

A Parametric Kernel Function Yielding the Best Known Iteration Bound of Interior-Point Methods for Semidefinite Optimization

Xiyao Luo^{*}, Gang Ma, Xiaodong Hu, Yuqing Fu

College of Fundamental Studies, Shanghai University of Engineering Science, Shanghai, China

Email address:

vicegod1996@163.com (Xiyao Luo)

^{*}Corresponding author

To cite this article:

Xiyao Luo, Gang Ma, Xiaodong Hu, Yuqing Fu. A Parametric Kernel Function Yielding the Best Known Iteration Bound of Interior-Point Methods for Semidefinite Optimization. *American Journal of Applied Mathematics*. Vol. 4, No. 6, 2016, pp. 316-323.

doi: 10.11648/j.ajam.20160406.18

Received: November 2, 2016; Accepted: December 6, 2016; Published: December 26, 2016

Abstract: In this paper, a class of large-update primal-dual interior-point methods for semidefinite optimization based on a parametric kernel function are presented. The proposed kernel function is not only used for determining the search directions but also for measuring the distance between the given iterate and the center for the algorithms. By means of the Nesterov and Todd scaling scheme, the currently best known iteration bounds for large-update methods is established.

Keywords: Interior-Point Methods, Semidefinite Optimization, Large-Update Methods, Polynomial Complexity

1. Introduction

In this paper, we focus on the primal problem of semidefinite optimization (SDO) in the standard form

$$(P) \quad \min \{C \cdot X : A_i \cdot X = b_i, i = 1, 2, \dots, m, X \succeq 0\},$$

and its dual problem

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

Here, each $A_i \in S^n, b \in R^m, C \in S^n$. Throughout the paper, we assume that the matrices A_i are linearly independent. Recently, (SDO) has been one of the most active research areas in mathematical programming.

Many interior-point methods (IPMs) for linear optimization (LO) are successfully extended to (SDO) due to their polynomial complexity and practical efficiency. For an overview of these results, we refer to [1, 2] and the references [3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13].

Kernel functions play an important role in the design and analysis of primal-dual (IPMs) for optimization and complementarity problems. They are not only used for

determining the search directions but also for measuring the distance between the given iterate and the μ -center for the algorithms. Currently, (IPM) based on kernel function is one of the most effective methods for (LO) and (SDO) and is a very active research areas in mathematical programming. Particularly, Bai et al. [14] introduced a variety of non-self-regular kernel functions, i.e., the so-called eligible kernel functions, which is defined by some simple conditions on the kernel functions and their derivatives. They provided a simple and unified computational scheme for the complexity analysis of primal-dual kernel function based (IPMs) for (LO). Consequently, a series of eligible kernel functions are considered for various optimization problems and complementarity problems, see, e.g., [15, 16, 17, 18]. For a survey, we refer to the monograph [19] on the subject and the references therein.

In this paper, we consider the following parametric kernel function [8]

$$\varphi(t) = \frac{t^2 - 1}{2} + \frac{q^{\frac{1}{t-1}}}{\log q}, \quad q > 1, t > 0, \quad (1)$$

which is a generalization of the finite kernel function considered in [15] for (LO), namely,

$$\varphi(t) = \frac{t^2 - 1}{2} + e^{\frac{1}{t-1}}, \quad t > 0. \quad (2)$$

The purpose of the paper is to extend the primal-dual large-update (IPMs) for (LO) based on the parametric function considered in [15] to (SDO) by using the NT-scaling scheme [11, 21]. Furthermore, the complexity results match the currently best result of iteration bounds for large-update methods is established, namely, $O\left(\sqrt{n} \log n \log \frac{n}{\varepsilon}\right)$, by choosing $q = 1 + O(n)$.

The outline of the rest of the paper is as follows. In Section 2, we recall some basic concepts and results on matrix theory, the properties of the parametric kernel (and barrier) function. Primal-dual kernel function-based (IPMs) for (SDO) are presented in Section 3. In Section 4, we give the complexity analysis of the primal-dual (IPMs) for (SDO). Finally, some concluding remarks are made in Section 5.

Some of the notations used throughout the paper are as follows. R^n , R_+^n and R_{++}^n denote the set of vectors with n components, the set of nonnegative vectors and the set of positive vectors, respectively. $R^{n \times n}$ denotes the set of $n \times n$ real matrices. E denotes the $n \times n$ identity matrix. $\|\cdot\|_F$ and $\|\cdot\|_2$ denote the Frobenius norm and the spectral norm for matrices, respectively. S^n , S_+^n and S_{++}^n denote the cone of symmetric, symmetric positive semidefinite and symmetric positive definite $n \times n$ matrices, respectively. $A \bullet B = \text{Tr}(A^T B)$ denotes the matrix inner product of two matrices A and B , respectively. The Löwner partial order \succeq (or \succ) on positive semidefinite (or positive definite) matrices means $A \succeq B$ (or $A \succ B$) if $A - B$ is positive semidefinite (or positive definite). Finally, if $g(x) \geq 0$ is a real valued function of a real nonnegative variable, the notation $g(x) = O(x)$ means that $g(x) \leq \bar{c}x$ for some positive constant \bar{c} and $g(x) = \Theta(x)$ that $c_1 x \leq g(x) \leq c_2 x$ for two positive constants c_1 and c_2 .

