A SIMPLE PROOF OF PRIMAL AFFINE SCALING METHOD

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

Technical Report 92-60

December 1992
Revised January 1993

A Simple Proof of Primal Affine Scaling Method

Romesh SAIGAL

Department of Industrial and Operations Engineering,

The University of Michigan,

Ann Arbor, Michigan 48109-2117, USA

December 17, 1992 (revised January 5, 1993)

Abstract

In this paper we present a simpler proof of the result of Tsuchiya and Muramatsu on the convergence of the primal affine scaling method. We show that the primal sequence generated by the method converges to the interior of the optimum face and the dual sequence to the analytic center of the optimal dual face, when the step size implemented in the procedure is bounded by $\frac{2}{3}$. This paper is in the spirit of Monterio, Tsuchiya and Wang [10].

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

Abbreviated title: Proof of affine scaling method

1 Introduction

We consider here the linear programming problem:

$$\begin{aligned} \text{minimize} \quad & c^T x \\ & Ax = b \qquad\qquad \text{LP} \\ & x \geq 0 \end{aligned}$$

where A is a $m \times n$ matrix and b and c are appropriate vectors. We also assume that

Assumption 1 *The linear program has an interior solution.*

Assumption 2 *The objective function is not constant on the feasible region.*

Assumption 3 *The matrix A has rank m.*

On the basis of the global convergence analysis by the use of a local potential fuction, developed by Tsuchiya in the papers [13] and [14] and the analysis of reduction of potential function developed by Dikin [7], Dikin [6] proved the global convergence of the long step affine scaling method when the step size is restricted by $\frac{1}{2}$. Stimulated by [7], but independently, Tsuchiya and Muramatsu[15] proved that the method converges to the optimum solution if the step size is restricted by $\frac{2}{3}$; and that, in this case, the primal sequence converges to the relative interior of the optimal face of the primal and that dual sequence to the analytic center of the optimal face of the dual. Hall and Vanderbei[8] present an example that shows the $\frac{2}{3}$ is tight in the sense that with a larger step size the dual sequence may not converge.

This note is written in the hope that it will bring out the basic idea of the proof of [15], and is written in the spirit of Monterio, Tsuchiya and Wang[10], and presents some of their arguments.

2 Primal Affine Scaling Method

We refer the reader to Barnes [2], Vanderbei, Meketon and Freedman [16] for more on this method. The iterative scheme of the method is the following and is known as the long step version:

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

Step 1 Tentative Solution to the Dual:

$$y^k = (AX_k^2A^T)^{-1}AX_k^2c$$

Step 2 Tentative Dual Slack:

$$s^k = c - A^Ty^k$$

If $s^k \leq 0$ then STOP. The solution is unbounded.

Step 3 Min-Ratio test:

$$\begin{aligned}\theta_k &= \min\{\frac{\|X_ks^k\|}{x_j^ks_j^k} : s_j^k > 0\} \\ &= \frac{\|X_ks^k\|}{\phi(X_ks^k)}\end{aligned}$$

where $\phi(x) = \max_j x_j$. If $\theta_k = 1$ set $\alpha = 1$.

Step 4 Next Interior Point:

$$x^{k+1} = x^k - \alpha\theta_k\frac{X_k^2s^k}{\|X_ks^k\|}$$

Step 5 Iterative Step: If $x_j^{k+1} = 0$ for some j, then STOP. x^{k+1} is an optimal solution. Otherwise set $k = k + 1$ and go to step 1.

It is shown in Vanderbei and Lagarius [17] that if the algorithm is finite, i. e., stops in Step 5, then an optimal solution has been found.

2.1 Ellipsoidal Approximating Problem

Barnes [2] has shown that this algorithm is generated by solving the following ellipsoidal approximating problem:

At the kth iterate, given an interior point $x^k > 0$, the direction $v^k = X_k^2s^k$ to the next iterate x^{k+1}, generated in Step 4, is obtained by scaling the solution to the following ellipsoidal approximating problem:

$$\begin{aligned}\text{maximize} \quad & c^Tv \\ Av &= 0 \qquad\qquad \text{EAP} \\ \|X_k^{-1}v\|^2 &\leq 1.\end{aligned}$$

2.2 Least Squares Problem

The dual estimate y^k, calculated in Step 1, can be readily seen as generated by the following "least squares" problem:

$$y^k = \operatorname{argmin} \|X_k s\| \qquad \text{LSP}$$
$$s = c - A^T y$$

and we note that this estimate, given a primal feasible solution x^k, is chosen to minimize the complementary slackness violation without the nonnegativity constraint on the dual variable s.

3 Properties of Sequences

Let $\{x^k\}$, $\{y^k\}$, $\{s^k\}$ and $\{v^k\} = \{X_k^2 s^k\}$ be the sequences generated by the primal affine scaling algorithm. If these sequences are finite, then the last point of the first sequence solves the linear program, or the problem has an unbounded solution. Thus, we henceforth assume that they are not finite. We now establish some important properties of these sequences.

Theorem 1 *$x^k > 0$ and $\theta_k \geq 1$ for all k. Also,*

$$c^T v^k = c^T X_k^2 s^k = \|X_k s^k\|^2 = \|X_k^{-1} v^k\|^2$$

for all k.

