IJAMML 3:2 (2015) 89-111 December 2015
ISSN: 2394-2258
Available at http://scientificadvances.co.in
DOI: http://dx.doi.org/10.18642/ijamml_7100121393

ITERATION BOUNDS OF INTERIOR-POINT METHODS FOR LINEAR OPTIMIZATION BASED ON A NEW KERNEL FUNCTION

Liying Liu[a], Tao Hua[b] and Bo Sun[a]

[a]College of Mathematics Science, Liaocheng University, Liaocheng 252000, P. R. China

[b]Network Information Center, Liaocheng University, Liaocheng 252000, P. R. China

Abstract

In this paper, a new kernel function, which yields an efficient primal-dual interior-point method, is introduced. Iteration bounds for large-update methods are derived, namely, $O(n^{\frac{2}{3}} \log \frac{n}{\epsilon})$, which significantly improve the classical iteration bounds. Iteration bounds for small-update methods are derived, namely, $O(\sqrt{n} \log \frac{n}{\epsilon})$, which are currently the best known bounds.

Keywords: linear optimization, primal-dual interior-point methods, large-update methods, small-update methods, kernel function.

*Corresponding author.

E-mail address: lyliu001@163.com (Liying Liu).

Project supported by the Development Project Foundation for Science Research of Shandong Education Department (No. J13LI10), and the Foundation of Liaocheng University (No. 318011303).

2010 Mathematics Subject Classification: 90C05, 90C33, 90C51.
Submitted by Jianqiang Gao.
Received September 4, 2014

1. Introduction

After the landmark paper of Karmarkar [1], linear optimization (LO) revitalized as an active area of research. The interior-point methods (IPMs) is one of the most efficient methods for solving LO problems. For a survey, we refer to recent books on the subject [2-4]. It is generally agreed that these IPMs are most efficient from a computational point of view (see Andersen et al. [5]). This is especially true for so-called primal-dual large-update methods, which are the most efficient methods in practice [5].

Most of the classical primal-dual interior point methods (IPMs) for LO are based on the use of the logarithmic barrier function [5]. If n denotes the number of inequalities in the problem, then the theoretical iteration bounds is, $O(\sqrt{n}\log\frac{n}{\epsilon})$ and $O(n\log\frac{n}{\epsilon})$, for the so-called small-update methods and the so-called large-update methods, where ϵ represents the desired accuracy of the solution. However, in practice, the so-called large-update methods are much more efficient than the so-called small-update methods. The significant gap between the practical behaviour and the theoretical performance results of the large- and small-update methods, which has been referred to as the irony of IPMs [6], is being an open difficult problem.

To narrow the gap, Peng et al. [7, 8] introduced self-regular barrier functions for primal-dual IPMs for LO and obtained so far the best complexity result, $O(\sqrt{n}\log n\log\frac{n}{\epsilon})$, for large-update primal-dual IPMs with some specific self-regular barrier functions. Then, Bai et al. [9-14] proposed new primal-dual IPMs based on various kernel functions to improve the iteration bounds for large-update methods from $O(n\log\frac{n}{\epsilon})$ to $O(\sqrt{n}\log n\log\frac{n}{\epsilon})$. Especially, Bai et al. [10] proposed a new class of barrier functions which are called eligible, neither logarithmic barrier

nor self-regular, and they presented a unified computational framework for the complexity analysis of the algorithm. And they greatly simplified the analysis of IMPs. Recently, in order to resolve the open difficulty, which is the gap between the practical behaviour and the theoretical performance results of the large-and small-update methods, many works have been done, and a large amount of results have been reported, see [15-18].

Motivated by their works, we define a new kernel function which yielding a new primal-dual interior point algorithm based on this kernel function for LO. It shows that the iteration bounds for the new small-update methods is $O(sqrtn \log \frac{n}{\epsilon})$, which is currently the best known bounds. The iteration bounds for the new large-update methods is $O(n^{\frac{2}{3}} \log \frac{n}{\epsilon})$, which significantly improve the classical iteration bounds.

The paper is organized as follows. In Section 2, we briefly review the basic concepts on IPMs for LO, such as the central path, the new search directions. The generic polynomial interior-point algorithm for LO is also presented. In Section 3, we define a new kernel function and give its properties, which play a crucial role in the complexity analysis of algorithm. The complexity analysis is also performed in Section 4. Finally, Section 5 contains some concluding remarks and directions for future research.