2. Preliminaries

2.1. Some Results on Matrices and Matrix Functions

In this section, some well known results on matrices and matrix functions from linear algebra are considered. Our presentation is mainly based on the monograph [20] and the references [8, 21].

Let $V \in S^n$ and

$$V = Q^T \text{diag}(\lambda_1(V), \lambda_2(V), \dots, \lambda_n(V))Q, \quad (3)$$

where Q is any orthonormal matrix ($Q^T = Q^{-1}$) that diagonalizes V . The matrix valued function $\varphi(V) : S^n \rightarrow S^n$ is defined by

$$\varphi(V) = Q^T \text{diag}(\varphi(\lambda_1(V)), \varphi(\lambda_2(V)), \dots, \varphi(\lambda_n(V)))Q. \quad (4)$$

Let $\varphi(t)$ be differentiable, i.e., the derivative $\varphi'(t)$ exists. Then the matrix valued function $\varphi'(V)$ is well defined, namely

$$\varphi'(V) = Q^T \text{diag}(\varphi'(\lambda_1(V)), \varphi'(\lambda_2(V)), \dots, \varphi'(\lambda_n(V)))Q. \quad (5)$$

Recall that a matrix $A(t)$ is said to be a matrix of functions if each entry of $A(t)$ is a function of t , i.e., $A(t) = [A_{ij}(t)]$. Let $A(t)$ and $B(t)$ be two matrices of functions. Then

$$\frac{d}{dt} A(t) = \left[\frac{d}{dt} A_{ij}(t) \right] = A'(t), \quad (6)$$

$$\frac{d}{dt} \text{Tr}(A(t)) = \text{Tr}(A'(t)), \quad (7)$$

$$\frac{d}{dt} (A(t)B(t)) = A'(t)B(t) + A(t)B'(t). \quad (8)$$

For any function $\varphi(t)$, let us denote by $\Delta\varphi$ the divided difference of $\varphi(t)$ as follows

$$\Delta\varphi(t_1, t_2) = \frac{\varphi(t_1) - \varphi(t_2)}{t_1 - t_2}, \quad t_1 \neq t_2 \in R. \quad (9)$$

If $t_1 = t_2$, we simply write $\Delta\varphi(t, t) = \varphi'(t)$.

The following theorem provides to measure the first-order directional derivative of a general function $\varphi(A(t))$ and bound its second-order derivative with respect to t .

Theorem 2.1 (Lemma 16 in [21]) Suppose that $A(t)$ is a matrix of functions such that the matrix $A(t)$ is positive definite with eigenvalues $\lambda_1(t) \geq \lambda_2(t) \geq \dots \geq \lambda_n(t) > 0$. If $A(t)$ is twice differentiable with respect to $t \in (l_i, u_i)$ and $\varphi(t)$ is twice continuously differentiable function in a domain that contains all the eigenvalues of $A(t)$, then

$$\frac{d}{dt} \text{Tr}(\varphi(A(t))) = \text{Tr}(\varphi'(A(t))A'(t)), \quad (10)$$

and

$$\frac{d^2}{dt^2} \text{Tr}(\varphi(A(t))) \leq \omega \|A'(t)\|^2 + \text{Tr}(\varphi''(A(t))A''(t)), \quad (11)$$

where

$$\omega = \max\{|\Delta\varphi'(\lambda_j(t), \lambda_k(t))| : t \in (l_i, u_i), j, k = 1, 2, \dots, n\} \quad (12)$$

is a number depending on $A(t)$ and $\varphi(t)$ with

$$\Delta\varphi'(t_1, t_2) = \frac{\varphi'(t_1) - \varphi'(t_2)}{t_1 - t_2}, \forall t_1, t_2 \in [l_t, u_t]. \quad (13)$$

2.2. The Parametric Kernel (Barrier) Function

The first three derivatives of $\varphi(t)$ defined by (1) with respect to t are given by

$$\varphi'(t) = t - \frac{1}{t^2} q^{\frac{1}{t-1}}, \quad (14)$$

$$\varphi''(t) = 1 + \frac{\log q + 2t}{t^4} q^{\frac{1}{t-1}}, \quad (15)$$

$$\varphi'''(t) = -\frac{\log^2 q + 6t \log q + 6t^2}{t^6} q^{\frac{1}{t-1}}. \quad (16)$$

In what follows, we recall some useful results in [15, 16] without proofs.

