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

Gyeong-Mi Cho and Yong-Hoon Lee*

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([2, 11]). 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 $\mathcal{O}((\log p)\sqrt{n}\log n\log\frac{n}{\epsilon})$ and $\mathcal{O}((q\log p)^{\frac{3}{2}}\sqrt{n}\log\frac{n}{\epsilon})$ iteration bound for large- and small-update methods, respectively. These are currently the best known complexity results.

1. Introduction

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

$$\min\{c^T x : Ax = b, \ x \geq 0\}, \tag{1}$$

where $A \in \mathbf{R}^{m \times n}$ with rank$(A) = m$ and $c, x \in \mathbf{R}^n$, $b \in \mathbf{R}^m$. We know that its dual problem can be given as

$$\max\{b^T y : A^T y + s = c, \ s \geq 0\}, \tag{2}$$

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

Since Karmarkar's paper ([7]) 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 primal-dual IPM which is the most efficient for a computational point of view([2]). It is generally agreed that the iteration complexity of the algorithm is an appropriate measure for its efficiency.

Peng et al.([9]) proposed new variants of IPMs based on self-regular kernel functions and obtained so far the best known complexity, e.g. $\mathcal{O}(\sqrt{n}\log n\log\frac{n}{\epsilon})$ for large-update IPMs with a specific self-regular kernel function. Recently,

Roos et al.([3, 4, 5]) proposed new primal-dual IPMs for LO problems based on eligible kernel functions to improve the iteration bound for large-update methods from $\mathcal{O}(n \log \frac{n}{\epsilon})$ to $\mathcal{O}(\sqrt{n} \log n \log \frac{n}{\epsilon})$. They also proposed the framework for analyzing the algorithm based on four conditions on the kernel function to obtain the iteration bounds of the algorithm([4]) and generalized those methods to the general optimization.([6])

Motivated by their works, in this paper, we define a new kernel function which includes the kernel functions in [1] and [4] as special cases and propose a new primal-dual IPM for LO problems based on this function. For the complexity analysis we follow the framework in [4]. We show that the algorithm has $\mathcal{O}(\sqrt{n} \log n \log \frac{n}{\epsilon})$ and $\mathcal{O}(\sqrt{n} \log \frac{n}{\epsilon})$ iteration complexity for large- and small-update methods, respectively. 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, on the other hand, $e = 2.719 \cdots$ is the Napier's constant (or Euler's number). For notational convenience we denote the natural logarithm by log. For $f(x), g(x) : \mathbf{R}_{++} \to \mathbf{R}_{++}$, $f(x) = \mathcal{O}(g(x))$ 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 geneic algorithm. Without loss of generality, we assume that both (1) and (2) satisfy the interior-point condition (IPC)([10]), 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 [10]), finding an optimal solution of (1) and (2) is equivalent to solving the following system:

$$Ax = b, \ x \geq 0, \ A^T y + s = c, \ s \geq 0, \ xs = 0. \tag{3}$$

The basic idea of primal-dual IPMs is to replace the third equation in (3) by the parameterized equation $xs = \mu \mathbf{e}$ with $\mu > 0$. Now we consider the following system:

$$Ax = b, \ x > 0, \ A^T y + s = c, \ s > 0, \ xs = \mu \mathbf{e}. \tag{4}$$

If the IPC holds, then the system (4) has a unique solution for each $\mu > 0$([8]). We denote this solution as $(x(\mu), y(\mu), s(\mu))$ and call $x(\mu)$ the μ-center of (1) and $(y(\mu), s(\mu))$ the μ-center of (2). The set of μ-centers $(\mu > 0)$ is the central path of (1) and (2). The limit of the central path (as μ goes to zero) exists and since the limit point satisfies (3), it naturally yields optimal solutions for (1) and (2) ([10]). Primal-dual IPMs follow the central path approximately and approach the solution of (1) and (2) as μ goes to zero.

