School of Operations Research and Industrial Engineering
Cornell University
Ithaca, New York 14853-3801

Technical Report No. 1004

May 1992

A NOTE ON THE PRIMAL-DUAL
AFFINE SCALING ALGORITHMS

Levent TUNÇEL[1]

A NOTE ON THE PRIMAL-DUAL AFFINE SCALING ALGORITHMS

*Levent TUNÇEL**

School of Operations Research and Industrial Engineering
Cornell University
Ithaca, NY 14853-3801

May, 1992

Abstract

We study the primal-dual affine scaling algorithms for linear programs. Using an idea of Mizuno and Nagasawa and a new potential function we achieve the same complexity bounds they give. Our proofs are simpler and shorter. Our potential function seems to be more natural for this algorithm than the Tanabe-Todd-Ye potential function used by Mizuno and Nagasawa.

Keywords: Linear programming, interior point methods, potential function, primal-dual algorithms.

1

1 Introduction

In this paper, we study the primal-dual affine scaling algorithms for linear programs (see Monteiro, Adler and Resende [MAR90], Kojima, Megiddo, Noma and Yoshise [KMNY91]). We follow up on the recent algorithm by Mizuno and Nagasawa [MN92]. They suggest starting from an interior primal-dual solution that lies in a one-sided infinity neighborhood of the central path and determine the next iterate based on the potential function (Tanabe [Ta87], Todd and Ye[TY90])

$$\phi_q(x, s) := (q + n) \log(x^T s) - \sum_{j=1}^{n} \log(x_j s_j),$$

where q is a positive constant. Given the affine scaling direction at the current iterate, Mizuno and Nagasawa propose choosing the next iterate such that the value of this potential function does not increase (rather than finding the point which minimizes the potential function along the given search direction). Here we use the same idea with a potential function proposed by Tunçel [Tu92] for $q > 0$:

$$\psi_q(x, s) := (q + 1) \log(\frac{x^T s}{n}) - \log(\min_j\{x_j s_j\}).$$

We show that one can achieve the same complexity bounds achieved by Mizuno and Nagasawa [MN92] by keeping the potential function ψ_q constant. Since the potential function ψ_q has very nice properties related to the infinity-norm neighborhoods of the central path (see Tunçel [Tu92]), using the contours of ψ_q we can find a simple expression relating the decrease in the duality gap to the distance of the next iterate from the boundary of the feasible region. As a result our proofs are shorter and simpler.

2 Primal-Dual Affine Scaling Algorithm

We consider linear programming problems in the following primal (P) and dual (D) forms:

(P)

$$
\begin{array}{rrl}
\min & c^T x & \\
& Ax & = & b \\
& x & \geq & 0,
\end{array}
$$

(D)

$$\begin{aligned} \max \quad & b^T y \\ A^T y \ + \ & s \ = \ c \\ & s \ \geq \ 0, \end{aligned}$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$. Without loss of generality, we will assume A has full row rank and that there exist interior solutions for both problems, i.e.,

$$\mathcal{F}_o := \{(x, s) > 0 : x \in F(P), s \in F(D)\} \neq \emptyset,$$

where $F(P)$ and $F(D)$ denote the set of feasible solutions for the primal and dual problems respectively. Most of the time we will deal only with s as a dual feasible solution. So, whenever we say $s \in F(D)$, we mean that $s \geq 0$ and there exists a $y \in \mathbb{R}^m$ such that $A^T y + s = c$. Given a vector x, X will denote the diagonal matrix whose entries are the components of x, and e will denote the vector of ones. We will denote the components of a vector using subscripts and the iterate numbers using superscripts. Whenever we ignore the iterate numbers it will be clear from the context what the iterate number is.

The central path is given by the set of solutions to the following system of equalities for $\mu > 0$ (see for instance Megiddo [Me88]):

$$\begin{aligned} Ax \ &= \ b, \ x \geq 0 & (1) \\ A^T y + s \ &= \ c, \ s \geq 0 & (2) \\ Xs \ &= \ \mu e & (3) \end{aligned}$$

Given $\pi \in (0, 1)$ a one-sided infinity-norm neighborhood of the central path is defined (see for instance Mizuno, Todd and Ye [MTY90]) as:

$$\mathcal{N}(\pi) := \{(x, s) \in \mathcal{F}_o : \|Xs - \mu e\|_\infty^- \leq (1 - \pi)\}.$$