Lemma 2.1 Let $\varrho: [0, \infty) \rightarrow [1, \infty)$ be the inverse function of $\varphi(t)$ for $t \geq 1$. Then

$$\sqrt{1+2s} \leq \varrho(s) \leq 1 + \sqrt{2s}.$$

Lemma 2.1 Let $\varphi_b(t) = \frac{q^t}{\log q}$ be the barrier term of $\varphi(t)$,

$\rho(s)$ be the inverse function of $-\frac{1}{2}\varphi'(t)$ for $t \in (0, 1]$ and $\bar{\rho}: [0, \infty) \rightarrow (0, 1]$ be the inverse function of the restriction of $-\varphi_b(t)$ for $t \in (0, 1]$, respectively. Then

$$\rho(s) \geq \bar{\rho}(1+2s); \quad (17)$$

$$\bar{\rho}(s) \geq \frac{1}{1 + \frac{\log(s)}{\log(q)}}; \quad (18)$$

$$\rho(s) \geq \frac{1}{1 + \frac{\log(1+2s)}{\log(q)}}. \quad (19)$$

The following property, i.e., the exponential convexity, which plays an important role in the analysis of kernel-function based (IPMs) [15, 21].

Lemma 2.3 (Lemma 2.1 in [15]) Let $t_1 > 0$ and $t_2 > 0$. Then

$$\varphi(\sqrt{t_1 t_2}) \leq \frac{1}{2}(\varphi(t_1) + \varphi(t_2)).$$

Now, we define the barrier function $\Psi(V): S_+^n \rightarrow R_+$ according to the kernel function $\varphi(t)$ as follows

$$\Psi(X, S, \mu) := \Psi(V) := \text{Tr}(\varphi(V)). \quad (20)$$

From (16), we have

$$\Psi(V) = \sum_{i=1}^n \varphi(\lambda_i(V)). \quad (21)$$

One can easily verify that the derivative of the barrier function exactly equal to $\varphi'(V)$, which is defined by (5). Furthermore, we know that $\Psi(V)$ is strictly convex with respect to $V \succ 0$ and vanishes at its global minimal point $V = E$, i.e., $\varphi(E) = \varphi'(E) = 0$, and $\Psi(E) = 0$.

We have the following theorem, by Lemma 2.3.

Theorem 2.2 (Proposition 3 (II) in [21]) Let $V_1, V_2 \in S_{++}^n$. Then

$$\Psi\left(\left[V_1^{\frac{1}{2}} V_2 V_1^{\frac{1}{2}}\right]^{\frac{1}{2}}\right) \leq \frac{1}{2}(\Psi(V_1) + \Psi(V_2)).$$

The following theorem provides an estimate for the effect of a μ -update on the value of $\Psi(V)$, which is a reformulation of Theorem 3.2 in [15].

Theorem 2.3 Let $V \in S_{++}^n$ and $\beta \geq 1$. Then

$$\Psi(\beta V) \leq n\varphi\left(\frac{\Psi(V)}{n}\right).$$

Corollary 2.1 Let $0 \leq \theta < 1$ and $V_+ = \frac{V}{\sqrt{1-\theta}}$. If $\Psi(V) \leq \tau$, then

$$\Psi(V_+) \leq n\varphi\left(\frac{\varrho\left(\frac{\tau}{n}\right)}{\sqrt{1-\theta}}\right).$$

Proof: With $\beta = \frac{1}{\sqrt{1-\theta}} \geq 1$ and $\Psi(V) \leq \tau$, the result follows immediately from Theorem 2.3. This completes the proof.

The norm-based proximity measure $\delta(V): S_+^n \rightarrow R_+$ is given by

$$\delta(V) := \frac{1}{2} \|\varphi'(V)\| = \frac{1}{2} \sqrt{\sum_{i=1}^n \varphi'(\lambda_i(V))^2}. \quad (22)$$

The lower bound on $\delta(V)$ in terms of $\Psi(V)$ can be obtained from the following theorem, which is a reformulation of Theorem 4.8 in [15].

Theorem 2.4 Let $V \in S_{++}^n$. Then

$$\delta(V) \geq \frac{1}{2} \varphi'(\varrho(\Psi(V))).$$

Corollary 2.2 Let $V \in S_{++}^n$. Then

$$\delta(V) \geq \frac{\sqrt{2}}{2} (\Psi(V))^{\frac{1}{2}}.$$

Proof: We have

$$\begin{aligned} \delta(V) &\geq \frac{1}{2} \psi'(\varrho(\Psi(V))) \geq \frac{1}{2} \left(\sqrt{1+2\Psi(V)} - \frac{q^{\frac{1}{\sqrt{1+2\Psi(V)}}-1}}{1+2\Psi(V)} \right) \\ &\geq \frac{1}{2} (\sqrt{1+2\Psi(V)} - 1) = \frac{\Psi(V)}{1+\sqrt{1+2\Psi(V)}} \geq \frac{\sqrt{2}}{2} (\Psi(V))^{\frac{1}{2}}. \end{aligned}$$

This completes the proof.

