A THREE STEP QUADRATICALLY CONVERGENT IMPLEMENTATION OF THE PRIMAL AFFINE SCALING METHOD

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

Technical Report 93-9

February 1993
Revised July 1993

A Three Step Quadratically Convergent Implementation of the Primal Affine Scaling Method

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

February 26, 1993 (revised)

Abstract

Recently, Tsuchiya and Monterio have proposed an implementation of the affine scaling method that attains , asymptotically, a two step superlinear convergence rate of 1.3. This is achieved by a step size selection rule under which the affine scaling method can be viewed as a predictor-corrector method. In this paper, we propose a different, but related, step size selection strategy that asymptotically, attains a three step quadratic convergence rate.

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

Abbreviated title: Quadratically convergent affine scaling method

1 Introduction

In this paper we consider the application of the (primal) affine scaling method for solving the following linear programming problem:

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

where we assume that the matrix A is an $m \times n$ matrix, b is an m vector and c is an n vector. In addition we will assume that

Assumption 1 *The linear program is feasible and has an interior point.*

Assumption 2 *A has full row rank m.*

Assumption 3 *The objective function is not constant on the feasible region.*

There have been some significant recent developments in the convergence theory of this method. A notable development is the series of papers Tsuchiya [17], [18], Dikin [6], [7] and Tsuchiya and Muramatsu [19]. The last of these papers shows that if the step-size at each iteration, is restricted to $\frac{2}{3}$ of the step to the boundary, the dual sequence converges to the analytic center of the optimal dual face, and the primal sequence to the interior of the optimal primal face. Prior to these results, the convergence of the primal sequence was known, but the optimality of its limit point was known only under the non-degeneracy assumption, Vanderbei and Lagarias [22]. Simpler proofs of the $\frac{2}{3}$ result can be found in Monterio, Tsuchiya and Wang [12] and Saigal [15]. Two-thirds is sharp, (as is shown by the example produced by Hall and Vanderbei [9]), in the sense that with a larger step size the dual sequence may not converge. The optimality of the limit of the primal sequence with larger step-sizes is still an open question. In Saigal [15] the convergence to optimality has been shown for a slightly larger step-size of $\frac{2q+2}{3q+2}$, where q is the number of variables in the limit at 0. This fraction can be easily increased to $\frac{2q}{3q-1}$, and it appears that this may be the largest constant step-size under which the convergence to optimality can be established.

In another recent paper, Tsuchiya and Monterio [20] have achieved another milestone. By establishing a connection between the affine scaling step and the Newton step, they have devised a strategy of adjusting step sizes under which the dual sequence converges to the analytic center of the optimal face and the primal sequence to the interior of the primal optimal face. Their method, asymptotically, attains a two step super-linear convergence rate of 1.3, and can be viewed as a predictor-corrector method. They show that asymptotically, one can take large step sizes (predictor steps) provided the step-size at the next step is restricted to $\frac{1}{2}$ (corrector steps). In this paper we build on their work and generate a different step selection strategy, which also can be viewed as a predictor-corrector method. This step selection strategy asymptotically attains a three step quadratic convergence rate. In this method, we take two corrector steps between each pair of predictor steps. We also show that the primal converges to the interior of the optimal primal face and the dual to the analytic center of the optimal dual face.

This paper is organized as follows. Besides the introduction it has 9 other sections. In Section 2, we introduce the affine scaling method, and state some of its known properties. In Section 3 we investigate the affine scaling step with a view to relating it to Newton's method. In Section 4, we study an analytic center problem, and the application of Newton's method for computing it. In Section 5, we relate the affine scaling and Newton directions for this center problem, and in Section 6 we the relate centers of two polytopes. We introduce the accelerated affine scaling method in Section 7, and in Section 8 we establish the preliminary results required for proving the three step quadratic convergence of the method, which is established in Sections 9 and 10. Finally in Section 11, we compare the theoretical efficiency of this method a two step quadratically convergent method and the superlinearly convergent method of Tsuchiya and Monterio [20].

2 The Affine Scaling Method

In this section we present the method and known results (generally without proof) about the sequences generated by this method. We now present the affine scaling method we will deal with in this paper:

Step 0 Let x^0 be an interior point solution 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 $\phi(x) = \max_j x_j$.

Step 4 Step Size Selection: Choose, by some rule, the next step size $0 < \alpha_k < 1$. Also, if $\theta_k = 1$ set $\alpha_k = 1$.

Step 5 Next Interior Point:

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

Step 6 Iterative Step: If $x_j^{k+1} = 0$ for some j, then STOP. x^{k+1} is an optimal solution to the primal, y^k the optimum solution to the dual. Otherwise set $k = k+1$ and go to step 1.

Assume that the method presented above generates infinite sequences $\{x^k\}$, $\{y^k\}$ and $\{s^k\}$. Then the following can be shown for these sequences, and the proofs can be found in the references cited:

Theorem 1 *Let the sequence $\{c^Tx^k\}$ be bounded.*

1. *$\{c^Tx^k\}$ is strictly monotone, decreasing, and thus converges, say to c^*.*

2. $X_ks^k \to 0$.

3. *$\{x^k\}$ converges, say to x^*.*

4. There is an $L \geq 1$ and a $\delta > 0$ such that for all $k \geq L$

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

Proof: The proof can be found in Tsuchiya [17], Monterio, Tsuchiya and Wang [12] and Saigal [15]. ∎

Given that the sequence $\{x^k\}$ converges, let the limit of this sequence be x^*. Then we can define:

$$N = \{j : x_j^* = 0\},\ B = \{j : x_j^* > 0\},\ q = |N|$$

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

The following is a corollary to Theorem 1(4).

Corollary 2 *There exists an $L \geq 1$ and $\delta_1 > 0$ and $\delta_2 > 0$ such that for all $k \geq L$*

$$\delta_1 \|x_N^k\| \leq c^T x^k - c^* \leq \delta_2 \|x_N^k\|.$$

Proof: Readily follows by applying Theorem 1(4). ∎

We now investigate the properties of the sequences $\{s^k\}$, $\{d^k\}$, $\{u^k\}$ and $\{v^k\}$.