Proof The first part readily follows. To see the second part first observe that, from the definition $c^T v^k = c^T X_k^2 s^k$, and

$$\begin{aligned} c^T X_k^2 s^k &= c^T X_k (I - X_k A^T (A X_k A^T)^{-1} A X_k) X_k c \\ &= \|P_k X_k c\|^2 \end{aligned}$$

where $P_k = I - X_k A^T (A X_k A^T)^{-1} A X_k$ is the projection matrix into the null space $\mathcal{N}(AX_k)$ of the matrix AX_k, with $P_k = P_k^2 = P_k^T$. Now $P_k X_k c = X_k(c - A^T(AX_kA^T)^{-1} AX_k^2 c) = X_k s^k$ and we are done. □

Theorem 2 $\{c^Tx^k\}$ *is a strictly decreasing sequence with*

$$c^Tx^{k+1} = c^Tx^k - \alpha\theta_k\|X_ks^k\|.$$

Either this sequence diverges to $-\infty$, or it is bounded, and thus converges with $\|X_ks^k\| \to 0$ as $k \to 0$, and $\infty > \sum_{k=1}^{\infty} c^T(x^k - x^{k+1})$.

Proof It is readily seen, from Step 4, and Theorem 1 that

$$\begin{aligned} c^Tx^{k+1} &= c^Tx^k - \alpha\theta_k\frac{c^TX_k^2s^k}{\|X_ks^k\|} \\ &= c^Tx^k - \alpha\theta_k\|X_ks^k\| \end{aligned}$$

From Assumption 2, $\|X_ks^k\| \neq 0$, and so $\{c^Tx^k\}$ is strictly decreasing. Hence, it either diverges to $-\infty$ or is bounded. In the later case, since every bounded monotone sequence converges and $\alpha\theta_k \geq \alpha > 0$, $X_ks^k \to 0$. Now

$$\sum_{k=1}^{\infty} c^T(x^k - x^{k+1}) = c^Tx^1 - c^*$$

where $\lim_{k\to 0} c^Tx^k = c^*$ and we are done. □

We now prove some preliminary results, and then establish the convergence of sequences $\{x^k\}$, $\{y^k\}$ and $\{s^k\}$.

4 Preliminary Results

In this section we prove some preliminary results and two theorems related to the ellipsoidal approximating problem discussed in section 2.1.

Lemma 3 *Let r_j and $s_j > 0$ for each $j = 1, \cdots, l$ be arbitrary real numbers. Then*

$$\frac{|\sum_{j=1}^{l} r_j|}{\sum_{j=1}^{l} s_j} \leq max_j\frac{|r_j|}{s_j}$$

Proof Let

$$\frac{|r_k|}{s_k} = \max_j\frac{|r_j|}{s_j}.$$

Then for each $j = 1, \cdots, l$, $s_j|r_k| \geq s_k|r_j|$. Adding all these l inequalities we have

$$|r_k|\sum_{j=1}^{l} s_j \geq s_k\sum_{j=1}^{l}|r_j| \geq s_k|\sum_{j=1}^{l} r_j|$$

and we are done. □

Theorem 4 *For every $x > 0$, the function $\|(AX^2A^T)^{-1}AX^2p\| \leq q(A,p)$, where $q(A,p) > 0$.*

Proof By the Cauchy-Binet Theorem and Cramer's rule we can write the ith component

$$((AX^2A^T)^{-1}AX^2p)_i = \frac{\sum_J (x_{j_1}\cdots x_{j_m})^2 \det(A_J)\det(A_J^{(i)})}{\sum_J (x_{j_1}\cdots x_{j_m})^2 (\det(A_J))^2}$$

where $A^{(i)}$ is the matrix obtained by replacing the ith row of A by p^T, and the sum is over all possible J which are ordered sets of m elements $1 \leq j_1 \leq \cdots \leq j_m \leq n$. Since a term in the denominator is zero only when the term in the numerator is zero, from Lemma 3 we can use

$$q(A,p) = \sqrt{n}\ \max_J \frac{|\det(A_J^{(i)})|}{|\det(A_J)|}$$

and we are done. □

Given a $k \times l$ matrix Q of rank k, a k vector $c \neq 0$ and a $l \times l$ diagonal matrix D with positive diagonal entries, consider the problem:

$$\begin{aligned} &\text{maximize } c^T x \\ &x^T Q D Q^T x \ \leq\ 1. \end{aligned} \tag{1}$$

The following result establishes a relation between the objective function value and the solution vector of this problem.

Theorem 5 *Let $\hat{x}$ solve the problem 1. There exists a constant $p(Q,c) > 0$ such that*

$$|c^T\hat{x}| \geq p(Q,c)\|\hat{x}\|.$$

Proof Let $c, B_1, \cdots, B_{k-1}$ be a basis for $\boldsymbol{R}^k$ with $B = (B_1, \cdots, B_{k-1})$ an orthnormal basis for the orthogonal complement of the one dimensional subspace spanned by c. Thus, $c^T B = 0$ and $B^T B = I$. Expressing the columns of Q in this basis we obtain

$$Q = ca^T + BR$$

where a is a l vector and R is a $(k-1) \times l$ matrix. Since Q has full rank k, so does (a, R). Also, let

$$x = uc + Bv$$

where u is a scalar. Rewriting the problem 1 in the variables u and v we get

$$\begin{aligned} c^T x &= uc^T c \\ x^T QDQ^T x &= (uc^T c)^2 a^T Da + 2(uc^T c)a^T DR^T v + v^T RDR^T v \end{aligned}$$

and if $w = c^T cu$, problem 1 is equivalent to

$$\begin{aligned} &\text{maximize } w \\ w^2 a^T Da + 2wa^T DR^T v + v^T RDR^T v &\leq 1. \end{aligned}$$

