PRIMAL-DUAL INTERIOR-POINT ALGORITHM FOR LO BASED ON A NEW KERNEL FUNCTION

Xin Li
Mingwang Zhang[1]
Ping Ji

College of Science
China Three Gorges University
Yi Chang – 443002
P.R. China
emails: sxlixin12@126.com
zmwang@ctgu.edu.cn
haojpkkll@163.com

Abstract. Based on a new kernel function, a large-update primal-dual interior-point algorithm for solving linear optimization is proposed. The kernel function is used both for determining the search directions and for measuring the distance between the given iterate and the μ-center for the algorithm. By using several new technical lemmas, the iteration complexity bound as $O(\sqrt{n}\log n \log \frac{n}{\varepsilon})$ is obtained, which coincides with the currently best iteration complexity bounds for large-update methods. In addition, we present some preliminary numerical results.

Keywords: linear optimization; kernel function; primal-dual interior-point algorithm; large-update methods; iteration complexity bound.

1. Introduction

In this paper, we consider the standard form of the linear optimization (LO) problem

$$(P) \qquad min\ \{c^T x : Ax = b, \quad x \geq 0\},$$

where $A \in \mathbb{R}^{m\times n}$ with $rank(A) = m \leq n,\ x, c \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$. The dual problem of (P) is given by

$$(D) \qquad \max\ \{b^T y : A^T y + s = c, \quad s \geq 0\},$$

where $y \in \mathbb{R}^m$ and $s \in \mathbb{R}^n$.

After the landmark paper of Karmarkar [9], interior point methods (IPMs) for solving LO attract much attention, since IPMs not only are the most effective methods in practice but also have polynomial complexity. The primal-dual IPMs for LO were first introduced by Kojima et al. in [11]. Lately, a large amount of

[1]Corresponding author.

beautiful results have been reported. For a survey, we refer to recent books on the subject [18], [22], [23].

In the IPMs, the choice of the barrier function is crucial not only for the quality and elegance of the analysis but also for the performance of the algorithm. Most IPMs for LO were based on the classical logarithmic barrier function. However, today there is still a gap between the practical behavior of these algorithms and the theoretical performance results. Especially, the so-called large-update methods for which the theoretical iteration complexity bound is $O(n\log\frac{n}{\varepsilon})$. In practice, the so-called large-update methods are much more efficient than the so-called small-update methods for which the theoretical iteration complexity bound is only $O(\sqrt{n}\log\frac{n}{\varepsilon})$. This significant gap between theory and practice has been referred to as the irony of IPMs [3].

Recently, a new variant of primal-dual IPM for LO and semi-definite optimization (SDO) based on the so-called self-regular kernel functions was proposed by Peng et al. in [15] and aforementioned gap was narrowed. They obtained the currently best iteration complexity bound, namely, $O(\sqrt{n}\log n\log\frac{n}{\varepsilon})$, for large-update IPMs for LO and SDO. Subsequently, Bai et al. [3] proposed a class of primal-dual IPMs for LO based on an eligible kernel function (neither logarithmic nor self-regular function). The algorithm derived the same favourable iteration complexity bound with large-update strategy as [15]. Kim et al. [10] proposed a new adaptive single-step primal-dual IPM with wide neighborhood for LO based on a simple kernel function, and showed that the algorithm has $O(q\sqrt{n\tau}\log\frac{n}{\varepsilon})$ iteration complexity bound for large-update methods. Later, Liu et al. [13] introduced a new kind of kernel function, and concluded that in some situations the iteration complexity bound as $O(m^{\frac{3m+1}{2m}}n^{\frac{m+1}{2m}}\log\frac{n}{\varepsilon})$. Very recently, Peyghami et al. [17] proposed a primal-dual IPM based on a trigonometric kernel function for LO, and obtained the iteration complexity bound as $O(n^{\frac{2}{3}}\log\frac{n}{\varepsilon})$ for large-update methods. Cai et. al [5] introduced a new parametric kernel function, which is a combination of the classic kernel function and a trigonometric barrier term, and obtained the iteration bounds for large-update methods the same as [17]. Li et al. [12] presented a new primal-dual interior-point algorithm for linear optimization based on a trigonometric kernel function, and derived the worst case complexity for a large-update primal-dual interior-point method based on the kernel function. For more studies with primal-dual IPMs based on kernel functions please refer to [1], [2], [4], [6], [7], [8], [16], [19], [21], [24].

Motivated by their works, we present a large-update primal-dual IPM for LO based on a new kernel function and the corresponding barrier function. Since our kernel function is exponentially convex and is not self-regular and logarithmic barrier function, which is different from the analysis provided to self-regular and logarithmic barrier kernel function. By using several technical lemmas, we derive that the algorithm has favorable iteration complexity iteration bound, namely, $O(\sqrt{n}\log n\log\frac{n}{\varepsilon})$, which is currently the best known iteration complexity bound for large-update IPMs. Furthermore, some preliminary numerical results are presented to show that the algorithm is efficient and reliable.

The paper is organized as follows. In Section 2, after recalling some basic concepts, we study the central path and the new search directions for LO. The generic primal-dual interior-point algorithm for LO is also presented. In Section 3, we define a new kernel function and the corresponding barrier function and present some properties of kernel function and the barrier function. The amount of decrease of the barrier function during an inner iteration and the default step size are given in Section 4. The complexity analysis of primal-dual IPMs based on our kernel function is given in Section 5. In Section 6, we give some preliminary numerical results. Finally, some concluding remarks are given in Section 7.