Some notation used throughout the paper is as follows. First, $\mathbf{R}^n$, $\mathbf{R}^n_+$, and $\mathbf{R}^n_{++}$ denote the set of vectors with n components, the set of nonnegative vectors, and the set of positive vectors, respectively. The 2-norm and the infinity norm are denoted by $\|\cdot\|$ and $\|\cdot\|_\infty$, respectively. If $x, s \in \mathbf{R}^n$, then xs denotes the componentwise (or Hadamard) product of the vectors x and s. Furthermore, e denotes the all-one vector of length n. If $z \in \mathbf{R}^n_+$ and $f : \mathbf{R}_+ \to \mathbf{R}_+$, then $f(z)$ denotes the vector in $\mathbf{R}^n_+$

whose i-th component is $f(z_i)$ with $1 \le i \le n$. We write $f(x) = O(g(x))$ if $f(x) \le cg(x)$ for some positive constant c and $f(x) = \Theta(g(x))$ if $c_1 g(x) \le f(x) \le c_2 g(x)$ for positive constants c_1 and c_2.

2. Preliminaries

2.1. The central path

We consider the standard linear optimization

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

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

$$\max\{b^T y : A^T y + s = c,\ s \ge 0\}. \tag{D}$$

It is well known that finding an optimal solution of (P) and (D) is equivalent to solving the following system:

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

The basic idea of primal-dual IPMs is to replace the third equation in (1), the so-called complementarity condition for (P) and (D), by the parameterized equation $xs = \mu e$, with $\mu > 0$. Thus, we consider the system

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

Due to the last equation, any solution (x, y, s) of (2) will satisfy $x > 0$ and $s > 0$. So a solution exists only if (P) and (D) satisfy the interior-point condition (IPC), i.e., there exists (x^0, s^0, y^0) such that

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

Surprisingly enough, if the IPC is satisfied, then there exists a solution, for each $\mu > 0$, and this solution is unique. It is denoted as $(x(\mu), y(\mu), s(\mu))$ and we call $x(\mu)$ the μ-center of (P) and $(y(\mu), s(\mu))$ the μ-center of (D). The set of μ-centers (with μ running through all positive real numbers) gives a homo type path, which is called the central path of (P) and (D). The relevance of the central path for LO was recognized first by Sonnevend [19] and Megiddo [20]. If $\mu \to 0$, then the limit of the central path exists and since the limit points satisfy the complementarity condition, the limit yields optimal solutions for (P) and (D).

From a theoretical point of view, the IPC can be assumed without loss of generality. In fact we may, and will, assume that $x^0 = s^0 = e$. In practice, this can be realized by embedding the given problems (P) and (D) into a homogeneous self-dual problem which has two additional variables and two additional constraints. For this and the other properties mentioned above, see [2].

2.2. The search directions

IPMs follow the central path approximately. We briefly describe the usual approach. Without loss of generality, we assume that $(x(\mu), y(\mu), s(\mu))$ is known for some positive μ. For example, due to the above assumption, we may assume this for $\mu = 1$, with $x(1) = s(1) = e$. We then decrease μ to $\mu := (1 - \theta)\mu$ for some fixed $\theta \in (0, 1)$ and we solve the following Newton system:

$$\begin{aligned} A\Delta x &= 0, \\ A^T \Delta y + \Delta s &= 0, \\ s\Delta x + x\Delta s &= \mu e - xs \end{aligned} \tag{4}$$

This system uniquely defines a search direction $(\Delta x, \Delta s, \Delta y)$. By taking a step along the search direction, with the step size defined by some line search rules, one constructs a new triple (x, y, s). If necessary, we repeat the procedure until we find iterates that are "close" to $(x(\mu), y(\mu), s(\mu))$. Then μ is again reduced by the factor $1-\theta$ and we apply Newton's method targeting at the new μ-centers, and so on. This process is repeated until μ is small enough, say until $n\mu \leq \epsilon$; at this stage we have found an ϵ-solution of the problems (P) and (D).

Let us mention that in practice many LO solvers use the ϵ- solution to construct a basic solution and then produce an optimal basic solution by crossing over to the simplex method. An alternative way is to apply a rounding procedure as described by Ye and Mehrotra [21, 22] and Roos et al. [2].

The result of a Newton step with step size α is denoted as

$$x_+ := x + \alpha\Delta x, \quad s_+ := s + \alpha\Delta s, \tag{5}$$

where the step size α satisfy $(0 \leq \alpha \leq 1)$.

