ON THE PRIMAL-DUAL AFFINE SCALING METHOD

Romesh Saigal
Department of Industrial and Operations Engineering
The University of Michigan
Ann Arbor, MI 48109-2117

Technical Report 94-10

March 1994

On the Primal-Dual Affine Scaling Method

Romesh SAIGAL

Department of Industrial and Operations Engineering,

The University of Michigan,

Ann Arbor, Michigan 48109-2117, USA

March 31, 1994

Abstract

We consider here the primal-dual affine scaling method with a fixed step size α to the boundary. We show that for all $\alpha < 1$, the primal and dual objective function values are monotone and the generated sequence of primal and dual solutions converge to a solution on the boundary of the respective polyhedrons. Also, when $\alpha < \frac{\sqrt{5}-1}{2}$, the limiting primal and dual solutions either converge to an optimum solution, or a non-optimum solution where no pair of variables satisfy the strict complementarity condition. The proof of this result uses a modified version of the potential function of Tanabe, Todd and Ye; and, Mizuno and Nagasawa used in the study of the potential reduction strategy for investigating this method.

Key words: Linear Programming, primal dual affine scaling methods, interior point methods.

Abbreviated title: Primal dual affine scaling method

1 Introduction

We consider here the linear programming problem:

$$\begin{array}{ll} \text{minimize} & c^T x \\ & Ax \quad = \quad b \\ & x \geq 0 \end{array} \tag{1}$$

where A is an $m \times n$ matrix and b and c are appropriate vectors; and its dual:

$$\begin{array}{ll} \text{maximize} & b^T y \\ & A^T y \;+\; s \quad = \quad c \\ & \qquad s \geq 0 \end{array} \tag{2}$$

We also assume that

Assumption 1 *The primal and the dual linear programs have interior solutions.*

Assumption 2 *The objective functions of the primal, (1); and, the dual, (2), are not constant on their respective feasible regions.*

Assumption 3 *The matrix A has rank m.*

In this paper we consider the primal-dual affine scaling method, and investigate its convergence properties. This method was presented in Section 3 of Monterio, Adler and Resende [9]. In their work they established its convergence for a small step size. Recently, Mizuno and Nagasawa [7] have developed a potential reduction method, and proved its convergence for a larger step size. The step size at each iteration is selected so that the potential function will not increase.

In this paper, we investigate the convergence properties of this method. We show that the objective functions are monotone along the generated sequence. This result has also been obtained by Todd, Tuncel, and Mizuno [13]. Also, the sequences are shown to converge for an arbitrary choice of step size, $\alpha < 1$. When the fixed step size α is strictly less than $\frac{\sqrt{5}-1}{2}$, we show here that the sequences converge to the respective optimum solutions or to non-optimum solutions for which no pair of variables satisfy the strict complementarity condition. It is generally believed that some "centering" is needed in the primal-dual methods. Our result does not contradict this. We give two proofs of the main result in the hope that these will help in the construction of an example to demonstrate this need.

Besides the introduction, this paper has 6 other sections. In section 2 we present the primal-dual affine scaling method. In section 3 we prove the convergence of the primal sequence, while in section 4 we investigate the dual sequence. In section 5, we prove some important properties of the primal-dual sequences, and in section 6 we prove that if the step size is required to be strictly less than $\frac{\sqrt{5}-1}{2}$, the sequences converge to optimum solutions or to non-optimum solutions for which no pair of variables satisfy the strict complementarity condition. Finally, we conclude the paper with a conclusions section and an appendix presenting another proof of the main theorem of section 6.

2 The Primal Dual Affine Scaling Method

From the complementary slackness theorem, vectors x, and y, s are optimal for the primal (1) and the dual (2) respectively, if and only if the following conditions hold:

$$\begin{array}{ccccc} Ax & & & = & b \\ & A^T y & + \quad s & = & c \\ & Xs & & = & 0 \\ (x \geq 0) & & (s \geq 0) & & \end{array} \tag{3}$$

At a given interior point $x > 0$ of the primal and (y, s), $s > 0$ of the dual, we derive the primal-dual affine scaling step by applying Newton's method to the nonlinear system (3). The Newton step $(\Delta x, \Delta y, \Delta s)$ is generated by solving the following system:

$$\begin{array}{ccccc} A\Delta x & & & = & 0 \\ & A^T \Delta y & + \quad \Delta s & = & 0 \\ S\Delta x & & X\Delta s & = & -Xs \end{array} \tag{4}$$

The primal-dual affine scaling method that thus results follows:

Step 0 Let (x^0, y^0, s^0) be an interior point solution, $0 < \alpha < 1$, and let $k = 0$.