We use the following notational conventions.

Throughout the paper, $\| x \| = \sqrt{x^T x}$ denotes the 2-norm of the vector x. The nonnegative orthant and positive orthant are denoted as $\mathbb{R}^n_+$ and $\mathbb{R}^n_{++}$, respectively. If $x, s \in \mathbb{R}^n$, then xs denotes the coordinatewise (or Hadamard) product of the vectors x and s and e denotes the all-one vector of length n, i.e., $e = (1, 1, ..., 1)^T$. The diagonal matrix with the vector x is denoted $X = diag(x)$. Furthermore, $x^T y = \sum_{i=1}^{n} x_i y_i$ for $x, y \in \mathbb{R}^n$. For any $x = (x_1, x_2, ..., x_n)^T \in \mathbb{R}^n$, $x_{\min}$ denotes the smallest and $x_{\max}$ denotes the largest value of the components of x. Finally, if $f(x) \geq 0$ is a real valued function of a real nonnegative variable, the notation $f(x) = O(x)$ means that $f(x) \leq cx$ for some positive constant 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

Without loss of generality, we may assume that both (P) and (D) satisfy the interior-point condition (IPC) [11], i.e., there exist x_0 and (y_0, s_0) such that

$$\begin{aligned} Ax_0 = b, &\quad x_0 > 0, \\ A^T y_0 + s_0 = c, &\quad s_0 > 0. \end{aligned}$$

The optimality condition for (P) and (D) is equivalent to solving the following nonlinear system

$$\begin{aligned} Ax = b, &\quad x \geq 0, \\ A^T y + s = c, &\quad s \geq 0, \\ xs = 0. & \end{aligned} \tag{1}$$

The basic idea of primal-dual IPMs is to replace the third equation in system (1) by a parametric equation $xs = \mu e$, where $\mu > 0$. This leads the system (1) to the following parametric system

$$\begin{aligned} Ax = b, &\quad x \geq 0, \\ A^T y + s = c, &\quad s \geq 0, \\ xs = \mu e. & \end{aligned} \tag{2}$$

Surprisingly enough, if the IPC is satisfied, the parameterized system (2) has a unique solution $(x(\mu), y(\mu), s(\mu))$ for each $\mu > 0$, which is called a μ-center of (P)

and (D). The set of μ-centers is said to be the central path of (P) and (D). The relevance of the central path for LO was recognized first by Sonnevend in [20] and Megiddo in [14]. If $\mu \to 0$, then the limit of the central path exists. Since the limit point satisfies the complementarity condition, the limit point yields optimal solutions for (P) and (D).

For fixed $\mu > 0$, by applying Newton's method to the parameterized system (2), we obtain the search direction $(\triangle x, \triangle y, \triangle s)$ from the following Newton system

$$\begin{aligned} A\Delta x &= 0, \\ A^T \Delta y + \Delta s &= 0, \\ s\Delta x + x\Delta s &= \mu e - xs. \end{aligned} \tag{3}$$

Since A has full row rank, the system (3) uniquely defines $(\triangle x, \triangle y, \triangle s)$ for any $x > 0$ and $s > 0$. By taking a step along the search direction $(\triangle x, \triangle y, \triangle s)$, we can construct a new tripe (x_+, y_+, s_+) according to

$$x_+ := x + \alpha \triangle x, \qquad y_+ := y + \alpha \triangle y, \qquad s_+ := s + \alpha \triangle s, \tag{4}$$

where $\alpha \in (0, 1]$ is obtained by using some rules so that the new iterate satisfies $(x_+, y_+, s_+) > 0$.

Introducing the notations

$$v := \sqrt{\frac{xs}{\mu}}, \quad d_x := \frac{v\triangle x}{x}, \quad d_s := \frac{v\triangle s}{s}, \tag{5}$$

then the Newton system (3) can be rewritten as

$$\begin{aligned} \overline{A} d_x &= 0, \\ \overline{A}^T \triangle y + d_s &= 0, \\ d_x + d_s &= v^{-1} - v, \end{aligned} \tag{6}$$

where $\overline{A} := \frac{1}{\mu} A V^{-1} X, \quad V := diag(v), \quad X := diag(x), \quad S := diag(s).$

Note that d_x and d_s are orthogonal vectors, since d_x belongs to the null space and d_s to the row space of the matrix $\bar{A}$. Hence, we will have

$$d_x = d_s = 0 \Leftrightarrow v - v^{-1} = 0 \Leftrightarrow x = x(\mu), s = s(\mu).$$

It is surprising that the right-side of the third equation in (6) is called the scaled centering equation, which equals the minus gradient of the barrier function

$$\Psi_c(v) := \sum_{i=1}^{n} \left(\frac{v_i^2 - 1}{2} - \log v_i \right).$$

In this paper, we replace $\Psi_c(v)$ by a strictly convex function $\Psi(v)$, $v \in \mathbb{R}^n_{++}$ such that $\Psi(v)$ is minimal at $v = e$ and $\Psi(e) = 0$. Thus the new scaled centering equation becomes

$$d_x + d_s = -\nabla \Psi(v).$$

As before, we will have

$$d_x = d_s = 0 \Leftrightarrow \nabla\Psi(v) = 0 \Leftrightarrow v = e \Leftrightarrow x = x(\mu), s = s(\mu).$$

