

AN ELIGIBLE KERNEL BASED PRIMAL-DUAL INTERIOR-POINT METHOD FOR LINEAR OPTIMIZATION

GYEONG-MI CHO

Abstract. It is well known that each kernel function defines a primal-dual interior-point method (IPM). Most of polynomial-time interior-point algorithms for linear optimization (LO) are based on the logarithmic kernel function ([9]). In this paper we define a new eligible kernel function and propose a new search direction and proximity function based on this function for LO problems. We show that the new algorithm has $\mathcal{O}((\log p)^{\frac{5}{2}}\sqrt{n}\log n\log\frac{n}{\epsilon})$ and $\mathcal{O}(q^{\frac{3}{2}}(\log p)^3\sqrt{n}\log\frac{n}{\epsilon})$ iteration complexity for large- and small-update methods, respectively. These are currently the best known complexity results for such methods.

1. Introduction

In this paper we propose a new primal-dual IPM for the following standard LO problem

$$(1.1) \quad \min\{c^T x : Ax = b, x \geq 0\},$$

where $A \in \mathbf{R}^{m \times n}$ with $\text{rank}(A) = m$, $c, x \in \mathbf{R}^n$, $b \in \mathbf{R}^m$, and its dual problem

$$(1.2) \quad \max\{b^T y : A^T y + s = c, s \geq 0\},$$

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

Since Karmarkar's paper ([5]) in 1984, interior-point methods (IPMs) have shown their efficiency in solving large-scale LO problems with a wide variety of successful applications. In this paper, we propose a new

Received March 27, 2013. Accepted April 22, 2013.

2010 Mathematics Subject Classification. 90C05, 90C51.

Key words and phrases. primal-dual interior point method, kernel function, complexity, polynomial algorithm, linear optimization.

This work was supported by Dongseo University "Dongseo Frontier Project" Research Fund of 2011.

primal-dual IPM based on an eligible kernel function. It is generally agreed that the iteration complexity of the algorithm is an appropriate measure for its efficiency.

Peng et al. ([7]) proposed new variants of IPMs based on self-regular kernel functions and obtained the best known complexity results for large- and small-update methods for LO with a specific self-regular kernel function and extended to more general optimization problems. Recently, Roos et al. ([1, 3]) proposed new primal-dual IPMs for LO problems based on eligible kernel functions and obtained the best known complexity results of large- and small-update methods. They also proposed the framework for analyzing the algorithm based on four conditions on kernel function and generalized those methods to the general optimization problem ([4]).

Motivated by their works, in this paper, we define a new kernel function which includes the kernel function in [2] as a special case and propose a new primal-dual IPM for LO problems based on this function and improves the complexity result of [2] for large-update method. For the complexity analysis we follow the framework in [2]. We show that the algorithm has $\mathcal{O}((\log p)^{\frac{5}{2}}\sqrt{n} \log n \log \frac{n}{\epsilon})$ and $\mathcal{O}(q^{\frac{3}{2}}(\log p)^3\sqrt{n} \log \frac{n}{\epsilon})$ iteration complexity for large- and small-update methods, respectively, where $p \geq e$, $q \geq 1$, ϵ is a desired accuracy, and n is a dimension of the problem. These bounds are currently the best known complexity results for such methods.

The paper is organized as follows. In Section 2, we recall the generic IPM. In Section 3, we define a new kernel function and give its properties which are essential for the complexity analysis. In Section 4, we propose the new algorithm and derive the complexity result for both large- and small-update methods. Finally, concluding remarks are given in Section 5.

We use the following notations throughout the paper. \mathbf{R}_+^n and \mathbf{R}_{++}^n denote the set of n -dimensional nonnegative and positive vectors, respectively. For $x, s \in \mathbf{R}^n$, x_{\min} and xs denote the smallest component of the vector x and the componentwise product of the vectors x and s , respectively. We denote X the diagonal matrix from a vector x , i.e. $X = \text{diag}(x)$. \mathbf{e} denotes the n -dimensional vector of ones and e , the Euler's number. For notational convenience we denote the natural logarithm by \log . For $f(x), g(x) : \mathbf{R}_{++} \rightarrow \mathbf{R}_{++}$, $f(x) = \mathcal{O}(g(x))$ if $f(x) \leq c_1 g(x)$ for some positive constant c_1 and $f(x) = \Theta(g(x))$ if $c_2 g(x) \leq f(x) \leq c_3 g(x)$ for some positive constants c_2 and c_3 .