Step 1 Compute the direction $(\Delta x^k, \Delta y^k, \Delta s^k)$ by solving the system (4). If $\Delta x^k = 0$, $\Delta y^k = 0$ and $\Delta s^k = 0$ then stop. The last solution found is an optimum solution.

Step 2 Minimum Ratio Test: Define

$$\phi_k = \max\{\phi(-X_k^{-1}\Delta x^k), \phi(-S_k^{-1}\Delta s^k)\}$$

where $\phi(u) = \max_j u_j$ for a given vector u.

Step 3 Next Interior Point:

$$\begin{aligned} x^{k+1} &= x^k + \frac{\alpha}{\phi_k}\Delta x^k \\ y^{k+1} &= y^k + \frac{\alpha}{\phi_k}\Delta y^k \\ s^{k+1} &= s^k + \frac{\alpha}{\phi_k}\Delta s^k \end{aligned}$$

Step 4 Iterative Step: Set $k = k+1$, and go to Step 1.

We now show that the minimum ratio test in Step 2 is well defined.

Theorem 1 *Let $\Delta x^k \neq 0$ and $\Delta s^k \neq 0$. Then $\phi_k > 1$.*

Proof: From system (4), it is easily established that $(\Delta x^k)^T \Delta s^k = 0$. Thus, under the hypothesis of the theorem, there is a j such that $\Delta x_j^k \Delta s_j^k < 0$, and so $\phi_k > 0$. Also, using the system (4), we obtain

$$\begin{aligned} x^k + \Delta x^k &= -S_k^{-1} X_k \Delta s^k \\ s^k + \Delta s^k &= -X_k^{-1} S_k \Delta x^k \end{aligned}$$

and thus $(x_j^k + \Delta x_j^k)(s_j^k + \Delta s_j^k) = \Delta x_j^k \Delta s_j^k < 0$. Thus, there is a $0 < \mu_k < 1$ such that $x^k + \mu_k \Delta x^k \geq 0$ and $s^k + \mu_k \Delta s^k \geq 0$. Our result follows by noting that $\frac{1}{\phi_k} = \mu_k$. ■

Since the conditions of Theorem 1 are satisfied at non - optimum solutions of the pair of linear programs, we will henceforth assume that the algorithm generates an infinite sequence satisfying the conditions of the above theorem.

3 The Primal Sequence $\{\mathbf{x}^k\}$

Define a diagonal matrix $D = S^{-1}X$. Then the system (3) can be written as

$$\begin{array}{lllll} A\Delta x & & & = & 0 \\ & A^T\Delta y & + \ \Delta s & = & 0 \\ D^{-1}\Delta x & & \ \ \Delta s & = & -s \end{array} \tag{5}$$

which is readily seen equivalent to the following system:

$$\begin{array}{rcl} -c \; - \; D^{-1}\Delta x \; + \; A^T\tilde{y} & = & 0 \\ A\Delta x & = & 0. \end{array} \tag{6}$$

where $\tilde{y} = y + \Delta y$. The system (6) represents the K.K.T. conditions of the following convex quadratic programming problem:

$$\begin{array}{rl} \text{minimize} & c^T\Delta x + \frac{1}{2}\Delta x D^{-1}\Delta x \\ & A\Delta x \quad = \; 0 \end{array} \tag{7}$$

where $\tilde{y}$ is the Lagrange multiplier associated with the constraint of the quadratic program (7). Also, consider the following ellipsoidal approximating problem:

$$\begin{array}{rrcl} \text{maximize} & -c^T v & & \\ & Av & = & 0 \\ & \|D^{-\frac{1}{2}}v\| & \leq & 1 \end{array} \tag{8}$$

We now prove a simple lemma:

Lemma 2 *The direction Δx^* generated by the quadratic program (7) is a positively scaled multiple of the direction v^* generated by the ellipsoidal approximating problem (8).*

Proof It is readily seen that the solution to the quadratic program (7) is

$$\begin{array}{rcl} \tilde{y} & = & (ADA^T)^{-1}ADc \\ \Delta x^* & = & -D(c - A^T\tilde{y}) \end{array}$$

and the solution to the ellipsoidal approximating problem is

$$\begin{array}{rcl} y' & = & -(ADA^T)^{-1}ADc \\ v^* & = & -\dfrac{D(c + A^T y')}{\|D^{\frac{1}{2}}(c + A^T y')\|} \end{array}$$

and we have our result. ■

We now show that the sequence $\{c^T x^k\}$ is strictly monotone decreasing. This result also appears in Todd, Tuncel and Mizuno [13].