Theorem 3 *The sequence $\{d^k\}$ converges to 0.*

Proof: Follows from Theorem 1. Can also be found in Vanderbei and Lagarias [22] and [15]. ∎

Theorem 4 *There is an $L \geq 1$ and a $\delta > 0$ such that for all $k \geq L$*

$$\|s_B^k\| \leq \delta (c^T x^k - c^*)^2.$$

Proof: The proof follows from Corollary 2 and results found in [12] and [15]. ∎

Theorem 5 *Consider the sequence $\{u^k\}$.*

1. It is bounded.

2. There is an $L \geq 1$ and $\delta > 0$ such that for all $k \geq L$

(a) $e^T u_N^k = 1 + \delta_k$ *where* $|\delta_k| \leq \delta(c^T x^k - c^*)^2$.

(b) $\phi(u^k) = \phi(u_N^k)$.

(c) $\|u^k\|^2 > \frac{1}{2q}$.

Proof: The proof can be found in [15]. ■

Theorem 6 *There is an* $L \geq 1$ *a* $\delta > 0$ *such that for every* $k \geq L$, $\|v_N^k\| \leq \delta$.

Proof: Readily follows from Theorem 1 part 4 and Corollary 2. ■

3 The Affine Scaling Step

We will study the affine scaling step with a view to relating it to a Newton step for computing an analytic center of a certain polyhedal set. Towards this end, define

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

as the set that represents all dual estimates complementary to x^*; i.e., $x_j^* > 0$ implies $s_j = 0$. Please note that in the definition of D, we have relaxed the non-negativity condition on s_N required for dual feasibility. The following can be readily established:

Theorem 7 *The sequences* $\{y^k\}$ *and* $\{s^k\}$ *are bounded, and thus* $D \neq \emptyset$. *Also, for every primal feasible pair* x *and* $\hat{x}$ *and every* $(\bar{y}, \bar{s}) \in D$

$$c^T x - c^T \hat{x} = \bar{s}_N^T (x_N - \hat{x}_N).$$

Proof: The proof readily follows by substituting $c = \bar{s} + A^T \bar{y}$ and noting that membership in D implies $\bar{s}_B = 0$; and $A(x - \hat{x}) = 0$. ■

As is now well known (see for example, Barnes [2]) the affine scaling direction

$$\frac{X_k^2 s^k}{\|X_k s^k\|},$$

used in Step 5 of the method, is generated by solving the following Ellipsoidal Approximating Problem:

$$\begin{aligned} \text{minimize} \quad & c^T x \\ Ax &= b \\ \|X_k^{-1}(x - x^k)\| &\leq 1. \end{aligned}$$

Substituting $p = x^k - x$ we obtain the equivalent problem:

$$\begin{array}{lrcl} \text{maximize} & c^T p & & \\ & Ap & = & 0 \\ & \|X_k^{-1} p\| & \leq & 1 \end{array}$$

Now $c^T p = \bar{s}^T p = \bar{s}_N^T p_N$ for $(\bar{y}, \bar{s}) \in D$. Thus, for some fixed $(\bar{y}, \bar{s})$, the above problem can be written as

$$\begin{array}{lrcrcl} \text{maximize} & \bar{s}_N^T p_N & & & & \\ & A_N p_N & + & A_B p_B & = & 0 \\ & p_N^T X_{k,N}^{-2} p_N & + & p_B^T X_{k,B}^{-2} p_B & \leq & 1 \end{array}$$

By noting that the solution of this problem is on the boundary of the ellipsoid (i.e., the second constraint is at equality), the K.K.T. conditions for this problem are

$$\bar{s}_N - A_N^T y - 2\theta X_{k,N}^{-2} p_N = 0 \qquad (1)$$

$$-A_B^T y - 2\theta X_{k,B}^{-2} p_B = 0 \qquad (2)$$

$$A_N p_N + A_B p_B = 0 \qquad (3)$$

$$\|X_k^{-1} p\| = 1 \qquad (4)$$

Now, let the sequences $\{u^k\}$ and $\{v^k\}$ be as defined in Section 2. Letting, for some given k, u and v represent u^k and v^k respectively, we define $\hat{A} = AV$ and $\hat{s} = V\bar{s}$, where V is the diagonal matrix whose jth diagonal entry is v_j for each $j = 1, \cdots, n$. Now consider the system:

$$\frac{u_N}{\|u\|^2} - \hat{A}_N^T \tilde{y} - \frac{\rho}{\|u\|^2} \hat{s}_N = 0 \qquad (5)$$

$$-A_B^T \tilde{y} = -\frac{s_B^k}{\|u\|^2} \qquad (6)$$

$$\hat{A}_N \frac{u_N}{\|u\|^2} + A_B p'_B = 0 \qquad (7)$$

$$\frac{\hat{s}_N^T u_N}{\|u\|^2} = 1 \qquad (8)$$

The following theorem establishes a connection between the conditions represented by the above systems.

Theorem 8 *Consider the systems represented by the equations (1) - (4) and (5) - (8).*

1. *(1) - (4) have a unique solution which generates a solution to (5) - (8).*

2. *The solution to (5) - (8) is unique upto a choice of p'_B; and, there is a value for p'_B for which the resulting solution also solves (1) - (4).*

3. *When A_B has full column rank, the two systems of equations are equivalent.*

Proof: Since the Equations (1) - (4) represent the solution to a strictly convex problem, they have a unique solution. Using this solution, define the vectors

$$\begin{aligned} \tilde{y} &= \frac{-y(c^T x^k - c^*)}{2\theta\|u\|} \\ \rho &= \frac{(c^T x^k - c^*)\|u\|}{2\theta} \\ p'_B &= \frac{p_B}{(c^T x^k - c^*)\|u\|}. \end{aligned}$$

It is now readily confirmed, by simple algebra, that u, $\tilde{y}$ ρ and p'_B solve the System (5) - (8). Thus we have proved Part 1.

From Part 1, it follows that (5) - (8) have a solution. Considering $q_N = \frac{u_N}{\|u\|^2}$, $\tilde{\rho} = \frac{\rho}{\|u\|^2}$, $\tilde{y}$ and p'_B as variables, this system is linear in these variables. If A_B has full column rank, the solution to (5) - (8) is unique, and Part 3 follows. Otherwise, since (5) - (8) can have a solution only if s_B lies in the row space $\mathcal{R}(A_B^T)$ of A_B, the third condition must have redundant constraints readily identified by choosing any full column rank submatrix of A_B.