To simplify matters, we restrict ourselves to the case where $\Psi(v)$ is separable with identical coordinate functions. Thus, letting $\Psi(v)$ denote the function on the coordinates, we write

$$\Psi(v) := \sum_{i=1}^{n} \psi(v_i),$$

where $\psi(t) : D \to R_+^n$ with $R_{++}^n \subseteq D$, is called the kernel function of barrier function $\Psi(v)$ if $\psi(t)$ is twice differentiable and satisfies the following condition

$$\psi(1) = \psi'(1) = 0, \quad \psi''(t) > 0, \quad \lim_{t\to 0^+} \psi(t) = \lim_{t\to\infty} \psi(t) = \infty. \tag{7}$$

Now the new search directions d_x, $\triangle y$, d_s are obtained by solving the following system so that $(\triangle x, \triangle y, \triangle s)$ is computed via (5)

$$\begin{aligned} \overline{A} d_x &= 0, \\ \overline{A}^T \triangle y + d_s &= 0, \\ d_x + d_s &= -\nabla\Psi(v). \end{aligned} \tag{8}$$

Thus, the generic primal-dual interior-point algorithm works as follows.

Algorithm 1: Generic Primal-Dual Algorithm for LO

Input:
  A threshold parameter $\tau \geq 1$;
  a barrier function $\Psi(v)$;
  a fixed barrier update parameter θ, $0 < \theta < 1$;
  an accuracy parameter $\varepsilon > 0$;
begin
  $x := x^0$; $s := s^0$; $\mu := \mu^0$;
  while $n\mu > \varepsilon$ do
  begin
    $\mu := (1-\theta)\mu$;
   while $\Psi(v) > \tau$ do
    begin
     $x := x + \alpha\triangle x$; $y := y + \alpha\triangle y$; $s := s + \alpha\triangle s$;
     $v := \sqrt{\frac{xs}{\mu}}$;
    end
  end
end

Remark 1 The choice of the barrier update parameter θ plays an important role in both theory and practice of IPMs. Usually, if θ is a constant which is independent of the dimension n of the problem, such as $\theta = \frac{1}{2}$, then the algorithm is called a large-update (or long-step) method. If θ depends on the dimension of the problem, for instance $\theta = \frac{1}{\sqrt{n}}$, then the algorithm is called a small-update (or short-step) method.

Remark 2 The choice of the step size α ($0 < \alpha \leq 1$) is another crucial issue in the analysis of the algorithm. It has to be taken such that the closeness of the iterates to the current μ-center improves by a sufficient amount. In the theoretical analysis the step size α is usually given a value that depends on the closeness of the current iterates to the μ-center.

3. The new kernel (barrier) function and its properties

This section is devoted to present a new kernel (barrier) function and give its properties, which are used in the complexity analysis of the algorithm. Let's define a univariate function $\psi(t) : D \to R^n_+$ with $R^n_{++} \subseteq D$ as follows

$$\psi(t) = \frac{t^2 - 1}{2} - \int_1^t \frac{1}{x^{2p}} e^{\frac{p}{x} - p} dx, \qquad p \geq 1. \tag{9}$$

The first three derivatives of the function $\psi(t)$ are given by

$$\psi'(t) = t - \frac{1}{t^{2p}} e^{\frac{p}{t} - p}, \tag{10}$$

$$\psi''(t) = 1 + \frac{2p}{t^{2p+1}} e^{\frac{p}{t} - p} + \frac{p}{t^{2p+2}} e^{\frac{p}{t} - p}, \tag{11}$$

$$\psi'''(t) = -\frac{2p(2p+1)}{t^{2p+2}} e^{\frac{p}{t} - p} - \frac{4p^2 + 2p}{t^{2p+3}} e^{\frac{p}{t} - p} - \frac{p^2}{t^{2p+4}} e^{\frac{p}{t} - p}. \tag{12}$$

Using (9)-(11), we can easily get

$$\psi(1) = \psi'(1) = 0, \psi''(t) > 1, \lim_{t \to 0^+} \psi(t) = \lim_{t \to \infty} \psi(t) = \infty,$$

which show that $\psi(t)$ is indeed a kernel function. According to the definition of self-regular function in [15], we can easily verify that the kernel function $\psi(t)$ does not belong to the family of self-regular functions. Due to conditions $\psi(1) = \psi'(1) = 0$, one can completely describe $\psi(t)$ by its second derivative

$$\psi(t) = \int_1^t \int_1^\xi \psi''(\zeta) d\zeta d\xi. \tag{13}$$

The next lemma serves to prove that the new kernel function $\psi(t)$ is an eligible kernel function according to [3].