3. Primal-Dual Kernel Function-Based (IPMs) for (SDO)

Without loss of generality, we assume that both the primal problem and its dual problem of (SDO) satisfy the interior-point condition (IPC), i.e., there exists (X^0, y^0, S^0) such that

$$A_i \cdot X^0 = b_i, i = 1, \dots, m, X^0 \succ 0, \sum_{i=1}^m y_i^0 A_i + S^0 = C^0, S^0 \succ 0. \quad (23)$$

The Karush-Kuhn-Tucker conditions for (P) and (D) are equivalent to the following system

$$A_i \cdot X = b_i, i = 1, \dots, m, X \succeq 0, \sum_{i=1}^m y_i A_i + S = C, S \succeq 0, XS = 0. \quad (24)$$

The standard approach is to replace the third equation in (24), i.e., the so-called complementarity condition for (P) and (D), by the parameterized equation $XS = \mu E$ with $\mu > 0$. This yields

$$A_i \cdot X = b_i, i = 1, \dots, m, X \succeq 0, \sum_{i=1}^m y_i A_i + S = C, S \succeq 0, XS = \mu E. \quad (25)$$

Under the assumption that (P) and (D) satisfy the (IPC), the system (25) has a unique solution, denoted by $(X(\mu), y(\mu), S(\mu))$. Let $X(\mu)$ be the μ -center of (P) and $(y(\mu), S(\mu))$ be the μ -center of (D). The set of μ -centers (with μ running through positive real numbers) gives a homotopy path, which is called the central path of (P) and (D).

If $\mu \rightarrow 0$, then the limit of the central path exists, and since the limit points satisfy the complementarity condition, i.e., $XS = 0$, it naturally yields an optimal solution for (P) and (D), see, e.g., [2].

In order to provide the scaled Newton system has a unique symmetric solution, Zhang [22] introduced the following symmetrization operator

$$H_p(M) := \frac{1}{2} (PMP^{-1} + (PMP^{-1})^T), \forall M \in R^{n \times n}. \quad (26)$$

One can easily verify that

$$H_p(M) = \mu E \Leftrightarrow M = \mu E, \quad (27)$$

for any nonsingular matrix P , any matrix M with real spectrum and any $\mu \in R$. For any given nonsingular matrix P , the system (25) is equivalent to

$$\begin{aligned} A_i \cdot X &= b_i, i = 1, \dots, m, X \succeq 0, \\ \sum_{i=1}^m y_i A_i + S &= C, S \succeq 0, \\ H_p(XS) &= \mu E. \end{aligned} \quad (28)$$

By using Newton method to the system (28), this yields

$$\begin{aligned} A_i \cdot \Delta X &= 0, i = 1, \dots, m, \\ \sum_{i=1}^m \Delta y_i A_i + \Delta S &= 0, \\ H_p(X\Delta S + \Delta XS) &= \mu E - H_p(XS). \end{aligned} \quad (29)$$

The search direction obtained through the system (29) is called the Monteiro-Zhang unified direction. Different choices of the matrix P result in different search directions (see, e.g., [2, 22]).

In this paper, we consider the so-called NT-symmetrization scheme [11, 21], which yields the NT search direction. Let

$$P := X^{\frac{1}{2}} (X^{\frac{1}{2}} S X^{\frac{1}{2}})^{-\frac{1}{2}} X^{\frac{1}{2}} = S^{-\frac{1}{2}} (S^{\frac{1}{2}} X S^{\frac{1}{2}})^{\frac{1}{2}} S^{-\frac{1}{2}}, \quad (30)$$

and $D := P^{\frac{1}{2}}$. The matrix D can be used to rescale X and S to the same matrix V , defined by

$$V := \frac{1}{\sqrt{\mu}} D^{-1} X D^{-1} = \frac{1}{\sqrt{\mu}} D S D. \quad (31)$$

From (31), after some elementary reductions, we have

$$\begin{aligned} \bar{A}_i \cdot D_X &= 0, i = 1, \dots, m, \\ \sum_{i=1}^m \Delta y_i \bar{A}_i + D_S &= 0, \\ D_X + D_S &= V^{-1} - V. \end{aligned} \quad (32)$$

Here

$$\bar{A}_i := \frac{1}{\sqrt{\mu}} D A_i D, i = 1, \dots, m,$$

and

$$D_X := \frac{1}{\sqrt{\mu}} D^{-1} \Delta X D^{-1}, D_S := \frac{1}{\sqrt{\mu}} D \Delta S D. \quad (33)$$

One can easily verify that

$$V^{-1} - V = -\nabla \Psi_c(V), \quad (34)$$

where $\nabla \Psi_c(V)$ denotes the gradient of $\Psi_c(V)$ is given by

$$\begin{aligned}\Psi_c(V) &:= \text{Tr}(\varphi_c(V)) = \sum_{i=1}^n (\varphi_c(\lambda_i(V))) \\ &= \sum_{i=1}^n \left(\frac{\lambda_i(V)^2 - 1}{2} - \log(\lambda_i(V)) \right).\end{aligned}\quad (35)$$

Hence, the system (29) is equivalent to

$$\begin{aligned}\overline{A}_i \cdot D_X &= 0, i = 1, \dots, m, \\ \sum_{i=1}^m \Delta y_i \overline{A}_i + D_S &= 0, \\ D_X + D_S &= -\nabla \Psi_c(V).\end{aligned}\quad (36)$$

This means that the logarithmic barrier function essentially determines the classical NT search direction.

