STUDIA UNIV. BABEŞ–BOLYAI, INFORMATICA, Volume **XLVII**, Number 2, 2002

A WEIGHTED-PATH-FOLLOWING METHOD FOR LINEAR OPTIMIZATION

ZSOLT DARVAY

Abstract. In a recent paper [4] we introduced a new method for finding search directions for interior point methods (IPMs) in linear optimization (LO), and we developed a new polynomial algorithm for solving LO problems. It is well-known that using the self-dual embedding we can find a starting feasible solution, and this point will be on the central path. We proved [3] that this initialization method can be applied for the new algorithm as well. However, practical implementations often don't use perfectly centered starting points. Therefore it is worth analysing the case when the starting point is not on the central path. In this paper we develop a new weighted-path-following algorithm for solving LO problems. We conclude that following the central path yields to the best iteration bound in this case as well.

1. Introduction

In this paper we discuss a generalized form of path-following IPMs. The field of IPMs is an active research area, since Karmarkar [8] has developed the first IPM in 1984. For a survey of results see the following books [1, 2, 6, 11, 13, 14]. In this paper we generalize the algorithm introduced in [4], and we develop a new weighted-path-following algorithm. It is well known that with every algorithm which follows the central path we can associate a target sequence on the central path. This observation led to the concept of *target-following* methods introduced by Jansen et al. [7]. A survey of target-following algorithms can be found in [11] and [6]. Weighted-path-following methods can be viewed as a particular case of target-following methods. These methods were studied by Ding and Li [5] for primal-dual linear complementarity problems, and by Roos and den Hertog [10] for primal problems. In this paper we consider the LO problem in the following

2000 *Mathematics Subject Classification.* 90C05.
1998 *CR Categories and Descriptors.* G.1.6. [**Mathematics of Computing**]: Numerical Analysis – *Linear programming*.

standard form

$$(P) \qquad \begin{aligned} & min\ c^T x \\ s.t. \quad & Ax = b, \\ & x \geq 0, \end{aligned}$$

where $A \in \Re^{m \times n}$ with $rank(A) = m$, $b \in \Re^m$ and $c \in \Re^n$. The dual of this problem can be written in the following form

$$(D) \qquad \begin{aligned} & max\ b^T y \\ s.t. \quad & A^T y + s = c, \\ & s \geq 0. \end{aligned}$$

We assume that the *interior point condition* (IPC) holds for these probelms.

Assumption 1 (Interior point condition). There exist (x^0, y^0, s^0) such that

$$\begin{aligned} Ax^0 &= b, \qquad x^0 > 0, \\ A^T y^0 + s^0 &= c, \qquad s^0 > 0. \end{aligned}$$

Using the self-dual embedding method a larger LO problem can be constructed in such a way that the IPC holds for that problem. Hence, the IPC can be assumed without loss of generality. Finding the optimal solutions of both the original problem and its dual, is equivalent to solving the following system

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

where xs denotes the coordinatewise product of the vectors x and s, hence

$$xs = [x_1 s_1, x_2 s_2, \ldots, x_n s_n]^T .$$

We mention that in this paper for an arbitrary function f, and an arbitrary vector x we will use the notation

$$f(x) = [f(x_1), f(x_2), \ldots, f(x_n)]^T .$$

The first and the second equations of system (1) serve for maintaining feasibility, hence we call them the *feasibility conditions*. The last relation is the *complementarity condition*, which in IPMs is generally replaced by a parameterized equation, thus we obtain

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

where $\mu > 0$, and e is the n-dimensional all-one vector, hence $e = [1, 1, \ldots, 1]^T$. If the IPC is satisfied, then for a fixed $\mu > 0$ the system (2) has a unique solution.