Lemma 1 *For the function $\psi(t)$, defined by (9), we have*

$$\text{(a)} \quad \psi''(t) > 1, \quad \forall t > 0, \tag{14}$$

$$\text{(b)} \quad t\psi''(t) + \psi'(t) > 0, \quad \forall t > 0, \tag{15}$$

$$\text{(c)} \quad t\psi''(t) - \psi'(t) > 0, \quad \forall t > 0, \tag{16}$$

$$\text{(d)} \quad \psi'''(t) < 0, \quad \forall t > 0. \tag{17}$$

Proof. Inequalities (a) and (d) immediately follows from (11) and (12), respectively. Next, we prove that (b) holds. Using (10), (11) and $p \geq 1$, we get

$$\begin{aligned} t\psi''(t) + \psi'(t) &= t + \frac{2p}{t^{2p}} e^{\frac{p}{t}-p} + \frac{p}{t^{2p+1}} e^{\frac{p}{t}-p} + t - \frac{1}{t^{2p}} e^{\frac{p}{t}-p} \\ &= 2t + \frac{2p-1}{t^{2p}} e^{\frac{p}{t}-p} + \frac{p}{t^{2p+1}} e^{\frac{p}{t}-p} > 0. \end{aligned}$$

Furthermore, for proving (c), we have

$$\begin{aligned} t\psi''(t) - \psi'(t) &= t + \frac{2p}{t^{2p}} e^{\frac{p}{t}-p} + \frac{p}{t^{2p+1}} e^{\frac{p}{t}-p} - t + \frac{1}{t^{2p}} e^{\frac{p}{t}-p} \\ &= \frac{2p+1}{t^{2p}} e^{\frac{p}{t}-p} + \frac{p}{t^{2p+1}} e^{\frac{p}{t}-p} > 0. \end{aligned}$$

This completes the proof. ■

Lemma 2 (Lemma 2.1.2 in [15]) *$\psi(t)$ is exponentially convex (e-convex), namely,*

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

Proof. This result follows easily by using Lemma 2.1 in [3], which states that the above inequality holds if and only if $t\psi''(t) + \psi'(t) \geq 0$ for all $t > 0$. From (15) the result holds. ■

In the analysis of the algorithm, we introduce the norm-based proximity measure $\delta(v)$ defined by

$$\delta(v) := \frac{1}{2} \parallel \nabla\Psi(v) \parallel = \frac{1}{2}\sqrt{\sum_{i=1}^{n} (\psi'(v_i))^2}, \qquad v \in \mathbb{R}^n_{++}. \tag{18}$$

Since $\psi(t)$ is an eligible kernel function, we can get the following lemma which will be used in next section. Its proof is similar to the proof of Lemma 2.1, Corollary 2.2 and 2.3 in [3], and is therefore stated without proof.

Lemma 3 *Let $\psi(t)$ be as defined in* (9)*, we have*

(a) $\frac{1}{2}(t-1)^2 \leq \psi(t) \leq \frac{1}{2}\psi'(t)^2, \quad if \;\; t > 0,$

(b) $\Psi(v) \leq 2\delta(v)^2,$

(c) $\parallel v \parallel \leq \sqrt{n} + \sqrt{2\Psi(v)} \leq \sqrt{n} + 2\delta(v).$

Corollary 1 *If $\Psi(v) \geq 1$, then we have*

$$\delta(v) \geq \frac{1}{\sqrt{2}}. \tag{19}$$

Proof. It is easily followed from the (b) of Lemma 3. ■

4. Analysis of the algorithm

In this section, we deal with the growth behavior of the barrier function. A default value for the step size is also presented in this section.

4.1. Growth behavior for the barrier function

For given threshold parameter τ, we have $\Psi(v) \leq \tau$ before the update of μ with the factor $1-\theta$. According to the update of μ, the vector v is divided by a factor $\sqrt{1-\theta}$, which leads to an increase of the value $\Psi(v)$ in general. The subsequent inner iterations are performed in order to bring the value of $\Psi(v)$ decrease until it passes the threshold parameter τ again. Therefore, the largest values of $\Psi(v)$ occur just after the update of μ. In order to investigate the growth behavior of $\Psi(v)$, we need the following lemmas which are important in deriving the iteration complexity bound.

Lemma 4 *Let $\beta \geq 1$. Then*

$$\psi(\beta t) \leq \psi(t) + \frac{t^2}{2}(\beta^2 - 1). \tag{20}$$

Proof. Defining

$$\psi_b(t) := -\int_1^t \frac{1}{x^{2p}} e^{\frac{p}{x}-p} dx,$$

we have

$$\psi_b'(t) = -\frac{1}{t^{2p}} e^{\frac{p}{t}-p} < 0,$$

i.e., $\psi_b(t)$ is a decreasing function with $t > 0$. Thus $\psi_b(\beta t) \leq \psi_b(t)$ for $\beta \geq 1$. So

$$\psi(\beta t) - \psi(t) = \frac{1}{2}(\beta^2 - 1)t^2 + \psi_b(\beta t) - \psi_b(t) \leq \frac{1}{2}(\beta^2 - 1)t^2.$$

That implies the lemma. ∎

Lemma 5 *Assume that $0 < \theta < 1$ and $v_+ := \frac{v}{\sqrt{1-\theta}}$, then*

$$\Psi(v_+) \leq \Psi(v) + \frac{\theta}{2(1-\theta)}(2\Psi(v) + 2\sqrt{2n\Psi(v)} + n).$$

Proof. We can conclude that the above relations by using Lemma 4, with $\beta = \frac{1}{\sqrt{1-\theta}}$, and Lemma 3. ∎

By the assumption $\Psi(v) \leq \tau$ just before the update of μ, we have

$$\Psi\left(\frac{v}{\sqrt{1-\theta}}\right) \leq \tau + \frac{\theta}{2(1-\theta)}(2\tau + 2\sqrt{2n\tau} + n).$$