To see Part 2, let $A_B = (A_C, A_D)$ where A_C has full column rank and spans the range (or column space) $\mathcal{R}(A_B)$ of A_B. Thus, for some unique matrix Λ, $A_D = A_C\Lambda$. Replacing equations (6) and (7) by

$$-A_C^T\tilde{y} = \frac{s_B^k}{\|u\|^2} \tag{9}$$

and

$$\hat{A}_N\frac{u_N}{\|u\|^2} + A_C p'_C = 0 \tag{10}$$

respectively, we obtain a new system that has a unique solution. By setting $p'_B = (p'_C, p'_D)$, and letting $p'_D = 0$, the solution to Equation (10) generates a solution to (7). Now, let $(q_N, \tilde{y}, \tilde{\rho}, p'_B)$ be any solution to (5) - (8). This then generates the unique solution $(q_N, \tilde{y}, \tilde{\rho}, p'_C - \Lambda p'_D)$ to (5) and (8) - (10). Since only the vector p'_B is modified in any solution to (5) - (8). Part 2 is established with the required $p'_B = \frac{p_B}{(c^T x^k - c^*)\|u\|^2}$. ■

4 Newton's Method and an Analytic Center Problem

In this section we consider the application of Newton's method for finding the analytic center of a certain polytope. By interpreting the affine scaling step with a step-size of $\frac{1}{2}$ on this polytope, a close connection can be established between the two steps. This connection plays a central role in the accelerated implementations of the affine scaling method. We now introduce this polytope.

Consider the sequence $\{x^k\}$ generated by the affine scaling method. From Theorem 1, it converges to, say x^*, and as in Section 2, let $N = \{j : x_j^* = 0\}$ and $B = \{j : x_j^* > 0\}$. Also, define the sequence

$$v^k = \frac{x^k - x^*}{c^T x^k - c^*}$$

and note that for every k, $Av^k = A_N v_N^k + A_B v_B^k = 0$. Also, if $(\bar{y}, \bar{s}) \in D$, then from Theorem 7:

$$\bar{s}_N^T x_N^k = c^T x^k - c^*$$

and thus $\bar{s}_N^T v_N^k = 1$. Now define the polytope:

$$P = \{v : A_N v_N + A_B v_B = 0, \bar{s}_N^T v_N = 1, v_N \geq 0\}$$

and its projection

$$P_N = \{v_N : v \in P\}.$$

Its analytic center is the solution v^* to the following problem:

$$\begin{array}{lrcl} \text{maximize} & \sum_{j \in N} \log(v_j) & & \\ & A_N v_N + A_B v_B & = & 0 \\ & \bar{s}_N^T v_N & = & 1. \\ & v_N & > & 0 \end{array}$$

The K.K.T. conditions for this problem are:

$$V_N^{-1} e - A_N^T y - \bar{s}_N \theta = 0 \tag{11}$$

$$-A_B^T y = 0 \tag{12}$$

$$A_N v_N + A_B v_B = 0 \tag{13}$$

$$\bar{s}_N^T v_N = 1 \tag{14}$$

$$v_N > 0$$

The solution to the K.K.T. conditions is unique, if it exists. The analytic center problem has a feasible solution, but the set may not be bounded, and thus the center may not exist. Indeed, it can be shown that, for a given N, the center exists if and only if x^* is an optimal solution of the primal linear program. Since we have not shown this fact, we cannot claim the K.K.T. conditions have a solution. However, we will, in this section, assume that the center exists. Thus the results established as a consequence of this assumption can only be used when this existence has been shown.

We now apply Newton's method to the system of equations (11) - (14) to determine its zero. The Newton direction $(\Delta v, \Delta y, \Delta\theta)$ at (v, y, θ) is obtained by solving the following system:

$$\begin{aligned} -V_N^{-2}\Delta v_N - A_N^T\Delta y - \bar{s}_N\Delta\theta &= -V_N^{-1}e + A_N^T y + \bar{s}_N\theta && (15)\\ -A_B^T\Delta y &= A_B^T y && (16)\\ A_N\Delta v_N + A_B\Delta v_B &= 0 && (17)\\ \bar{s}_N^T\Delta v_N &= 0 && (18)\end{aligned}$$

Defining $\hat{y} = y + \Delta y$, $\hat{\theta} = \theta + \Delta\theta$ and substituting $w_N = V_N^{-1}\Delta v_N$, $\hat{A}_N = A_N V_N$ and $\hat{s}_N = V_N\bar{s}_N$ we can rewrite the system (15) - (18) as:

$$\begin{aligned} w_N + \hat{A}_N^T\hat{y} + \hat{s}_N\hat{\theta} &= e && (19)\\ A_B^T\hat{y} &= 0 && (20)\\ \hat{A}_N w_N + A_B\Delta v_B &= 0 && (21)\\ \hat{s}_N^T w_N &= 0 && (22)\end{aligned}$$

We are now ready to prove an important result:

Theorem 9 *Let w_N solve the system (19) - (22). Then*

1. $e^T w_N = w_N^T w_N$.

2. $\|w_N\| \le \sqrt{q}$, *where q is the number of elements in N.*

Proof: (1) is obtained by multiplying equation (19) by w_N^T and substituting $w_N^T\hat{A}_N^T\hat{y} = 0$ obtained from equations (20) and (21), and equation (22). (2) can be obtained by considering the problem

$$\begin{aligned} \text{maximize} \quad & z^T z \\ z^T z - e^T z &= 0 \end{aligned}$$

and noting that $z = e$ solves this problem. ∎

The system (19) - (22) is linear in the variables w_N, Δv_B, $\hat{y}$ and $\hat{\theta}$, with the underlying matrix:

$$M(v) = \begin{bmatrix} I & \hat{A}_N^T & \hat{s}_N & 0 \\ 0 & A_B^T & 0 & 0 \\ \hat{A}_N & 0 & 0 & A_B \\ \hat{s}_N^T & 0 & 0 & 0 \end{bmatrix}.$$

It can be readily verified that, when A_B has full column rank, this matrix is non-singular for every $v_N > 0$. In the other case, just as was done in Theorem 8, it can be shown that the solution to the system (19) - (22) is unique upto a choice of Δv_N.

Now, consider a sequence $\{v_N^k\}$ in P_N that converges to v_N^*, the analytic center of P_N. Also, let

$$\beta = \|M(v_N^*)^{-1}\|.$$

We state, without proof, the following standard result on the convergence and convergence rate of Newton's method.

Lemma 10 *There is an $L \geq 1$ and $\rho_1 > 0$, $\rho_2 > 0$ such that for all $k \geq L$*

1. $\|M(v_N^k)^{-1}\| \leq 2\beta$.
2. $\frac{\|\Delta v_N^k\|}{\|v_N^k - v_N^*\|} = 1 + \delta_k$ *where* $|\delta_k| \leq \rho_1 \|v_N^k - v_N^*\|$.
3. $\|v_N^k + \Delta v_N^k - v_N^*\| \leq \rho_2 \|v_N^k - v_N^*\|^2$.