2. Preliminaries

In this section, we recall the basic concepts and propose the generic algorithm. Without loss of generality, we assume that both problem (1.1) and (1.2) satisfy the interior-point condition (IPC) ([8]), i.e. there exists (x^0, y^0, s^0) such that

$$Ax^0 = b, x^0 > 0, A^T y^0 + s^0 = c, s^0 > 0.$$

By the duality theorem (Theorem II.2 in [8]), finding an optimal solution of problem (1.1) and (1.2) is equivalent to solving the following system:

$$(2.1) \quad Ax = b, x \geq 0, A^T y + s = c, s \geq 0, xs = 0.$$

The basic idea of primal-dual IPMs is to replace the third equation in (2.1) by the parameterized equation $xs = \mu \mathbf{e}$ with $\mu > 0$. Now we consider the following system:

$$(2.2) \quad Ax = b, x > 0, A^T y + s = c, s > 0, xs = \mu \mathbf{e}.$$

If the IPC holds, then the system (2.2) has a unique solution for each $\mu > 0$ ([6]). We denote this solution as $(x(\mu), y(\mu), s(\mu))$ and call $x(\mu)$ the μ -center of (1.1) and $(y(\mu), s(\mu))$ the μ -center of (1.2). The set of μ -centers ($\mu > 0$) is the central path of (1.1) and (1.2). The limit of the central path (as μ goes to zero) exists and since the limit point satisfies (2.1), it naturally yields optimal solutions for (1.1) and (1.2) ([8]). Primal-dual IPMs follow the central path approximately and approach the solution of (1.1) and (1.2) as μ goes to zero.

For given $(x, y, s) := (x^0, y^0, s^0)$ by applying Newton's method to the system (2.2) we have the following Newton system

$$(2.3) \quad A\Delta x = 0, A^T \Delta y + \Delta s = 0, s\Delta x + x\Delta s = \mu \mathbf{e} - xs.$$

Since A has full row rank, the system (2.3) has a unique search direction vector $(\Delta x, \Delta y, \Delta s)$. By taking a step along the search direction $(\Delta x, \Delta y, \Delta s)$, one constructs a new iteration (x_+, y_+, s_+) with

$$x_+ := x + \alpha \Delta x, y_+ := y + \alpha \Delta y, s_+ := s + \alpha \Delta s,$$

for some $\alpha > 0$.

For the motivation of the new algorithm we define the scaled vectors as follows:

$$(2.4) \quad v := \sqrt{\frac{xs}{\mu}}, d_x := \frac{v\Delta x}{x}, d_s := \frac{v\Delta s}{s}.$$

Using (2.4), we can rewrite the system (2.3) as follows:

$$(2.5) \quad \bar{A}d_x = 0, \bar{A}^T \Delta y + d_s = 0, d_x + d_s = v^{-1} - v,$$

where $\bar{A} := \frac{1}{\mu}AV^{-1}X$, $V := \text{diag}(v)$, and $X := \text{diag}(x)$. Note that the right hand side of the third equation in (2.5) equals the negative gradient of the logarithmic barrier function $\Psi_l(v)$, i.e.

$$(2.6) \quad d_x + d_s = -\nabla\Psi_l(v),$$

where

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

We call ψ_l the kernel function of the logarithmic barrier function $\Psi_l(v)$. We call $\psi : \mathbf{R}_{++} \rightarrow \mathbf{R}_+$ a kernel function if ψ is twice differentiable and satisfies the following conditions:

$$(2.7) \quad \psi'(1) = \psi(1) = 0, \quad \psi''(t) > 0, \quad \forall t > 0, \quad \lim_{t \rightarrow 0^+} \psi(t) = \lim_{t \rightarrow \infty} \psi(t) = \infty.$$

For a kernel function $\psi(t)$, we define $\Psi(v) := \sum_{i=1}^n \psi(v_i)$. In this paper we replace $\Psi_l(v)$ in (2.6) by $\Psi(v)$.

Note that d_x and d_s are orthogonal because the vector d_x belongs to null space and d_s to the row space of the matrix \bar{A} . Since d_x and d_s are orthogonal, we have

$$\begin{aligned} d_x = d_s = 0 &\Leftrightarrow \nabla\Psi(v) = 0 \Leftrightarrow v = \mathbf{e} \\ &\Leftrightarrow \Psi(v) = 0 \Leftrightarrow x = x(\mu), \quad s = s(\mu). \end{aligned}$$

We use $\Psi(v)$ as the proximity function which measures the distance between current iteration and corresponding μ -center. Also, we define the norm-based proximity measure $\delta(v)$ as follows: for $v \in \mathbf{R}_{++}^n$,

$$(2.8) \quad \delta(v) := \frac{1}{2} \|\nabla\Psi(v)\| = \frac{1}{2} \|d_x + d_s\|.$$

The generic IPM works as follows: Assume that we are given a strictly feasible point (x, y, s) which is in a τ -neighborhood of the given μ -center. Then we decrease μ to $\mu_+ := (1 - \theta)\mu$, for some fixed $\theta \in (0, 1)$ and then we solve the Newton system (2.3) to obtain the unique search direction. The positivity condition of a new iteration is ensured with the right choice of the step size α which is defined by some line search rule. This procedure is repeated until we find a new iteration (x_+, y_+, s_+) that is in a τ -neighborhood of the μ_+ -center and then we let $\mu := \mu_+$ and $(x, y, s) := (x_+, y_+, s_+)$. Then μ is again reduced by the factor $1 - \theta$ and we solve the Newton system targeting at the new μ_+ -center, and so on. This process is repeated until μ is small enough, i.e. $n\mu < \varepsilon$.