Here, for $v \in \mathbb{R}^n$, $\|v\|_\infty^- := -\min\{0, \min\{v_j\}\}$. Note that $\mathcal{N}(\pi)$ is a level set of ψ_o:

$$\mathcal{N}(\pi) = \{(x, s) \in \mathcal{F}_o : \psi_0(x, s) \leq \log(\frac{1}{\pi})\}.$$

Suppose we have an initial interior point solution $(x, s) \in \mathcal{F}_o$. Then the affine scaling direction $(\bar{d}x, \bar{d}s)$ can be generated by solving the following set of equalities (see Kojima, Mizuno and Yoshise [KMY88]):

3

$$
\begin{align}
A\bar{dx} &= 0 \tag{4}\\
A^T\bar{dy} + \bar{ds} &= 0 \tag{5}\\
S\bar{dx} + X\bar{ds} &= -Xs \tag{6}
\end{align}
$$

which is equivalent to solving

$$
\begin{align}
\bar{A}dx &= 0 \tag{7}\\
\bar{A}^T dy + ds &= 0 \tag{8}\\
dx + ds &= -X^{1/2}S^{1/2}e \tag{9}
\end{align}
$$

where $\bar{A} := AX^{1/2}S^{-1/2}$. The equivalence of the above two systems can be easily checked by substituting $dx = X^{-1/2}S^{1/2}\bar{dx}$, $ds = X^{1/2}S^{-1/2}\bar{ds}$. The solution of (7)-(9) is

$$
\begin{align}
dx &= -P_{\bar{A}}(X^{1/2}S^{1/2}e)\\
ds &= -(I - P_{\bar{A}})(X^{1/2}S^{1/2}e)
\end{align}
$$

where $P_{\bar{A}} := I - \bar{A}^T(\bar{A}\bar{A}^T)^{-1}\bar{A}$, the projection matrix into the null space of $\bar{A}$. Now, we describe a primal-dual affine scaling algorithm based on the potential function ψ_q:

Algorithm:

Given $(x^o, s^o) \in \mathcal{N}(\pi_0)$ with $(x^o)^T s^o \leq 2^t$, set $k := 0$.

While $((x^k)^T s^k > 2^{-t})$ do

    $(x, s) := (x^k, s^k)$

    compute $(\bar{dx}, \bar{ds})$ from (4)-(6)

    choose step size $\alpha_k \in (0, 1)$ such that $\psi_q(x + \alpha_k \bar{dx}, s + \alpha_k \bar{ds}) = \psi_q(x, s)$

4

$$(x^{k+1}, s^{k+1}) := (x + \alpha_k \bar{d}x, s + \alpha_k \bar{d}s)$$

$$k := k + 1$$

end

3 Analysis and Convergence Results

We define

$$\mu_k := \frac{(x^k)^T s^k}{n},$$

$$\pi_k := \frac{\min\{x_j^k s_j^k\}}{\mu_k}.$$

We assume that $\pi_0 \in (0, 1)$ is a constant independent of n and t. Let (x, s) be the current iterate with $(x, s) \in \mathcal{N}(\pi)$, and let $(x(\alpha), s(\alpha))$ denote the next iterate for a given step size $\alpha \in [0, 1]$, so that

$$x(\alpha) = x - \alpha X^{1/2} S^{-1/2} P_{\bar{A}} v \tag{10}$$
$$s(\alpha) = s - \alpha X^{-1/2} S^{1/2} (I - P_{\bar{A}}) v, \tag{11}$$

where $v := X^{1/2} S^{1/2} e$. We will denote $v_p := P_{\bar{A}} v$. From (10)-(11) we obtain

$$(x(\alpha))^T s(\alpha) = x^T s - \alpha e^T X^{1/2} S^{1/2} [v_p + (v - v_p)] + \alpha^2 v_p^T (v - v_p) \tag{12}$$
$$= x^T s - \alpha x^T s \tag{13}$$
$$= (1 - \alpha) x^T s. \tag{14}$$

We also have

$$x(\alpha)_j s(\alpha)_j = x_j s_j - \alpha \sqrt{x_j s_j} [(v_p)_j + (v - v_p)_j] + \alpha^2 (v_p)_j (v - v_p)_j \tag{15}$$
$$= x_j s_j - \alpha(x_j s_j) + \alpha^2 (v_p)_j (v - v_p)_j \tag{16}$$
$$= (1 - \alpha) x_j s_j + \alpha^2 (v_p)_j (v - v_p)_j \tag{17}$$