Theorem 3 *The sequence $\{c^T x^k\}$ is strictly monotone decreasing and is bounded, and thus converges, to say c^*.*

Proof: Note that

$$c^T x^{k+1} = c^T x^k + \frac{\alpha}{\phi_k} c^T \Delta x^k$$

and

$$\begin{aligned} c^T \Delta x^k &= -c^T D_k^{\frac{1}{2}}(I - D_k^{\frac{1}{2}} A^T (A D_k A^T)^{-1} A D_k^{\frac{1}{2}}) D_k^{\frac{1}{2}} c \\ &= -\|P_k D_k^{\frac{1}{2}} c\|^2 \\ &= -\|D_k^{\frac{1}{2}}(c - A^T \tilde{y}^k)\|^2, \end{aligned}$$

where P_k is the projection matrix into the null space $\mathcal{R}(AD_k^{\frac{1}{2}})$. From Assumption (2), c does not belong to the row space $\mathcal{R}(A^T)$ of A, and so $D_k^{\frac{1}{2}} c$ does not belong to the row space $\mathcal{R}(D_k^{\frac{1}{2}} A^T)$ of $AD_k^{\frac{1}{2}}$; thus $P_k D_k^{\frac{1}{2}} c$ is not zero, and we are done. ■

Before we prove the convergence of the primal sequence we will need the following result related to the ellipsoidal approximating problem (8).

Theorem 4 *There exists a constant $\rho > 0$ such that for every $k = 1, 2, \cdots$*

$$\|v^k\| \leq -\rho c^T v^k$$

where v^k is the solution to (8) with $D = S_k^{-1} X_k$.

Proof: See Corollary 6, Saigal [10]. ■

We are now ready to show that the sequence $\{x^k\}$ converges.

Theorem 5 *The sequence $\{x^k\}$ converges, say to x^*.*

Proof: From Step 3, Lemma 2 and Theorem 4 we note that for each $k = 1, 2, \cdots$, there is $\mu_k > 0$ such that

$$x^k - x^{k+1} = -\mu_k v^k.$$

Thus

$$\infty > c^T x^0 - c^* = \sum_{k=0}^{\infty} c^T (x^k - x^{k+1}) = -\sum_{k=0}^{\infty} \mu_k c^T v^k \geq \frac{1}{\rho} \sum_{k=0}^{\infty} \mu_k \|v^k\| = \frac{1}{\rho} \sum_{k=0}^{\infty} \|x^{k+1} - x^k\|$$

and thus the sequence $\{x^k\}$ is a Cauchy sequence and converges, say to x^*. ■

We now prove an important result that is needed in the proof of convergence to optimality.

Theorem 6 *There exists a $\delta > 0$ such that for each j such that $x_j^* = 0$, and every $k = 1, 2, \cdots$,*

$$\frac{c^T x^k - c^*}{x_j^k} \geq \delta.$$

Proof: Let k be arbitrary. Then, from Step 3, Lemma 2 and Theorem 4

$$\infty > c^T x^k - c^* = \sum_{l=k}^{\infty} c^T (x^{k+l+1} - x^{k+l}) \geq \frac{1}{\rho} \sum_{l=0}^{\infty} \|(x^{k+l+1} - x^{k+l})\| \geq \frac{1}{\rho} \| \sum_{l=0}^{\infty} (x^{k+l+1} - x^{k+l}) \| = \frac{1}{\rho} \|x^k - x^*\|$$

and our result follows. ■

4 The Dual Sequence $(\{\mathrm{y}^\mathrm{k}\}, \{\mathrm{s}^\mathrm{k}\})$

From system (4), it is readily seen that

$$\Delta y = (ADA^T)^{-1} b.$$

Δy can also be obtained by solving the following quadratic program:

$$\begin{array}{ll} \text{maximize} & b^T \Delta y \\ & \Delta y^T ADA^T \Delta y \leq 1. \end{array} \tag{9}$$

The following theorem relates the solution vector of this quadratic program to its objective function value.

Theorem 7 *There exists a positive constant, $\rho > 0$, which is only determined by A and b, such that*

$$\|\Delta y\| \leq \rho b^T \Delta y.$$

Proof: See Theorem 5, Saigal [10]. ■

We can now prove that the dual objective function is strictly monotone increasing on the sequence $\{y^k\}$. This result also appears in [13].

Theorem 8 *The sequence $\{b^T y^k\}$ is strictly monotone increasing and thus converges, say to b^*.*

Proof: The result follows from the fact that

$$b^T(y^{k+1}-y^k)=\frac{\alpha}{\phi_k}b^T\Delta y^k=\frac{\alpha}{\phi_k}b^T(ADA^T)^{-1}b$$

and that ADA^T is a positive definite matrix and from assumption (2), $b\neq 0$. Our result now follows since the sequence $\{b^Ty^k\}$ is bounded (by the weak duality theorem) and every bounded monotone sequence converges. ■

We are now ready to show that the dual sequence also converges.