Let λ be the multiplier on the constraint. Then the optimality conditions for this problem can be stated as follows:

$$1 + \lambda(2wa^T Da + 2a^T DR^T v) = 0 \tag{2}$$

$$wRDa + RDR^T v = 0 \tag{3}$$

and we note that Equation 3 has the solution

$$\hat{v} = -w(RDR^T)^{-1} RDa$$

and let $\hat{w}$ be the solution obtained by substituting $\hat{v}$ into Equation 2. Then, after computing $\hat{x}$ we note that

$$\begin{aligned} c^T \hat{x} &= \hat{w} \\ \hat{x} &= \hat{u}c + B\hat{v} \\ &= \frac{\hat{w}}{c^T c} c - \hat{w}B(RDR^T)^{-1} RDa. \end{aligned}$$

Thus

$$\frac{\hat{x}}{c^T \hat{x}} = \frac{c}{c^T c} - B(RDR^T)^{-1} RDa,$$

and, noting that $\|B\| \leq 1$ we get

$$\frac{\|\hat{x}\|}{|c^T \hat{x}|} \leq \frac{1}{\|c\|} + \|(RDR^T)^{-1} RDa\|$$

and if, from Theorem 4, $q(R, a)$ bounds the second term of the right hand side above, since both R and a are only functions of Q and c, we get our result with $p(Q, c) = 1/(\frac{1}{\|c\|} + q(R, a))$.

□

We now establish an important corollary to the preceding theorem. Given an $x > 0$, consider the ellipsoidal approximating problem (EAP) of Section (2.1).

Corollary 6 *There exists a constant $p(A,c) > 0$ such that for every $x > 0$, if $\hat{v}$ solves the above problem,*

$$\|\hat{v}\| \leq p(A,c)c^T\hat{v}.$$

Proof Let the set of vectors $\{Q_1, \cdots, Q_{n-m}\}$ be an orthonormal basis for the null space $\mathcal{N}(A)$ of A; i.e., for $Q = (Q_1, \cdots, Q_{n-m})$, $Q^TQ = I$. Since, in the above problem, $v \in \mathcal{N}(A)$, we note that $c^Tv = c_n^Tv$ where c_n is the projection of c into the null space of A. Define $\bar{c}$ and u such that

$$\begin{aligned} c_n &= Q\bar{c} \\ v &= Qu. \end{aligned}$$

Expressing the problem in variables u we obtain an equivalent problem:

$$\begin{aligned} &\text{maximize } \bar{c}^Tu \\ &u^TQ^TX_k^{-2}Qu \;\leq\; 1. \end{aligned}$$

Let u^k solve this problem. From Theorem 5, there is a constant $p(Q,\bar{c}) > 0$ such that

$$\begin{aligned} c^Tv^k &= \bar{c}^Tu^k \\ &\geq p(Q,\bar{c})\|u^k\| \\ &= p(Q,\bar{c})\|v^k\| \end{aligned}$$

since $\|v^k\|^2 = (u^k)^TQ^TQu^k = \|u^k\|^2$. Since $Q, \bar{c}$ are only functions of A, c, we are done. □

Let R be an arbitrary $r \times l$ matrix, with rank $\leq r$ and let $b \in \mathcal{R}(R)$. Also let $z \in \boldsymbol{R}^n$ be any vector. Consider the following problem:

$$\begin{aligned} \text{minimize } &\|z - x\|^2 \\ &Rx = b \end{aligned}$$

We can then prove:

Theorem 7 *There is a constant $q(R) > 0$ such that for every z, for the solution $z(x)$ of the above problem we have*

$$\|z - z(x)\| \leq q(R)\|Rz - b\|$$

Proof Let $\bar{R}$ be a $\bar{r} \times l$ submatrix of R which has full row rank. Since $b \in \mathcal{R}(R)$,

$$\{x : Rx = b\} = \{x : \bar{R}x = \bar{b}\}$$

for some $\bar{b}$. Then, it is readily seen that $z(x)$ solves the problem if and only if for some y,

$$\begin{aligned} 2(z - z(x)) - \bar{R}^T y &= 0 \\ \bar{R} z(x) &= \bar{b} \end{aligned}$$

and the result is readily seen by setting $q(R) = \bar{R}^T(\bar{R}\bar{R}^T)^{-1}$. □

The following simple lemma on the natural logarithmic function is important.

Lemma 8 *Let $w \in \mathbf{R}^q$ and $0 < \lambda < 1$ be such that $w_j \leq \lambda$. Then*

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

Proof The result follows as a consequence of the following facts: When $a < 0$, by considering the function $g(x) = \log(1 + x) - x + \frac{x^2}{2(1-\lambda)}$, it is readily confirmed that its derivative is positive when $x > 0$, and thus $g(|a|)$ is an increasing function of $|a|$ and thus

$$\log(1 - a) \geq -a - \frac{a^2}{2(1-\lambda)}.$$

To see that second fact, let $0 \leq a \leq \lambda$. Then

$$\begin{aligned} \log(1 - a) &= -a - \frac{a^2}{2} - \frac{a^3}{3} - \\ &\geq -a - \frac{|a|^2}{2} - \frac{|a|^3}{2} - \\ &= -a - \frac{a^2}{2(1-a)} \\ &\geq -a - \frac{a^2}{2(1-\lambda)}. \end{aligned}$$

The result now readily follows from the above two facts. □

5 Convergence of Primal Sequence

We assume here that the sequence $\{c^T x^k\}$ is bounded. We now prove that the primal sequence $\{x^k\}$ converges. We prove this in the next theorem.