This solution is called the *μ-center* (Sonnevend [12]), and the set of μ-centers for $\mu > 0$ formes the *central path.* The target-following approach starts from the observation that the system (2) can be generalized by replacing the vector μe with an arbitrary positive vector w^2. Thus we obtain the following system

$$\begin{aligned} Ax &= b, \qquad x \geq 0, \\ A^T y + s &= c, \qquad s \geq 0, \\ xs &= w^2, \end{aligned} \tag{3}$$

where $w > 0$. If the IPC holds then the system (3) has a unique solution. This feature was first proved by Kojima et al. [9]. Hence we can apply Newton's method for the system (3) to develop a primal-dual target-following algorithm. In the following section we present a new method for finding search directions by applying Newton's method for an equivalent form of system (3).

2. New Search-Directions

In this section we introduce a new method for constructing search directions by using the system (3). Let $\Re^+ = \{x \in \Re \mid x \geq 0\}$, and consider the function

$$\varphi \in C^1, \qquad \varphi : \Re^+ \to \Re^+.$$

Furthermore, suppose that the inverse function φ^{-1} exists. Then, the system (3) can be written in the following equivalent form

$$\begin{aligned} Ax &= b, \qquad x \geq 0, \\ A^T y + s &= c, \qquad s \geq 0, \\ \varphi(xs) &= \varphi(w^2), \end{aligned} \tag{4}$$

and we can apply Newton's method for the system (4) to obtain a new class of search directions. We mention that a direct generalization of the approach defined in [4] would be the following variant. The system (3) is equivalent to

$$\begin{aligned} Ax &= b, \qquad x \geq 0, \\ A^T y + s &= c, \qquad s \geq 0, \\ \varphi\left(\frac{xs}{w^2}\right) &= \varphi(e), \end{aligned} \tag{5}$$

and using Newton's method for the system (5) yields new search directions. For our purpose it is more convenient the first approach, hence in this paper we use the system (4). Let us introduce the vectors

$$v = \sqrt{xs} \quad \text{and} \quad d = \sqrt{xs^{-1}},$$

and observe that these notations lead to

$$d^{-1}x = ds = v. \tag{6}$$

Suppose that we have $Ax = b$, and $A^T y + s = c$ for a triple (x, y, s) such that $x > 0$ and $s > 0$, hence x and s are strictly feasible. Applying Newton's method for the system (4) we obtain

$$\begin{array}{rcl} A\Delta x &=& 0, \\ A^T \Delta y + \Delta s &=& 0, \\ s\varphi\prime\,(xs)\,\Delta x + x\varphi\prime\,(xs)\,\Delta s &=& \varphi(w^2) - \varphi\,(xs)\,. \end{array} \tag{7}$$

Furthermore, denote

$$d_x = d^{-1}\Delta x, \qquad d_s = d\Delta s,$$

and observe that we have

$$v(d_x + d_s) = s\Delta x + x\Delta s, \tag{8}$$

and

$$d_x d_s = \Delta x \Delta s. \tag{9}$$

Hence the linear system (7) can be written in the following equivalent form

$$\begin{array}{rcl} \bar{A} d_x &=& 0, \\ \bar{A}^T \Delta y + d_s &=& 0, \\ d_x + d_s &=& p_v, \end{array} \tag{10}$$

where

$$p_v = \frac{\varphi(w^2) - \varphi(v^2)}{v\varphi\prime(v^2)}, \tag{11}$$

and $\bar{A} = A diag(d)$. We also used the notation

$$diag(\xi) = \begin{bmatrix} \xi_1 & 0 & \dots & 0 \\ 0 & \xi_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \xi_n \end{bmatrix},$$

for any vector ξ. In the following section we will develop a new primal-dual weighted-path-following algorithm based on one particular search direction.

3. The Algorithm

In this section we let $\varphi(x) = \sqrt{x}$, and we develop a new primal-dual weighted-path-following algorithm based on the appropriate search directions. Thus, making the substitution $\varphi(x) = \sqrt{x}$ in (11) we get

$$p_v = 2(w - v). \tag{12}$$

Now for any positive vector v, we define the folowing proximity measure

$$\sigma(v, w) = \frac{\|p_v\|}{2\min(w)} = \frac{\|w - v\|}{\min(w)}, \tag{13}$$

where $\|\cdot\|$ is the Euclidean norm (l_2 norm), and for every vector ξ we denote $\min(\xi) = \min\{\xi_i \mid 1 \leq i \leq n\}$. We introduce another measure