Generic Primal-Dual Algorithm for LO

Input:
 a threshold parameter $\tau \geq 1$;
 an accuracy parameter $\varepsilon > 0$;
 a fixed barrier update parameter θ , $0 < \theta < 1$;
 (x^0, s^0) and $\mu^0 := 1$ such that $\Psi_l(x^0, s^0, \mu^0) \leq \tau$.
 begin
 $x := x^0$; $s := s^0$; $\mu := \mu^0$;
 while $n\mu \geq \varepsilon$ do
 begin
 $\mu := (1 - \theta)\mu$;
 while $\Psi_l(v) > \tau$ do
 begin
 Solve the system (2.3) for $\Delta x, \Delta y, \Delta s$,
 Determine a step size α ;
 $x := x + \alpha\Delta x$;
 $s := s + \alpha\Delta s$;
 $y := y + \alpha\Delta y$;
 $v := \sqrt{\frac{xs}{\mu}}$;
 end
 end
 end
 end

Remark 2.1. *If θ is a constant independent of the dimension of the problem n , e.g. $\theta = \frac{1}{2}$, then we call the algorithm a large-update method. If θ depends on n , e.g. $\theta = \frac{1}{\sqrt{n}}$, then the algorithm is called a small-update method.*

3. The kernel function

In this section, we define a new kernel function and give its properties which are essential to the complexity analysis.

Consider a function $\psi(t)$ as follows:

$$(3.1) \quad \psi(t) = \frac{(\log p)(t^2 - 1)}{2} + \frac{1}{q}(p^{q(\frac{1}{t}-1)} - 1), \quad p \geq e, \quad q \geq 1, \quad t > 0.$$

Then we have the following:

$$\begin{aligned}
 \psi'(t) &= (\log p)t - (\log p)t^{-2}p^{q(\frac{1}{t}-1)}, \\
 (3.2) \quad \psi''(t) &= (\log p) + (\log p)t^{-4}(q(\log p) + 2t)p^{q(\frac{1}{t}-1)}, \\
 \psi^{(3)}(t) &= -(\log p)t^{-6}(q^2(\log p)^2 + 6q(\log p)t + 6t^2)p^{q(\frac{1}{t}-1)}.
 \end{aligned}$$

From (3.2) and (2.7), $\psi(t)$ is a kernel function and for $p \geq e$, $q \geq 1$,

$$(3.3) \quad \psi''(t) > 1, \quad t > 0.$$

When $q = 1$ and $p = e$, $\psi(t)$ is a kernel function in [2].

Lemma 3.1. For $\psi(t)$ as in (3.1), we have for $p \geq e$ and $q \geq 1$,

- (i) $t\psi''(t) + \psi'(t) > 0$, $t > 0$, i.e. $\psi(t)$ is exponentially convex for $t > 0$,
- (ii) $\psi^{(3)}(t) < 0$, $t > 0$,
- (iii) $t\psi''(t) - \psi'(t) > 0$, $t > 0$,
- (iv) $2(\psi''(t))^2 - \psi'(t)\psi^{(3)}(t) > 0$, $t > 0$.

Proof: For (i), using (3.2), we have for $p \geq e$, $q \geq 1$ and $t > 0$,

$$t\psi''(t) + \psi'(t) = 2(\log p)t + (\log p)t^{-3}(q(\log p) + t)p^{q(\frac{1}{t}-1)} > 0.$$

Hence $\psi(t)$ is exponentially convex, $t > 0$.

For (ii), from (3.2), $\psi^{(3)}(t) < 0$.

For (iii), using (3.2), we have

$$t\psi''(t) - \psi'(t) = (\log p)t^{-3}(q(\log p) + 3t)p^{q(\frac{1}{t}-1)} > 0, \quad t > 0.$$

For (iv), using (3.2), we have for $t > 0$,

$$\begin{aligned}
 &2(\psi''(t))^2 - \psi'(t)\psi^{(3)}(t) \\
 &= (\log p)^2 t^{-8} (2t^8 + (14t^2 + 10q(\log p)t + q^2(\log p)^2) t^3 p^{q(\frac{1}{t}-1)} \\
 &\quad + (2t^2 + 2q(\log p)t + q^2(\log p)^2) p^{2q(\frac{1}{t}-1)}) > 0.
 \end{aligned}$$

This completes the proof. \square

Note that we call $\psi(t)$ an eligible function if $\psi(t)$ satisfies four conditions of Lemma 3.1.