Theorem 9 *Let the assumptions 1 through 3 hold, $\{c^T x^k\}$ be bounded, and the sequence $\{x^k\}$ be generated by the primal affine scaling algorithm. Then $x^k \to x^*$.*

Proof Consider the approximating problem of Section 2.1. Note that

$$\alpha\theta_k \hat{v}^k = x^k - x^{k+1}$$

where $\hat{v}^k = \frac{X_k^2 s^k}{\|X_k s^k\|}$ is the solution to the ellipsoidal approximating problem (EAP) of Section (2.1). We will now show that $\sum_{k=1}^{\infty} \alpha\theta_k \|\hat{v}^k\| < \infty$ and thus the sequence $\{x^k\}$ converges. But from Theorem 2 and Corollary 6

$$\infty > c^T x^1 - c^* = \sum_{k=1}^{\infty} c^T(x^k - x^{k+1}) = \sum_{k=1}^{\infty} \alpha\theta_k c^T \hat{v}^k \geq p(A,c) \sum_{k=1}^{\infty} \alpha\theta_k \|\hat{v}^k\|$$

and thus $\{x^k\}$ converges, say to x^*, and we are done. □

6 More on the Sequences

We have seen in the previous section that the primal sequence converges. Let

$$\lim_{k\to\infty} x^k = x^*.$$

Let

$$N = \{j : x_j^* = 0\}$$

$$B = \{j : x_j^* > 0\}$$

and let F be the face of $P = \{x : Ax = b, x \geq 0\}$ such that x^* lies in the relative interior of the face F; i.e., it is the smallest face of P containing x^*. We now prove an important property of the sequence $\{x^k\}$.

Theorem 10 *There is a $\delta > 0$ such that for each $k = 1, 2, \cdots$*

$$\frac{c^T x^k - c^*}{\sum_{j\in N} x_j^k} \geq \delta, \quad and \quad \frac{c^T x^k - c^*}{\sum_{j\in B} |x_j^k - x_j^*|} \geq \delta.$$

Proof Note that using the results of Theorem 2 and Corollary 6 with $q = |N|$, the number of elements in N, and $M = q(A,c)$,

$$\begin{aligned} c^T x^k - c^* &= \sum_{j=0}^{\infty} c^T(x^{k+j} - x^{k+j+1}) \\ &\geq \frac{1}{M}\sum_{j=0}^{\infty} \|x^{k+j} - x^{k+j+1}\| \\ &\geq \frac{1}{M}\|\sum_{j=0}^{\infty}(x^{k+j} - x^{k+j+1})\| \\ &= \frac{1}{M}\|x^k - x^*\| \end{aligned}$$

Thus $c^T x^k - c^* \geq \frac{\|x_N^k\|}{M} \geq \frac{\sum_{j\in N} x_j^k}{\sqrt{q}M}$. Also $c^T x^k - c^* \geq \frac{\|x_B^k - x_B^*\|}{M} \geq \frac{\sum_{j\in B}|x_j^k - x_j^*|}{\sqrt{n-q}M}$ and we have our result with $\delta = \min\{\frac{1}{\sqrt{q}M}, \frac{1}{\sqrt{n-q}M}\}$. □

The following is a well known result on the convergence rate of the sequence $\{c^T x^k - c^*\}$.

Theorem 11 *There exists an $L \geq 1$ such that for all $k \geq L$*

$$c^T x^{k+1} - c^* \leq (1 - \frac{\alpha}{\sqrt{n}})(c^T x^k - c^*).$$

Proof Consider the problem (EAP) of section 2.1. We note that $\frac{X_k^2 s^k}{\|X_k s^k\|}$ solves this problem and that $\frac{(x^k - x^*)}{\|X_k^{-1}(x^k - x^*)\|}$ is feasible for it. Thus, from the above fact and Theorem 1,

$$\frac{c^T x^k - x^*}{\|X_k^{-1}(x^k - x^*)\|} \leq \frac{c^T X_k^2 s^k}{\|X_k s^k\|} = \|X_k s^k\|. \tag{4}$$

Now $\|X_k^{-1}(x^k - x^*)\|^2 = \|e_B - X_{B,k}^{-1} x_B^*\|^2 + \|e_N\|^2$, and since $X_{B,k}^{-1} x_B^* \to e_B$, there is an $L \geq 1$ such that for all $k \geq L$, $\|e_B - X_{B,k}^{-1} x_B^*\|^2 \leq n - q$. Thus, for each $k \geq L$, $\|X_k^{-1}(x^k - x^*)\| \leq \sqrt{n}$. Now, as $c^T x^{k+1} - c^* = c^T x^k - c^* - \frac{\alpha}{\phi(X_k s^k)} c^T X_k^2 s^k$, we have

$$\frac{c^T x^{k+1} - c^*}{c^T x^k - c^*} = 1 - \alpha \frac{\|X_k s^k\|}{c^T x^k - c^*}$$

and the result follows from the equation (4). □

Now define

$$D = \{(y,s) : A_B^T y = c_B, A_N^T y + s_N = c_N, s_B = 0\}$$

to be the set of all tentative dual solutions that are complementary to each solution in F. We can readily establish the following:

Lemma 12 $D \neq \emptyset$

Proof From Step 1 and Theorem 4, the sequences $\{y^k\}$ and $\{s^k\}$ are bounded, and thus have cluster points y^* and s^* respectively. From Theorem 2, since $X_k s^k \to 0$, $s_B^* = 0$. The result now follows since $A^T y^* + s^* = c$. □

Consider the sequence $\{u^k\}$ defined by

$$u^k = \frac{X_k s^k}{c^T x^k - c^*}.$$

We now state an important result about this sequence. Its proof is given in the appendix.