$$\sigma_c(w) = \frac{max(w^2)}{min(w^2)},$$

where for any vector ξ we denote $\max(\xi) = \max\{\xi_i \mid 1 \leq i \leq n\}$. Observe that $\sigma_c(w)$ can be used to measure the distance of w^2 to the central path. Furthermore, let us introduce the notation

$$q_v = d_x - d_s,$$

observe that from (10) we get $d_x^T d_s = 0$, hence the vectors d_x and d_s are orthogonal, and thus we find that

$$\|p_v\| = \|q_v\|.$$

Consequently, the proximity measure can be written in the following form

$$\sigma(v, w) = \frac{\|q_v\|}{2\min(w)}, \tag{14}$$

thus we obtain

$$d_x = \frac{p_v + q_v}{2}, \qquad d_s = \frac{p_v - q_v}{2},$$

and

$$d_x d_s = \frac{p_v^2 - q_v^2}{4}. \tag{15}$$

Making the substitution $\varphi(x) = \sqrt{x}$ in (7) yields

$$\begin{array}{rcl} A\Delta x &=& 0, \\ A^T\Delta y + \Delta s &=& 0, \\ \sqrt{\dfrac{s}{x}}\Delta x + \sqrt{\dfrac{x}{s}}\Delta s &=& 2(w - \sqrt{xs}). \end{array} \tag{16}$$

Now we can define the algorithm.

Algorithm 3.1 *Suppose that for the triple (x^0, y^0, s^0) the interior point condition holds, and let $w^0 = \sqrt{x^0 s^0}$. Let $\epsilon > 0$ be the accuracy parameter, and $0 < \theta < 1$ the update parameter (default $\theta = \frac{1}{5\sqrt{\sigma_c(w^0)n}}$),*

begin
  $x := x^0;\ y := y^0;\ s := s^0;$
  $w := w^0;$
  while $x^T s > \epsilon$ **do begin**
    $w := (1-\theta)w;$
    Compute $(\Delta x, \Delta y, \Delta s)$ *from (16)*
    $x := x + \Delta x;$
    $y := y + \Delta y;$
    $s := s + \Delta s;$

end
end.

In the next section we shall prove that this algorithm is well defined for the default value of θ, and we will also give an upper bound for the number of iterations performed by the algorithm.

4. Convergence Analysis

In the first lemma of this section we prove that if the proximity measure is small enough, then the Newton process is strictly feasible. Denote $x_+ = x + \Delta x$ and $s_+ = s + \Delta s$ the vectors obtained by a full Newton step, and let $v = \sqrt{xs}$ as usual.

Lemma 4.1 *Let $\sigma = \sigma(v, w) < 1$. Then the full Newton step is strictly feasible, hence*

$$x_+ > 0 \quad \textit{and} \quad s_+ > 0.$$

Proof: For every $0 \leq \alpha \leq 1$ let $x_+(\alpha) = x + \alpha\Delta x$ and $s_+(\alpha) = s + \alpha\Delta s$. Hence

$$x_+(\alpha)s_+(\alpha) = xs + \alpha(s\Delta x + x\Delta s) + \alpha^2\Delta x\Delta s$$

Now using (8) and (9) we find that

$$x_+(\alpha)s_+(\alpha) = v^2 + \alpha v(d_x + d_s) + \alpha^2 d_x d_s,$$

and from (10) and (15) we obtain

$$x_+(\alpha)s_+(\alpha) = (1-\alpha)v^2 + \alpha(v^2 + vp_v) + \alpha^2\left(\frac{p_v^2}{4} - \frac{q_v^2}{4}\right).$$