Remark 3.2. (i) By Lemma 2.4 in [2], if $\psi(t)$ satisfies Lemma 3.1 (ii) and (iii), then $\psi(t)$ satisfies

$$\psi''(t)\psi'(\beta t) - \beta\psi'(t)\psi''(\beta t) > 0, \quad t > 1, \quad \beta > 1.$$

(ii) By Lemma 2.1.2 in [7], Lemma 3.1 (i) is equivalent to

$$(3.4) \quad \psi(\sqrt{t_1 t_2}) \leq \frac{1}{2}(\psi(t_1) + \psi(t_2)), \quad t_1 > 0, \quad t_2 > 0.$$

Lemma 3.3. For $\psi(t)$ with $p \geq e$ and $q \geq 1$, we have

- (i) $\frac{1}{2}(t-1)^2 \leq \psi(t) \leq \frac{1}{2}(\psi'(t))^2$, $t > 0$,
- (ii) $\psi(t) \leq \frac{(\log p)(3+q \log p)}{2}(t-1)^2$, $t \geq 1$,
- (iii) $\psi(t) \leq \frac{\log p}{2}(t^2-1)$, $t \geq 1$.

Proof: For (i), using the first condition of (2.7) and (3.3), we have

$$\psi(t) = \int_1^t \int_1^\xi \psi''(\zeta) d\zeta d\xi \geq \int_1^t \int_1^\xi d\zeta d\xi = \frac{1}{2}(t-1)^2$$

which proves the first inequality. The second inequality is obtained as follows:

$$\begin{aligned} \psi(t) &= \int_1^t \int_1^\xi \psi''(\zeta) d\zeta d\xi \leq \int_1^t \int_1^\xi \psi''(\xi) \psi''(\zeta) d\zeta d\xi \\ &= \int_1^t \psi''(\xi) \psi'(\xi) d\xi = \int_1^t \psi'(\xi) d\psi'(\xi) = \frac{1}{2}(\psi'(t))^2. \end{aligned}$$

For (ii), using Taylor's Theorem, $\psi(1) = \psi'(1) = 0$, $\psi^{(3)} < 0$, and $\psi''(1) = (\log p)(3+q \log p)$, we have for $p \geq e$, $q \geq 1$ and $t \geq 1$,

$$\begin{aligned} \psi(t) &= \psi(1) + \psi'(1)(t-1) + \frac{1}{2}\psi''(1)(t-1)^2 + \frac{1}{3!}\psi^{(3)}(\xi)(t-1)^3 \\ &= \frac{1}{2}\psi''(1)(t-1)^2 + \frac{1}{3!}\psi^{(3)}(\xi)(t-1)^3 \\ &\leq \frac{1}{2}\psi''(1)(t-1)^2 = \frac{(\log p)(3+q \log p)}{2}(t-1)^2, \end{aligned}$$

for some ξ , $1 \leq \xi \leq t$.

For (iii), by the definition of $\psi(t)$, $\psi(t) \leq \frac{\log p}{2}(t^2-1)$, $t \geq 1$. This completes the proof. \square

Let $\varrho : [0, \infty) \rightarrow [1, \infty)$ be the inverse function of $\psi(t)$ for $t \geq 1$ and $\rho : [0, \infty) \rightarrow (0, 1]$, the inverse function of $-\frac{1}{2}\psi'(t)$ for $t \in (0, 1]$. Then we have the following lemma.

Lemma 3.4. For $\psi(t)$ with $p \geq e$, $q \geq 1$, we have

- (i) $\varrho(s) \leq 1 + \sqrt{2s}$, $s \geq 0$,
- (ii) $\rho(z) \geq \frac{\log p}{(\log p)+q^{-1} \log(\frac{2z+\log p}{\log p})}$, $z \geq 0$.

Proof: For (i), using Lemma 3.3 (i), we have $s = \psi(t) \geq \frac{(t-1)^2}{2}$. Then we have

$$t = \varrho(s) \leq 1 + \sqrt{2s}, \quad s \geq 0.$$

For (ii), let $z = -\frac{1}{2}\psi'(t)$, for all $t \in (0, 1]$. Then by the definition of ρ , $\rho(z) = t$, $t \in (0, 1]$ and $2z = -\psi'(t)$. So we have $2z = (\log p)t^{-2}p^{q(\frac{1}{t}-1)} - (\log p)t$.

$$p^{q(\frac{1}{t}-1)} \leq \frac{2z + \log p}{\log p}, \quad p \geq e, \quad 0 < t \leq 1.$$

Hence we have

$$\rho(z) = t \geq \frac{\log p}{(\log p) + q^{-1} \log\left(\frac{2z + \log p}{\log p}\right)}, \quad z \geq 0. \quad \square$$

4. Complexity analysis

In this section, we compute a growth bound due to the update of the barrier parameter during an outer iteration, a default step size and the decrease of the proximity function during an inner iteration and give the complexity results of the algorithm. In this paper, we replace the logarithmic barrier function $\Psi_l(v)$ in (2.6) with the eligible barrier function $\Psi(v)$ as follows:

$$(4.1) \quad d_x + d_s = -\nabla \Psi(v),$$

where $\Psi(v) = \sum_{i=1}^n \psi(v_i)$, where $\psi(t)$ is defined in (3.1).