Proof: Can be found in standard texts in numerical mathematics. ∎

Now consider the affine scaling step as determined by the system (5) - (8). We note that if we consider $\frac{u_N}{\|u\|^2}$, p_B', $\tilde{y}$ and $\frac{\rho}{\|u\|^2}$ as variables, this system is also linear with the underlying matrix $M(v_N)$. Thus we can prove the following theorem:

Theorem 11 *There exists an $L \geq 1$ and $\rho > 0$ such that for all $k \geq L$,*

$$e - w_N^k - \frac{u_N^k}{\|u^k\|^2} = \Delta^k$$

with $\|\Delta^k\| \leq \rho\|c^T x^k - c^*\|^2$.

Proof: Consider the systems (5) - (8) and (19) - (22). By selecting possibly a submatrix of A_B, we assume that A_B has full column rank in these systems. In the later system, make the change of variable $w'_N = e - w_N$. Then $\hat{s}_N^T w'_N = 1$. Now, by defining $v'_B = v_B - \Delta v_B$ and

$$\begin{aligned} V_N^{-1} a_1 &= \frac{u_N}{\|u\|^2} - w'_N \\ a_2 &= \tilde{y} - \hat{y} \\ a_3 &= \frac{\rho}{\|u\|^2} - \hat{\theta} \\ a_4 &= p'_B - v'_B \end{aligned}$$

generate the system $M(v_N)a = (0, -\frac{s_B}{\|u\|^2}, 0, 0)^T$. This system is readily seen, with $s'_B = \frac{-s_B}{\|u\|^2}$, as the K. K. T. conditions of the following optimization problem:

$$\begin{array}{lllll} \text{minimize} & & s_B'^T a_4 & & \\ & A_N a_1 & + \; A_B a_4 & = & 0 \\ & s_N^T a_1 & & = & 0 \\ & \|V_N^{-1} a_1\|^2 & & \leq & r^2. \end{array}$$

for some $r > 0$. Consider the first constraint. Since A_B has full column rank, by using the formula

$$a_4 = -(A_B^T A_B)^{-1} A_B^T A_N a_1$$

we can eliminate a_4 to generate the following equivalent problem:

$$\begin{array}{llll} \text{maximize} & s_B^{*T} a_1 & & \\ & A_N^* a_1 & = & 0 \\ & \bar{s}_N^T a_1 & = & 0 \\ & \|V_N^{-1} a_1\|^2 & \leq & 1 \end{array}$$

where $s_B^* = A_N^T A_B (A_B^T A_B)^{-1} s'_B$ and $A_N^* = A_N - A_B (A_B^T A_B)^{-1} A_B^T A_N$. This is an affine scaling problem. It is well known that the K.K.T. multipliers (a_2, a_3) of this problem are bounded by a function of the form $q(A_N^*, \bar{s}_N).\|s_B^*\|$ independent of the diagonal matrix $V_N > 0$ and r. (See for example, Theorem 4, Saigal [15]). Thus, for some $\bar{q}(A, \bar{s}_N)$,

$$\|(a_2, a_3)\| \leq \frac{\bar{q}(A, \bar{s}_N)\|s_B\|}{\|u\|^2}.$$

Thus $\|V_N^{-1}a_1\| \leq \|V_N(A_N^T a_2 + \bar{s}_N a_3)\| \leq \frac{\beta'\|s_B\|\|v_N\|}{\|u\|^2}$. By using Theorems 4 and 5 part 2(c), and that the bound is independent of a_4. ∎

5 Affine Scaling and Newton Directions in P

In this section, we show the connection between the Newton direction $\Delta v_N^k = V_{k,N} w_N^k$ at v_N^k determined in Theorem 11, and the affine scaling direction as interpreted in P. Consider the sequence $\{v^k\}$ in P generated by the affine scaling algorithm. Then we can show the following result:

Theorem 12 *The affine scaling direction at v_N^k in P is*

$$v_N^{k+1} - v_N^k = \frac{\alpha_k \delta(u^k)}{1 - \alpha_k \delta(u^k)}(v_N^k - V_{k,N}\frac{u_N^k}{\|u^k\|^2})$$

where $\delta(u^k) = \frac{\|u^k\|^2}{\phi(u^k)}$. Also, the Newton direction Δv_N^k at v_N^k in P is:

$$\Delta v_N^k = v_N^k - V_{k,N}\frac{u_N^k}{\|u^k\|^2} + V_{k,N}\Delta^k$$

where Δ^k is as in Theorem 11.

Proof: Since $w_N^k = V_{k,N}^{-1}\Delta v_N^k$ the formula for the Newton direction follows from Theorem 11. To see the affine direction note that:

$$\begin{aligned} v_N^{k+1} - v_N^k &= \frac{x_N^{k+1}}{c^T x_N^{k+1} - c^*} - \frac{x_N^k}{c^T x_N^k - c^*} \\ &= \frac{v_N^k - \frac{\alpha_k V_{k,N} u_N^k}{\phi(u^k)}}{1 - \frac{\alpha_k \|u_N^k\|^2}{\phi(u^k)}} - v_N^k \\ &= \frac{\frac{\alpha_k \|u_N^k\|^2}{\phi(u^k)}}{1 - \frac{\alpha_k \|u_N^k\|^2}{\phi(u^k)}}(v_N^k - V_{N,k}\frac{u_N^k}{\|u^k\|^2}) \end{aligned} \quad (23)$$

and we are done. ∎

Using Theorem 11, if α_k is chosen such that the scaler term in the formula for the affine scaling direction is 1, the directions taken by Newton's method for computing the analytic center of P will be very close to the affine scaling direction. It is this observation that allows the interpretation of this step as a corrector step. We now present a relationship between two analytic center problems, and the version that exploits this property of the affine scaling step.

6 Relation between two analytic centers

Consider the analytic center problem for $D \cap \{s : s_N \geq 0\}$

$$\begin{array}{lrcl} \text{maximize} & \sum_{j \in N} \log(s_j) & & \\ & A_N^T y + s_N & = & c_N \\ & A_B^T y & = & c_B \\ & s_N & > & 0 \end{array}$$

and its K.K.T. conditions:

$$S_N^{-1} e - v_N = 0 \tag{24}$$

$$A_N v_N + A_B v_B = 0 \tag{25}$$

$$A_N^T y + s_N = c_N \tag{26}$$

$$A_B^T y = c_B \tag{27}$$

$$s_N > 0$$

The following relates the two analytic centers defined by systems (11) - (14), and (24) - (27).