Moreover (12) yields

$$v + \frac{p_v}{2} = w,$$

and thus

$$v^2 + vp_v = w^2 - \frac{p_v^2}{4}.$$

Consequently

$$x_+(\alpha)s_+(\alpha) = (1-\alpha)v^2 + \alpha\left(w^2 - (1-\alpha)\frac{p_v^2}{4} - \alpha\frac{q_v^2}{4}\right), \tag{17}$$

thus the inequality $x_+(\alpha)s_+(\alpha) > 0$ certainly holds if

$$\left\|(1-\alpha)\frac{p_v^2}{4} + \alpha\frac{q_v^2}{4}\right\|_\infty < min(w^2),$$

where $\|\cdot\|_\infty$ denotes the Chebychev norm (l_∞ norm). Using (13) and (14) we get

$$\left\|(1-\alpha)\frac{p_v^2}{4}+\alpha\frac{q_v^2}{4}\right\|_\infty \leq (1-\alpha)\frac{\|p_v^2\|_\infty}{4}+\alpha\frac{\|q_v^2\|_\infty}{4} \leq$$

$$\leq (1-\alpha)\frac{\|p_v\|^2}{4}+\alpha\frac{\|q_v\|^2}{4} = \sigma^2\min(w^2) < min(w^2).$$

Hence, for any $0 \leq \alpha \leq 1$ we have $x_+(\alpha)s_+(\alpha) > 0$. As a consequence we observe that the linear functions of α, $x_+(\alpha)$ and $s_+(\alpha)$ do not change sign on the interval $[0,1]$. For $\alpha = 0$ we have $x_+(0) = x > 0$ and $s_+(0) = s > 0$ thus we obtain $x_+(1) = x_+ > 0$ and $s_+(1) = s_+ > 0$, and this implies the lemma. ■

In the next lemma we prove that the same condition, namely $\sigma < 1$ is sufficient for the quadratic convergence of the Newton process.

Lemma 4.2 *Let $x_+ = x + \Delta x$ and $s_+ = s + \Delta s$ be the vectors obtaind after a full Newton step, $v = \sqrt{xs}$ and $v_+ = \sqrt{x_+ s_+}$. Suppose $\sigma = \sigma(v,w) < 1$. Then*

$$\sigma(v_+, w) \leq \frac{\sigma^2}{1+\sqrt{1-\sigma^2}}.$$

Thus $\sigma(v_+,w) < \sigma^2$, which means quadratic convergence of the Newton step.

Proof: From Lemma 4.1 we get $x_+ > 0$ and $s_+ > 0$. Now substitute $\alpha = 1$ in (17) and get

$$v_+^2 = w^2 - \frac{q_v^2}{4}. \tag{18}$$

Using (18) we obtain

$$\min(v_+^2) \geq \min(w^2) - \frac{\|q_v^2\|_\infty}{4} \geq \min(w^2) - \frac{\|q_v\|^2}{4} = \min(w^2)(1-\sigma^2),$$

and this realtion yields

$$\min(v_+) \geq \min(w)\sqrt{1-\sigma^2}. \tag{19}$$

Furthermore, from (18) and (19) we get

$$\sigma(v_+,w) = \frac{1}{\min(w)}\left\|\frac{w^2-v_+^2}{w+v_+}\right\| \leq \frac{\|w^2-v_+^2\|}{\min(w)\left(\min(w)+\min(v_+)\right)} \leq$$

$$\leq \frac{\|q_v^2\|}{\left(2\min(w)\right)^2\left(1+\sqrt{1-\sigma^2}\right)} \leq \frac{1}{1+\sqrt{1-\sigma^2}}\left(\frac{\|q_v\|}{2\min(w)}\right)^2 = \frac{\sigma^2}{1+\sqrt{1-\sigma^2}}.$$

Consequently, we have $\sigma(v_+,w) < \sigma^2$, and this implies the lemma. ■

In the following lemma we give an upper bound for the duality gap obtained after a full Newton step.

Lemma 4.3 *Let $\sigma = \sigma(v, w)$. Moreover, let $x_+ = x + \Delta x$ and $s_+ = s + \Delta s$. Then*

$$(x_+)^T s_+ = \|w\|^2 - \frac{\|q_v\|^2}{4},$$

hence $(x_+)^T s_+ \leq \|w\|^2$.