Theorem 9 *The sequences $\{y^k\}$ and $\{s^k\}$ converge, say to y^* and s^* respectively.*

Proof: From Step 3 and Theorems 7 and 8, we see that

$$\infty > b^* - b^Ty^k=\sum_{k=0}^{\infty} b^T(y^{k+1}-y^k)=\sum_{k=0}^{\infty}\frac{\alpha}{\phi_k}b^T\Delta y^k\geq\sum_{k=0}^{\infty}\frac{\alpha}{\phi_k}\frac{1}{\rho}\|\Delta y^k\|=\frac{1}{\rho}\sum_{k=0}^{\infty}\|y^{k+1}-y^k\|$$

and thus $\{y^k\}$ is a Cauchy sequence and thus converges, say to y^*. But $s^k=c-A^Ty^k$, and thus $\{s^k\}$ also converges, to $s^*=c-A^Ty^*$. ■

The following theorem plays an important role in proving convergence to optimality.

Theorem 10 *There exists a $\delta>0$ such that for each j with $s_j^*=0$, and every $k=1,2,\cdots$*

$$\frac{b^*-b^Ty^k}{s_j^k}\geq\delta.$$

Proof: Follows exactly as the proof of Theorem 6 and the fact that $\|s^k-s^*\|\leq\|A^T(y^*-y^k)\|\leq\mu\|y^k-y^*\|$. ■

5 The Primal-Dual Sequences

We have seen that there exist vectors $x^*\geq 0$, y^* and $s^*\geq 0$ such that

$$\begin{aligned} x^k &\longrightarrow x^* \\ y^k &\longrightarrow y^* \\ s^k &\longrightarrow s^*. \end{aligned}$$

Define

$$\begin{aligned} B &= \{j : x_j^*>0\} \\ N &= \{j : s_j^*>0\}. \end{aligned} \tag{10}$$

We do not assume here that $B \cap N = \emptyset$ nor that $B \cup N = \{1, \cdots, n\}$. We will obtain some important results in this section. These will be used in the next section to establish the optimality of the limit points.

We now prove two simple lemmas:

Lemma 11 *The sequence* $\{(x^k)^T s^k\}$ *is monotone decreasing.*

Proof We note that using system (4) and Step 3, we obtain

$$\begin{aligned} (x^{k+1})^T s^{k+1} &= (x^k + \frac{\alpha}{\phi_k}\Delta x^k)^T (s^k + \frac{\alpha}{\phi_k}\Delta s^k) \\ &= (1 - \frac{\alpha}{\phi_k})(x^k)^T s^k \end{aligned} \tag{11}$$

and our result follows from Theorem 1. ■

Lemma 12 *For every* $k = 1, 2, \cdots$

$$\|X_k^{-\frac{1}{2}} S_k^{\frac{1}{2}} \Delta x^k\| \leq \sqrt{(x^k)^T s^k}$$
$$\|X_k^{\frac{1}{2}} S_k^{-\frac{1}{2}} \Delta s^k\| \leq \sqrt{(x^k)^T s^k}.$$

Proof: This lemma readily follows from the following facts: for $\bar{D}_k = (X_k S_k^{-1})^{\frac{1}{2}}$, from system (4), we obtain

$$\bar{D}_k^{-1} \Delta x^k + \bar{D}_k \Delta s^k = -(X_k S_k)^{\frac{1}{2}} e,$$

and

$$(\bar{D}_k^{-1} \Delta x^k)^T (\bar{D}_k \Delta s^k) = (\Delta x^k)^T \Delta s^k = 0.$$

Thus $\|\bar{D}_k^{-1} \Delta x^k\| \leq \|(X_k S_k)^{\frac{1}{2}} e\| = \sqrt{(x^k)^T s^k}$. We are done, as the other result follows by the same argument. ■

We now establish an important property of the sequences $\{\Delta x^k\}$, $\{\Delta s^k\}$ and $\{(x^k)^T s^k\}$. This theorem also appears in Ye, Guler, Tapia and Zhang [16].

Theorem 13 *There exist constants* $\beta_1 > 0$ *and* $\beta_2 > 0$ *such that for all* $k = 1, 2, \cdots$

$$\|\Delta x^k\| \leq \beta_1 (x^k)^T s^k$$
$$\|\Delta s^k\| \leq \beta_2 (x^k)^T s^k.$$

Proof: From definitions and weak duality theorem we derive the following inequalities:

$$\begin{aligned}(x^k)^T s^k &= c^T x^k - b^T y^k \\ &\geq c^T x^k - c^T x^* \\ &= (s^* - A^T y^*)^T (x^k - x^*) \\ &= (s_N^*)^T x_N^k - (x^*)^T s^*.\end{aligned}$$