Theorem 13 *The sequence $\{u^k\}$ has the following properties:*

1. *It is bounded.*
2. *There is an $L \geq 1$ such that for each $k \geq L$,*
   (a) $\|u^k\|^2 = \|u_N^k\|^2 + \gamma_k$ *and* $\sum_{k=L}^{\infty} |\gamma_k| < \infty$.
   (b) $e^T u_N^k = 1 + \delta_k$ *and* $\sum_{k=L}^{\infty} |\delta_k| < \infty$
   (c) $\frac{1}{\alpha} \geq \phi(u_N^k) \geq \frac{1}{2q}$
   (d) $\phi(u^k) = \phi(u_N^k)$.

7 Local Potential Function

To show the convergence of the dual sequence, it is necessary to control the step size. In particular, the step size $\alpha \leq \frac{2}{3}$. This result is achieved by the use of a local version of the potential function, which we now introduce.

For any $x > 0$ with $c^T x - c^* > 0$ and $N \subset \{1, \cdots, n\}$ with $q = |N|$, define

$$F_N(x) = q\log(c^T x - c^*) - \sum_{j \in N} \log(x_j).$$

For the sequence $\{x^k\}$ we can prove that

Theorem 14 *There exists an $L \geq 1$ such that for all $k \geq L$*

$$F_N(x^{k+1}) - F_N(x^k) = q\log(1 - \theta\|w_N^k\|^2 + \theta(\frac{2\delta_k}{q} + \gamma_k)) - \sum_{j \in N} \log(1 - \theta w_j^k)$$

where $w_N^k = u_N^k - \frac{1}{q}e$, $\tilde{\alpha} = \frac{\alpha}{\phi(u_N^k)}$, $\theta = \frac{q\tilde{\alpha}}{q - \tilde{\alpha}}$ *and* $\sum_{k=L}^{\infty} |\delta_k| < \infty$, $\sum_{k=L}^{\infty} |\gamma_k| < \infty$.

Proof Since $c^T x^{k+1} - c^* = c^T x^k - c^* - \frac{\alpha}{\phi(X_k s^k)} c^T X_k^2 s^k$, from Theorem 1 and simple algebra we see that

$$\frac{c^T x^{k+1} - c^*}{c^T x^k - c^*} = 1 - \frac{\alpha}{\phi(u^k)} \|u^k\|^2 \tag{5}$$

and, for each $j \in N$

$$\frac{x_j^{k+1}}{x_j^k} = 1 - \frac{\alpha}{\phi(u^k)} u_j^k.$$

From Theorem 13 part 2(a), there is an $L \geq 1$ such that for all $k \geq L$,

$$\frac{c^T x^{k+1} - c^*}{c^T x^k - c^*} = 1 - \frac{\alpha}{\phi(u_N^k)} \|u_N^k\|^2 - \frac{\alpha \gamma_k}{\phi(u_N^k)}.$$

Now let $k \geq L$, and note that from Theorem 13 part 2(b),

$$\|w_N^k\|^2 = \|u_N^k\|^2 - \frac{2e^T u_N^k}{q} + \frac{1}{q} = \|u_N^k\|^2 - \frac{1}{q} - \frac{2\delta_k}{q},$$

and we have our result by observing that

$$\frac{q}{q - \tilde{\alpha}}(1 - \tilde{\alpha}\|u_N^k\|^2 - \tilde{\alpha}\gamma_k) = 1 - \theta\|w_N^k\|^2 - \theta(\frac{2\delta_k}{q} + \gamma_k)$$

and

$$\frac{q}{q - \tilde{\alpha}}(1 - \tilde{\alpha} u_j^k) = 1 - \theta w_j^k$$

and the constant $\frac{q}{q - \tilde{\alpha}}$ cancels. □

8 Convergence of Dual Sequence

We are now ready to prove that with the step size of $\alpha \leq \frac{2}{3}$ the dual sequence converges to the analytic center of the optimal face of the dual polytope. Throughout this section we will assume that the sequence $\{c^T x^k\}$ is bounded and we now define the analytic center of the optimal dual face as the solution (y^*, s^*) to the following problem:

$$\begin{array}{lll} \text{maximize} & \sum_{j \in N} \log s_j & \\ & (y, s) \in D & \text{ACP} \\ & s_N > 0. & \end{array}$$

The solution to (ACP) must satisfy the following K.K.T. conditions:

$$\begin{aligned} A_B x_B + A_N x_N &= 0 \\ s_N^{-1} + x_N &= 0 \\ A_B^T y &= c_B \\ A_N^T y + s_N &= c_N \\ s_N &> 0 \\ s_B &= 0 \end{aligned}$$

We now establish a lemma.

Lemma 15 *If the analytic center defined by (ACP) exists, it is unique.*

Proof The uniqueness follows from the strict concavity of the log function. □

We are now ready to prove our main theorem.