Proof: From

$$x_+ s_+ = w^2 - \frac{q_v^2}{4},$$

we obtain

$$(x_+)^T s_+ = e^T (x_+ s_+) = e^T w^2 - \frac{e^T q_v^2}{4} = \|w\|^2 - \frac{\|q_v\|^2}{4},$$

and this proves the lemma. ■

In the following lemma we discuss the influence on the proximity measure of the Newton process followed by a step along the weighted-path. We assume that each component of the vector w will be reduced by a constant factor $1 - \theta$.

Lemma 4.4 *Let $\sigma = \sigma(v, w) < 1$ and $w_+ = (1 - \theta)w$, where $0 < \theta < 1$. Then*

$$\sigma(v_+, w_+) \leq \frac{\theta}{1-\theta}\sqrt{\sigma_c(w)n} + \frac{1}{1-\theta}\sigma(v_+, w).$$

Furthermore, if $\sigma \leq \frac{1}{2}$, $\theta = \frac{1}{5\sqrt{\sigma_c(w)n}}$ and $n \geq 4$ then we get $\sigma(v_+, w_+) \leq \frac{1}{2}$.

Proof: We have

$$\sigma(v_+, w_+) = \frac{1}{\min(w_+)} \|w_+ - v_+\| \leq \frac{1}{\min(w_+)}\|w_+ - w\| + \frac{1}{\min(w_+)}\|w - v_+\| =$$

$$= \frac{1}{(1-\theta)min(w)}\|\theta w\| + \frac{1}{1-\theta}\sigma(v_+, w) \leq \frac{\theta}{1-\theta}\sqrt{\sigma_c(w)n} + \frac{1}{1-\theta}\sigma(v_+, w).$$

Thus the first part of the lemma is proved. Now let $\theta = \frac{1}{5\sqrt{\sigma_c(w)n}}$, observe that $\sigma_c(w) \geq 1$, and for $n \geq 4$ we obtain $\theta \leq \frac{1}{10}$. Furthermore, if $\sigma \leq \frac{1}{2}$ then from Lemma 4.2 we deduce $\sigma(v_+, w) \leq \frac{1}{4}$. Finally, the above relations yield $\sigma(v_+, w_+) \leq \frac{1}{2}$. The proof of the lemma is complete. ■

Observe that $\sigma_c(w) = \sigma_c(w^0)$ for all iterates produced by the algorithm. Thus, an immediate result of Lemma 4.4 is that for $\theta = \frac{1}{5\sqrt{\sigma_c(w^0)n}}$ the conditions $(x, s) > 0$ and $\sigma(v, w) \leq \frac{1}{2}$ are maintained throughout the algorithm. Hence the algorithm is well defined. In the next lemma we calculate an upper bound for the total number of iterations performed by the algorithm.

Lemma 4.5 *Assume that x^0 and s^0 are strictly feasible, an let $w^0 = \sqrt{x^0 s^0}$. Moreover, let x^k and s^k be the vectors obtained after k iterations. Then the inequality $(x^k)^T s^k \leq \epsilon$ is satisfied for*

$$k \geq \left\lceil \frac{1}{2\theta} \log \frac{(x^0)^T s^0}{\epsilon} \right\rceil .$$

Proof: After k iterations we get $w = (1-\theta)^k w^0$. Using Lemma 4.3 we find that

$$(x^k)^T s^k \leq \|w\|^2 = (1-\theta)^{2k} \|w^0\|^2 = (1-\theta)^{2k} (x^0)^T s^0,$$

hence $(x^k)^T s^k \leq \epsilon$ holds if

$$(1-\theta)^{2k} (x^0)^T s^0 \leq \epsilon.$$

Taking logarithms, we obtain

$$2k \log(1-\theta) + \log((x^0)^T s^0) \leq \log \epsilon.$$

Using the inequality $-\log(1-\theta) \geq \theta$ we deduce that the above relation holds if

$$2k\theta \geq \log((x^0)^T s^0) - \log \epsilon = \log \frac{(x^0)^T s^0}{\epsilon}.$$

The proof is complete. ■

For the default value of θ specified in Algorithm 3.1 we obtain the following theorem.