Theorem 13 *(y^*, s^*) is an analytic center of $D \cap \{s : s_N \geq 0\}$ if and only if v^* is an analytic center of P_N and*

$$q v_N^* = S_N^{*^{-1}} e$$

Proof: It is readily confirmed that if v_N^* is the analytic center of P_N, then for some $\theta = \theta^* > 0$ and $y = u^*$, equations (11) -(14) are satisfied. Thus

$$s_N = \frac{1}{\theta^*} A_N^T u^* + \bar{s}_N, v = \frac{1}{\theta^*} v^*, y = \bar{y} - \frac{1}{\theta^*} u^*$$

satisfy the equations (24) - (27). Also, if (y^*, s^*) is the analytic center of $D \cap \{s : s_N \geq 0\}$, then for some v^* equations (24) - (27) are satisfied. Thus

$$v = \frac{1}{q} v^*, y = -q(y^* - \bar{y}), \theta = q$$

solve equations (11) - (14). We are now done by equation (24). ■

7 Accelerated Primal Affine Scaling Method

In this section we present a step-size selection strategy that asymptotically behaves as a predictor-corrector strategy for computing the analytic center of P. This method follows:

Step 0 Let x^0 be an interior point solution, $0 < \alpha < 1$ (normally between 0.95 and 0.99) 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. Otherwise, define

$$d^k = X_ks^k$$

Step 3 Min-Ratio test:

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

Also, if $\theta_k = 1$ set $\alpha_k = 1$, and go to Step 7.

Step 4 Tentative Non-Basic set:

$$N_k = \{j : x_j^k \leq \phi(d^k)\}$$

$$\gamma_k = |e^Td_{N_k}^k|$$

Step 5 Estimate of Newton Step:

$$\begin{aligned}h_{N_k}^k &= \frac{x_{N_k}^k}{\gamma_k} - X_{k,N_k}\frac{d_{N_k}^k}{\|d^k\|^2} \\ \epsilon_k &= \|h_{N_k}^k\|\end{aligned}$$

Step 6 Step-size calculation:

1. Normal Step: If $\gamma_k \geq 1$ then set

$$\alpha_k = \alpha$$

2. Predictor Step: If $\gamma_k < 1$ and $\epsilon_k^{0.75} \leq \gamma_k$ then set:

$$\alpha_k = 1 - \max\ \{\gamma_k, \epsilon_k^{0.50}\}$$

3. Corrector Step: Otherwise, set

$$\alpha_k = \min\{\frac{\gamma_k}{2\theta_k \|d^k\|}, \frac{2}{3}\}$$

Step 7 Next Interior Point:

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

Step 8 Iterative Step: If $x_j^{k+1} = 0$ for some j, then STOP. x^{k+1} is an optimal solution to the primal, y^k the optimum solution to the dual. Otherwise set $k = k + 1$ and go to step 1.

We will show that the above implementation is three-step quadratically convergent. Before we prove this, some comments are in order.

As a consequence of Theorem 3, $\phi(d^k)$ converges to 0. Thus, asymptotically, N_k will become some constant N. Also, at Step 5

$$\begin{aligned} h_N^k &= \frac{x_N^k}{e^T d_N^k} - X_{k,N} \frac{d_N^k}{\|d^k\|^2} \\ &= \frac{v_N^k}{e^T u_N^k} - V_{k,N} \frac{u_N^k}{\|u^k\|^2} \end{aligned} \tag{28}$$

and using the Theorems 5 and 11, is readily seen as a very good estimate of Δv_N^k.

We wish to apply the predictor step when $\|h_N^k\| = O(c^T x^k - c^*)^2$. Since $e^T d_N^k$ goes to zero; asymptotically the acceptance condition for the predictor step will be satisfied. Since the resulting constant in this relation is unknown, a predictor step may be performed with a lower order of $c^T x^k - c^*$. This will result in a lower convergence rate than two. We will show that this error can be corrected by taking three corrector steps following the first predictor step selected by the rule. After this, two corrector steps after each predictor step can be taken to achieve two-step quadratic convergence. The corrector step chooses the step size

$$\frac{\gamma_k}{2\theta_k\|d^k\|} = \frac{\phi(u_N^k)e^T u_N^k}{2\|u^k\|^2} \tag{29}$$

which approaches $\frac{1}{2}$. We are now ready to establish the preliminary results before proving the two step quadratic convergence of the sequence.

8 Preliminary Results

In this section we will establish the preliminary results for the proof of the convergence of the sequence generated by the above implementation.

Proposition 14 *There exists an $L \geq 1$ such that for every $k \geq L$, $N_k = N$.*

Proof: As $\phi(d^k) \to 0$, there exists an $L \geq 1$ such that for all $k \geq L$, $x_j^k \geq \phi(d^k)$ for each $j \in B$ and we are done. ■

Let $\hat{L}$ be generated by Proposition 14. We henceforth restrict our attention to the sequence $\{x^k\}_{k=\hat{L}}^{\infty}$.

Proposition 15 *There is an $L \geq \hat{L}$ and an $\alpha > 0$ such that for all $k \geq L$,*

$$\|h_N^k - \Delta v_N^k\| \leq \alpha(c^T x^k - c^*)^2.$$

Proof: From equation (28) and Theorems 12 and 5 we obtain

$$h_N^k - \Delta v_N^k = \frac{\delta_k v_N^k}{1+\delta_k} - V_{k,N}\Delta^k$$

and the result follows from the above theorems and Theorem 6. ■

Proposition 16 *There is an $L \geq \hat{L}$ and $\beta > 0$ such that for every $k \geq L$*

1. $\frac{c^T x^{k+1} - c^*}{c^T x^k - c^*} = 1 - \alpha_k \delta(u_N^k)$ *where* $\delta(u_N^k) = \frac{\|u^k\|^2}{\phi(u_N^k)}$.

2. $1 - \|w_N^k\| - |\delta_k| \leq \delta(u_N^k) \leq 1$ *with* $|\delta_k| \leq \beta(c^T x^k - c^*)^2$.