In this paper, we replace the right-hand side $-\nabla \Psi_c(V)$ in the third equation in (36) by $-\nabla \Psi(V)$, i.e., $-\varphi'(V)$. This yields

$$\begin{aligned}\overline{A}_i \cdot D_X &= 0, i = 1, 2, \dots, m, \\ \sum_{i=1}^m \Delta y_i \overline{A}_i + D_S &= 0, \\ D_X + D_S &= -\nabla \Psi(V).\end{aligned}\quad (37)$$

The scaled new search direction $(D_X, \Delta y, D_S)$ is computed by solving the system (37) so that ΔX and ΔS are obtained through (33). If $(X, y, S) \neq (X(\mu), y(\mu), S(\mu))$, then $(\Delta X, \Delta y, \Delta S)$ is nonzero.

By taking a default step size along the search directions, we get the new iterate (X_+, y_+, S_+) according to

$$X_+ := X + \alpha \Delta X, \quad y_+ := y + \alpha \Delta y, \quad S_+ := S + \alpha \Delta S. \quad (38)$$

One can easily verify that

$$XS = \mu E \Leftrightarrow V = E \Leftrightarrow \varphi'(V) = 0 \Leftrightarrow \Psi(V) = 0. \quad (39)$$

Hence, the value of $\Psi(V)$ can be considered as a measure for the distance between the given iterate (X, y, S) and the μ -center $(X(\mu), y(\mu), S(\mu))$.

The generic form of primal-dual kernel function-based (IPMs) for (SDO) is shown in Algorithm 1.

Algorithm 1 Primal-Dual Interior-Point Algorithm for (SDO)

Input: a threshold parameter $\tau \geq 1$;
an accuracy parameter $\varepsilon > 0$;
a fixed barrier update parameter $\theta, 0 < \theta < 1$;
a strictly feasible pair (X^0, y^0, S^0) and $\mu^0 = 1$ such that $\Psi(X^0, S^0, \mu^0) \leq \tau$.
begin $X := X^0; y := y^0; S := S^0; \mu := \mu^0$;
while $n\mu \geq \varepsilon$
begin

$$\mu := (1 - \theta)\mu;$$

while $\Psi(X, S, \mu) > \tau$ do

begin

computer the search directions $(\Delta X, \Delta y, \Delta S)$;

choose a suitable step size α ;

update $(X, y, S) := (X, y, S) + \alpha(\Delta X, \Delta y, \Delta S)$.

end

end

end

4. Complexity Analysis of Large-Update Methods

In each inner iteration the search direction $(\Delta X, \Delta y, \Delta S)$ is obtained by solving the system (37) and via (33). After a step with size α the new iterate is given by

$$X_+ = X + \alpha \Delta X, \quad y_+ = y + \alpha \Delta y, \quad S_+ = S + \alpha \Delta S. \quad (40)$$

Then, we have

$$X_+ = X + \alpha \Delta X = X + \alpha \sqrt{\mu} D D_X D^{-1} = \sqrt{\mu} D (V + \alpha D_X) D^{-1} \quad (41)$$

and

$$S_+ = S + \alpha \Delta S = S + \alpha \sqrt{\mu} D^{-1} D_S D^{-1} = \sqrt{\mu} D^{-1} (V + \alpha D_S) D^{-1}. \quad (42)$$

It follows from (31) that

$$V_+ = \frac{1}{\sqrt{\mu}} D^{-1} X_+ S_+ D^{\frac{1}{2}}. \quad (43)$$

One can easily verify that V_+^2 is unitarily similar to

the matrix $X_+^{\frac{1}{2}} S_+ X_+^{\frac{1}{2}}$ and thus to

$$(V + \alpha D_X)^{\frac{1}{2}} (V + \alpha D_S) (V + \alpha D_X)^{\frac{1}{2}}. \quad (44)$$

This implies that the eigenvalues of V_+ are precisely the same as those of the matrix

$$\bar{V}_+ := \left((V + \alpha D_X)^{\frac{1}{2}} (V + \alpha D_S) (V + \alpha D_X)^{\frac{1}{2}} \right)^{\frac{1}{2}}. \quad (45)$$

From the definition of $\Psi(V)$, one obtains $\Psi(V_+) = \Psi(\bar{V}_+)$. Hence, by Theorem 2.2, we have

$$\Psi(V_+) = \Psi(\bar{V}_+) \leq \frac{1}{2} (\Psi(V + \alpha D_X) + \Psi(V + \alpha D_S)). \quad (46)$$

Now, we consider the decrease in $\Psi(V)$ as a function of α and define

$$f(\alpha) := \Psi(V_+) - \Psi(V) = \Psi(\bar{V}_+) - \Psi(V). \quad (47)$$

Let define

$$f_1(\alpha) := \frac{1}{2}(\Psi(V + \alpha D_X) + \Psi(V + \alpha D_S)) - \Psi(V). \quad (48)$$

It follows that $f(\alpha) \leq f_1(\alpha)$ and $f(0) = f_1(0) = 0$.