We define

$$L(n, \theta, \tau) := \tau + \frac{\theta}{2(1-\theta)}(2\tau + 2\sqrt{2n\tau} + n).$$

Since $\tau = O(n)$ and $\theta = \Theta(1)$, we have

$$L := L(n, \theta, \tau) = O(n).$$

4.2. A default value for the step size

After a damped step, using (5), the new point is computed as

$$x_+ := x + \alpha \triangle x = \frac{x}{v}(v + \alpha d_x), \qquad y_+ := y + \alpha \triangle y,$$
$$s_+ := s + \alpha \triangle s = \frac{s}{v}(v + \alpha d_s),$$

where α is a step size which is obtained by using a line search strategy. Thus, we get

$$v_+^2 = \frac{x_+ s_+}{\mu} = (v + \alpha d_x)(v + \alpha d_s).$$

From Lemma 2, we have

$$\Psi(v_+) = \Psi(\sqrt{(v + \alpha d_x)(v + \alpha d_s)}) \leq \frac{1}{2}(\Psi(v + \alpha d_x) + \Psi(v + \alpha d_s)).$$

By defining

$$f(\alpha) := \Psi(v_+) - \Psi(v) = \Psi(\sqrt{(v + \alpha d_x)(v + \alpha d_s)}) - \Psi(v),$$

and

$$f_1(\alpha) := \frac{1}{2}(\Psi(v + \alpha d_x) + \Psi(v + \alpha d_s)) - \Psi(v), \tag{21}$$

we have $f(\alpha) \leq f_1(\alpha)$ and $f_1(\alpha)$ is a convex function with $f(0) = f_1(0) = 0$. Taking the derivative with respect to α, we further get

$$f_1'(\alpha) := \frac{1}{2}\sum_{i=1}^{n}(\psi'(v_i + \alpha d_{xi})d_{xi} + \psi'(v_i + \alpha d_{si})d_{si}),$$
$$f_1''(\alpha) = \frac{1}{2}\sum_{i=1}^{n}(\psi''(v_i + \alpha d_{xi})d_{xi}^2 + \psi''(v_i + \alpha d_{si})d_{si}^2).$$

By using the third equation of the system (8), we obtain

$$f_1'(0) = \frac{1}{2}\nabla\Psi^T(v)(d_x + d_s) = -\frac{1}{2}\nabla\Psi^T(v)\nabla\Psi(v) = -2\delta^2(v).$$

For simplicity, in the sequel we use the following notations

$$v_1 := \min(v), \qquad \delta := \delta(v).$$

Since our kernel function is an eligible function, so we only state the following important results without proof.

Lemma 6 (Lemma 4.1 in [3]) *Let $f_1(\alpha)$ be defined as* (21). *Then*

$$f_1''(\alpha) \leq 2\delta^2\psi''(v_1 - 2\alpha\delta).$$

Lemma 7 (Lemma 3.4 in [2]) *One has $f_1(\alpha) \leq 0$, if α satisfies the inequality*

$$-\psi'(v_1 - 2\alpha\delta) + \psi'(v_1) \leq 2\delta. \tag{22}$$

Lemma 8 (Lemma 4.3 in [16]) *Let $\rho : [0, \infty) \to (0, 1]$ denote the inverse function of $-\frac{1}{2}\psi'(t)$ restricted to the interval* $(0, 1]$. *Then, the step size*

$$\bar{\alpha} := \frac{1}{2\delta}(\rho(\delta) - \rho(2\delta)), \tag{23}$$

is the largest possible solution of inequality (22).

Lemma 9 (Lemma 4.4 in [6]) *Let ρ and $\bar{\alpha}$ be as define in Lemma 8. Then*

$$\bar{\alpha} \geq \frac{1}{\psi''(\rho(2\delta))}. \tag{24}$$

We are using the following step size as our default step size in Algorithm 1

$$\tilde{\alpha} := \frac{1}{\psi''(\rho(2\delta))}. \tag{25}$$

By (24) and (25), we have $\tilde{\alpha} \leq \bar{\alpha}$.

Lemma 10 (Lemma 4.6 in [16]) *If the step size α is such that $\alpha \leq \bar{\alpha}$, then*

$$f(\alpha) \leq -\alpha\delta^2. \tag{26}$$

Theorem 1 *Let ρ be as defined in Lemma* 8, *$\tilde{\alpha}$ as define in* (25) *and* $\Psi(v) \geq 1$. *Then*

$$f(\tilde{\alpha}) \leq -\frac{\delta^2}{\psi''(\rho(2\delta))} \leq -\frac{\delta}{(18p + 2)(1 + \frac{1}{p}\log(4\delta + 1))^2}. \tag{27}$$

Proof. Lemma 10 and the fact $\tilde{\alpha} \leq \bar{\alpha}$ imply that $f(\tilde{\alpha}) \leq -\tilde{\alpha}\delta^2$, which follows the first inequality.

