Journal of the Operations Research
Society of Japan
Vol. 33, No. 1, March 1990

Local Convergence Properties of New Methods in Linear Programming*

Takashi Tsuchiya Kunio Tanabe
The Institute of Statistical Mathematics

(Received November 17, 1988; Revised December 27, 1989)

Abstract Asymptotic behavior of some of the interior point methods for Linear Programming is investigated without assuming nondegeneracy of constraints. A detailed analysis is given to the fundamental pair of vectors, the Newton direction leading to the center of the problem "d_{C}" and the direction of the affine-scaling method "d_{AF}". The quadratic convergence property of the iteration by $-d_{\mathrm{C}}$ is demonstrated. Step-sizes for the direction $-d_{\mathrm{C}}$ are also given to maintain feasibility without sacrificing the quadratic convergence. The sequence generated by $-d_{\mathrm{AF}}$ is pulled towards the central trajectory while it converges linearly to the optimal solution. Local convergence properties of Iri and Imai's Multiplicative Barrier Function method and Yamashita's method (a variation of the projective scaling methods) are also discussed. It is shown that the search direction of the method by Iri and Imai converges to $-d_{\mathrm{C}}$, whereas the direction by Yamashita converges to d_{AF}. A proof is given for the quadratic convergence property of Iri and Imai's method with an exact line search procedure in the case where constraint is degenerate.

1. Introduction

Since Karmarkar's method for Linear Programming was presented in 1984 [4], various new approaches have been attempted to get another method having the same property of polynomial time order of convergence. In 1985, Iri and Imai [3] proposed "Multiplicative Barrier Function" for the dual problem of the standard form Linear Programming, which is a natural affine variation of Karmarkar's potential function, to be minimized for solving the problem. They showed that the function is strictly convex under mild assumptions and that it can be minimized easily by the Newton method. They also found various interesting properties of the Newton method when applied to Multiplicative Barrier Function. Specifically, if the problem is nondegenerate, the Newton method attains the quadratic convergence even though the Hessian matrix does not exist at the optimal solution.

However, they did not give a proof for the property of polynomial time order of convergence. The method of minimizing Multiplicative Barrier Function which has the property of polynomial order of convergence was first proposed by Yamashita [11]. One of the important observations in his paper is that the search directions of the two methods are written as linear combinations of the same pair of vectors, the search direction "d_{AF}" of the dual affine scaling method [1], and the Newton direction "d_{C}" leading to the center of the system of linear inequalities of the problem (e. g., see [2, 6, 7]). Yamashita's method belongs to the group of the projective scaling methods [8].

In this paper we give a local analysis on those interior point methods. Under the assumption of the uniqueness of the optimal solution, we show the following properties.

(1) The negative centering displacement vector $-d_{\mathrm{C}}$, which is exactly "opposite" to the

* This is a revised version of the paper, "Local Analysis of the Methods by Iri and Imai, Yamashita and Adler-Karmarkar" (in Japanese) appeared in The Institute of Statistical Mathematics Cooperative Research Report 5 *"New Methods for Linear Programming"*, pp.99-117, March, 1987.

Newton direction leading to the "center" of the system of linear inequalities, has the property of the Newton displacement vector to find the optimal solution. If we adopt appropriate step-sizes, the iteration by $-d_{\mathrm{C}}$ produces a sequence which converges quadratically to the optimal solution from the interior of the feasible region.

(2) The sequence generated by the dual affine scaling method converges linearly to the optimal solution. In particular, if appropriate step-sizes are adopted in the iteration, the generated sequence is forced to approach the optimal solution from the unique direction which is the limiting tangential direction of the "central trajectory" at the optimal solution. Simultaneously, the "dual variable" converges to an optimal solution of the standard form linear programming problem which is the dual problem.

(3) The displacement vector of Iri and Imai's method d_{N} converges to the vector $-d_{\mathrm{C}}/(m-k)$ as the sequence approaches the opitmal solution, where m is the number of constraints of the problem and k is the number of active constraints at the optimal solution. The method with an exact line search procedure has the property of quadratic convergence even if the constraint is degenerate at the optimal solution.

(4) The search direction of Yamashita's method d_{P} converges to the search direction d_{AF} of the dual affine scaling method as the sequence approaches the optimal solution.