By Theorem 2.1, one has

$$f_1'(\alpha) = \frac{1}{2} \text{Tr}(\psi'(V + \alpha D_X) D_X + \psi'(V + \alpha D_S) D_S) \quad (49)$$

and

$$\begin{aligned} f_1''(\alpha) &= \frac{1}{2} \frac{d^2}{d\alpha^2} \text{Tr}(\phi(V + \alpha D_X) + \phi(V + \alpha D_S)) \\ &\leq \frac{1}{2}(\omega_1 \|D_X\|^2 + \omega_2 \|D_S\|^2), \end{aligned} \quad (50)$$

where

$$\omega_1 = \max\{|\Delta\phi'(\lambda_j(V + \alpha D_X), \lambda_k(V + \alpha D_X))| : j, k = 1, 2, \dots, n\}, \quad (51)$$

and

$$\omega_2 = \max\{|\Delta\phi'(\lambda_j(V + \alpha D_S), \lambda_k(V + \alpha D_S))| : j, k = 1, 2, \dots, n\}. \quad (52)$$

Hence, using the third equation of the system (37), one has

$$f_1'(0) = \frac{1}{2} \text{Tr}(\phi'(V)(D_X + D_S)) = \frac{1}{2} \text{Tr}(-\phi'(V)^2) = -2\delta(V)^2. \quad (53)$$

In order to facilitate discussion, we denote $\delta := \delta(V)$, and we have the following result [8].

Theorem 3.1 One has $f_1''(\alpha) \leq 2\delta^2 \phi''(\lambda_{\min}(V) - 2\alpha\delta)$.

The default step size for the algorithm should be chosen such that X_+ and S_+ are feasible and $\Psi(V_+) - \Psi(V)$ decreases sufficiently. For the details we leave it for the interested readers (see, e.g., [8,15]). Following the strategy considered in [8], we briefly recall how to choose the default step size. Suppose that the step size α satisfies

$$-\phi'(\lambda_{\min}(V) - 2\alpha\delta) + \phi'(\lambda_{\min}(V)) \leq 2\delta. \quad (54)$$

Then $f_1(\alpha) \leq 0$. The largest possible value of the step size of α satisfying (53) is given by

$$\bar{\alpha} := \frac{1}{2\delta}(\rho(\delta) - \rho(2\delta)). \quad (55)$$

Furthermore, we can conclude that

$$\frac{1}{\phi''(\rho(2\delta))} \leq \bar{\alpha} \leq \frac{1}{\phi''(\rho(\delta))}. \quad (56)$$

After some elementary reductions, we have

$$\bar{\alpha} \geq \frac{1}{1 + (1 + 4\delta)(2 + \log q) \left(1 + \frac{\log(1 + 4\delta)}{\log q}\right)^2}. \quad (57)$$

In the sequel, let

$$\tilde{\alpha} := \frac{1}{1 + (1 + 4\delta)(2 + \log q) \left(1 + \frac{\log(1 + 4\delta)}{\log q}\right)^2} \quad (58)$$

be the default step size.

As a consequence of Lemma A.1 and the fact that $f(\alpha) \leq f_1(\alpha)$, which is a twice differentiable convex function with $f_1(0) = 0$, and $f_1'(0) = -2\delta^2 < 0$, the following lemma is obtained.

Lemma 4.1 Let the step size α be such that $\alpha \leq \tilde{\alpha}$. Then

$$f(\alpha) \leq -\alpha\delta^2.$$

The following theorem shows that the default step size yields sufficient decrease of the barrier function during each inner iteration.

Theorem 4.2 Let $\Psi_0 \geq \tau \geq 3$ and $\tilde{\alpha}$ be the default step size as given by (57). Then

$$f(\tilde{\alpha}) \leq -\frac{1}{20(2 + \log q) \left(1 + \frac{\log(1 + \sqrt{\Psi_0})}{\log q}\right)^2} \Psi(V)^{\frac{1}{2}}.$$

Proof: From Lemma 4.1 with (56) and Corollary 2.2, we have

$$\begin{aligned} f(\tilde{\alpha}) &\leq -\tilde{\alpha}\delta^2 \leq -\frac{\delta^2}{1 + (1 + 4\delta)(2 + \log q) \left(1 + \frac{\log(1 + 4\delta)}{\log q}\right)^2} \\ &\leq -\frac{\delta^2}{10\delta(2 + \log q) \left(1 + \frac{\log(1 + 4\delta)}{\log q}\right)^2} \\ &\leq -\frac{\delta^2}{10(2 + \log q) \left(1 + \frac{\log(1 + 4\delta)}{\log q}\right)^2} \\ &\leq -\frac{1}{10\sqrt{2}(2 + \log q)\delta \left(1 + \frac{\log(1 + \sqrt{\Psi_0})}{\log q}\right)^2} \Psi(V)^{\frac{1}{2}}. \end{aligned}$$

This completes the proof.