We define

$$\mu(\alpha) := \frac{x(\alpha)^T s(\alpha)}{n},$$

$$\pi(\alpha) := \frac{\min\{x_j(\alpha)s_j(\alpha)\}}{\mu(\alpha)}.$$

Then from (14) and (17) we have

$$\pi(\alpha) = \min\{\frac{x_j s_j}{\mu} + \frac{\alpha^2}{1-\alpha}\frac{(v_p)_j(v-v_p)_j}{\mu}\} \qquad (18)$$

Proposition 3.1. *Let* (x,s), $(x(\alpha),s(\alpha))$, π, $\pi(\alpha)$ *be defined as above. Then*

$$\psi_q(x(\alpha),s(\alpha)) = \psi_q(x,s) \iff (1-\alpha)^q = \frac{\pi(\alpha)}{\pi} \qquad (19)$$

Proof:

$$\psi_q(x(\alpha),s(\alpha)) = \psi_q(x,s)$$

$$\iff q\log(\mu(\alpha)) - \log(\frac{\min\{x_j(\alpha)s_j(\alpha)\}}{\mu(\alpha)}) = q\log(\mu) - \log(\frac{\min\{x_j s_j\}}{\mu})$$

$$\iff q\log((1-\alpha)\mu) - \log(\pi(\alpha)) = q\log(\mu) - \log(\pi)$$

$$\iff q\log(1-\alpha) = \log(\frac{\pi(\alpha)}{\pi})$$

$$\iff (1-\alpha)^q = \frac{\pi(\alpha)}{\pi}.$$

$\square$

Corollary 3.1. *After* k *iterations of the affine scaling algorithm described in section 2, we have,*

$$(x^k)^T s^k = (\frac{\pi_k}{\pi_o})^{1/q}(x^o)^T s^o. \qquad (20)$$

Proof: Follows from (14) and (19). $\square$

6

Corollary 3.1 provides a very compact description of the contours of the potential function ψ_q and it relates the distance between the current iterate and the boundary of the feasible region to the duality gap directly. If π_k is very small relative to π_o (i.e. (x_k, s_k) is close to the boundary of the feasible region) then the duality gap must also be relatively small.

Lemma 3.1.
$$\frac{\pi(\alpha)}{\pi} \geq 1 - (\frac{\alpha^2}{1-\alpha})\frac{n}{\pi}.$$

Proof: Using (18) we get

$$\pi(\alpha) \geq \frac{\min\{x_j s_j\}}{\mu} - (\frac{\alpha^2}{1-\alpha})\frac{\max\{|(v_p)_j(v-v_p)_j|\}}{\mu} \tag{21}$$

$$= \pi - (\frac{\alpha^2}{1-\alpha})\frac{\max\{|(v_p)_j(v-v_p)_j|\}}{\mu} \tag{22}$$

Note that $\|v\|_2 = (x^T s)^{1/2}$, and for any $v \in \mathbb{R}^n$ we have $\|v_p\|_\infty \leq \|v_p\|_2 \leq \|v\|_2$ and $\|v - v_p\|_\infty \leq \|v - v_p\|_2 \leq \|v\|_2$.
So, $\max\{|(v_p)_j(v-v_p)_j|\} \leq x^T s$. Hence, from (21)-(22) we get

$$\frac{\pi(\alpha)}{\pi} \geq 1 - (\frac{\alpha^2}{1-\alpha})\frac{n}{\pi}.$$

$\square$

Now, we settle a technical point showing that the contours of the potential function are well-defined.

Lemma 3.2. *Unless $(x(\alpha), s(\alpha))$ is feasible for $\alpha = 1$, (19) has exactly one root for $\alpha \in (0,1)$.*

Proof: We define
$$g(\alpha) := (1 - \alpha)^{q+1},$$

$$h_j(\alpha) := \frac{x_j s_j}{\pi \mu}(1 - \alpha) + \alpha^2 \frac{(v_p)_j(v - v_p)_j}{\pi \mu},$$

$$f(\alpha) := g(\alpha) - \min\{h_j(\alpha)\}.$$

Using (18) we have
$$(1 - \alpha)^q = \frac{\pi(\alpha)}{\pi} \iff f(\alpha) = 0.$$

7

We partition the indices as follows: $J^- := \{j : (v_p)_j(v - v_p)_j < 0\}$, $J^0 := \{j : (v_p)_j(v - v_p)_j = 0\}$, $J^+ := \{j : (v_p)_j(v - v_p)_j > 0\}$. Note that J^- and J^+ are non-empty, because if they both are empty then $(x(\alpha), s(\alpha))$ is feasible for $\alpha = 1$ and we have the optimal solution. Since $\sum_j (v_p)_j(v - v_p)_j = 0$, if one of them is non-empty then they both are.