The "centering property" of the dual affine scaling method described in (2) was already observed by Megiddo and Shub [5] for the continuous version of the primal affine scaling method in the case where both primal and dual problems are nondegenerate. However, no analysis seems to be given so far on how this property holds in the discrete version of the algorithm. Our analysis is new in that it shows the same behavior of a discrete version of the algorithm and that it requires only the uniqueness of the optimal solution.

2. Problem

Here we consider the following dual standard form linear programming problem $<\mathrm{D}>$ with n variables and m constraints:

$$\begin{aligned}&\text{minimize } c^t x, \text{ subject to } A^t x - b \geq 0,\\ &A = (a_1, ..., a_m) \in R^{n\times m},\ x, c \in R^n,\ b \in R^m.\end{aligned}$$

As is well-known, the dual problem of $<\mathrm{D}>$ is the standard form linear programming problem $<\mathrm{P}>$:

$$\text{maximize } b^t y, \text{ subject to } Ay = c,\ y \geq 0,\ y \in R^m.$$

We assume the following conditions on $<\mathrm{D}>$:

(i) The feasible region of the problem $<\mathrm{D}>$ is not empty and it has at least one interior point.

(ii) Rank $A = n$.

(iii) The objective function is nondegenerate, i.e., the problem $<\mathrm{D}>$ has the unique optimal solution x^*.

(iv) The minimum value of the objective function is known in advance. Without loss of generality, we assume that the value of the objective function is 0 at x^*.

(v) There exists at least one inactive constraint at x^*.

Throughout this paper we put no assumption on degeneracy of constraints. We allow any number of constraints ($\geq n$) to be active at the optimal solution x^*, and denote by k the number of active constraints at x^*.

We introduce the following notations:

$$\begin{aligned} D(x) &= \operatorname{diag}[a_1^t x - b_1, ..., a_m^t x - b_m], \\ B(x) &= AD(x)^{-2}A^t, \\ \eta(x) &= AD(x)^{-1}\mathbf{1}, \\ \mathbf{1} &= (1, ..., 1)^t \in R^m. \end{aligned}$$

Without loss of generality, we assume that the first n constraints in A^t are active and independent at the optimal solution, while consecutive $(k-n)$ constraints are active but redundant ones. Then we introduce

$$\begin{aligned} &A_{A1} = (a_1, ..., a_n), A_{A2} = (a_{n+1}, ..., a_k), \\ &A_A = (A_{A1}, A_{A2}), A_I = (a_{k+1}, ..., a_m), \\ &D_{A1}(x) = \operatorname{diag}[a_1^t x - b_1, ..., a_n^t x - b_n], \\ &D_{A2}(x) = \operatorname{diag}[a_{n+1}^t x - b_{n+1}, ..., a_k^t x - b_k], \\ &D_A(x) = \operatorname{diag}[D_{A1}(x), D_{A2}(x)], \\ &D_I(x) = \operatorname{diag}[a_{k+1}^t x - b_{k+1}, ..., a_m^t x - b_m], \\ &B_{A1}(x) = A_{A1}D_{A1}(x)^{-2}A_{A1}^t, B_{A2}(x) = A_{A2}D_{A2}(x)^{-2}A_{A2}^t, \\ &B_A(x) = A_A D_A(x)^{-2}A_A^t = A_{A1}D_{A1}(x)^{-2}A_{A1}^t + A_{A2}D_{A2}(x)^{-2}A_{A2}^t \\ &\qquad = B_{A1}(x) + B_{A2}(x), \\ &A_{A1}, D_{A1}(x), B_{A1}(x), B_{A2}(x), B_A(x) \in R^{n\times n}, \\ &A_{A2} \in R^{n\times(k-n)}, D_{A2}(x) \in R^{(k-n)\times(k-n)}, \\ &A_A \in R^{n\times k}, D_A(x) \in R^{k\times k}, \\ &A_I \in R^{n\times(m-k)}, D_I(x) \in R^{(m-k)\times(m-k)}. \end{aligned}$$

The matrix $D_A(x)$ converges to 0 as $x \to x^*$, whereas $D_I(x)$ converges to an invertible matrix. Since the rank of A_{A1} is n, the matrix A_{A2} and A_I are represented as $A_{A2} = A_{A1}U$ and $A_I = A_{A1}W$ with appropriate matrices $U \in R^{n\times(k-n)}$ and $W \in R^{n\times(m-k)}$. For a vector $z \in R^m$, we denote the vector consists of the first n components of z by z_{A1}, the vector consists of the consecutive $k-n$ components by z_{A2}, and $(z_{A1}^t\ z_{A2}^t)^t$ by z_A. The vector consists of the remaining components of z is denoted by z_I. We use $\|\cdot\|$ for the l_2 norm.