At the start of an outer iteration and just before updating the parameter μ , one has $\Psi(V) \leq \tau$.

It follows that the value of $\Psi(V)$ exceeds from the threshold τ after updating of μ . Therefore, one need to count how many inner iterations are required to return to the situation where $\Psi(V) \leq \tau$. Let denote the value of $\Psi(V)$ after the μ -update be Ψ_0 , the subsequent values in the same outer

iteration are denoted as Ψ_k , $k=1, \dots, K$, where K denotes the total number of inner iterations in the outer iteration.

Since $\varphi(t) \leq \frac{t^2-1}{2}$ for $t \geq 1$, we have

$$\begin{aligned} \Psi_0 &\leq n\Psi \left(\frac{\rho\left(\frac{\tau}{n}\right)}{\sqrt{1-\theta}} \right) \leq n\Psi \left(\frac{1+\sqrt{\frac{2\tau}{n}}}{\sqrt{1-\theta}} \right) \\ &\leq \frac{n\left(\sqrt{\frac{2\tau}{n}} + \frac{2\tau}{n}\right)}{2(1-\theta)} = \frac{\tau + \sqrt{2n\tau}}{1-\theta} = O(n). \end{aligned} \quad (59)$$

According to the decrease of $f(\tilde{\alpha})$ in Lemma 4.2, we have

$$\Psi_{k+1} \leq \Psi_k - \beta(\Psi_k)^{1-\gamma}, k=0, 1, \dots, K-1, \quad (60)$$

where

$$\beta = \frac{1}{10\sqrt{2}(2+\log q) \left(1 + \frac{\log(1+\sqrt{\Psi_0})}{\log q}\right)^2}, \text{ and } \gamma = \frac{1}{2}. \quad (61)$$

The following lemma provides an estimate for the number of inner iterations between two successive barrier parameter updates.

Lemma 4.2 One has

$$K \leq 20\sqrt{2}(2+\log q) \left(1 + \frac{\log(1+\sqrt{\Psi_0})}{\log q}\right)^2 (\Psi_0)^{\frac{1}{2}}.$$

Proof: From Lemma A.2 and (59), the result of the lemma follows. This completes the proof.

It is well known that an upper bound of the number of outer iterations is bounded above by [23]

$$\left\lceil \frac{1}{\theta} \log \frac{n}{\varepsilon} \right\rceil. \quad (62)$$

By multiplying the number of outer iterations and the number of inner iterations, we get an upper bound for the total number of iterations, namely,

$$\frac{20\sqrt{2}(2+\log q)}{\theta} \left(1 + \frac{\log(1+\sqrt{\frac{\tau+\sqrt{2n\tau}}{1-\theta}})}{\log q}\right)^2 \left(\frac{\tau+\sqrt{2n\tau}}{1-\theta}\right)^{\frac{1}{2}} \log \frac{n}{\varepsilon}. \quad (63)$$

Note that $\Psi_0 \leq O(n)$. By choosing $q = 1 + O(n)$, the best total iteration bound is obtained.

Theorem 4.3 For large-update methods, we set $\theta = \Theta(1)$, and $\tau = O(n)$. Then the iteration bound becomes

$$O\left(\sqrt{n} \log n \log \frac{n}{\varepsilon}\right),$$

which matches the currently best well-known complexity for large-update methods.

5. Conclusion

In this paper, we have investigated a class of large-update primal-dual (IPMs) for (LO) based on a parametric kernel function presented in [16] can be extended to the context of (SDO). Furthermore, the best result of iteration bounds for large-update methods is derived. In our future study, the generalizations of the primal-dual (IPMs) for (LO) to symmetric cone optimization (SCO) and symmetric cone complementarity problems (SCCP) are interesting.

Acknowledgements

The authors would like to thank the editor and the anonymous referees for their useful comments and suggestions, which helped to improve the presentation of this paper. This work was supported by University Students' Innovative Training Program of Shanghai (No. CS1621001).

Appendix(Some Technical Lemmas)

Lemma A.1 (Lemma 12 in [21]) Let $h(t)$ be a twice differentiable convex function with $h(0)=0$, $h'(0)<0$ and let $h(t)$ attain its (global) minimum at $t^*>0$. If $h''(t)$ is increasing for $t \in [0, t^*]$, then

$$h(t) \leq \frac{th'(0)}{2}, \quad 0 \leq t \leq t^*.$$

Lemma A.2 (Lemma 14 in [21]) Suppose t_0, t_1, \dots, t_K is a sequence of positive numbers such that

$$t_{k+1} \leq t_k - \beta t_k^{1-\gamma}, \quad k=0, 1, \dots, K-1,$$

where $\beta > 0$ and $0 < \gamma \leq 1$. Then

$$K \leq \left\lceil \frac{t_0^\gamma}{\beta\gamma} \right\rceil.$$

References

- [1] Anjos M. F., Lasserre, J. B.: Handbook on Semidefinite, Conic and Polynomial Optimization: Theory, Algorithms, Software and Applications. International Series in Operational Research and Management Science. Volume 166, Springer, New York, USA (2012)
- [2] De Klerk, E.: Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications. Kluwer Academic Publishers, Dordrecht, The Netherlands (2002)