Theorem 16 *Let the sequence generated by the affine scaling algorithm be infinite, and the step size* $\alpha \leq \frac{2}{3}$. *Then*

1. $x^k \to x^*$
2. $y^k \to y^*$
3. $s^k \to s^*$, *and* $s^* \geq 0$.

The limits are optimal solutions of their respective problems, and satisfy the strict complementarity condition. In addition, the dual solution (y^*, s^*) *converges to the analytic center of the optimal dual face, and the primal solution to the relative interior of the optimal primal face.*

Proof From Theorems 8, 13 and 14 and the fact that for any $a < 1$, $\log(1-a) < -a$, we obtain an L such that for each $k \geq L$

$$\begin{aligned} F_N(x^{k+1}) - F_N(x^k) &\leq -q\theta(\|w_N^k\|^2 + \frac{2\delta_k}{q} + \gamma_k) + \theta(e^T w_N^k + \theta \frac{\|w_N^k\|^2}{2(1-\theta\phi(w_N^k))}) \\ &= \theta\|w_N^k\|^2(-q + \frac{\theta}{2(1-\theta\phi(w_N^k))}) - \theta(q\gamma_k + \delta_k) \\ &= \theta\|w_N^k\|^2(-q + \frac{\alpha}{2(1-\alpha)}\frac{1}{\phi(u_N^k)}) - \theta(q\gamma_k + \delta_k). \end{aligned}$$

From Theorem 13 part 2(c) it can be shown that $\theta \geq \alpha^2$. Now, as $\|w_N^k\| = \|u_N^k - \frac{1}{q}e\|$, we note that

$$\phi(u_N^k) \geq \frac{1 + \|w_N^k\|}{q}.$$

Thus, we have

$$F_N(x^{k+1}) - F_N(x^k) \leq \theta \|w_N^k\|^2 (-q + \frac{\alpha}{2(1-\alpha)} \frac{q}{1 + \|w_N^k\|}) - \theta(q\gamma_k + \delta_k). \tag{6}$$

For each $\alpha \leq \frac{2}{3}$, $\frac{\alpha}{2(1-\alpha)} \leq 1$. Hence

$$F_N(x^{k+1}) - F_N(x^k) \leq -\frac{q\theta \|w_N^k\|^3}{1 + \|w_N^k\|} - \theta(q\gamma_k + \delta_k).$$

From Theorem 10, $\sum_{k=L}^{\infty}(F_N(x^{k+1}) - F_N(x^k)) > -\infty$, and, from Theorem 13 parts 2(a) and (b), $0 \leq \sum_{k=L}^{\infty}(q|\gamma_k| + |\delta_k|) < \infty$ Thus, we must have $\|w_N^k\| \to 0$, and we get

$$\lim_{k \to \infty} u_N^k = \frac{1}{q}e. \tag{7}$$

Now, consider the sequences $\{y^k\}$, $\{s^k\}$, $\{\frac{x_j^k}{c^T x^k - c^T x^*}\}$ for each $j \in N$ and $\{\frac{x_j^k - x_j^*}{c^T x^k - c^T x^*}\}$ for each $j \in B$. Let $s^k \to s^*$ on some subsequence $K' \subset \{1, 2, \cdots\}$. From Theorem 4, Steps 1 and 2 and Theorem 10, these are bounded for each $1 \leq j \leq n$. Thus, on some common subsequence $K \subset K'$

$$y^k \to y^* \ , \ s^k \to s^*$$

$$\frac{q x_j^k}{c^T x^k - c^T x^*} \to a_j \text{ for each } j \in N$$

and

$$\frac{q(x_j^k - x_j^*)}{c^T x^k - c^T x^*} \to b_j \text{ for each } j \in B.$$

In view of (7), $a_j > 0$ for each $j \in N$. Also, since $s_j^k = \frac{c^T x^k - c^T x^*}{q x_j^k}$, we note that $\lim_{k \in K} s_j^k = a_j^{-1}$.

Now, $A_N x_N^k + A_B x_B^k = A_B x_B^*$, we see that

$$A_N a + A_B b = 0$$

and thus $s_j = a_j^{-1}$ for each $j \in N$ and $x_N = -a$, $x_B = -b$ and $s_B = 0$, $s_N = s_N^*$, $y = y^*$ solve the K.K.T.conditions for the analytic center problem (ACP). Thus s_N^k converges to a^{-1}, the analytic center, on each convergent subsequence, and thus parts (2) and (3) of the theorem are established by Lemma 15. We finish the proof by observing that x^* is primal feasible, (y^*, s^*) is dual feasible and the pair satisfies the complementary slackness theorem (by Theorem 2) and thus are optimal solutions for the respective problems. The strict complementarity holds as $s_N^* > 0$. □

The following is a sharp bound on the convergence rate of $\{c^T x^k - c^*\}$.

Theorem 17 *Let $\alpha \leq \frac{2}{3}$. Then*

$$\lim_{k \to \infty} \frac{c^T x^{k+1} - c^*}{c^T x^k - c^*} = 1 - \alpha$$

Proof Follows readily from Theorem 13 parts 2 (a), (c) and (d), equation (5) and equation (7). □

We now prove another theorem which shows the convergence to optimality for a slightly larger stepsize than $\frac{2}{3}$.