For given $(x, y, s) := (x^0, y^0, s^0)$ by applying Newton method to the system (4) we have the following Newton system

$$A\Delta x = 0, \ A^T\Delta y + \Delta s = 0, \ s\Delta x + x\Delta s = \mu\mathbf{e} - xs. \tag{5}$$

Since A has full row rank, the system (5) has a unique solution $(\Delta x, \Delta y, \Delta s)$ which is called the search direction. 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:

$$v := \sqrt{\frac{xs}{\mu}}, \ d_x := \frac{v\Delta x}{x}, \ d_s := \frac{v\Delta s}{s}. \tag{6}$$

Using (6), we can rewrite the system (5) as follows:

$$\bar{A}d_x = 0, \ \bar{A}^T\Delta y + d_s = 0, \ d_x + d_s = v^{-1} - v, \tag{7}$$

where $\bar{A} := \frac{1}{\mu}AV^{-1}X$, $V := \mathrm{diag}(v)$, and $X := \mathrm{diag}(x)$. Note that the right-side of the third equation in (7) equals the negative gradient of the logarithmic barrier function $\Psi_l(v)$, i.e.

$$d_x + d_s = -\nabla\Psi_l(v), \tag{8}$$

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:

$$\psi'(1) = \psi(1) = 0, \ \psi''(t) > 0, \ \forall t > 0, \ \lim_{t\to 0^+}\psi(t) = \lim_{t\to\infty}\psi(t) = \infty. \tag{9}$$

In this paper, for a kernel function $\psi(t)$, we define $\Psi(v) := \sum_{i=1}^{n}\psi(v_i)$ and replace $\Psi_l(v)$ in (8) 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

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

We use $\Psi(v)$ as the proximity function. Also, we define the norm-based proximity measure $\delta(v)$ as follows:

$$\delta(v) := \frac{1}{2}||\nabla\Psi(v)|| = \frac{1}{2}||d_x + d_s||. \tag{10}$$

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 (5) 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 (5) 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

Remark 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 new kernel function

In this section, we define a new kernel function and give its properties. Consider a function $\psi(t)$ as follows:

$$\psi(t) = \frac{t^2 - 1}{2} - \int_1^t p^{q(\frac{1}{\xi}-1)} d\xi, \ p \geq e, \ q \geq 1, \ t > 0. \tag{11}$$

Then we have the following:

$$\psi'(t) = t - p^{\ q(\frac{1}{t}-1)},$$

$$\psi''(t) = 1 + \frac{q \log p}{t^2} p^{q(\frac{1}{t}-1)}, \tag{12}$$

$$\psi'''(t) = -\frac{q(\log p)\ (q \log p + 2t)}{t^4} p^{q(\frac{1}{t}-1)}.$$

From (12) and (9), $\psi(t)$ is a kernel function and for $p \geq e, \ q \geq 1$,

$$\psi''(t) > 1, \ t > 0. \tag{13}$$

When $q = 1$, $\psi(t)$ is a kernel function in [1] and when $q = 1$ and $p = e$, $\psi(t)$ is a kernel function in [4].

Lemma 3.1. *For $\psi(t)$ as in (11), we have*
(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) $2(\psi''(t))^2 - \psi'(t)\psi^{(3)}(t) > 0$, $0 < t < 1$.

Proof: For (i), using (12), we have for $p \geq e, \ q \geq 1$ and $t > 0$,

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

Hence $\psi(t)$ is exponentially convex for $t > 0$.
For (ii), from (12), we easily see $\psi^{(3)}(t) < 0$.
For (iii), using (12), we have

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

For (iv), using (12), we have

$$2(\psi''(t))^2 - \psi'(t)\psi^{(3)}(t)$$
$$= t^{-4}[2t^4 + qt(\log p)(q(\log p) + 6t)p^{q(\frac{1}{t}-1)} + q(\log p)(q \log p - 2t)p^{\ 2q(\frac{1}{t}-1)}].$$

Since $t > 0$ and $q(\log p)p^{q(\frac{1}{t}-1)} > 0$, it is sufficient to show that for $0 < t < 1$, $q \geq 1$ and $p \geq e$,

$$t(q(\log p) + 6t) + (q \log p - 2t)p^{q(\frac{1}{t}-1)} > 0. \tag{14}$$

If $q \geq \frac{2}{\log p}$, then the relation (14) is satisfied for $0 < t < 1$ and $p \geq e$.