Clearly, there exists a small enough $\epsilon > 0$ such that $f(\alpha) < 0$ for $\alpha \in (0, \epsilon)$ (e.g. $\epsilon = \frac{\pi q}{2n}$). It is also clear that if $(x(\alpha), s(\alpha))$ is not feasible for $\alpha = 1$, then $f(1) > 0$. By the continuity of f we conclude that there exists at least one root in $(0, 1)$. For $j \in J^0 \cup J^+$ we have

$$h_j(\alpha) \geq \frac{x_j s_j}{\pi \mu}(1 - \alpha) \geq (1 - \alpha) > (1 - \alpha)^{q+1}.$$

Hence, for $j \in J^0 \cup J^+$, $h_j(\alpha)$ does not intersect $g(\alpha)$ in $(0, 1)$. For $j \in J^-$ we have h_j as a concave function, so $\min_{j \in J^-}\{h_j(\alpha)\}$ is concave. Since $g(\alpha)$ is convex, the solution to (19) is unique. $\square$

Lemma 3.3. *If α^* solves (19) for $\frac{1}{2t} \leq q \leq \sqrt{n}$ $(n \geq 4)$ then $\alpha^* \geq \frac{\pi q}{2n}$.*

Proof: We want $q \log(1 - \alpha) = \log(\frac{\pi(\alpha)}{\pi})$. Note that if for $\bar{\alpha} \in (0, 1)$ we have

$$q \log(1 - \bar{\alpha}) \leq \log(\frac{\pi(\bar{\alpha})}{\pi}) \tag{23}$$

then $\alpha^* \geq \bar{\alpha}$ (from Lemma 3.2). We will show that $\bar{\alpha} = \frac{\pi q}{2n}$ satisfies (23) which will prove the lemma.

Using the linear approximation to the logarithm we have

$$q \log(1 - \bar{\alpha}) \leq -q\bar{\alpha} \quad \text{for} \quad \bar{\alpha} \in (0, 1) \tag{24}$$

By Lemma 3.1 we have

$$\log(\frac{\pi(\bar{\alpha})}{\pi}) \geq \log[1 - (\frac{\bar{\alpha}^2}{1 - \bar{\alpha}})(\frac{n}{\pi})].$$

Note that for $\alpha = \frac{\pi q}{2n}$ and $\frac{1}{2t} \leq q \leq \sqrt{n}$, from which

$$(\frac{\bar{\alpha}^2}{1 - \bar{\alpha}})\frac{n}{\pi} = \frac{\pi q^2/4n}{1 - \pi q/2n} \leq \frac{\pi q^2}{2n} \leq \frac{1}{2}.$$

8

Using the approximation $\log(1 + \lambda) \geq \lambda - \frac{\lambda^2}{2(1-|\lambda|)}$ for $|\lambda| < 1$, we get

$$\log(\frac{\pi(\bar{\alpha})}{\pi}) \geq -(\frac{1}{1 - \bar{\alpha}})[\bar{\alpha}^2\frac{n}{\pi} + \frac{\bar{\alpha}^4 n^2/\pi^2}{2(1 - \bar{\alpha} - \bar{\alpha}^2 n/\pi)}]. \tag{25}$$

From (24) and (25) we conclude that if

$$-q\bar{\alpha} + q\bar{\alpha}^2 \leq -\bar{\alpha}^2\frac{n}{\pi} - \frac{\bar{\alpha}^4 n^2/\pi^2}{2(1 - \bar{\alpha} - \bar{\alpha}^2 n/\pi)} \tag{26}$$

then (23) holds. Letting $\bar{\alpha} = \frac{\pi q}{2n}$ and dividing both sides by $q\bar{\alpha}$ in (26) we get

$$-1 + \frac{\pi q}{2n} \leq -\frac{1}{2} - \frac{\pi q^2/8n}{2(1 - \pi q/2n - \pi q^2/4n)} \tag{27}$$

For given values of q and $n \geq 4$, the right hand side of (27) is at least $-\frac{1}{2} - \frac{\pi q^2}{8n} \geq -5/8$. The left hand side of (27) is at most $-1 + 1/4 = -6/8$. Therefore, (23) holds. $\square$

Now, we can state a theorem about the complexity of the algorithm.