Theorem 18 *Let $\frac{2}{3} < \alpha \leq \frac{2+2q}{2+3q}$. Then, there exists a subsequence $K \subset \{1, 2, \cdots\}$ such that*

1. $\lim_{k \in K} x^k \to x^*$

2. $\lim_{k \in K} y^k \to y^*$

3. $\lim_{k \in K} s^k \to s^*$, *and* $s^* \geq 0$.

Thus x^* and (y^*, s^*) are optimum solutions of the dual pair. Also, they satisfy the strict complementarity condition.

Proof As a result of Theorem 10

$$\liminf(F(x^{k+1}) - F(x^k)) \geq 0.$$

Thus, either on some subsequence K', $F(x^{k+1}) - F(x^k) \geq 0$ for each $k \in K'$ or $\lim(F(x^{k+1}) - F(x^k)) \to 0$. Thus, on some subsequence, from Equation 6

$$\liminf_{k \in K'}(-q + \frac{\alpha}{2(1-\alpha)} \frac{q}{1 + \|w_N^k\|}) \geq 0.$$

Let $K \subset K'$ be such that as $k \in K$ goes to infinity, $x^k \to x^*$, $y^k \to y^*$, $s^k \to s^*$, $u_N^k \to u_N^*$ and $\|w_N^k\| \to \theta \geq 0$. Then

$$-q + \frac{\alpha}{2(1-\alpha)} \frac{q}{1+\theta} \geq 0.$$

Letting $\alpha \leq \frac{2+2q}{2+3q}$, it is readily confirmed that $\theta \leq \frac{1}{q}$. Since $\|u_N^* - \frac{1}{q} e\| = \theta$ and $e^T u_N^* = 1$, we can conclude that $u_N^* > 0$, and we are done. □

9 Appendix

We now prove the Theorem 13.

Proof To see (1), note that from Theorem 1, for any $(\tilde{y}, \tilde{s}) \in D$, $\|X_k s^k\|^2 = c^T v^k = (A^T\tilde{y} + \tilde{s})^T v^k = \tilde{s}_N v_N^k = (X_{N,k}\tilde{s}_N)^T(X_{N,k}^{-1} v_N^k) \leq \phi(x_N^k)\|\tilde{s}_N\|\|X_k s^k\|$. From Theorem 10 we obtain that $\|u^k\| \leq M\|\tilde{s}_N\|$ for $M = p(A, c)$.

To see 2(a), from Corollary 6 and for any $(\tilde{y}, \tilde{s}) \in D$, $\|v_B^k\| \leq \|v^k\| \leq Mc^T v^k = M\tilde{s}_N v_N^k \leq M\|\tilde{s}_N\|\|v_N^k\| \leq M\phi(x_N^k)\|\tilde{s}_N\|\|u_N^k\|$. Also, $\|u_B^k\| \leq \|X_{B,k}^{-1}\|\|v_B^k\| \leq \|X_{B,k}^{-1}\|.M\phi(x_N^k)\|\tilde{s}_N\|\|u_N^k\|$. From Theorem 9, $\phi(x_N^k) \to 0$, and $x_B^k \to x_B^* > 0$. Thus there exists an $\hat{M}$ and an $\bar{L} \geq 1$ such that for every $k \geq \bar{L}$

$$\gamma_k = \|u_B^k\|^2 \leq (M\|X_{B,k}^{-1}\|\phi(x_N^k)\|\tilde{s}_N\|\|u_N^k\|)^2 < 1, M\|X_{B,k}^{-1}\|\|\tilde{s}_N\|\|u_N^k\| < \hat{M}.$$

Since, $\|u^k\|^2 = \|u_N^k\|^2 + \|u_B^k\|^2$ we are done if $\sum_{k=\bar{L}}^{\infty} \phi(x_N^k) < \infty$. But from Theorem 10 and Theorem 11, there is an $\tilde{L} \geq 1$ such that for all $k \geq \tilde{L}$,

$$\begin{aligned}
\phi(x_N^k) &\leq \|x^k - x^*\| \\
&\leq M(c^T x^k - c^*) \\
&\leq M(1 - \frac{\alpha}{\sqrt{n}})^{k-\tilde{L}}(c^T x^{\tilde{L}} - c^*)
\end{aligned}$$

and thus the sum is finite.

To see part 2(b), let $(\tilde{y}^k, \tilde{s}^k)$ solve the following problem:

$$\begin{array}{ll}
\text{minimize} & \|s^k - s\| \\
(y, s) \in D &
\end{array}$$

From Theorem 7, there is a constant $\tilde{N} > 0$ such that $\|\tilde{s}^k - s^k\| \leq \tilde{N}\|s_B^k\|$. Since $\{s^k\}$ is bounded, so is $\{\tilde{s}^k\}$, and let $\|s^k\| \leq \hat{N}$ and $\|\tilde{s}^k\| \leq \hat{N}$ for each k. Also, since s^k solves the least squares problem (LSQ), we get $\|X_{B,k}s_B^k\|^2 + \|X_{N,k}s_N^k\|^2 \leq \|X_{N,k}\tilde{s}_N^k\|^2$.