Proof: From Steps 3 and 7 we obtain

$$c^T x^{k+1} = c^T x^k - \alpha_k \frac{c^T X_k^2 s^k}{\phi(X_k s^k)}$$

and thus

$$\frac{c^T x^{k+1} - c^*}{c^T x^k - c^*} = 1 - \alpha_k \frac{\|u^k\|^2}{\phi(u^k)}. \tag{30}$$

Because of Theorem 5, we have (1). Since (1) is true for every $\alpha_k \leq 1$, by setting $\alpha_k = 1$ we obtain the upper bound of (2). Now, from Theorem 11,

$$\begin{aligned} \frac{\phi(u_N^k)}{\|u^k\|^2} &= \phi(\frac{u_N^k}{\|u^k\|^2}) \\ &= \phi(e + w_N^k - \Delta^k) \\ &\leq 1 + \|w_N^k\| + \delta_k \end{aligned} \tag{31}$$

where $\|\Delta^k\| = \delta_k$; and we have our result from Theorem 11. ■

Proposition 17 *Let $w_N^k \to 0$ on some subsequence K. Then, on K*

$$\rho_k = \frac{e^T u_N^k \phi(u_N^k)}{2\|u^k\|^2} \to \frac{1}{2}$$

Proof: From Theorem 5 and Proposition 16, $2\rho_k \geq e^T u_N^k = 1 - \delta_k$ From equation (31)

$$\rho_k \leq \frac{1 + \delta_k}{2}(1 + \|w_N^k\| + \|\Delta^k\|)$$

and we get the result as δ_k and $\|\Delta^k\|$ go to zero. ■

Proposition 18 *There is an $L \geq \hat{L}$ such that for all $k \geq L$*

$$0.50(c^T x^k - c^*) \leq e^T d_N^k \leq 1.5(c^T x^k - c^*)$$

Proof: Note that $e^T d_N^k = (c^T x^k - c^*) e^T u_N^k$, and so the result follows from Theorem 5.

■

We are now ready to prove the optimality of the limit point x^* of the primal sequence generated by the accelerated method. The convergence of the primal sequence follows from the Theorem 1.

Theorem 19 *x^* lies in the relative interior of the optimal primal face.*

Proof Let L be as in Proposition 14. In case the predictor step is only taken a finite number of times, there is an $\hat{L} \geq L$ such that for all $k \geq \hat{L}$, $\alpha_k \leq \frac{2}{3}$. Thus the result follows from Theorem 1.1 of Tsuchiya and Maramatsu [19] (see also Theorem 16 of Saigal [15]).

Now, assume the predictor step is taken an infinite number of times, and let K be the subsequence of such iterates. Let $k \in K$. Then, from Steps 4, 5 and 6 Part 2,

$$\|h_N^k\| \leq (e^T X_{k,N} s_N^k)^{\frac{4}{3}} \leq (e^T X_k s^k)^{\frac{4}{3}}$$

and, from Theorem 1, Part 2

$$h_N^k \longrightarrow 0 \text{ for } k \in K.$$

Also, from Proposition 15, $\Delta v_N^k \to 0$, for $k \in K$ as

$$\|\Delta v_N^k\| \leq \|h_N^k\| + \|\Delta v_N^k - h_N^k\|.$$

But, $v_N^* \in \mathcal{P}_N$ is the only vector for which $\Delta v_N^* = 0$. Thus

$$v_N^k \longrightarrow v_N^* \text{ for } k \in K.$$

As $\Delta v_N^k \to 0$, $w_N^k \to 0$ for $k \in K$. Thus from Theorem 11

$$\frac{u_N^k}{\|u^k\|^2} \longrightarrow e \text{ for } k \in K,$$

and, from Theorem 5 Part 2(a),

$$\|u^k\|^2 \longrightarrow \frac{1}{q} \text{ for } k \in K.$$

Thus

$$u_N^k \longrightarrow \frac{1}{q} e \text{ for } k \in K.$$

The sequences $\{y^k\}$ and $\{s^k\}$ are bounded and so, on some subsequence $K' \subset K$, they converge, say to (y^*, s^*). So

$$V_{*,N} s_N^* = \frac{1}{q}.$$

Also, as v_N^* is the analytic center of $\mathcal{P}_N$, from Theorem 13, (y^*, s^*) is an analytic center of $\mathcal{D} \cap \{s : s_N \geq 0\}$, and we are done. ∎

9 On Predictor and Corrector Steps

In this section we will establish the asymptotic rate of convergence of the sequence $\{c^T x^k - c^*\}$ generated by the accelerated affine scaling method. We will show that this sequence converges three-step quadratically to 0. Because of Theorem 19, the analytic center v_N^* of $\mathcal{P}_N$ exists, and thus Newton's method, when initiated close to the center, will converge to it. Thus, there is a sufficiently large $\hat{L} \geq 1$ for which the following properties hold:

For all $k \geq \hat{L}$,

1. $v_i^* \geq (c^T x^k - c^*)^{0.25}$ for all $i \in N$.

2. From Lemma 10 Part 2, related to Newton's method, there is a $\theta > 0$ such that $(c^T x^k - c^*)^{0.25} \leq \theta$; and, for all $\|v_N - v_N^*\| < \theta$

$$\frac{1}{2}\|v_N - v^*\| \leq \|\Delta v_N\| \leq \frac{3}{2}\|v_N - v_N^*\|.$$

3. The conditions of Lemma 10 related to Newton's method; and, Proposition 14, are satisfied.

Such an $\hat{L}$ exists as $v_i^* > 0$ for each i, $\theta > 0$ and $c^T x^k - c^* \to 0$. We can then prove the following straight forward lemmas:

Lemma 20 *There is an $L \geq \hat{L}$ such that for all $k \geq L$*

$$0.50(c^T x^k - c^*) \leq e^T d_N^k \leq 1.5(c^T x^k - x^*).$$

Proof From Theorem 5 Part 2(a), there is a $L' \geq \hat{L}$ and a $\theta > 0$ such that for all $k \geq L'$

$$e^T d_N^k = (1 + \delta_k)(c^T x^k - c^*)$$

where $|\delta_k| \leq \theta(c^T x^k - c^*)^2$. Thus, there is a $L \geq L'$ such that $|\delta_k| \leq 0.50$ and we are done.■

Lemma 21 *Let $\|v_N^k - v_N^*\| \leq \beta(c^T x^k - c^*)^p$ for some $\beta > 0$ and $p > 0.25$. There is an $L \geq \hat{L}$ and a $\theta > 0$ such that for all $k \geq L$*

$$\|w_N^k\| \leq \theta(c^T x^k - c^*)^{p-0.25}.$$

Proof Consider

$$v_i^k \geq v_i^* - |v_i^k - v_i^*| \geq (c^T x^k - c^*)^{0.25} - \beta(c^T x^k - c^*)^p$$

and thus there is a $L \geq \hat{L}$ such that for all $k \geq L$

$$v_i^k \geq 0.5(c^T x^k - c^*)^{0.25}.$$

Thus, $\|V_{N,k}^{-1}\| \leq \frac{2}{(c^T x^k - c^*)^{0.25}}$, and since $w_N^k = V_{N,k}^{-1} \Delta v_N^k$, we obtain our result from Lemma 10. ■

We are now ready to investigate the predictor step and the corrector step of the accelerated method. The next theorem investigates the predictor step, Step 6, part 2 of the accelerated method.