- [3] Cai X. Z., Wang G. Q., Zhang Z. H.: Complexity analysis and numerical implementation of primal-dual interior-point methods for convex quadratic optimization based on a finite barrier. *Numer. Algorithms* 62(2), 289-306 (2013)
- [4] Cho G. M.: Large-update primal-dual interior-point algorithm for semidefinite optimization. *Pac. J. Optim.* 11(1), 29-36 (2015)
- [5] El Ghami M., Bai Y. Q., Roos C.: Kernel-function based Algorithms for semidefinite optimization. *RAIRO Oper. Res.* 43(2), 189-199 (2009)
- [6] Lee Y. H., Cho Y. Y., Jin J. H., Cho G. M.: Interior-point algorithms for LO and SDO based on a new class of kernel functions. *J. Nonlinear Convex Anal.* 13(3), 555-573 (2012)
- [7] Liu, H. W., Liu, C. H., Yang X. M.: New complexity analysis of a Mehrotra-type predictor-corrector algorithm for semidefinite programming. *Optim. Methods Softw.* 28(6), 1179-1194 (2013)
- [8] Wang G. Q., Bai Y. Q.: A new primal-dual path-following interior-point algorithm for semidefinite optimization. *J. Math. Anal. Appl.* 353(1), 339-349 (2009)
- [9] Wang G. Q., Bai Y. Q.: Primal-dual interior-point algorithms for convex quadratic semidefinite optimization. *Nonlinear Anal.* 71(7-8), 3389-3402 (2009)
- [10] Wang G. Q., Bai Y. Q., Gao X. Y., Wang D. Z.: Improved complexity analysis of full Nesterov-Todd step interior-point methods for semidefinite optimization. *J. Optim. Theory Appl.* 165(1), 242-262 (2015)
- [11] Wang G. Q., Bai Y. Q., Roos C.: Primal-dual interior-point algorithms for semidefinite optimization based on a simple kernel function. *J. Math. Model. Algorithms* 4(4), 409-433 (2005)
- [12] Wang G. Q., Zhu D. T.: A unified kernel function approach to primal-dual interior-point algorithms for convex quadratic SDO. *Numer. Algorithms* 57(4), 537-558 (2011)
- [13] Yang X. M., Liu H. W., Zhang Y. K.: A second-order Mehrotra-type predictor-corrector algorithm with a new wide neighbourhood for semidefinite programming. *Int. J. Comput. Math.* 91(5), 1082-1096 (2014)
- [14] Zhang M. W.: A large-update interior-point algorithm for convex quadratic semidefinite optimization based on a new kernel function. *Acta Math. Sin. (Engl. Ser.)* 28(11), 2313-2328 (2012)
- [15] Bai, Y. Q., El Ghami, M., Roos, C.: A comparative study of kernel functions for primal-dual interior-point algorithms in linear optimization. *SIAM J. Optim.* 15(1), 101-128 (2004)
- [16] Achache M.: A new parameterized kernel function for LO yielding the best known iteration bound for a large-update interior point algorithm. *Afrika Mat.* 27(3), 591-601 (2015)
- [17] Bai, Y. Q., El Ghami, M., Roos, C.: A new efficient large-dual interior-point method based on a finite barrier. *SIAM J. Optim.* 13(3), 766-782 (2003)
- [18] Cai X. Z., Wang G. Q., El Ghami M., Yue Y. J.: Complexity analysis of primal-dual interior-point methods for linear optimization based on a parametric kernel function with a trigonometric barrier term. *Abstr. Appl. Anal.* 2014, 710158 (2014)
- [19] Wang G. Q., Bai Y. Q.: Interior-Point Methods for Symmetric Cone Complementarity Problems: Theoretical Analysis and Algorithm Implementation. Harbin Institute of Technology Press, Harbin (2014)
- [20] Horn, R. A., Johnson, C. R.: Topics in Matrix Analysis. Cambridge University Press, UK (1991)
- [21] Peng, J., Roos, C., Terlaky, T.: Self-regular functions and new search directions for linear and semidefinite optimization. *Math. Program.* 93(1), 129-171 (2002)
- [22] Zhang, Y.: On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming. *SIAM J. Optim.* 8(2), 365-386 (1998)
- [23] Roos, C., Terlaky, T., Vial, J.-Ph.: Theory and Algorithms for Linear Optimization. Springer, Chichester, UK (1st Edition, Theory and Algorithms for Linear Optimization. An Interior-Point Approach. John Wiley & Sons, 1997) (2005)