Now,

$$\begin{aligned}
\|s_B^k\|^2 &\leq \|X_{B,k}^{-1}\|^2\|X_{B,k}s_B^k\|^2 \\
&\leq \|X_{B,k}^{-1}\|^2(\|X_{N,k}\tilde{s}_N^k\|^2 - \|X_{N,k}s_N^k\|^2) \\
&= \|X_{B,k}^{-1}\|^2(X_{N,k}\tilde{s}_N^k - X_{N,k}s_N^k)^T(X_{N,k}\tilde{s}_N^k + X_{N,k}s_N^k)
\end{aligned}$$

$$\leq \|X_{B,k}^{-1}\|^2\phi(x_N^k)^2\|\tilde{s}_N^k - s_N^k\|\|\tilde{s}_N^k + s_N^k\|$$
$$\leq \|X_{B,k}^{-1}\|^2\phi(x_N^k)^2 2\hat{N}\tilde{N}\|s_B^k\|.$$

Thus, $\|\tilde{s}_N^k - s_N^k\| \leq 2\hat{N}\tilde{N}^2\|X_{B,k}^{-1}\|^2\phi(x_N^k)^2$. Now, $c^Tx^k - c^* = c^T(x^k - x^*) = (\tilde{s}^k + A^T\tilde{y}^k)^T(x^k - x^*) = \tilde{s}_N^k x_N^k$ Thus, from Theorem 10,

$$\begin{aligned}|e^Tu_N^k - 1| &= \frac{|(s_N^k)^Tx_N^k - (\tilde{s}_N^k)^Tx_N^k|}{c^Tx^k - c^*} \\ &\leq \frac{\|s_N^k - \tilde{s}_N^k\|\|x_N^k\|}{M^{-1}\|x_N^k\|} \\ &\leq 2M\hat{N}\tilde{N}^2\|X_{B,k}^{-1}\|^2\phi(x_N^k)^2.\end{aligned}$$

Thus, there exists an $\hat{L} \geq 1$ such that for all $k \geq \hat{L}$,

$$|\delta_k| = |e^Tu_N^k - 1| \leq 2M\hat{N}\tilde{N}^2\|X_{B,k}^{-1}\|^2\phi(x_N^k)^2 < 1.$$

The required sum can be shown to be finite by the same argument as for the sum in 2(a).

To see part 2(c), using the result of part 2(b), there exists an $L^* \geq \tilde{L}$ such that for all $k \geq L^*$,

$$\phi(u_N^k) \geq \frac{1+\delta_k}{q} \geq \frac{1}{2q}.$$

Also, since $1 - \frac{\alpha}{\phi(u_N^k)}\|u_N^k\|^2 > 0$, we have

$$\phi(u_N^k) \leq \frac{\|u_N^k\|^2}{\phi(u_N^k)} \leq \frac{1}{\alpha}.$$

To see part 2(d), using the results of 2(a) and 2(c), we note that there is an $L' \geq L^*$ such that for all $k \geq L'$,

$$\phi(u_B^k) \leq \|u_B^k\| \leq \hat{M}\phi(x_N^k) < \frac{1}{2q} \leq \phi(u_N^k)$$

and we are done with $L = \max\{\bar{L}, \hat{L}, L'\}$, since $\phi(u^k) = \max\{\phi(u_B^k), \phi(u_N^k)\}$. □

References

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

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

[3] J. R. Birge and C. Rosa, "A simplified proof of the convergence of affine scaling," Technical Report, Department of Industrial and Operations Engineering, University of Michigan, (Ann Arbor, Michigan 1991).

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

[5] I. I. Dikin, " On the convergence of an iterative process," *Upravlyaemye Sistemi* 12 (1974) 54-60. (In Russian) .

[6] I. I. Dikin, "The convergence of dual variables," Technical Report, Siberian Energy Institute, Irkutsk, Russia, December 1991.

[7] I. I. Dikin, "Determination of interior point of one system of linear inequalities" *Kibernetica and system analysis* 1 (1992).

[8] L. A. Hall and R. J. Vanderbei " Two-thirds is sharp for affine scaling," Technical Report SOR-92/9, Department of Civil Engineering and Operations Research, Princeton University, Princeton, New Jersey, 08544, September 1992.

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

[10] R. D. C. Monteiro, T. Tsuchiya, and Y. Wang, " A simplified global convergence proof of the affine scaling algorithm," Technical Report, Dept. of Systems and Industrial Engineering, University of Arizona, Tucson, AZ 85721, USA, October 1992.

[11] M. Muramatsu and T. Tsuchiya, " A convergence analysis of a long-step variant of the projective scaling algorithm," Research Memorandum 454, The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106, Japan, November 1992.

[12] P. Tseng and Z. Q. Luo, " On the convergence of the affine-scaling algorithm," *Mathematical Programming* 56 (1992) 301-319.

[13] T. Tsuchiya, " Global convergence of the affine-scaling methods for degenerate linear programming problems," *Mathematical Programming* 52 (1991) 377-404.

[14] T. Tsuchiya, " Global convergence property of the affine scaling method for primal degenerate linear programming problems'" *Mathematics of Operations Research* 17 (1992) 527- 557.

[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).

[16] R. J. Vanderbei, M. J. Meketon and B. A. Freedman, "A Modification of Karmarkar's linear programming algorithm," *Algorithmica* 1 (1986) 395 - 407.

[17] R. J. Vanderbei and J. C. Lagarias, " I. I. Dikin's convergence result for the affine-scaling algorithm," In J. C. Lagarias and M. J. Todd, editors, *Mathematical Developments Arising from Linear Programming: Proceedings of a Joint Summer Research Conference held at Bowdoin College, Brunswick, Maine, USA, June/July 1988* volume 114 of Contemporary Mathematics, pages 109-119. American Mathematical Society, Providence, Rhode Island, USA, 1990.