Now we consider the case of $1 \leq q < \frac{2}{\log p}$. When $0 < t \leq \frac{q \log p}{2}$, the relation (14) is clearly satisfied. Hence it is enough to show that the relation (14) is satisfied for $\frac{q \log p}{2} < t < 1$ and $1 \leq q < \frac{2}{\log p}$. For $\frac{q \log p}{2} < t < 1$ and $1 \leq q < \frac{2}{\log p}$,

$$t(q(\log p) + 6t) > (2t - q \log p)p^{q(\frac{1}{t}-1)}.$$

Since $2t - q \log p > 0$,

$$p^{q(\frac{1}{t}-1)} < \frac{qt \log p + 6t^2}{2t - q \log p}. \tag{15}$$

Since $t > 0$, the relation (15) is satisfied if

$$p^{\frac{q}{t}-q} < \frac{q \log p}{2 - \frac{q}{t} \log p}. \tag{16}$$

Let $z := \frac{q}{t}$. Then $\frac{1}{\log p} \leq z < \frac{2}{\log r}$ and the relation (16) can be written as follows:

$$p^{z-q} < \frac{q \log p}{2 - z \log p}. \tag{17}$$

The left-hand side of the relation (17) is monotone decreasing in q and the right-hand side is monotone increasing in q for $1 \leq q < \frac{2}{\log p}$,

$$1 - (\frac{2}{\log p} - z)p^{z-1} > 0. \tag{18}$$

Since the left-hand side of the relation (18) is monotone increasing in z and $\frac{1}{\log p} \leq z < \frac{2}{\log r}$,

$$1 - (\frac{2}{\log p} - z)p^{z-1} > 1 - \frac{e}{p \log p} > 0. \tag{19}$$

This completes the proof. $\square$

Remark 2. (*i*) By Lemma 2.4 in [4], if $\psi(t)$ satisfies Lemma 3.1 (*ii*) and (*iii*), then $\psi(t)$ satisfies

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

(*ii*) By Lemma 2.1.2 in [9], Lemma 3.1 (*i*) is equivalent to

$$\psi(\sqrt{t_1 t_2}) \leq \frac{1}{2}(\psi(t_1) + \psi(t_2)), \ t_1 > 0, \ t_2 > 0. \tag{20}$$

                                                                             $\square$

Lemma 3.2. *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) $\psi(t) \leq \frac{1+q\log p}{2}(t-1)^2$, $t \geq 1$,
(iii) $\psi(t) \leq \frac{t^2-1}{2}$, $t \geq 1$.

Proof: For (i), using the first condition of (9) and (13), we have

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

for $t \in (0,1)$ and

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

for $t \in (1,\infty)$ which proves the first inequality. The second inequality can be obtained in a similar fashion.

$$\psi(t) = \int_t^1 \int_\xi^1 \psi''(\zeta)d\zeta d\xi \leq \int_t^1 \int_\xi^1 \psi''(\xi)\psi''(\zeta)d\zeta d\xi$$
$$= -\int_t^1 \psi''(\xi)\psi'(\xi)d\xi = -\int_t^1 \psi'(\xi)d\psi'(\xi) = \frac{1}{2}(\psi'(t))^2,$$

for $t \in (0,1)$ and

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

for $t \in (1,\infty)$.

For (ii), using Taylor's Theorem, $\psi(1) = \psi'(1) = 0$, $\psi''' < 0$, and $\psi''(1) = 1 + q\log p$, we have for $p \geq e$, $q \geq 1$,

$$\psi(t) = \psi(1) + \psi'(1)(t-1) + \frac{1}{2}\psi''(1)(t-1)^2 + \frac{1}{3!}\psi'''(\xi)(t-1)^3$$
$$= \frac{1}{2}\psi''(1)(t-1)^2 + \frac{1}{3!}\psi'''(\xi)(t-1)^3$$
$$< \frac{1}{2}\psi''(1)(t-1)^2 = \frac{1+q\log p}{2}(t-1)^2,$$

for some ξ, $1 \leq \xi \leq t$.

For (iii), by the definition of $\psi(t)$ as in (11), $\psi(t) \leq \frac{t^2-1}{2}$, $t \geq 1$. This completes the proof. $\square$

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

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

(i) $\varrho(s) \leq 1 + \sqrt{2s}$, $s \geq 0$,

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

Proof: For (i), using Lemma 3.2 (i), we have $s = \psi(t) \geq \frac{(t-1)^2}{2}$. Then we have

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

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 $2z = p^{\ q(\frac{1}{t}-1)} - t$.