From Lemma 11 $(x^k)^T s^k \geq (x^*)^T s^*$, Thus, there is a $\mu_1 > 0$ such that

$$(x^k)^T s^k \geq \mu_1 \|x_N^k\|.$$

Also, from Lemma 12,

$$\begin{aligned}\|\Delta x_N^k\| &\leq \|X_{k,N}^{\frac{1}{2}} S_{k,N}^{-\frac{1}{2}}\| \|S_{k,N}^{\frac{1}{2}} X_{k,N}^{-\frac{1}{2}} \Delta x_N^k\| \\ &\leq \mu_2 \|X_{k,N}^{\frac{1}{2}} e\| \|S_{k,N}^{\frac{1}{2}} X_{k,N}^{-\frac{1}{2}} \Delta x_N^k\| \\ &\leq \mu_3 (x^k)^T s^k\end{aligned}$$

and; from Theorem 4, for some $\rho > 0$,

$$\begin{aligned}\|\Delta x^k\| &\leq -\rho c^T \Delta x^k \\ &= -\rho (s^* + A^T y^*)^T \Delta x^k \\ &= -\rho (s_N^*)^T \Delta x_N^k \\ &\leq \rho \|s_N^*\| \|\Delta x_N^k\|.\end{aligned}$$

For some $\beta_1 > 0$ we have our first result. The second result follows in a similar manner, with the use of Theorem 7. ■

The next theorem characterizes the convergence property of the sequences:

Theorem 14 $(x^k)^T s^k \longrightarrow 0$ *if and only if*

$$X_{k,N}^{-1} \Delta x_N^k \longrightarrow -e \tag{12}$$

$$S_{k,B}^{-1} \Delta s_B^k \longrightarrow -e. \tag{13}$$

Proof: Let $(x^k)^T s^k \to 0$. Then $B \cap N = \emptyset$. From Theorem 13, $\Delta x^k \to 0$, and $\Delta s^k \to 0$. Also, from system (4), $X_{k,B}^{-1} \Delta x_B^k + S_{k,B}^{-1} \Delta s_B^k = -e$. Thus $x_B^* > 0$ implies (13). Similarly, from (4), $X_{k,N}^{-1} \Delta x_N^k + S_{k,N}^{-1} \Delta s_N^k = -e$. Thus $s_N^* > 0$ implies (12).

We complete the proof by noting that if the conditions (12) and (13) hold, then $\Delta s_N^k \to 0$ and $\Delta x_B^k \to 0$. Thus, if there is a $j \in B \cap N$ then

$$\frac{\Delta x_j^k}{x_j^k} + \frac{\Delta s_j^k}{s_j^k} \to 0$$

which is a contradiction. ■

We now establish an important property:

Theorem 15 *For each $j = 1, \cdots, n$ and $k = 1, 2, \cdots$*

$$-1 \leq \frac{1}{\phi_k^2} \frac{\Delta x_j^k}{x_j^k} \frac{\Delta s_j^k}{s_j^k} \leq 1.$$

Proof: Let k be arbitrary. From system (4), for each $j = 1, \cdots, n$

$$\frac{\Delta x_j^k}{x_j^k} + \frac{\Delta s_j^k}{s_j^k} = -1. \tag{14}$$

Thus, either $\Delta x_j^k < 0$ and $\Delta s_j^k < 0$; or, $\Delta x_j^k \Delta s_j^k < 0$. In the first case, by the definition of ϕ_k, we have

$$0 \geq \frac{\Delta x_j^k}{\phi_k x_j^k} \geq -1 \text{ and } 0 \geq \frac{\Delta s_j^k}{\phi_k s_j^k} \geq -1.$$

Thus their product satisfies the required inequality. To see the second case, without loss of generality assume that $\Delta x_j^k > 0$. Then $0 \geq \frac{\Delta s_j^k}{\phi_k s_j^k} \geq -1$, and so

$$0 \leq \frac{\Delta x_j^k}{\phi_k x_j^k} = -\frac{1}{\phi_k} - \frac{\Delta s_j^k}{\phi_k s_j^k} \leq 1$$

and we are done. ■

We now state a technical lemma on the natural logarithm function.

Lemma 16 *Let w an n vector and $0 < \lambda < 1$ be such that $w_j \leq \lambda$. Then*

$$\sum_{j=1}^{n} log(1 - w_j) \geq -e^T w - \frac{\|w\|^2}{2(1-\lambda)}.$$

Proof: See Lemma 8, Saigal [10]. ■

6 On Convergence to Optimality

We are now ready to investigate the limits x^*, y^* and s^* as optimal solutions of their respective problems. We note that there is a j such that either $x_j^* = 0$ or $s_j^* = 0$ or both. With such a j, we define the following potential function:

$$F_r(x, s, j) = (2r + 1)\log(x^T s) - \log(x_j s_j).$$

where r is a large positive constant. This function is a specialization of the one considered by Tanabe [12], Todd and Ye [14] and Mizuno and Nagasawa [7]. We can then prove the following about this function.