Theorem 4.6 *Suppose that the pair (x^0, s^0) is strictly feasible, an let $w^0 = \sqrt{x^0 s^0}$. If $\theta = \frac{1}{5\sqrt{\sigma_c(w^0) n}}$ then Algorithm 3.1 requires at most*

$$\left\lceil \frac{5}{2} \sqrt{\sigma_c(w^0) n} \log \frac{(x^0)^T s^0}{\epsilon} \right\rceil$$

iterations. For the resulting vectors we have $x^T s \leq \epsilon$. ■

5. Conclusion

In this paper we have developed a new weighted-path-following algorithm for solving LO problems. Our approach is a generalization of [4] for weighted-paths. We have transformed the system (3) in an equivalent form by introducing a function φ. We have defined a new class of search directions by applying Newton's method for that form of the weighted-path. Using $\varphi(x) = \sqrt{x}$ we have developed a new primal-dual weighted-path-following algorithm, and we have proved that this algorithm performs no more than

$$\left\lceil \frac{5}{2} \sqrt{\sigma_c(w^0) n} \log \frac{(x^0)^T s^0}{\epsilon} \right\rceil$$

iterations. Observe, that this means that the best bound is obtained by following the central path. Indeed, we have $\sigma_c(w^0) = 1$ in this case, and we get the well-known iteration bound

$$O\left(\sqrt{n}\log\frac{(x^0)^T s^0}{\epsilon}\right).$$

If the starting point is not perfectly centered, then $\sigma_c(w^0) > 1$ and thus the iteration bound is worse.

References

[1] N. Andrei. *Mathematical Programming. Interior Point Methods.* Editura Tehnică, Bucureşti, 1999. (In Romanian).

[2] Zs. Darvay. *Interior Point Methods in Linear Programming.* ELTE, Budapest, 1997. (In Hungarian).

[3] Zs. Darvay. A new algorithm for solving self-dual linear optimization problems. *Studia Universitatis Babeş-Bolyai, Series Informatica*, 47(1):15–26, 2002.

[4] Zs. Darvay. A new class of search directions for linear optimization. In *Proceedings of Abstracts, McMaster Optimizations Conference: Theory and Applications held at McMaster University Hamilton, Ontario, Canada*, page 18, August 1-3, 2002. Submitted to *European Journal of Operational Research.*

[5] J. Ding and T.Y. Li. An algorithm based on weighted logarithmic barrier functions for linear complementarity problems. *Arabian Journal for Science and Engineering*, 15(4):679–685, 1990.

[6] B. Jansen. *Interior Point Techniques in Optimization. Complexity, Sensitivity and Algorithms.* Kluwer Academic Pubishers, 1997.

[7] B. Jansen, C. Roos, T. Terlaky, and J.-Ph. Vial. Primal-dual target-following algorithms for linear programming. *Annals of Operations Research*, 62:197–231, 1996.

[8] N.K. Karmarkar. A new polinomial-time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.

[9] M. Kojima, N. Megiddo, T. Noma, and A. Yoshise. *A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems*, volume 538 of *Lecture Notes in Computer Science.* Springer Verlag, Berlin, Germany, 1991.

[10] C. Roos and D. den Hertog. A polinomial method of approximate weighted centers fo linear programming. Technical Report 89-13, Faculty of Technical Mathematics and Informatics, TU Delft, NL-2628 BL Delft, The Netherlands, 1994.

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

[12] Gy. Sonnevend. An "analytic center" for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming. In A. Prékopa, J. Szelezsán, and B. Strazicky, editors, *System Modelling and Optimization: Proceedings of the 12th IFIP-Conference held in Budapest, Hungay, September 1985*, volume 84 of *Lecture Notes in Control and Information Sciences*, pages 866–876. Springer Verlag, Berlin, West-Germany, 1986.

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

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

Department of Computer Science, Babeş-Bolyai University, 1 M. Kogălniceanu St., RO-3400 Cluj-Napoca, Romania

E-mail address: darvay@cs.ubbcluj.ro