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

Hence we have

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

$\square$

4. Complexity analysis

In this paper, we replace the logarithmic barrier function $\Psi_l(v)$ in (8) with the function $\Psi(v)$ as follows:

$$d_x + d_s = -\nabla\Psi(v), \tag{21}$$

where $\Psi(v) = \sum_{i=1}^{n} \psi(v_i)$, where $\psi(t)$ is defined in (11). In the following, we compute a growth bound due to the update of the barrier paramter, a default step size and the decrease of the proximity function during an inner iteration and give the complexity results of the algorithm.

Using Remark 2 (i), we have the following lemma. The reader can refer to Theorem 3.2 in [4] for the proof.

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

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

$\square$

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

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

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 3.2 (ii), Lemma 3.3 (i), and $\Psi(v) \leq \tau$, we have

$$\Psi(v_+) \leq n\psi\left(\frac{1}{\sqrt{1-\theta}}\varrho\left(\frac{\Psi(v)}{n}\right)\right) \leq \frac{(1+q\log p)n}{2}\left(\frac{\varrho\left(\frac{\Psi(v)}{n}\right)}{\sqrt{1-\theta}} - 1\right)^2$$

$$= \frac{(1+q\log p)n}{2}\left(\frac{\varrho\left(\frac{\Psi(v)}{n}\right) - \sqrt{1-\theta}}{\sqrt{1-\theta}}\right)^2 \leq \frac{(1+q\log p)n}{2}\left(\frac{1+\sqrt{\frac{2\tau}{n}} - \sqrt{1-\theta}}{\sqrt{1-\theta}}\right)^2$$

$$\leq \frac{(1+q\log p)n}{2}\left(\frac{\theta + \sqrt{\frac{2\tau}{n}}}{\sqrt{1-\theta}}\right)^2 = \frac{1+q\log p}{2(1-\theta)}\left(\sqrt{n}\theta + \sqrt{2\tau}\right)^2,$$

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 3.3 (i), $\Psi(v) \leq \tau$, and Lemma 3.2 (iii),

$$\Psi(v_+) \leq n\psi\left(\frac{1}{\sqrt{1-\theta}}\varrho\left(\frac{\Psi(v)}{n}\right)\right) \leq n\psi\left(\frac{1+\sqrt{\frac{2\Psi(v)}{n}}}{\sqrt{1-\theta}}\right)$$

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

This completes the proof. $\hfill\square$

Denote

$$\tilde{\Psi}_0 := \frac{1+q\log p}{2(1-\theta)}\left(\sqrt{n}\theta + \sqrt{2\tau}\right)^2, \ \bar{\Psi}_0 := \frac{2\tau + 2\sqrt{2n\tau} + n\theta}{2(1-\theta)}. \tag{22}$$

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 Ψ_k, $k = 1, 2, \cdots$. Then we have

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

where $\tilde{\Psi}_0$ and $\bar{\Psi}_0$ are defined in (22). 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. Let K denote the total number of inner iterations per outer iteration. Then we have

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

Remark 3. For large-update method with $\tau = \mathcal{O}(n)$ and $\theta = \Theta(1)$, $\bar{\Psi}_0 = \mathcal{O}(n)$ and for small-update method with $\tau = \mathcal{O}(1)$ and $\theta = \Theta(\frac{1}{\sqrt{n}})$, $\tilde{\Psi}_0 = \mathcal{O}(q\log p)$.

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 (6), 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 (20), we have

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

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

$$f_1(\alpha) := \frac{1}{2}\left(\Psi(v + \alpha d_x) + \Psi(v + \alpha d_s)\right) - \Psi(v). \qquad (23)$$

Obviously, we have

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

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

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

where $[d_x]_i$ and $[d_s]_i$ denote the ith components of the vectors d_x and d_s, respectively. Using (21) and (10), we have

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

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

$$f_1''(\alpha) = \frac{1}{2}\sum_{i=1}^{n}\left(\psi''(v_i + \alpha[d_x]_i)[d_x]_i^2 + \psi''(v_i + \alpha[d_s]_i)[d_s]_i^2\right). \qquad (24)$$

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

Lemma 4.3. *Let $\delta(v)$ be as defined in (10). Then we have*

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

Proof: Using Lemma 3.2 (i), we have

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

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

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

Lemma 4.4. (Lemma 4.2 in [4]) If the step size α satisfies the inequality

$$-\psi'(v_{min} - 2\alpha\delta) + \psi'(v_{min}) \leq 2\delta, \tag{25}$$

then we have

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

Lemma 4.5. (Lemma 4.3 in [4]) Let $\rho : [0,\infty) \to (0,1]$ 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 (25) is given by