Theorem 17 *Let $(x^k)^T s^k \not\to 0$. Then $F_r(x^k, s^k, j) \longrightarrow \infty$.*

Proof: Since $x_j^* s_j^* = 0$ from Theorems 5 and 9

$$x_j^k s_j^k \longrightarrow 0$$

and thus, under the hypothesis of the theorem, $F_r(x^k, s^k, j)$ cannot be bounded above. ■

We now establish our main theorem.

Theorem 18 *Let $\alpha < \frac{\sqrt{5}-1}{2}$. Then, there exist vectors x^*, y^* and s^* such that*

1. $x^k \longrightarrow x^*$
2. $y^k \longrightarrow y^*$
3. $s^k \longrightarrow s^*$,

where exactly one of the following holds:

1. *x^* is an optimum solution of the primal, and (y^*, s^*) is an optimum solution of the dual.*
2. *$B = N$, i.e., x^* and (y^*, s^*) are non-optimum solutions where no pair of variables satisfy the strict complementarity condition.*

Proof: The convergence of the sequences, and thus the existence of the required vectors follows from Theorems 5 and 9. Let α satisfy the hypothesis of the theorem, and let $\alpha_k = \frac{\alpha}{\phi_k}$. Also, let

$$\xi(\alpha) = \frac{\alpha}{1-\alpha} + \frac{\alpha^3}{2(1-\alpha)(1-\alpha-\alpha^2)}.$$

Either $B = N$ or there is a j such that either $x_j^* = 0$ and $s_j^* > 0$ or $s_j^* = 0$ and $x_j^* > 0$. Let the potential function $F_r(x^k, s^k, j)$ be defined for this j, with some r. Now

$$\begin{aligned} x_j^{k+1} s_j^{k+1} &= (x_j^k + \frac{\alpha}{\phi_k}\Delta x_j^k)(s_j^k + \frac{\alpha}{\phi_k}\Delta s_j^k) \\ &= (1 - \frac{\alpha}{\phi_k}) x_j^k s_j^k + (\frac{\alpha}{\phi_k})^2 \Delta x_j^k \Delta s_j^k. \end{aligned}$$

First assume that $x_j^* = 0$ and $s_j^* > 0$. The analysis of the case when $x_j^* > 0$ and $s_j^* = 0$ is similar. Let $r > 0$ be sufficiently large such that there is an $L \geq 1$ and for every $k \geq L$,

$$\max\{\frac{|\Delta s_j^k|}{s_j^k}\xi(\alpha), \frac{|\Delta s_j^k|^2}{(s_j^k)^2}\xi(\alpha)\} < 2r.$$

Such an r exists from our choice of α and Theorems 9 and 13. Define $w_k = \frac{\Delta x_j^k \Delta s_j^k}{x_j^k s_j^k}$. By a simple substitution of above and Identity 11, we obtain

$$\begin{aligned} F_r(x^{k+1}, s^{k+1}, j) - F_r(x^k, s^k, j) &= (2r+1)\log(1-\alpha_k) - \log(1 - \alpha_k + \alpha_k^2 w_k) \\ &= 2r\log(1-\alpha_k) - \log(1 + \frac{\alpha_k^2}{1-\alpha_k} w_k). \end{aligned}$$

Using the fact that $\log(1-a) \leq -a$ and Lemma 16, we obtain

$$F_r(x^{k+1}, s^{k+1}, j) - F_r(x^k, s^k, j) \leq -2r\alpha_k + \frac{\alpha_k^2}{1-\alpha_k}|w_k| + \frac{\alpha_k^4}{(1-\alpha_k)^2}\frac{w_k^2}{2(1 - \frac{\alpha_k^2}{1-\alpha_k}|w_k|)}. \quad (15)$$

From Theorem 15 we have

$$\frac{|w_k|}{\phi_k} \leq \frac{|\Delta s_j^k|}{s_j^k} \text{ and } \frac{|w_k|}{\phi_k^2} \leq 1$$

and, from Theorem 1, $\alpha_k \leq \alpha$. Thus, substituting the above expressions in (15), we obtain

$$F_r(x^{k+1}, s^{k+1}, j) - F_r(x^k, s^k, j) \leq \alpha_k(-2r + \frac{\alpha|\Delta s_j^k|}{(1-\alpha)s_j^k} + \frac{\alpha^3|\Delta s_j^k|^2}{2(1-\alpha)(1-\alpha-\alpha^2)(s_j^k)^2})$$

and for the chosen r, the expression in the bracket remains strictly less than zero. Thus, $\{F_r(x^k, s^k, j)\}$ cannot increase, so from Theorem 17 $(x^k)^T s^k \to 0$ and our theorem follows. ■

7 Concluding Remarks

It is generally believed that the primal-dual methods will not converge to optimality without "centering" steps. Our Theorem 18 does not contradict this. Another proof of this theorem is given in the appendix, which gives more insight into the non-convergence to optimality. Also, convergence to optimality follows readily if the sequence $\{\Delta y^k\}$ converges.

