, we recall some concepts of IPMS for SDO problems. The new kernel function along with its analytical properties is introduced in Section 3. An

* Corresponding author at: Faculty of Mathematics, K.N. Toosi University of Tech., P.O. Box 16315-1618, Tehran, Iran.

E-mail addresses: peyghami@kntu.ac.ir (M.R. Peyghami), s.fathi@sutech.ac.ir (S. Fathi-Hafshejani), chensy@mathstat.yorku.ca (S. Chen).

http://dx.doi.org/10.1016/j.orl.2016.02.013
0167-6377/

estimation of the step size and its default value are provided in Section 4. Finally, the worst case iteration complexity is given in Section 5.

Notations. $\mathbb{R}^n_+$ and $\mathbb{R}^n_{++}$ denote the nonnegative and positive orthants, respectively. $\|\cdot\|$ stands for the Frobenius norm for matrices and the Euclidean norm for vectors. $\mathbb{R}^{m\times n}$ is the space of all real $m\times n$ matrices. S^n, S^n_+ and S^n_{++} denote the cone of symmetric, symmetric positive semidefinite and symmetric positive definite matrices, respectively. For any $A\in\mathbb{R}^{n\times n}$, $Tr(A)$ stands for the trace of A. For $A, B\in\mathbb{R}^{m\times n}$, the inner product is defined by $A\bullet B = Tr(AB^T)$. For any $Q\in S^n_{++}$, $Q^{\frac{1}{2}}$ stands for the symmetric square root of Q. For a vector $x\in\mathbb{R}$, diag(x) is a diagonal matrix with x_i's as its diagonal entries. For any $V\in S^n_{++}$, $\lambda(V)$ denotes the vector of eigenvalues of V arranged in non-increasing order.

2. Preliminaries

In this section, we briefly describe the idea behind IPMs for primal and dual SDO problems as below:

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

$$(D)\quad \max\left\{b^Ty : \sum_{i=1}^m y_iA_i + S = C,\ S\succeq\mathbf{0}\right\},$$

where C and A_i, for $1\le i\le m$, are symmetric $n\times n$ matrices, and $b, y\in\mathbb{R}^m$. Furthermore, $X\succeq\mathbf{0}$ ($\succ\mathbf{0}$) means that X is a symmetric positive semidefinite (positive definite) matrix. We assume that A_i's are linearly independent. Besides, it is assumed that both problems (P) and (D) satisfy the Interior Point Condition (IPC), i.e., there exist $X^0\succ\mathbf{0}$ and (y^0, S^0) with $S^0\succ\mathbf{0}$ so that $A_i\bullet X^0 = b_i$, for $1\le i\le m$, and $\sum_{i=1}^m y_i^0A_i + S^0 = C$.

Definition 2.1 (*Definition 2.2 in [16]*)**.** Let $V\in S^n_{++}$ and $Q\in\mathbb{R}^{n\times n}$ be an orthogonal matrix that diagonalizes V, i.e., $V = Q^T\text{diag}(\lambda(V))Q$. Then, for a given real function $\psi(t)$, $t\ge 0$, the matrix function is defined by

$$\psi(V) = Q^T\text{diag}(\psi(\lambda_1(V)), \psi(\lambda_2(V)), \dots, \psi(\lambda_n(V)))Q. \tag{1}$$

Moreover, the real valued matrix function $\Psi(V) : S^n_{++}\to\mathbb{R}_+$ induced by $\psi(V)$ is given by $\Psi(V) := \sum_{i=1}^n\psi(\lambda_i(V))$.

Let $\psi(t)$ be a twice differentiable function for $t>0$. Then, for the matrix function $\psi(V)$, the first and second order derivatives are defined by replacing $\psi(\lambda_i(V))$'s in (1) with $\psi'(\lambda_i(V))$'s and $\psi''(\lambda_i(V))$'s, respectively. A matrix $M(t)$ is said to be a matrix of functions if each of its entry is a function of t. The usual concepts of continuity, differentiability, and integrability can be naturally extended to matrices of function, by interpreting them as entry-wise.