To obtain the inverse function of the restriction of $-\frac{1}{2}\psi'(t)$ in the interval $t \in (0, 1]$, we need to solve the equation $-\frac{1}{2}\psi'(t) = s$ for t. To do so, we have

$$-(t - \frac{1}{t^{2p}}e^{\frac{p}{t}-p}) = 2s.$$

This implies that

$$\frac{1}{t^{2p}}e^{\frac{p}{t}-p} = 2s + t \leq 2s + 1, \tag{28}$$

where the last inequality is obtained from the fact that $t \le 1$. Hence, putting $t = \rho(2\delta)$, which is equivalent to $4\delta = -\psi'(t)$, we get

$$\begin{aligned}\frac{1}{t^{2p}} e^{\frac{p}{t}-p} = 4\delta + t \le 4\delta + 1, &\Rightarrow e^{\frac{p}{t}-p} \le t^{2p}(4\delta+1) \le 4\delta + 1,\\ &\Rightarrow \frac{p}{t} - p \le \log(4\delta+1),\\ &\Rightarrow \frac{1}{t} \le 1 + \frac{1}{p}\log(4\delta+1).\end{aligned} \tag{29}$$

Furthermore, using (11) and (29), we conclude that

$$\begin{aligned}\tilde{\alpha} = \frac{1}{\psi''(t)} &= \frac{1}{1 + \frac{2p}{t^{2p+1}} e^{\frac{p}{t}-p} + \frac{p}{t^{2p+2}} e^{\frac{p}{t}-p}}\\ &\ge \frac{1}{1 + 2p(4\delta+1)(1+\frac{1}{p}\log(4\delta+1)) + p(4\delta+1)(1+\frac{1}{p}\log(4\delta+1))^2}\\ &\ge \frac{1}{1 + 2p(4\delta+1)(1+\frac{1}{p}\log(4\delta+1))^2 + p(4\delta+1)(1+\frac{1}{p}\log(4\delta+1))^2}\\ &\ge \frac{1}{1 + 12p\delta(1+\frac{1}{p}\log(4\delta+1))^2 + 3p(1+\frac{1}{p}\log(4\delta+1))^2}\\ &= \frac{1}{1 + (12p\delta + 3p)(1+\frac{1}{p}\log(4\delta+1))^2}.\end{aligned}$$

Using Corollary 1, one has

$$\begin{aligned}\tilde{\alpha} &\ge \frac{1}{\sqrt{2}\delta + (12p\delta + 3\sqrt{2}p\delta)(1+\frac{1}{p}\log(4\delta+1))^2}\\ &\ge \frac{1}{(2\delta + 12p\delta + 6p\delta)(1+\frac{1}{p}\log(4\delta+1))^2}\\ &\ge \frac{1}{(18p+2)\delta(1+\frac{1}{p}\log(4\delta+1))^2}.\end{aligned}$$

This implies that

$$f(\tilde{\alpha}) \le -\frac{\delta^2}{\psi''(\rho(2\delta))} \le -\frac{\delta}{(18p+2)(1+\frac{1}{p}\log(4\delta+1))^2}.$$

Thus the theorem follows. ∎

A direction consequence of applying the Lemma 3 to (27) is

$$\begin{aligned}f(\tilde{\alpha}) &\le -\frac{\delta}{(18p+2)(1+\frac{1}{p}\log(4\delta+1))^2}\\ &\le -\frac{\Psi(v)^{\frac{1}{2}}}{\sqrt{2}(18p+2)(1+\frac{1}{p}\log(2\sqrt{2\Psi(v)}+1))^2}\\ &\le -\frac{\Psi(v)^{\frac{1}{2}}}{\sqrt{2}(18p+2)(1+\frac{1}{p}\log(2\sqrt{2\Psi_0(v)}+1))^2},\end{aligned} \tag{30}$$

where the last inequality follows from $\Psi_0(v) \ge \Psi(v) \ge \tau \ge 1$.

5. Iteration complexity

We need to count how many inner iterations are required to return to the situation where $\Psi(v) \leq \tau$ after a μ-update. We define the value of $\Psi(v)$ after μ-update as $\Psi_0(v)$ and the subsequent values as Ψ_k, for $k = 1, 2, \ldots, K$. Let K stands for the total number of inner iterations in an outer iteration. Then we have

$$\Psi_0(v) \leq L = O(n), \quad \Psi_{K-1} > \tau, \quad 0 \leq \Psi_K \leq \tau. \tag{31}$$

To get the iteration complexity bound of algorithm, we need the following technical result.

Lemma 11 (Lemma 4.7 in [21]) *Let $t_0, t_1, \ldots, t_K$ be a sequence of positive numbers such that*

$$t_{k+1} \leq t_k - \beta t_k^{1-\gamma}, \quad k = 0, \ldots, K-1, \tag{32}$$

where $\beta > 0$ and $0 < \gamma \leq 1$. Then $K \leq \lfloor t_0^{\gamma}/\beta\gamma \rfloor$.

Lemma 12 *If K denote the total number of inner iterations in the outer iteration. Then we have*

$$K \leq 2\sqrt{2}(18p+2)\left(1 + \frac{1}{p}\log(2\sqrt{2\Psi_0(v)}+1)\right)^2 \Psi_0(v)^{\frac{1}{2}}.$$

Proof. From (30), we have

$$\Psi_{k+1}(v) \leq \Psi_k(v) - \frac{\Psi_k(v)^{\frac{1}{2}}}{\sqrt{2}(18p+2)(1+\frac{1}{p}\log(2\sqrt{2\Psi_0(v)}+1))^2}.$$

Letting $t_k = \Psi_k(v)$, $\beta = \frac{1}{\sqrt{2}(18p+2)(1+\frac{1}{p}\log(2\sqrt{2\Psi_0(v)}+1))^2}$ and $\gamma = \frac{1}{2}$, we can get the result of the lemma from Lemma 11. ∎

The following theorem gives the iteration bound for the algorithm with large-update methods.