Using Remark 3.2 (i), we have the following lemma. The reader can refer to Theorem 3.2 in [2] for the proof.

Lemma 4.1. *Let $\varrho : [0, \infty) \rightarrow [1, \infty)$ be the inverse function of $\psi(t)$, $t \geq 1$. Then we have*

$$\Psi(\beta v) \leq n\psi\left(\beta\varrho\left(\frac{\Psi(v)}{n}\right)\right), \quad v \in \mathbf{R}_{++}, \quad \beta \geq 1.$$

In the following theorem we obtain an estimate for the effect of a μ -update on the value of $\Psi(v)$.

Theorem 4.2. *Let $0 \leq \theta < 1$ and $v_+ = \frac{v}{\sqrt{1-\theta}}$. If $\Psi(v) \leq \tau$, then we have*

$$(i) \quad \Psi(v_+) \leq \frac{(\log p)(3+q \log p)}{2(1-\theta)} (\sqrt{n}\theta + \sqrt{2\tau})^2,$$

$$(ii) \quad \Psi(v_+) \leq \frac{\log p}{2(1-\theta)} (2\tau + 2\sqrt{2\tau n} + \theta n).$$

Proof: For (i), since $\frac{1}{\sqrt{1-\theta}} \geq 1$ and $\varrho\left(\frac{\Psi(v)}{n}\right) \geq 1$, we have $\frac{\varrho\left(\frac{\Psi(v)}{n}\right)}{\sqrt{1-\theta}} \geq 1$. Using Lemma 4.1 with $\beta = \frac{1}{\sqrt{1-\theta}}$, Lemma 3.3 (ii), Lemma 3.4 (i), and

$\Psi(v) \leq \tau$, we have

$$\begin{aligned} \Psi(v_+) &\leq n\psi\left(\frac{\varrho\left(\frac{\Psi(v)}{n}\right)}{\sqrt{1-\theta}}\right) \\ &\leq \frac{n(\log p)(3+q\log p)}{2}\left(\frac{\varrho\left(\frac{\Psi(v)}{n}\right)-\sqrt{1-\theta}}{\sqrt{1-\theta}}\right)^2 \\ &\leq \frac{n(\log p)(3+q\log p)}{2}\left(\frac{1+\sqrt{\frac{2\tau}{n}}-\sqrt{1-\theta}}{\sqrt{1-\theta}}\right)^2 \\ &\leq \frac{(\log p)(3+q\log p)}{2(1-\theta)}\left(\sqrt{n\theta}+\sqrt{2\tau}\right)^2, \end{aligned}$$

where the last inequality holds from $1-\sqrt{1-\theta}=\frac{\theta}{1+\sqrt{1-\theta}}\leq\theta, 0\leq\theta<1$. For (ii), using Lemma 4.1, Lemma 3.4 (i), $\Psi(v)\leq\tau$, and Lemma 3.3 (iii),

$$\begin{aligned} \Psi(v_+) &\leq n\psi\left(\frac{\varrho\left(\frac{\Psi(v)}{n}\right)}{\sqrt{1-\theta}}\right)\leq n\psi\left(\frac{1+\sqrt{\frac{2\tau}{n}}}{\sqrt{1-\theta}}\right) \\ &\leq \frac{\log p}{2(1-\theta)}(2\tau+2\sqrt{2\tau n}+\theta n). \end{aligned}$$

This completes the proof. \square

Denote

$$(4.2) \quad \begin{aligned} \tilde{\Psi}_0 &= \frac{(3+q\log p)\log p}{2(1-\theta)}\left(\sqrt{n\theta}+\sqrt{2\tau}\right)^2, \\ \bar{\Psi}_0 &= \frac{(2\tau+2\sqrt{2\tau n}+\theta n)\log p}{2(1-\theta)}. \end{aligned}$$