Now we introduce the scaled vector v and the scaled search directions d_x and d_s as follows:

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

The system (4) can be rewritten as follows:

$$\begin{aligned} \overline{A}d_x &= 0, \\ \overline{A}^T\Delta y + d_s &= 0, \\ d_x + d_s &= v^{-1} - v, \end{aligned} \tag{7}$$

where $\overline{A} := \frac{1}{\mu} AV^{-1}X$, $V := \text{diag}(v)$, $X := \text{diag}(x)$. 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 the barrier function $\Psi_l(v) : \mathbf{R}^n_{++} \to \mathbf{R}_+$ is defined as follows:

$$\Psi_l(v) := \Psi_l(x, s; \mu) := \sum_{i=1}^{n} \psi_l(v_i), \tag{9}$$

$$\psi_l(v_i) = \frac{v_i^2 - 1}{2} - \log v_i. \tag{10}$$

We call $\psi_l(t)$ the kernel function of the logarithmic barrier function $\Psi_l(v)$. In this paper, we replace $\psi_l(t)$ with a new kernel function $\psi(t)$ which will be defined in Section 3. Note that the pair (x, s) coincides with the μ-center $(x(\mu), s(\mu))$ if and only if $v = \mathbf{e}$. One can easily verify that the kernel function $\psi(t)$ as defined by (10) is a strictly convex function which is defined for any $t \in \mathbf{R}_{++}$ and which is minimal at $t = 1$, whereas the minimal value equals 0.

2.3. The generic interior-point algorithm for LO

The generic form of this algorithm is shown in Figure 1.

It is clear from the description that the closeness of (x, s) to $(x(\mu), s(\mu))$ is measured by the value of $\Psi_l(v)$, with $\tau > 0$ as a threshold value. If $\Psi_l(v) \leq \tau$, then we start a new *outer iteration* by performing a μ-update, otherwise we enter an *inner iteration* by computing the search directions at the current iterates with respect to the current value of μ and apply (5) to get new iterates. If it is necessary, we repeat the procedure until we find iterates that are in the neighbourhood of $(x(\mu), s(\mu))$. Then μ is again reduced by the factor $1 - \theta$ with $0 < \theta < 1$ and we apply Newton's method targeting at the new μ-centers, and so on. This process is repeated until μ is small enough, say until $n\mu < \epsilon$, at this stage we have found an ϵ-apporximate solution of LO.

The parameters τ, θ and the step size α should be chosen in such a way that the algorithm is optimized in the sense that the number of iterations required by the algorithm is as small as possible. The choice of the so-called barrier update parameter θ plays an important role both in theory and practice of IPMs. Usually, if θ is a constant independent of the dimension n of the problem, for instance $\theta = \frac{1}{2}$, then we call the algorithm a *large-update* (or *long-step*) method. If θ depends on the dimension of the problem, such as $\theta = \frac{1}{\sqrt{n}}$, then the algorithm is named a *small-update* (or *short-step*) method.

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 made, 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.

Generic Interior-Point Algorithm for LO

Input:

A threshold parameter $\tau > 0$;

an accuracy parameter $\varepsilon > 0$;

a fixed barrier update parameter θ, $0 < \theta < 1$;

begin

$x := e$; $s := e$; $\mu := 1$;

while $n\mu \geq \varepsilon$ **do**

begin

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

while $\Psi_l(x, s; \mu) > \tau$ **do**

begin

solve system (7) via (6) to obtain $(\Delta x, \Delta y, \Delta s)$;

choose a suitable step size α;

$x := x + \alpha\Delta x$;

$y := y + \alpha\Delta y$;

$s := s + \alpha\Delta s$;

$v := \sqrt{\frac{xs}{\mu}}$;

end

end

end

Figure 1. Generic algorithm.

3. The New Kernel Function and its Properties

In this section, we define a new kernel function and give its properties which are essential to our complexity analysis. We call $\psi : R_{++} \to R_+$ a kernel function if ψ is twice differentiable and satisfies the following conditions [10]:

$$\begin{aligned} \psi'(1) = \psi(1) = 0, \\ \psi''(t) > 0, \\ \lim_{t\to 0^+} \psi(t) = \lim_{t\to\infty} \psi(t) = 0. \end{aligned} \tag{11}$$

Now, we define a new function $\psi(t)$ as follows:

$$\psi(t) = mt^2 + mt^{-2} - 2m, \quad t > 0. \tag{12}$$

For the convenience of reference, we gives the first three derivatives with respect to t as follows:

$$\begin{aligned} \psi'(t) &= 2mt - 2mt^{-3}, \\ \psi''(t) &= 2m + 6mt^{-4}, \\ \psi'''(t) &= -24mt^{-5}. \end{aligned} \tag{13}$$