Theorem 3.1. *The primal-dual affine scaling algorithm as stated in section 2 converges in $O(\frac{n 2^{2tq}t}{q})$ iterations.*

Proof: From corollary 3.1 we have $(x^k)^T s^k = (\frac{\pi_k}{\pi_o})^{1/q}(x^o)^T s^o$. Since $(x^o)^T s^o \leq 2^t$, if at the Kth iteration $\frac{\pi_K}{\pi_o} < 2^{-2tq}$ then we must have $(x^K)^T s^K < 2^{-t}$ and the algorithm will stop. Otherwise, we get a uniform bound on the distances of the iterates from the boundary:

$$\pi_k \geq \pi_o 2^{-2tq}, \quad \text{for all} \quad k \leq K.$$

By lemma 3.3, this uniform bound yields

$$\alpha^* > \frac{\pi q}{2n} \geq \frac{\pi_o q}{n 2^{2tq+1}}. \tag{28}$$

From equality (14) and inequality (28) we conclude that we must have $(x^k)^T s^k < 2^{-t}$ in $O(\frac{n 2^{2tq}t}{q})$ iterations. $\square$

Using the above theorem, for $q = 1/2t$, $q = 1$, and $q = \sqrt{n}$ we get the bounds $O(n\,t^2)$, $O(n\,2^{2t}\,t)$ and $O(\sqrt{n}\,2^{2t\sqrt{n}}\,t)$ respectively.

9

4 Conclusion

Using a different potential function we provided the same complexity bounds as those obtained by Mizuno and Nagasawa [MN92]. Since the potential function used by our algorithm fits in nicely with the one-sided infinity-norm neighborhoods, we could get a compact and useful description of the contours of the potential function for the affine scaling direction (one may also speculate that for the same reason potential function ψ_q can be more effective for the infinity neighborhoods than the Tanabe-Todd-Ye potential function). The contours of the potential function provide very useful information about the relation between the duality gap and the distance to the boundary which makes the convergence proof arguments simpler.

References:

[KMNY91] M. Kojima, N. Megiddo, T. Noma and A. Yoshise, A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems, *Lecture Notes in Computer Science*, Springer-Verlag, NY.

[KMY88] 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, pp. 29-47.

[Me88] N. Megiddo, Pathways to the Optimal Set in Linear Programming, in: N. Megiddo, ed., *Progress in Mathematical Programming, Interior Point and Related Methods*, Springer-Verlag, New York, pp.131-158.

[MAR90] R.C. Monteiro, I. Adler and M.G.C. Resende, A Polynomial Time Primal-Dual Affine Scaling Algorithm for Linear and Convex Quadratic Programming and Its Power Series Extension, *Math. of OR* 15, pp. 191-214.

[MN92] S. Mizuno and A. Nagasawa, A Primal-Dual Affine Scaling Potential Reduction Algorithm for Linear Programming, Research memorandum No. 427, Department of Industrial Engineering anf Management, Tokyo Institute of Technology, Japan.

[MTY90] S. Mizuno, M.J. Todd, and Y. Ye, On Adaptive-Step Primal-Dual Interior-Point Algorithms for Linear Programming, *TR-944*, School of OR and IE, Cornell University, Ithaca, NY, to appear in *Math. of Op. Res.*.

[Ta87] K. Tanabe, Centered Newton Method for Mathematical Programming, *in System Modeling and Optimization* (eds. M. Iri and K. Yajima), Springer-Verlag, The Proceedings of the 13th IFIP Conference, Tokyo,Aug. 31 - Sept. 4, pp. 197-206.

[TY90] M.J. Todd and Y. Ye, A Centered Projective Algorithm for Linear Programming, *Mathematics of Operations Research* 15, pp. 508-529.

[Tu92] L. Tunçel, On the Convergence of Primal-Dual Interior Point Methods with Wide Neighborhoods, *TR 1003*, School of OR and IE, Cornell University, Ithaca, NY.