We will use $\tilde{\Psi}_0$ and $\bar{\Psi}_0$ as upper bounds of small- and large-update methods for $\Psi(v)$ during the process of the algorithm. We define the value of $\Psi(v)$ after the μ -update as Ψ_0 and the subsequent values in the same outer iteration are denoted as $\Psi_k, k=1, 2, \dots$. Then we have

$$\Psi_0 \leq \min\{\tilde{\Psi}_0, \bar{\Psi}_0\}.$$

Let K denote the total number of inner iterations per outer iteration. Then we have

$$\Psi_{K-1} > \tau, \quad 0 \leq \Psi_K \leq \tau.$$

Remark 4.3. For small-update method with $\tau = \mathcal{O}(1)$ and $\theta = \Theta(\frac{1}{\sqrt{n}})$, $\tilde{\Psi}_0 = \mathcal{O}(q(\log p)^2)$ and for large-update method with $\tau = \mathcal{O}(n)$ and $\theta = \Theta(1)$, $\tilde{\Psi}_0 = \mathcal{O}((\log p)n)$.

In the following we compute a default step size.

For fixed μ , if we take a step size α , then we have new iterations $x_+ := x + \alpha\Delta x$, $s_+ := s + \alpha\Delta s$. Using (2.4), we have

$$x_+ = x \left(e + \alpha \frac{\Delta x}{x} \right) = x \left(e + \alpha \frac{d_x}{v} \right) = \frac{x}{v} (v + \alpha d_x)$$

and

$$s_+ = s \left(e + \alpha \frac{\Delta s}{s} \right) = s \left(e + \alpha \frac{d_s}{v} \right) = \frac{s}{v} (v + \alpha d_s).$$

Thus we have

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

Define for $\alpha > 0$,

$$f(\alpha) := \Psi(v_+) - \Psi(v).$$

Then $f(\alpha)$ is the difference of proximities between a new iteration and a current iteration for fixed μ . From (3.4), we have

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

Hence we have $f(\alpha) \leq f_1(\alpha)$, where

$$(4.3) \quad f_1(\alpha) := \frac{1}{2} (\Psi(v + \alpha d_x) + \Psi(v + \alpha d_s)) - \Psi(v).$$

Obviously, we have

$$f(0) = f_1(0) = 0.$$

By taking the derivative of $f_1(\alpha)$ with respect to α , we have

$$f_1'(\alpha) = \frac{1}{2} \sum_{i=1}^n (\psi'(v_i + \alpha[d_x]_i)[d_x]_i + \psi'(v_i + \alpha[d_s]_i)[d_s]_i),$$

where $[d_x]_i$ and $[d_s]_i$ denote the i th components of the vectors d_x and d_s , respectively. Using (4.1) and (2.8), we have

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

Differentiating $f_1'(\alpha)$ with respect to α , we have

$$(4.4) \quad f_1''(\alpha) = \frac{1}{2} \sum_{i=1}^n (\psi''(v_i + \alpha[d_x]_i)[d_x]_i^2 + \psi''(v_i + \alpha[d_s]_i)[d_s]_i^2).$$

Since $f_1''(\alpha) > 0$, $f_1(\alpha)$ is strictly convex in α unless $d_x = d_s = 0$.

Lemma 4.4. *Let $\delta(v)$ be as defined in (2.8). Then we have*

$$\delta(v) \geq \sqrt{\frac{\Psi(v)}{2}}.$$

Proof: Using Lemma 3.3 (i), we have

$$\Psi(v) = \sum_{i=1}^n \psi(v_i) \leq \frac{1}{2} \sum_{i=1}^n (\psi'(v_i))^2 = \frac{1}{2} \|\nabla \Psi(v)\|^2 = 2\delta^2(v).$$

Hence we have $\delta(v) \geq \sqrt{\frac{\Psi(v)}{2}}$. □

For notational convenience we denote $\delta := \delta(v)$ and $\Psi := \Psi(v)$.

Lemma 4.5. (Lemma 4.2 in [2]) If the step size α satisfies the inequality

$$(4.5) \quad -\psi'(v_{min} - 2\alpha\delta) + \psi'(v_{min}) \leq 2\delta,$$

then we have

$$f_1'(\alpha) \leq 0.$$

Lemma 4.6. (Lemma 4.3 in [2]) Let $\rho : [0, \infty) \rightarrow (0, 1]$ denote the inverse function of $-\frac{1}{2}\psi'(t)$ for all $t \in (0, 1]$. Then, in the worst case, the largest step size $\hat{\alpha}$ satisfying (4.5) is given by

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

Lemma 4.7. (Lemma 4.4 in [2]) Let ρ and $\hat{\alpha}$ be as defined in Lemma 4.6. Then we have

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

Define

$$(4.6) \quad \bar{\alpha} := \frac{1}{\psi''(\rho(2\delta))}.$$

Then $\bar{\alpha} \leq \hat{\alpha}$ and we will use $\bar{\alpha}$ as the default step size.

In the following we compute the decrement of the proximity function during an inner iteration.