Obviously, $\psi(t)$ is a kernel function and

$$\psi''(t) = 2m + 6mt^{-4} > 2m. \tag{14}$$

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

$$d_x + d_s = -\nabla\Psi(v). \tag{15}$$

where $\Psi(v) = \sum_{i=1}^{n} \psi(v_i)$, $\psi(t)$ is defined in (12). Hence the new search direction $(\Delta x, \Delta y, \Delta s)$ is obtained by solving the following modified Newton system:

$$A\Delta x = 0,$$

$$A^T\Delta y + \Delta s = 0, \tag{16}$$

$$s\Delta x + x\Delta s = -\mu v\nabla\Psi(v).$$

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 $\overline{A}$. Since d_x and d_s are orthogonal, we have

$$d_x = d_s = 0 \Leftrightarrow \nabla\Psi(v) = 0 \Leftrightarrow v = e \Leftrightarrow \Psi(v) = 0 \Leftrightarrow x = x(\mu),\ s = s(\mu). \tag{17}$$

We use $\Psi(v)$ as the proximity function to measure the distance between the current iterate and the μ-center for given $\mu > 0$. We also define the norm-based proximity measure $\delta(v) : \mathbf{R}^n_{++} \to \mathbf{R}_+$, as follows:

$$\delta(v) := \frac{1}{2}\|\nabla\Psi(v)\| = \frac{1}{2}\|d_x + d_s\|. \tag{18}$$

Lemma 3.1. *For $\psi(t)$, we have*

(i) *$\psi(t)$ is exponentially convex for all $t > 0$, that is,*

$$\psi(\sqrt{t_1t_2}) \le \frac{1}{2}(\psi(t_1) + \psi(t_2)).$$

(ii) *$\psi''(t)$ is monotonically decreasing for all $t > 0$.*

(iii) *$t\psi''(t) - \psi'(t) > 0$ for all $t > 0$.*

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

Proof. For (i), using (13), we have

$$t\psi''(t) + \psi'(t) = 2mt + 6mt^{-3} + 2mt - 2mt^{-3} = 4mt + 4mt^{-3}, \quad t > 0.$$

Let

$$g(t) = 4mt + 4mt^{-3}.$$

Then

$$g'(t) = 4m - 12mt^{-4},$$

$$g''(t) = 48mt^{-5} > 0, \quad t > 0.$$

Let $g'(t) = 0$, we get $t = \sqrt[4]{3}$. Since $g(t)$ is strictly convex and has a global minimum, $g(\sqrt[4]{3}) > 0$. And by Lemma 2.1.2 in [8], we have the result.

For (ii), using (13), we have $\psi'''(t) < 0$, so we have the result.

For (iii), using (13), we have $t\psi''(t) - \psi'(t) = 2mt + 6mt^{-3} - 2mt + 2mt^{-3} = 8mt^{-3}$, $t > 0$.

For (iv), using Lemma 2.4 in [10], (ii) and (iii), we have the result. This completes the proof.

Lemma 3.2. *For $\psi(t)$, we have*

$$m(t-1)^2 \le \psi(t) \le \frac{1}{4m}\psi'(t)^2, \quad t > 0, \tag{19}$$

$$\psi(t) \le 4m(t-1)^2, \quad t > 0. \tag{20}$$

Proof. For (19), using (11) and (14), we have

$$\psi(t) = \int_1^t \int_1^\xi \psi''(\zeta)d\zeta d\xi \ge 2m\int_1^t \int_1^\xi d\zeta d\xi = m(t-1)^2$$

$$\psi(t) = \int_1^t \int_1^\xi \psi''(\zeta)d\zeta d\xi$$

$$\le \frac{1}{2m}\int_1^t \int_1^\xi \psi''(\xi)\psi''(\zeta)d\zeta d\xi$$

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

$$= \frac{1}{2m}\int_1^t \psi'(\xi)d[\psi'(\xi)]$$

$$= \frac{1}{4m}\psi'(t)^2.$$

For (20), since $\psi(1) = \psi'(1) = 0$, $\psi'''(t) < 0$, $\psi''(1) = \frac{5}{2}$, and by using Taylor's theorem, we have

$$\psi(t) = \psi(1) + \psi'(1)(t-1) + \frac{1}{2}\psi''(1)(t-1)^2 + \frac{1}{6}\psi'''(\xi)(\xi-1)^3$$

$$= \frac{1}{2}\psi''(1)(t-1)^2 + \frac{1}{6}\psi'''(\xi)(\xi-1)^3$$

$$\leq \frac{1}{2}\psi''(1)(t-1)^2$$

$$= 4m(t-1)^2,$$

for some ξ, $1 \leq \xi \leq t$. This completes the proof.