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

Lemma 4.6. (Lemma 4.4 in [4]) Let ρ and $\hat{\alpha}$ be as defined in Lemma 4.5. Then we have

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

Define

$$\bar{\alpha} := \frac{1}{\psi''(\rho(2\delta))}. \tag{26}$$

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

Lemma 4.7. (Theorem 4.6 in [4]) If the step size $\bar{\alpha}$ is as in (26), then

$$f(\bar{\alpha}) \leq -\frac{\delta^2}{\psi''(\rho(2\delta))}. \tag{27}$$

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

Lemma 4.8. (Lemma 4.7 in [4]) The right-hand side of (27) is monotonically decreasing in δ.

Theorem 4.9. *Let $\bar{\alpha}$ be as defined in (26) and $\tau \geq 1$. Then*

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

Proof: Using Lemma 4.7, Lemma 4.8 and Lemma 4.3, we have

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

By Lemma 3.3 (ii), Lemma 3.1 (ii) and (12),

$$\psi''(\rho(\sqrt{2\Psi}))$$

$$\leq \quad \psi''\left(\frac{q\log p}{q\log p + \log(1 + 2\sqrt{2\Psi}\,)}\right)$$

$$= \quad 1 + q(\log p)\left(1 + \frac{1}{q\log p}\log(1 + 2\sqrt{2\Psi}\,)\right)^2 (1 + 2\sqrt{2\Psi}\,).$$

Since $\Psi \geq \tau \geq 1$, $1 + 2\sqrt{2\Psi} \leq 3\sqrt{2\Psi}$. Hence we have

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

$$\leq \quad -\frac{1}{2}\frac{\sqrt{\Psi}}{(1 + 3\sqrt{2})q(\log p)\left(1 + \frac{1}{q\log p}\log(1 + 2\sqrt{2\Psi_0}\,)\right)^2}.$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Lemma 4.10. (Lemma 1.3.2 in [9]) Let $t_0, t_1, \cdots, t_{\bar{K}}$ be a sequence of positive numbers such that

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

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

Lemma 4.11. *Let K be the total number of inner iterations in the outer iteration. Then we have*

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

Proof: Using Theorem 4.9 and Lemma 4.10 with

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

we have

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

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Theorem 4.12. *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*

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

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 $n\mu < \epsilon([10])$. 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

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

Remark 4. By Remark 3, for large-update methods with $\tau = \mathcal{O}(n)$ and $\theta = \Theta(1)$, by taking $q = \log(1 + 2\sqrt{2\bar{\Psi}_0}\ )$, the algorithm has $\mathcal{O}((\log p)\sqrt{n}\ (\log n)\ \log \frac{n}{\epsilon})$ iteration complexity for $p \geq e$.

For small-update methods, we have $\mathcal{O}((q\log p)^{\frac{3}{2}}\sqrt{n}\ \log \frac{n}{\epsilon})$ 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.([4]) we propose a new kernel function which generalizes the kernel function in [4] and define a primal-dual IPM for LO problems and improves the iteration complexity of the algorithm in [4]. We have $\mathcal{O}((\log p)\sqrt{n}\log n\log \frac{n}{\epsilon})$ and $\mathcal{O}((q\log p)^{\frac{3}{2}}\sqrt{n}\ \log \frac{n}{\epsilon})$ 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] M. Achache, *The application of a new class of parameterized kernel functions for monotone LCP*, preprint, 2013.

[2] E.D. Andersen, J. Gondzio, Cs. Meszaros, and X. Xu, *Implementation of interior point methods for large scale linear programming*, in: T. Terlaky(Ed.), Interior point methods of mathematical programming, Kluwer Academic Publisher, The Netherlands, 189-252, 1996.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

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

[8] 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, 29-47, 1989.

[9] J. Peng, C. Roos, and T. Terlaky, *Self-Regularity, A new paradigm for primal-dual interior-point algorithms*, Princeton University Press, 2002.

[10] C. Roos, T. Terlaky, and J. Ph. Vial, *Theory and algorithms for linear optimization, An interior approach*, John Wiley & Sons, Chichester, U.K., 1997.

[11] 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, 109-157, 1989.

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

Yong-Hoon Lee
Department of Mathematics, Pusan National University, Busan 609-735,
Republic of Korea
E-mail address: yhlee@pusan.ac.kr