Lemma 4.8. (Theorem 4.6 in [2]) Let $\bar{\alpha}$ be as defined in (4.6). Then we have

$$(4.7) \quad f(\bar{\alpha}) \leq -\frac{\delta^2}{\psi''(\rho(2\delta))}.$$

By Lemma 3.1 (iv), we have the following lemma.

Lemma 4.9. (Lemma 4.7 in [2]) The right-hand side of (4.7) is monotonically decreasing in δ .

Theorem 4.10. Let $\bar{\alpha}$ be as defined in (4.6) and $\tau \geq 1$. Then

$$f(\bar{\alpha}) \leq -\frac{\sqrt{\Psi}}{2(\log p) \left(1 + 3q(\log p)(1 + 2\sqrt{2}) \left(1 + \frac{1}{q \log p} \log \left(\frac{2\sqrt{2\Psi_0} + \log p}{\log p}\right)\right)^4\right)}.$$

Proof: Using Lemma 4.8, Lemma 4.9 and Lemma 4.4, we have

$$(4.8) \quad f(\bar{\alpha}) \leq -\frac{\delta^2}{\psi''(\rho(2\delta))} \leq -\frac{1}{2} \frac{\Psi}{\psi''(\rho(\sqrt{2\Psi}))}.$$

Since $\Psi \geq \tau \geq 1$, $1 + 2\sqrt{2\Psi} \leq (1 + 2\sqrt{2})\sqrt{\Psi}$. By Lemma 3.4 (ii), Lemma 3.1 (ii), (3.2), $p \geq e$ and $q \geq 1$,

$$\begin{aligned} \psi''(\rho(\sqrt{2\Psi})) &\leq \psi''\left(\frac{q \log p}{q \log p + \log\left(\frac{2\sqrt{2\Psi} + \log p}{\log p}\right)}\right) \\ &\leq (\log p) + 3q(\log p)^2 \left(1 + \frac{1}{q \log p} \log\left(\frac{2\sqrt{2\Psi} + \log p}{\log p}\right)\right)^4 (1 + 2\sqrt{2})\Psi \\ &\leq (\log p) \left(1 + 3q(\log p)(1 + 2\sqrt{2}) \left(1 + \frac{1}{q \log p} \log\left(\frac{2\sqrt{2\Psi} + \log p}{\log p}\right)\right)^4\right) \Psi^{\frac{1}{2}}. \end{aligned}$$

Using (4.8) and $\Psi \leq \Psi_0$, we have

$$\begin{aligned} &f(\bar{\alpha}) \\ &\leq -\frac{\Psi}{2(\log p) \left(1 + 3q(\log p)(1 + 2\sqrt{2}) \left(1 + \frac{1}{q \log p} \log\left(\frac{2\sqrt{2\Psi} + \log p}{\log p}\right)\right)^4\right)} \Psi^{\frac{1}{2}} \\ &\leq -\frac{\sqrt{\Psi}}{2(\log p) \left(1 + 3q(\log p)(1 + 2\sqrt{2}) \left(1 + \frac{1}{q \log p} \log\left(\frac{2\sqrt{2\Psi_0} + \log p}{\log p}\right)\right)^4\right)}. \end{aligned}$$

This completes the proof. \square

In the following we compute the complexity bounds of the algorithm.

Lemma 4.11. (Lemma 1.3.2 in [7]) Let $t_0, t_1, \dots, t_{\bar{K}}$ be a sequence of positive numbers such that

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

where $\gamma > 0$ and $0 < \tilde{\beta} \leq 1$. Then $\bar{K} \leq \left\lceil \frac{t_0^{\tilde{\beta}}}{\gamma^{\tilde{\beta}}} \right\rceil$.

Lemma 4.12. Let K be the total number of inner iterations in an outer iteration. Then we have

$$K \leq 4(\log p) \left(1 + 3q(\log p)(1 + 2\sqrt{2}) \left(1 + \frac{1}{q \log p} \log \left(\frac{2\sqrt{2\Psi_0} + \log p}{\log p} \right) \right)^4 \right) \Psi_0^{\frac{1}{2}}.$$

Proof: Using Theorem 4.10 and Lemma 4.11 with

$$\frac{1}{\gamma} := 2(\log p) \left(1 + 3q(\log p)(1 + 2\sqrt{2}) \left(1 + \frac{\log \left(\frac{2\sqrt{2\Psi_0} + \log p}{\log p} \right)}{q \log p} \right)^4 \right),$$

$$\frac{1}{\tilde{\beta}} := 2,$$

we have

$$K \leq 4(\log p) \left(1 + 3q(\log p)(1 + 2\sqrt{2}) \left(1 + \frac{1}{q \log p} \log \left(\frac{2\sqrt{2\Psi_0} + \log p}{\log p} \right) \right)^4 \right) \Psi_0^{\frac{1}{2}}.$$