Theorem 2 *For large-update methods, one takes $\theta = \Theta(1)$ and $\tau = O(n)$, then Algorithm 1 requires at most*

$$O\big(p\sqrt{n}(1 + \frac{1}{p}\log n)^2 \log \frac{n}{\varepsilon}\big)$$

iterations. The output gives an ε-approximate solution of (P) and (D).

Proof. It is well known that the number of outer iterations is bounded above by

$$\frac{1}{\theta}\log\frac{n}{\varepsilon}$$

(see, e.g.,[11]). By multiplying this result and the upper bound for the number of inner iterations per outer iteration, we can get an upper bound for the total number of iterations, namely,

$$\frac{2\sqrt{2}(18p+2)(1+\frac{1}{p}\log(2\sqrt{2\Psi_0(v)}+1))^2\Psi_0(v)^{\frac{1}{2}}}{\theta}\log\frac{n}{\varepsilon}.$$

Considering that $\theta=\Theta(1)$, $\tau=O(n)$ and $\Psi_0^{\frac{1}{2}}=O(\sqrt{n})$, some elementary transformation reduce this iteration complexity bound in the theorem. ■

Corollary 2 *If $p=\log n$, the iteration bound for large-update methods reduces to $O(\sqrt{n}\log n\log\frac{n}{\varepsilon})$, which matches the currently best known iteration bound for large-update IPMs.*

6. Numerical results

In this section, we present some numerical results. Number results are obtained by using Matlab 2012b. We consider the following simple LO problem

$$A=\begin{bmatrix}-3 & 2 & 1 & 0 & 0\\ -1 & 2 & 0 & 1 & 0\\ 1 & 1 & 0 & 0 & 1\end{bmatrix},\qquad b=\begin{bmatrix}0\\2\\3\end{bmatrix},\qquad C=\begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix}.$$

The main goal of this section is to compare iteration numbers of the algorithm for the following kernel functions

$$\psi_0(t)=\frac{t^2-1}{2}-\int_1^t\frac{1}{x^{2p}}e^{\frac{p}{x}-p}dx,\ p\geq 1,\quad \psi_1(t)=\frac{t^2-1}{2}-(t-1)e^{\frac{1}{t}-1},$$
$$\psi_2(t)=\frac{t^2-1}{2}-\log t,\qquad\qquad \psi_3(t)=t+t^{-1}-2,$$

$\psi_1(t)$ is the new kernel function proposed in [24], $\psi_2(t)$ is the classical logarithmic kernel function in [18], and $\psi_3(t)$ is the non-self-regular kernel function in [21]. We take $p\in\{1,2,2.5\}$ and the barrier update parameter $\theta\in\{0.1,0.3,0.7\}$, the step size $\alpha\in\{0.1,0.3,0.5\}$, the threshold parameter $\tau=3$, and the accuracy parameter $\varepsilon=10^{-8}$ in all experiments. Table 1 shows that the iteration numbers of the algorithm based on the above kernel functions. From Table 1, it is found that all cases our new kernel function $\psi_0(t)$ can produce better iteration numbers than those by the kernel functions $\psi_1(t)$, $\psi_2(t)$ and $\psi_3(t)$. Obviously, $\psi_3(t)$ gives iteration numbers that are definitely worst compared to the other kernel functions, especially compared to our kernel function. The results in table 1 also show that the algorithm based on our new kernel function $\psi_0(t)$ is effective and the iteration numbers of the algorithm depend on the values of the parameter θ and step size α. In fact, for each θ that considered, larger values of α gives better iteration numbers. However, the step size α should have an upper bound in practical

Table 1: The number of iterations for kernel function $\psi_i(t)$, i=0,1,2,3.

α	θ	$\psi_0(t)$ $p=1$	$\psi_0(t)$ $p=2$	$\psi_0(t)$ $p=2.5$	$\psi_1(t)$	$\psi_2(t)$	$\psi_3(t)$
$\alpha=0.1$	0.1	258	218	209	308	334	617
	0.3	243	210	204	292	279	718
	0.7	217	199	196	259	263	987
$\alpha=0.3$	0.1	80	66	63	96	105	294
	0.3	74	63	60	91	84	332
	0.5	66	59	59	80	82	432
$\alpha=0.5$	0.1	44	35	34	53	62	201
	0.3	40	34	32	56	47	223
	0.5	33	32	31	43	44	289

computation. In most cases, for each α, larger θ gives better iteration numbers for $\psi_0(t)$, $\psi_1(t)$ and $\psi_2(t)$, while for $\psi_3(t)$, smaller θ gives better results.

7. Concluding remarks

In this paper, we have presented a primal-dual large-update IPM for LO based on a new kernel function and derived favorable complexity bound for the algorithm, namely, $O(\sqrt{n}\log n \log \frac{n}{\varepsilon})$. This bound is the currently best known bound for primal-dual IPMs. Some preliminary numerical results are presented, which show that by using new kernel function, the best iteration numbers were achieved in most cases.

Some interesting topics remain for the further research. Firstly, finding a suitable kernel function for which the complexity of large-update methods is equal to (or even better than) $O(\sqrt{n}\log \frac{n}{\varepsilon})$ or showing that such a kernel function does not exist. Secondly, the extensions to cone optimization seems an interesting topic. Furthermore, the numerical test is an interesting topic for investigating the behavior of the algorithm so as to be compared with other approaches.