So long as at least one pair of complementary variables satisfy the strict complementarity condition, convergence to optimality is guaranteed. Adler and Monterio [2] show that for an infinitesimal step size, Part 2 of Theorem 18 cannot occur. This paper is written in the hope that strong believers in centering will look for an example where convergence to optimality for large steps sizes does not occur.

8 The Appendix

We now give another proof of Theorem 18, which does not make use of a potential function, and gives further insights into non-convergence to optimality. It also indicates that convergence to optimality occurs if the convergence rate of the sequence $\{c^T x^k - c^*\}$ or $\{b^* - b^T y^k\}$ is at least linear. We now prove a simple lemma before proving the main theorem in this section.

Lemma 19 *Let $\alpha_j = \frac{\alpha}{\phi_j}$ for each $j = 1, 2, \cdots$. Then*

1. $(x^k)^T s^k \not\to 0$ *if and only if* $\sum_{j=1}^{\infty} log(1 - \alpha_j) > -\infty$.

2. $\sum_{j=1}^{\infty} log(1 - \alpha_j) > -\infty$ *if and only if* $\sum_{j=1}^{\infty} \alpha_j < \infty$.

Proof: Note that

$$(s^{k+1})^T x^{k+1} = (1 - \alpha_k)(s^k)^T x^k = \prod_{j=0}^{k} (1 - \alpha_j)(s^0)^T x^0$$

and our Part 1 follows by taking logs of the above expression. To see the only if part of Part 2, since $\log(1 - \alpha_j) \leq -\alpha_j$, note that

$$-\infty < \sum_{j=1}^{\infty} \log(1 - \alpha_j) < -\sum_{j=1}^{\infty} \alpha_j$$

To see the if part, we note that from Lemma 16,

$$\sum_{j=0}^{\infty} \log(1-\alpha_j) \geq -\sum_{j=0}^{\infty} \alpha_j - \sum_{j=0}^{\infty} \frac{\alpha_j^2}{2(1-\alpha_j)} > -\infty$$

and we are done. ■

The next theorem gives another proof of Theorem 18.

Theorem 20 *Let $\sum_{j=1}^{\infty} log(1-\alpha_j) > -\infty$ and $x_j^k s_j^k \to 0$ for some j. Then, for that j*

1. $x_j^k \longrightarrow 0$
2. $s_j^k \longrightarrow 0$
3. $\frac{\Delta x_j^k \Delta s_j^k}{x_j^k s_j^k} \longrightarrow -\infty$

Proof: Let $a_j^k = \frac{\Delta x_j^k \Delta s_j^k}{x_j^k s_j^k}$. Then $a_j^k \not\to \infty$. Thus assume the sequence $\{a_j^k\}$ is bounded. Then, for some $R > 0$, we have $|a_j^k| < R$. Thus, there exists an $L \geq 1$ such that for all $k \geq L$, $\frac{\alpha_k^2}{1-\alpha_k} R < 1$. Now

$$s_j^{k+1} x_j^{k+1} = (1-\alpha_k + \alpha_k^2 a_j^k) x_j^k s_j^k.$$

Let $b_{j,k}^r = \frac{x_j^{k+r} s_j^{k+r}}{x_j^k s_j^k}$. Then

$$b_{j,k}^{r+1} = \prod_{i=0}^{r} (1-\alpha_{k+i} + \alpha_{k+i}^2 a_j^{k+i}).$$

Thus

$$\begin{aligned}
\log(b_{j,k}^{r+1}) &= \sum_{j=0}^{r} \log(1-\alpha_{k+i} + \alpha_{k+i}^2 a_j^{k+1}) \\
&\geq \sum_{i=k}^{\infty} \log(1-\alpha_i) + \sum_{i=k}^{r} \log(1 + \frac{\alpha_i^2}{1-\alpha_i} a_j^i) \\
&= -M + N
\end{aligned}$$

where,

$$N \geq \sum_{i=k}^{r} \log(1 - \frac{\alpha_i^2}{1-\alpha_i} R)$$

and using Lemma 16,

$$N \geq -\sum_{i=k}^{\infty} (\frac{\alpha_i^2}{1-\alpha_i} R + \frac{R^2 \alpha_i^4}{(1-\alpha_i)^2} \frac{1}{2(1-\frac{\alpha_i^2}{1-\alpha_i} R)}).$$