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

Lemma 3.3. *For $\psi(t)$, we have*

$$\sqrt{\frac{s}{m} + 1} \leq \varrho(s) \leq 1 + \sqrt{\frac{s}{m}}, \quad s \geq 0, \tag{21}$$

$$\rho(s) \geq \sqrt[3]{\frac{m}{s+m}}, \quad s \geq 0. \tag{22}$$

Proof. For (21), let $s = \psi(t)$, $t \geq 1$, i.e., $\varrho(s) = t$, $t \geq 1$. By the definition of $\psi(t)$, we have

$$mt^2 = s - mt^{-2} + 2m.$$

Because $-mt^{-2}+2m$ is monotone increasing with respect to $t \geq 1$, we have

$$mt^2 \geq s+m,$$

this implies that

$$\varrho(s) = t \geq \sqrt{\frac{s}{m}+1}.$$

By (19), we have $s = \psi(t) \geq m(t-1)^2$, so

$$\varrho(s) = t \leq 1+\sqrt{\frac{s}{m}}.$$

For (22), let $z = -\frac{1}{2}\psi'(t)$, $t \in (0, 1]$. By the definition of $\rho : \rho(z) = t$, $t \in (0, 1]$, and $2z = -\psi'(t)$, we have

$$\frac{2m}{t^3} = 2z + 2mt.$$

Because $2mt$ is monotone increasing with respect to $t \in (0, 1]$, we have

$$\frac{2m}{t^3} \leq 2z + 2m,$$

this implies that

$$\sigma(z) = t \geq \sqrt[3]{\frac{m}{z+m}}.$$

This completes the proof.

Lemma 3.4. *Let* $\varrho : [0, +\infty) \to [1, +\infty)$ *be the inverse function of* $\psi(t)$, $t \geq 1$. *Then we have*

$$\Psi(\beta v) \leq n\psi(\beta\varrho(\frac{\Psi(v)}{n})), \quad v \in R_{++}, \beta \geq 1.$$

Proof. Using Lemma 3.1 (iv), and Theorem 3.2 in [10], we can get the result. This completes the proof.

Lemma 3.5. *Let* $0 \le \theta < 1$, $v_+ = \frac{v}{\sqrt{1-\theta}}$. *If* $\Psi(v) \le \tau$. *Then we have*

$$\Psi(v_+) \le \frac{4m}{1-\theta}\left(\sqrt{n}\theta + \sqrt{\frac{\tau}{m}}\right)^2$$

Proof. Since $\frac{1}{\sqrt{1-\theta}} \ge 1$ and $\varrho\left(\frac{\Psi(v)}{n}\right) \ge 1$, we have $\frac{\varrho\left(\frac{\Psi(v)}{n}\right)}{\sqrt{1-\theta}} \ge 1$. Using Lemma 3.4 with $\beta = \frac{1}{\sqrt{1-\theta}}$, (20), (21), and $\Psi(v) \le \tau$, we have

$$\begin{aligned}
\Psi(v_+) &\le n\psi\left(\frac{1}{\sqrt{1-\theta}}\varrho\left(\frac{\Psi(v)}{n}\right)\right) \\
&\le 4nm\left(\frac{1}{\sqrt{1-\theta}}\varrho\left(\frac{\Psi(v)}{n}\right) - 1\right)^2 \\
&= \frac{4mn}{1-\theta}\left(\varrho\left(\frac{\Psi(v)}{n}\right) - \sqrt{1-\theta}\right)^2 \\
&\le \frac{4mn}{1-\theta}\left(1 + \sqrt{\frac{\Psi(v)}{mn}} - \sqrt{1-\theta}\right)^2 \\
&\le \frac{4mn}{1-\theta}\left(\theta + \sqrt{\frac{\tau}{mn}}\right)^2 \\
&= \frac{4m}{1-\theta}\left(\sqrt{n}\theta + \sqrt{\frac{\tau}{m}}\right)^2,
\end{aligned}$$

where the last inequality holds from $1 - \sqrt{1-\theta} = \frac{\theta}{1+\sqrt{1-\theta}} \le \theta$, $0 \le \theta < 1$. This completes the proof.