Acknowledgements. The research of this paper is supported by National Natural Science Foundation of China (No.71471102).

References

[1] Bai, Y.Q., Lesaja, G., Roos, C., Wang, G.Q., Ghami, M.E., *A class of large-update and small-update primal-dual interior-point algorithms for linear optimization*, Journal of Optimization Theory and Applications, 138 (2008), 341-359.

[2] Bai, Y.Q., Roos, C., *A primal-dual interior-point method based on a new kernel function with linear growth rate*, in: Proceedings of Industrial Optimization Symposium and Optimization Day, Nov., 2002.

[3] Bai, Y.Q., Roos, C., Ghami, M.E., *A comparative study of kernel functions for primal-dual interior-point algorithms in linear optimization*, SIAM Journal on Optimization, 15 (1) (2004), 101-128.

[4] Bai, Y.Q., Xie, W., Zhang, J., *New parameterized kernel functions for linear optimization*, Journal of Global Optimization, 54 (2012), 353-366.

[5] Cai, X.Z., Wang, G.Q., Ghami, M.E., Yue, Y.J., *Complexity analysis of interior-point methods for linear optimization based on a new parametric kernel function with a trigonometric barrier term*, Abstract and Applied Analysis, 2014, 710158 (2014)2.

[6] Cho, G.M., *An interior-point algorithm for linear optimization based on a new barrier function*, Applied Mathematics and Computation, 218 (2011), 386-395.

[7] Ghami, M.E., Guennounb, Z.A., Bouali c, S., Steihaug, T., *Interior-point methods for linear optimization based on a kernel function with a trigonometric barrier term*, Journal of Computational and Applied Mathematics, 236 (2012), 3613-3623.

[8] Ghami, M.E., Ivanov, I., Melissen, J.B., Roos, C., Steihaug, T., *A polynomial-time algorithm for linear optimization based on a new class of kernel functions*, Journal of Computational and Applied Mathematics, 224 (2009), 500-513.

[9] Karmarkar, N.K., *A new polynomial-time algorithm for linear programming*, Combinatorica, 4 (1984), 373-395.

[10] Kim, M.K., Lee, Y.H., Cho, G.M., *An adaptive-step primal-dual interior point algorithm for linear optimization*, Nonlinear Analysis: Theory Methods & Applications, 71 (2009), 2305-2315.

[11] Kojima, M., Mizuno, S., Yoshise, A., *A primal-dual interior point algorithm for linear programming*, in: N. Megiddo (Ed.), Progress in Mathematical Programming: Interior Point and Related Methods, Springer Verlag, New York, 1989, 29-47.

[12] Li, X., Zhang, M.W., *Interior-point algorithm for linear optimization based on a new trigonometric kernel function*, Operations Research Letters, 43 (2015), 471-475.

[13] Liu, L.Y., Li, S.Y., *A new kind of simple kennel function yielding good iteration bounds for primal-dual interior-point methods*, Journal of Computational and Applied Mathematics, 235 (2011), 2944-2955.

[14] Megiddo, M., *Pathways to the optimal set in linear programming*, in Progress in Mathematical Programming: Interior Point and Related Methods, N. Megiddo, ed., Springer-Verlag, New York, 1989, 131-158.

[15] Peng, J., Roos, C., Terlaky, T., *Self-regular functions and new search directions for linear and semi-definite optimization*, Mathematical Programming, 93 (2002), 129-171.

[16] Peyghami, M.R., Hafshejani, S.F., *Complexity analysis of an interior-point algorithm for linear optimization based on a new proximity function*, Numerical Algorithm, DOI 10.1007/s11075-013-9772-1.

[17] Peyghami, M.R., Hafshejani, S.F., Shirvani, L., *Complexity of interior-point methods for linear optimization based on a new trigonometric kernel function*, Journal of Computational and Applied Mathematics, 255 (2014), 74-85.

[18] Ross, C., Terlaky, T., Vial, J.P., *Theory and algorithm for linear optimization: An interior point approach*, John Wiley & Sons, Chichester, UK, 1997.

[19] Salahi, M., Terlarky, T., *An adaptive self-regular proximity based large-update IPM for linear optimization*, Optimization Methods and Software, 20 (1) (2005), 169-185.

[20] Sonnevend, G., *An analytic center for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming*, in: A. Prkopa, J. Szelezsn, B. Strazicky (Eds.), Lecture Notes in Control and Information Sciences, vol. 84, Springer Verlag, Berlin, 1986, 866-876.

[21] Wang, G.Q., Bai, Y.Q., Roos, C., *Primal-dual interior-point algorithms for semi-definite optimization based on a simple kernel function*, Journal of Mathematical Modelling and Algorithms, 4 (4) (2005), 409-433.

[22] Wright, S.J., *Primal-dual Interior-Point Methods*, SIAM, Philadelphia, USA, 1997.

[23] Ye, Y., *Interior point algorithms: Theory and Analysis*, Wiley-Interscience, 1997.

[24] Zhao, D.Q., Zhang, M.W., *A primal-dual large-update interior-point algorithm for semi-definite optimization based on new kernel functions*, Statistics, Optimization and Information Computing, 1 (2013), 41-61.

Accepted: 20.11.2015