This completes the proof. \square

Theorem 4.13. Let a LO problem be given and $\tau \geq 1$. Then the total number of iterations to have an approximate solution with $n\mu < \epsilon$ is bounded by

$$\left\lceil \frac{4(\log p) \left(1 + 3q(\log p)(1 + 2\sqrt{2}) \left(1 + \frac{1}{q \log p} \log \left(\frac{2\sqrt{2\Psi_0} + \log p}{\log p} \right) \right)^4 \right) \Psi_0^{\frac{1}{2}} \log \frac{n}{\epsilon}}{\theta} \right\rceil.$$

Proof: If the central path parameter μ has the initial value $\mu^0 := 1$ and is updated by multiplying $1 - \theta$ with $0 \leq \theta < 1$, then after at most

$$\left\lceil \frac{1}{\theta} \log \frac{n}{\epsilon} \right\rceil$$

iterations we have $n\mu < \epsilon$ ([8]). For the total number of iterations, we multiply the number of inner iterations by that of outer iterations. Hence

the total number of iterations is bounded by

$$\left\lceil \frac{4(\log p) \left(1 + 3q(\log p)(1 + 2\sqrt{2}) \left(1 + \frac{1}{q \log p} \log \left(\frac{2\sqrt{2}\Psi_0 + \log p}{\log p}\right)\right)^4\right)}{\theta} \Psi_0^{\frac{1}{2}} \log \frac{n}{\epsilon} \right\rceil.$$

□

Remark 4.14. By Remark 4.3, for large-update methods with $\tau = \mathcal{O}(n)$ and $\theta = \Theta(1)$, by taking $q := \log \left(\frac{1+2\sqrt{2}\Psi_0}{\log p}\right)$, the algorithm has $\mathcal{O}((\log p)^{\frac{5}{2}} \sqrt{n} \log n \log \frac{n}{\epsilon})$ iteration complexity for $p \geq e$. For small-update methods with $\tau = \mathcal{O}(1)$ and $\theta = \Theta(\frac{1}{\sqrt{n}})$, we have $\mathcal{O}(q^{\frac{3}{2}} (\log p)^3 \sqrt{n} \log \frac{n}{\epsilon})$ iteration complexity for $q \geq 1$ and $p \geq e$. These are the best known complexity results for such methods.

5. Concluding remarks

Motivated by recent works of Roos et al. ([2]) we propose a new eligible kernel function which generalizes the kernel function in [2] and define a primal-dual IPM for LO problems and improves the iteration complexity of the algorithm in [2]. Furthermore we have $\mathcal{O}((\log p)^{\frac{5}{2}} \sqrt{n} \log n \log \frac{n}{\epsilon})$ and $\mathcal{O}(q^{\frac{3}{2}} (\log p)^3 \sqrt{n} \log \frac{n}{\epsilon})$ complexity bound for large- and small-update methods which are the best known iteration bounds for such methods.

Future research might focus on the extension to semidefinite optimization and symmetric cone optimization. Numerical tests will be another topic for future research.

References

- [1] Y.Q. Bai, M. El Ghami, and C. Roos, *A new efficient large-update primal-dual interior-point method based on a finite barrier*, Siam J. on Optimization **13** (2003), 766-782.
- [2] Y.Q. Bai, M. El Ghami, and C. Roos, *A comparative study of kernel functions for primal-dual interior-point algorithms in linear optimization*, Siam J. on Optimization **15** (2004), 101-128.
- [3] Y.Q. Bai, J. Guo, and C. Roos, *A new kernel function yielding the best known iteration bounds for primal-dual interior-point algorithms*, ANZIAM J. **49** (2007), 259-270.

- [4] Y.Q. Bai, G.Q. Wang, and C. Roos, *Primal-dual interior point algorithms for second-order cone optimization based on kernel functions*, *Nonlinear Analysis* **70** (2009), 3584-3602.
- [5] N.K. Karmarkar, *A new polynomial-time algorithm for linear programming*, *Combinatorica* **4** (1984), 373-395.
- [6] M. Kojima, S. Mizuno, and A. Yoshise, *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.
- [7] J. Peng, C. Roos, and T. Terlaky, *Self-Regularity, A new paradigm for primal-dual interior-point algorithms*, Princeton University Press, 2002.
- [8] C. Roos, T. Terlaky, and J. Ph. Vial, *Theory and algorithms for linear optimization, An interior approach*, John Wiley & Sons, Chichester, U.K., 1997.
- [9] N.J. Todd, *Recent developments and new directions in linear programming*, in: M. Iri and K. Tanabe(Ed.), *Mathematical Programming : Recent developments and applications*, Kluwer Academic Publishers, Dordrecht, 1989, 109-157.

Gyeong-Mi Cho
Department of Software Engineering, Dongseo University,
Busan 617-716, Korea.
E-mail: gcho@dongseo.ac.kr