Denote

$$\widetilde{\Psi}_0 = \frac{4m}{1-\theta}\left(\sqrt{n}\theta + \sqrt{\frac{\tau}{m}}\right)^2 = L(n, \theta, \tau), \tag{23}$$

then $\widetilde{\Psi}_0$ is an upper bound for $\Psi(v)$ during the process of the algorithm.

Remark 3.6. For large-update method by taking $\tau = O(n)$, $\theta = \Theta(1)$, $\widetilde{\Psi}_0 = O(n)$. For small-update method with $\tau = O(1)$, $\theta = \Theta(\frac{1}{\sqrt{n}})$, $\widetilde{\Psi}_0 = O(1)$.

4. Analysis of Algorithm

In this section, we compute a proper step size and the decrease of the proximity function during an inner iteration and give the complexity results of the algorithm. For fixed μ taking a step size α, we have new iterates

$$x_+ := x + \alpha\Delta x, \quad s_+ := s + \alpha\Delta s.$$

Using (6), we have

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

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

So 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 iterate and a current iterate for fixed μ. By Lemma 3.1 (i), we have

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

Therefore, we have $f(\alpha) \le f_1(\alpha)$, where

$$f_1(\alpha) = \frac{1}{2}(\Psi(v + \alpha d_x) + \Psi(v + \alpha d_s)) - \Psi(v). \tag{24}$$

Obviously $f(0) = f_1(0) = 0$. Taking the first two derivative of $f_1(\alpha)$ with respect to α, we have

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

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

Using (15) and (18), we have

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

For convenience we denote

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

Lemma 4.1. *Let* $\delta(v)$ *be as defined in* (18). *Then we have*

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

Proof. Using (19), we have

$$\Psi(v) = \sum_{i=1}^{n}\psi(v_i) \leq \sum_{i=1}^{n}\frac{1}{4m}\psi'(t)^2 = \frac{1}{4m}\|\nabla\Psi\|^2 = \frac{1}{m}\delta(v)^2,$$

so

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

This completes the proof.

Remark 4.2. Throughout the paper, we assume that $\tau \geq 1$. Using Lemma 4.1 and the assumption $\Psi(v) \geq \tau$, we have

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

From the Lemmas 4.1-4.3 in [10], we have the following Lemmas 4.3-4.5.

Lemma 4.3. *Let* $f_1(\alpha)$ *be as defined in* (24) *and* $\delta(v)$ *be as defined in* (18). *Then we have*

$$f_1''(\alpha) \le 2\delta^2 \psi''(v_{\min} - 2\alpha\delta).$$

Lemma 4.4. *If the step size* α *satisfies the inequality*

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

we have

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

Lemma 4.5. *Let* $\rho : [0, +\infty) \to (0, 1]$ *be the inverse function of* $-\frac{1}{2}\psi'(t)$ *for all* $t \in (0, 1]$. *Then the largest step size* $\overline{\alpha}$ *satisfying* (25) *is given by*

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

Lemma 4.6. *Let* ρ *and* $\overline{\alpha}$ *be as defined in Lemma* 4.5. *If* $\Psi(v) \ge \tau \ge 1$, *then we have*

$$\overline{\alpha} \ge \frac{1}{2(1 + 6\sqrt[3]{2})m\delta^{\frac{4}{3}}}.$$

Proof. Using Lemma 4.4 in [10], Lemma 4.2, and (22), we have

$$\begin{aligned}
\overline{\alpha} &\ge \frac{1}{\psi''(\rho(2\delta))} \\
&= \frac{1}{2m + 6m\dfrac{1}{\rho(2\delta)^2}} \\
&\ge \frac{1}{2m + 6m\left(\dfrac{2\delta + m}{m}\right)^{\frac{4}{3}}} \\
&\ge \frac{1}{2m\delta + 6m\left(\dfrac{m\delta + m}{m}\right)^{\frac{4}{3}}}
\end{aligned}$$

$$\geq \frac{1}{2m\delta^{\frac{4}{3}} + 12\sqrt[3]{2}m\delta^{\frac{4}{3}}}$$

$$= \frac{1}{2(1+6\sqrt[3]{2})m\delta^{\frac{4}{3}}}.$$

This completes the proof.

Denote

$$\widetilde{\alpha} = \frac{1}{2(1+6\sqrt[3]{2})m\delta^{\frac{4}{3}}}, \tag{26}$$

we have $\widetilde{\alpha}$ is the default step size and $\widetilde{\alpha} \leq \overline{\alpha}$.