Theorem 22 *Assume that for some* $k \geq \hat{L}$ *and* $\beta > 0$, $\|v_N^k - v_N^*\| \leq \beta(c^T x^k - c^*)^{2p}$ *for some* $0.50 < p \leq 1$. *Then*

1. *There is a* $\theta > 0$ *such that*

$$0.50(c^T x^k - c^*)^2 \leq c^T x^{k+1} - c^* \leq \theta(c^T x^k - c^*)^{2p}.$$

2. *There is a* $\hat{\theta} > 0$ *such that*

$$\|v_N^{k+1} - v_N^*\| \leq \hat{\theta}(c^T x^{k+1} - c^*)^{\frac{p}{2}}.$$

Proof To see (1), note that from Step 6, part 2; Lemmas 20, 21 and Proposition 16, we get

$$\begin{aligned}
0.50(c^T x^k - c^*) &\leq e^T d_N^k \\
&\leq 1 - \alpha_k \\
&\leq 1 - \alpha_k \frac{\|u^k\|^2}{\phi(u_N^k)} \\
&= \frac{c^T x^{k+1} - c^*}{c^T x^k - c^*} \\
&\leq 1 - (1 - \max\{\epsilon_k^{0.50}, \gamma_k\})(1 - 2\|w_N^k\|) \\
&\leq \max\{\epsilon_k^{0.50}, \gamma_k\} + 2\|w_N^k\| \\
&\leq \theta(c^T x^k - c^*)^p.
\end{aligned}$$

To see (2), from Theorem 12 Part 3

$$v_N^{k+1} - v_N^* = v_N^k + \Delta v_N^k - v_N^* + \frac{2\alpha_k \delta(u^k) - 1}{1 - \alpha_k \delta(u^k)}(v_N^k - V_{N,k}\frac{u_N^k}{\|u^k\|^2}) - V_{N,k}\Delta^k.$$

From Propositions 15, 16 and Theorem 8, after some simple algebra, we obtain

$$\|v_N^{k+1} - v_N^*\| \leq \|v_N^k + \Delta v_N^k - v_N^*\| + \frac{2\alpha_k - 1}{1-\alpha_k}\|h_N^k\| + \beta(c^T x^k - c^*)^2.$$

Choosing $k \geq \hat{L}$ and substituting the formula for α_k from Step 6, Part 2; Part 1 of this theorem, and the Lemma 10

$$\begin{aligned} \|v_N^{k+1} - v_N^*\| &\leq \delta_1 \|v_N^k - v_N^*\|^2 + \theta(c^T x^k - c^*)^p \\ &\leq \hat{\beta}(c^T x^k - c^*)^p \\ &\leq \hat{\theta}(c^T x^{k+1} - c^*)^{\frac{p}{2}} \end{aligned}$$

where $\delta_1 > 0$, $\hat{\theta} > 0$, and $\hat{\beta} > 0$ are appropriate constants. We have used

$$\begin{aligned} \frac{2\alpha_k - 1}{1-\alpha_k}\epsilon_k &\leq \frac{1 - 2\max\{\epsilon_k^{0.50}, \gamma_k\}}{\max\{\epsilon_k^{0.50}, \gamma_k\}}\epsilon_k \\ &\leq \epsilon_k^{0.50}. \end{aligned}$$

■

We now investigate the corrector step.

Theorem 23 *There is an $L \geq \hat{L}$ and $\alpha > 0$ such that for some $k \geq L$, assume $\|v_N^k - v_N^*\| \leq \alpha(c^T x^k - c^*)^p$ where $0.25 < p \leq 0.50$.*

1. *Let $p = 0.50$. Then two iterations of the corrector step, Step 6, Part 3, can be taken, and for some $\theta > 0$,*

$$\|v_N^{k+2} - v_N^*\| \leq \theta(c^T x^{k+2} - c^*)^2.$$

2. *Let $0.25 < p < 0.50$. Then three iterations of the corrector step can be taken, and for some $\theta > 0$,*

$$\|v_N^{k+3} - v_N^*\| \leq \theta(c^T x^{k+3} - c^*)^2.$$

Proof From Lemma 21, we note that for all $k \geq \hat{L}$

$$\|w_N^k\| \leq \theta(c^T x^k - c^*)^{p-0.25}.$$

Let $L \geq \hat{L}$ be such that for all $k \geq L$, $\theta(c^T x^k - c^*)^{p-0.25} \leq 0.25$. Now, consider $k \geq L$. Then from Proposition 16

$$\begin{aligned}
\delta_k = \frac{\gamma_k}{2\theta_k \|d^k\|} &= \frac{e^T u_N^k \phi(u_N^k)}{\|u^k\|} \\
&\leq \frac{1}{2}(1+\delta_k)(1+\|w_N^k\|+\gamma_k) \\
&< \frac{2}{3}
\end{aligned}$$

where δ_k and γ_k are positive and of the order $O(c^T x^k - c^*)^2$.

Thus, during the corrector step, $\alpha_k = \delta_k$, and thus, from Proposition 16 Part 1 we get

$$\frac{c^T x^{k+1} - c^*}{c^T x^k - c^*} = 1 - \frac{e^T u_N^k}{2}$$

and thus from Theorem 5, Part 2(a), for all $k \geq L$

$$0.25(c^T x^k - c^*) \leq c^T x^{k+1} - c^* \leq 0.75(c^T x^k - c^*). \tag{32}$$

For the first corrector step, after substituting α_k Theorem 12, we get

$$v_N^{k+1} = v_N^k + \frac{e^T u_N^k}{2 - e^T u_N^k}(v_N^k - V_{k,N}\frac{u_N^k}{\|u^k\|^2}).$$