The optimality conditions for the problems (P) and (D) are as follows:

$$\begin{aligned} &A_i\bullet X = b_i,\quad i=1,\dots,m, \qquad \sum_{i=1}^m y_iA_i + S = C,\\ &XS = \mathbf{0},\quad X\succeq\mathbf{0},\ S\succeq\mathbf{0}.\end{aligned} \tag{2}$$

The core idea of primal–dual IPMs is to replace the equation $XS=\mathbf{0}$ in (2) by the parameterized equation $XS=\mu E$, for $\mu>0$, where E is the identity matrix. This leads system (2) to the following parameterized system:

$$\begin{aligned} &A_i\bullet X = b_i,\quad i=1,\dots,m, \qquad \sum_{i=1}^m y_iA_i + S = C,\\ &XS = \mu E,\quad X, S\succ\mathbf{0}.\end{aligned} \tag{3}$$

Assuming IPC, system (3) has a unique solution for each $\mu>0$, denoted by $(X(\mu), y(\mu), S(\mu))$. $X(\mu)$ is called the μ-center of (P) and $(y(\mu), S(\mu))$ is known as the μ-center of (D). The central path for SDO is the set of all μ-centers for $\mu>0$. As $\mu\to 0$, the limit of the central path exists and tends to the analytic center of the optimal solutions set. Most of IPMs follow the central path approximately to get close enough to the optimal solution, see e.g. [7,15,17].

An application of Newton's method on system (3) yields the following system for the search direction $(\Delta X, \Delta y, \Delta S)$:

$$\begin{aligned} &A_i\bullet\Delta X = 0,\quad 1\le i\le m, \qquad \sum_{i=1}^m\Delta y_iA_i + \Delta S = \mathbf{0},\\ &X\Delta S + S\Delta X = \mu E - XS.\end{aligned} \tag{4}$$

Again, this system has a unique solution in which ΔX is not necessarily symmetric and requires some symmetrization techniques [17]. Here, we use the Nesterov–Todd symmetrization scheme which leads to NT direction [17]. Let us define $P := X^{\frac{1}{2}}(X^{\frac{1}{2}}SX^{\frac{1}{2}})^{-\frac{1}{2}}X^{\frac{1}{2}} = S^{-\frac{1}{2}}(S^{\frac{1}{2}}XS^{\frac{1}{2}})^{\frac{1}{2}}S^{-\frac{1}{2}}$ and $D = P^{\frac{1}{2}}$. The matrix D is used to scale X and S in order to define a symmetric and positive definite matrix $V := \frac{1}{\sqrt{\mu}}D^{-1}XD^{-1} = \frac{1}{\sqrt{\mu}}DSD$. Thus, we have

$$V^2 := \frac{1}{\mu}D^{-1}XSD. \tag{5}$$

Let us further define

$$\begin{aligned} &\bar{A}_i := DA_iD,\quad 1\le i\le m,\\ &D_X := \frac{1}{\sqrt{\mu}}D^{-1}(\Delta X)D^{-1}, \qquad D_S := \frac{1}{\sqrt{\mu}}D(\Delta S)D.\end{aligned} \tag{6}$$

Now, by scaling (4), the (scaled) NT direction $(D_X, \Delta y, D_S)$ can be computed by solving the following system:

$$\begin{aligned} &\bar{A}_i\bullet D_X = 0,\quad 1\le i\le m, \qquad \sum_{i=1}^m\Delta y_i\bar{A}_i + D_S = \mathbf{0},\\ &D_X + D_S = V^{-1} - V.\end{aligned} \tag{7}$$

A crucial observation in (7) is that the right hand side of the last equation is $-\psi_c'(V)$, where the matrix function ψ_c is induced from the strongly convex function $\psi_c(t) = \frac{t^2-1}{2} - \log t$, for $t>0$, with $\psi_c'(1) = \psi_c(1) = 0$. Let $\psi(t)$ be any strictly convex function on $\mathbb{R}_{++}$ with $\psi'(1) = \psi(1) = 0$. The univariate function $\psi(t)$ is called the kernel function. Now, replacing the right-hand side of the last equation in (7) by $-\psi'(V)$ yields the direction $(D_X, \Delta y, D_S)$ as below:

$$\begin{aligned} &\bar{A}_i\bullet D_X = 0,\quad 1\le i\le m, \qquad \sum_{i=1}^m\Delta y_i\bar{A}_i + D_S = \mathbf{0},\\ &D_X + D_S = -\psi'(V).\end{aligned} \tag{8}$$

The unique solution of (8) is used to compute ΔX and ΔS from (6). Note that D_X and D_S are orthogonal and

$$\begin{aligned} &D_X = D_S = \mathbf{0}_{n\times n} \Leftrightarrow \psi'(V) = \mathbf{0}_{n\times n} \Leftrightarrow V = E \Leftrightarrow \Psi(V) = 0\\ &\quad\Leftrightarrow XS = \mu E \Leftrightarrow (X, S) = (X(\mu), S(\mu)).\end{aligned}$$

By taking an appropriate step size, one can construct a new triple (X_+, y_+, S_+) according to

$$X_+ = X + \alpha\Delta X, \qquad y_+ = y + \alpha\Delta y, \qquad S_+ = S + \alpha\Delta S. \tag{9}$$

The procedure of an IPM for finding ϵ-approximate solutions of (P) and (D) based on kernel functions has been outlined in Algorithm 1 in [10].