From the Lemma 1.3.3 in [8], we can get the following lemma:

Lemma 4.7. *Suppose that* $h(t)$ *is a twice differentiable convex function with*

$$h(0) = 0,\ h'(0) > 0,$$

and $h(t)$ *attains its global minimum at* $t^* > 0$ *and* $h''(t)$ *is increasing with respect to* t. *Then for any* $t \in [0, t^*]$, *we have*

$$h(t) \leq \frac{th'(0)}{2}.$$

Let the univariate function h be such that

$$h(0) = f_1(0) = 0, \quad h'(0) = f_1'(0) = -2\delta^2, \quad h''(\alpha) = 2\delta^2\psi''(v_{\min} - 2\alpha\delta).$$

Lemma 4.8. *Let* $\widetilde{\alpha}$ *is the default step size as defined in* (26) *and* $\Psi(v) \geq 1$. *Then*

$$f(\overline{\alpha}) \leq -\frac{m^{-\frac{2}{3}}}{2(1+6\sqrt[3]{2})}\Psi(v)^{\frac{1}{3}}. \tag{27}$$

Proof. Using the Lemma 4.5 in [10], if the step size α satisfies $\alpha \le \widetilde{\alpha}$, then $f(\alpha) \le -\alpha\delta^2$. So, for $\widetilde{\alpha} \le \overline{\alpha}$, we have

$$\begin{aligned} f(\overline{\alpha}) &\le -\overline{\alpha}\delta^2 \\ &= -\frac{1}{2(1+6\sqrt[3]{2})m\delta^{\frac{4}{3}}}\delta^2 \\ &= -\frac{1}{2(1+6\sqrt[3]{2})m}\delta^{\frac{2}{3}} \\ &\le -\frac{m^{-\frac{2}{3}}}{2(1+6\sqrt[3]{2})}\Psi(v)^{\frac{1}{3}}. \end{aligned}$$

This completes the proof.

This expresses the decrease in $\Psi(v)$ during an inner iteration completely in ψ, its first and second derivations, and the inverse functions ρ and ϱ.

After the update of μ to $1-\mu$, we have $\Psi(v_+) \le \frac{4m}{1-\theta}(\sqrt{n}\theta + \sqrt{\frac{\tau}{m}})^2 = L(n, \theta, \tau)$. We need to count how many inner iterations are required to return to the situation where $\Psi(v) \le \tau$. We denote the value of $\Psi(v)$ after the μ update as Ψ_0, the subsequent values in the same outer iteration are denoted as Ψ_k, $k = 1, 2, \cdots, K$, where K denotes the total number of inner iterations in the outer iteration. The decrease in each inner iteration is given by (27). In [10], we can find the appropriate values of κ and $\gamma \in (0, 1]$,

$$\kappa = \frac{m^{-\frac{2}{3}}}{2(1+6\sqrt[3]{2})}, \quad \gamma = 1 - \frac{1}{3} = \frac{2}{3}.$$

Lemma 4.9. *Let K be the total number of inner iterations in the outer iteration. Then we have*

$$K \leq 3(1 + 6\sqrt[3]{2})m^{\frac{2}{3}}\Psi_0^{\frac{2}{3}}.$$

Proof. By Lemma 1.3.2 in [8], we have

$$K \leq \frac{\Psi_0^{\gamma}}{\kappa\gamma} = 3(1 + 6\sqrt[3]{2})m^{\frac{2}{3}}\Psi_0^{\frac{2}{3}}.$$

This completes the proof.

Theorem 4.10. *Let a LO problem be given, Ψ_0 as defined in* (23) *and $\tau \geq 1$. Then the total number of iterations to have an approximate solution with $n\mu < \epsilon$ is bounded by*

$$3(1 + 6\sqrt[3]{2})m^{\frac{2}{3}}\Psi_0^{\frac{2}{3}}\frac{\log\frac{n}{\epsilon}}{\theta}.$$

Proof. Recall that Ψ_0 is the upper bound according to (23). The number of outer iterations is bounded above by $\frac{1}{\theta}\log\frac{n}{\epsilon}$ (see [2], Section 17, page 116). Through multiplying the number of outer iterations by the number of inner iterations we get an upper bound for the total number of iterations, namely,

$$3(1 + 6\sqrt[3]{2})m^{\frac{2}{3}}\Psi_0^{\frac{2}{3}}\frac{\log\frac{n}{\epsilon}}{\theta}.$$

This completes the proof.