Thus, by simple algebra and substitutions we get:

$$\begin{aligned}
\|v_N^{k+1} - v_N^*\| &\leq \|v_N^k + \Delta v_N^k - v_N^*\| + \frac{2|\delta_k|}{1-|\delta_k|}\|V_{k,N} w_N^k\| + \frac{1+|\delta_k|}{1-|\delta_k|}\|V_{k,N}\Delta^k\| \\
&\leq \rho\|v_N^k - v_N^*\|^2 + \beta(c^T x^k - c^*)^2 \\
&\leq \rho\alpha^2(c^T x^k - c^*)^{2p} + \beta(c^T x^k - c^*)^2 \\
&\leq \rho_1(c^T x^k - c^*)^{2p} \\
&\leq (4)^{2p}\rho(c^T x^{k+1} - c^*)^{2p}.
\end{aligned} \tag{33}$$

Applying the same argument as above to one more corrector step, we obtain

$$\|v_N^{k+2} - v_N^*\| \leq \alpha_2(c^T x^{k+2} - c^*)^{4p} \tag{34}$$

Now, if $p = 0.50$, then $4p = 2$ and we have part (1). Otherwise after one more corrector step, we obtain

$$\|v_N^{k+3} - v_N^*\| \leq \alpha_3(c^T x^{k+3} - c^*)^{8p} + \beta(c^T x^{k+3} - c^*)^2. \tag{35}$$

Since $8p > 2$ for $0.25 < p < 0.50$, we obtain part (2) and we are done. ■

10 Rate of Convergence of the Accelerated Method

We are now ready to prove the main theorem.

Theorem 24 *Let the sequences $\{x^k\}$, $\{y^k\}$ and $\{s^k\}$ be generated by the accelerated affine scaling method. Then*

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

where x^* lies in the relative interior of the optimal face of the primal, (y^*, s^*) is the analytic center of the optimal face of the dual. In addition, asymptotically, the sequence $\{c^T x^k - c^T x^*\}$ converges three-step quadratically to zero; i.e. , there is an $L \geq 1$, $\theta > 0$ and $K = \{L, L+3, L+6, \cdots\}$ such that the subsequence $\{c^T x^k - c^T x^*\}_{k \in K}$ converges quadratically to zero, or

$$c^T x^{(k+1)'} - c^T x^* \leq \theta (c^T x^{k'} - c^T x^*)^2 \textit{ for } k' \in K.$$

Proof Using equation (32) and Lemma 20 we note that, during corrector steps, asymptotically the sequence $\{\gamma_k\}$ decreases linearly in $c^T x^k - c^*$. Also, using equations (34), (35), Lemma 10 and Proposition 15 asymptotically, if no predictor step is taken, the sequence $\{\epsilon_k\}$ decreases at least as fast as $O(c^T x^k - c^*)^2$. Thus, for some large k, we must have

$$\epsilon_k^{0.75} \leq \gamma_k \tag{36}$$

and a predictor step must be taken. The first time the condition of equation (36) occurs, if $\epsilon_k > \gamma_k^2$ then we enter the predictor step with $1 > p > 0.50$; and, after the predictor step we enter the corrector step with $p < 0.50$ and thus three steps may be required to guarantee, by Lemma 10 and Proposition 15 that we enter the next predictor step with $p = 1$. Then every subsequent corrector step will be entered with $p = 0.50$, and will thus require at most two corrector iterations to satisfy the conditions for the predictor step, i.e, enter the predictor step with $p = 1$. Thus, if during the first corrector step, three corrector iterations are taken, from Theorem 23, we guarantee that the predictor step is initiated with $p = 1$, and the three step quadratic convergence of $\{c^T x^k - c^*\}$ follows.

It is readily confirmed from Theorems 22 and 23 that

$$\|v_N^k - v_N^*\| \longrightarrow 0$$

and thus $v_N^k \longrightarrow v_N^*$. But this can only happen, by Theorem 13, if (y^k, s^k) converges to the analytic center of $D \cap \{s : s_N \geq 0\}$ and we are done. ∎

11 Relative efficiency of Three Step Method

Ostrowski [14] section 6.11, introduced the following measure to compare algorithms achieving different rates of convergence. As is evident, greater rates of convergence can be achieved by increasing the work. The simplest way to get order four convergence of a sequence generated by quadratically convergent Newton's method is to drop each odd element of the sequence. Ostrowski's measure is invariant under such manipulations. For a method which requires w units of work per iteration and achieves a convergence rate of p, the measure of efficiency is defined as

$$\frac{\log(p)}{w}.$$

In the table below, we summarize the efficiency of three methods.

Algorithm	Rate	Work/Iter	Efficiency	Factor
Tsuchiya and Monterio	1.3	$2w$	$\frac{0.13118}{w}$	0.5678
Saigal	2	$3w$	$\frac{0.23105}{w}$	1.00
Two step Quadratic	2	$2w$	$\frac{0.34657}{w}$	1.50

The two step superlinear algorithm of Tsuchiya and Monterio [20] is about 44% slower than the algorithm presented here, which is 50% slower that a two step quadratic algorithm.

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] C. C. Gonzaga, " Convergence of the large step primal affine-scaling algorithm for primal nondegenerate linear programs," Technical Report ES-230/90, Dept. of Systems Engineering and Computer Science, COPPE Federal University of Rio de Janeiro, 21941 Rio de Janeiro, RJ, Brazil, September 1990.

[9] 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.

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

[11] S. Mehrotra, " Quadratic convergence in a primal-dual method," Technical Report 91-15, Dept. of Industrial Engineering and Management Science, Northwestern University, Evanston, IL 60208, USA, September 1991.

[12] 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.

[13] 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.

[14] A. M. Ostrowski, *Solution of Equations and Systems of Equations* , Academic Press, New York, (1960).

[15] R. Saigal " A simple proof of the primal affine scaling method," Technical Report No. 92-60, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI 48109, December 1992.

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

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

[18] T. Tsuchiya, " Global convergence property of the affine scaling method for primal degenerate linear programming problems," *Mathematics of Operations Research* 17 (1992) 527- 557.

[19] 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).

[20] T. Tsuchiya and R. D. C. Monterio " Superlinear Convergence of the Affine Scaling Algorithm," Research Memorandum, November 28, 1992.

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

[22] 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 Aris- ing 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.

[23] C. Witzgall, P. T. Boggs, and P. D. Domich, " On the convergence behavior of trajectories for linear programming," 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 161-187. American Mathematical Society, Providence, Rhode Island, USA, 1990.

[24] Y. Ye, O. Guler, R. A. Tapia, and Y. Zhang, " A quadratically convergent $O(\sqrt{n}L)$- iteration algorithm for linear programming," Technical Report TR-91-26, Dept. of Mathematical Sciences, Rice University, Houston, TX 77251, USA, August 1991.