Since $\alpha_i \to 0$, there is an $L \geq 1$ such that for all $i \geq L$ $\alpha_i < \frac{1}{2}$ and $1 - \alpha_i - \alpha_i^2 R > \frac{1}{2}$. Then

$$\frac{\alpha_i}{1-\alpha_i} < \frac{1}{2} \text{ and } \frac{\alpha_i^2}{1-\alpha_i-\alpha_i^2 R} < \frac{1}{2}.$$

$$\begin{aligned} N &\geq -\sum_{i=k}^{L-1}\left(\frac{\alpha_i^2}{1-\alpha_i}R + \frac{R^2\alpha_i^4}{(1-\alpha_i)^2}\frac{1}{2(1-\frac{\alpha_i^2}{1-\alpha_i}R)}\right) - \sum_{i=L}^{\infty}\left(\frac{\alpha_i^2}{1-\alpha_i}R + \frac{R^2\alpha_i^4}{(1-\alpha_i)^2}\frac{1}{2(1-\frac{\alpha_i^2}{1-\alpha_i}R)}\right) \\ &\geq -N_1 - RN_2 - R^2 N_3 \\ &= -\bar{M} \\ &> -\infty. \end{aligned}$$

Thus, for every $r \geq k$,

$$\log(b_{j,k}^{r+1}) \geq -M - \bar{M} > -\infty.$$

Thus, as $r \to \infty$, $b_{j,k}^{r+1} \not\to 0$, and this contradicts the fact that $x_j^k s_j^k \to 0$. Thus 2(c) follows.

To see 2(a) and 2(b) note that

$$\frac{\Delta x_i^k}{x_i^k} + \frac{\Delta s_i^k}{s_i^k} = -1.$$

Thus, $\frac{|\Delta x_j^k|}{x_j^k} \to \infty$ if and only if $\frac{|\Delta s_j^k|}{s_j^k} \to \infty$. Since the numerator of these expressions is bounded, the denominators must go to zero, and we have parts 2(a) and 2(b) from part 2(c). ■

References

[1] I. Adler, N. K. Karmarkar, M. G. C. Resende and G. Veiga, " An implementation of Karmarkar's algorithm for linear programming," *Mathematical Programming* 44 (1989) 297 - 335.

[2] I. Adler and R. D. C. Monterio, "Limiting behavior of the affine scaling continuous trajectories for linear programming," *Mathematical Programming* 51 (1991) 29-51.

[3] E. R. Barnes, " A Variation on Karmarkar's Algorithm for solving linear programming problems," *Mathematical Programming* 36 (1986) 174 - 182.

[4] I. I. Dikin, " Iterative solution of problems of linear and quadratic programming," *Soviet Mathematics Doklady* 8 (1967) 674 - 675.

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

[6] M. Kojima, S. Mizuno and A. Yoshise, "A primal-dual interior point algorithm for linear programming," in: N. Meggido ed., *Progress in Mathematical Programming, Interior Points and Related Methods* (Springer, New York, 1989) pp 29 - 47.

[7] S. Mizuno and A. Nagasawa, "A primal-dual affine scaling potential reduction algorithm for Linear Programming," *Mathematical Programming* 62 (1993) 119-131.

[8] R. D. C. Monterio and I. Adler, "Interior path following primal-dual algorithms. Part 1," *Mathematical Programming* 44 (1989) 27 - 41.

[9] R. D. C. Monterio, 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," *Mathematics of Operations Research* 15 (1990) 191 - 214.

[10] R. Saigal, "A simple proof of the primal affine scaling algorithm," Technical Report No 92-60, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, Michigan, December 1992.

[11] R. Saigal, "Efficient variants of the affine scaling method," Technical Report No 93-21, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, Michigan, August 20, 1992.August 20, 1993 (Revised October 22, 1993).

[12] K. Tanabe, "Centered Newton method for linear programming and quadratic programming," *Proceedings of the 8th Mathematical Programming Symposium, Japan* (1987) 131 - 152.

[13] M. Todd, L. Tuncel and S. Mizuno, "Monotonicity of Primal-Dual Objective Values in Primal-Dual interior-point algorithms," Research Report 1014, ORIE, Cornell University (1992).

[14] M. Todd and Y. Ye, "A centered projective algorithm for linear programming," *Mathematics of Operations Research* 15 (1990) 508 - 529.

[15] T. Tsuchiya and M. Muramatsu, "Global convergence of a long-step affine scaling algorithm for degenerate linear programming problems," Research Memorandum 423, The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106, Japan (January 1992 revised, September, 1992). (To appear in *SIAM Journal on Optimization*).

[16] Y. Ye, O. Guler, R. A. Tapia and Y. Zhang, " A quadratically convergent $O(\sqrt{n}L)$-iteration algorithm for linear programming," *Mathematical Programming* 59 (1993) 151 - 162.