Remark 4.11. Taking $\tau = O(n)$, $\theta = \Theta(1)$ and $\widetilde{\Psi}_0 = O(n)$, we have $O(n^{\frac{2}{3}}\log\frac{n}{\epsilon})$ iterations complexity for large-update IPMs. Taking $\tau = O(1)$, $\theta = \Theta(\frac{1}{\sqrt{n}})$ and $\widetilde{\Psi}_0 = O(1)$, we have $O(n^{\frac{1}{2}}\log\frac{n}{\epsilon})$ iterations complexity for small-update IPMs, which as same as the iterations complexity for small-update IPMs.

Acknowledgement

The authors are sincerely grateful to the editors and the anonymous referees for the careful reading and useful suggestions, which have improved the manuscript of the present version.

References

[1] N. K. Karmarkar, A new polynomial-time algorithm for linear programming, Proceedings of the 16th Annual ACM Symposium on Theory of Computing 4 (1984), 373-395.

[2] C. Roos, T. Terlaky and J. Ph. Vial, Theory and Algorithms for Linear Optimization, An Interior-Point Approach, John Wiley & Sons, Chichester, UK, 1997.

[3] S. J. Wright, Primal-Dual Interior-Point Methods, SIAM, Philadelphia, 1997.

[4] Y. Ye, Interior Point Algorithms, Theory and Analysis, John Wiley & Sons, Chichester, UK, 1997.

[5] E. D. Andersen, J. Gondzio, Cs. Mészáros and X. Xu, Implementation of interior point methods for large scale linear programming, T. Terlaky, ed., Kluwer Academic Publishers, Dordrecht, The Netherlands, (1996), 189-252.

[6] Y. Ye, A Mathematical View of Interior-Point Methods in Convex Optimization, MPS/SIAM Ser. Optim. 3, SIAM, Philadelphia, 2001.

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

[8] J. Peng, C. Roos and T. Terlaky, Self-Regularity: A New Paradigm for Primal-Dual Interior-Point Algorithms, Princeton University Press, Princeton, NJ, 2002.

[9] 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 Journal on Optimization 13(3) (2003), 766-782.

[10] 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 Journal on Optimization 15(1) (2004), 101-128.

[11] Y. Q. Bai and C. Roos, A Primal-Dual Interior Point Method Based on a New Kernel Function with Linear Growth Rate, Proceedings of the 9th Australian Optimization Day, Perth, Australia, 2002.

[12] Y. Q. Bai and C. Roos, A polynomial-time algorithm for linear optimization based on a new simple kernel function, Optimization Methods and Software 18 (2003), 631-646.

[13] Y. Q. Bai, G. Lesaja, C. Roos, G. Q. Wang and M. El Ghami, A class of large-update and small-update primal-dual interior-point algorithms for linear optimization, J. Optim. Theory and Appl., DOI: 10.1007/s10957-008-9389-z., 2008.

[14] Y. Q. Bai, J. Guo and C. Roos, A new kernel function yielding the best known iteration bounds for primal-dual interior-point algorithms, Acta Mathematica Sinica, English Series 49 (2007), 259-270.

[15] 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 (2008), DOI: 10.1016/j.na.2008.07.016.

[16] M. El Ghamia, I. Ivanov, J. B. M. Melissen, C. Roos and T. Steihaug, A polynomial-time algorithm for linear optimization based on a new class of kernel functions, Journal of Computational and Applied Mathematics 224(2) (2009), 500-513.

[17] L. Liu and Sh. Li, A new kind of kernel function yielding good iteration bounds for primal-dual interior-point methods, Journal of Computational and Applied Mathematics 235 (2011), 2944-2955.

[18] M. R. Peyghami, S. F. Hafshejani and L. Shirvani, 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.

[19] G. Sonnevend, An analytic center for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming, System Modelling and Optimization: Proceedings of the 12th IFIP-Conference, Budapest, Hungary, 1985, Lecture Notes in Control and Inform. Sci. 84, A. Prekopa, J. Szelezsan and B. Strazicky, eds., Springer-Verlag, Berlin, (1986), 866-876.

[20] N. Megiddo, 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.

[21] Y. Ye, On the finite convergence of interior-point algorithms for linear programming, Math. Program. 57 (1992), 325-335.

[22] S. Mehrotra and Y. Ye, On finding the optimal facet of linear programs, Math. Program. 62 (1993), 497-515.

[23] C. F. Ma and X. H. Chen, The convergence of a one-step smoothing Newton method for P_0-NCP based on a new smoothing NCP-function, Journal of Computational and Applied Mathematics 216(1) (2008), 1-13.

■