Concerning the relation between the magnitudes of two quantities $f(x)$ and $g(x)$, we introduce the notations "$g(x) = O(f(x))$" and "$g(x) = \Theta(f(x))$". The notation $g(x) = O(f(x))$ means $\limsup_{x\to\alpha} |g(x)/f(x)| < +\infty$. When $g(x) = O(f(x))$ and $f(x) = O(g(x))$, we denote this relation by $g(x) = \Theta(f(x))$. We use the analogous notations for given two sequences $\{p^{(i)}\}$ and $\{q^{(i)}\}$ as well.

3. Search Directions of the Interior Point Methods

To solve the problem < D >, Iri and Imai [3] introduced the Multiplicative Barrier Function as follows:

$$F(x) = \frac{(c^t x)^{m+1}}{\prod_{\kappa=1}^{m}(a_\kappa^t x - b_\kappa)}. \tag{3.1}$$

The function has the following nice properties for solving the problem < D >:

(i) Solving the problem < D > is equivalent to minimizing the function (3.1). If $F(x^{(i)})$ converges to 0 along the sequence $\{x^{(i)}\}$, the distance between $x^{(i)}$ and the optimal solution set converges to zero as $i \to \infty$.

(ii) The function $F(x)$ is strictly convex in the interior of the feasible region.

In order to minimize the function, Iri and Imai used the Newton method with an exact line search procedure. The quadratic convergence of the method is not trivial in this case because the Hessian matrix is not well-defined at the optimal solution. They demonstrated the quadratic convergence of their method in the case where both primal and dual are nondegenerate at the optimal solution. However, they did not give a proof for the polynomiality of the method.

A search direction with a choice of step-size which reduces Multiplicative Barrier Function by a constant factor per iteration was first proposed by Yamashita [11]. Yamashita compared the search direction of his method with the search direction of Iri and Imai's method, and observed that the displacement vector of Iri and Imai's method d_{N} and his method d_{P} can be written as linear combinations of the same pair of vectors $d_{\mathrm{C}}(x) \equiv B(x)^{-1}\eta(x)$ and $d_{\mathrm{AF}}(x) \equiv -B(x)^{-1}c$, i.e.,

$$\begin{aligned} d_{\mathrm{N}}(x) &= \rho_0(x)(-\rho_1(x)B(x)^{-1}c - B(x)^{-1}\eta(x)),\\ \rho_0(x) &= \frac{1}{1 - g(x)^t(-\rho_1(x)B(x)^{-1}c + B(x)^{-1}\eta(x))},\\ \rho_1(x) &= \frac{c^t x - c^t B(x)^{-1}\eta(x)}{\dfrac{(c^t x)^2}{m+1} - c^t B(x)^{-1}c},\\ g(x) &= (m+1)\frac{c}{c^t x} - \eta(x), \end{aligned} \tag{3.2}$$

and

$$\begin{aligned} d_{\mathrm{P}}(x) &= \xi_0(x)(-B(x)^{-1}c + \xi_1(x)B(x)^{-1}\eta(x)),\\ \xi_0(x) &= \frac{c^t x}{\xi_2(x)},\ \xi_1(x) = \xi_3(x)(c^t x - c^t B(x)^{-1}\eta(x)),\\ \xi_2(x) &= (c^t x - c^t B(x)^{-1}\eta(x))^2\xi_3(x) + c^t B(x)^{-1}c,\\ \xi_3(x) &= \frac{1}{m+1-\eta(x)^t B(x)^{-1}\eta(x)}. \end{aligned} \tag{3.3}$$

(We represent the search direction of Yamashita's method in a simpler form than the original one [11] assuming that $c^t x^* = 0$.) The vector d_{C}, which is called "d_2" in [11], is the vector obtained when we apply the Newton method to get the so-called the "center" of the system of linear inequalities of the problem [2, 6, 7], whereas d_{AF}, which is called "d_1" in [11], is the direction of the dual affine scaling method proposed by Adler, Karmarkar et al. [1].

4. Basic Lemmas

In this section we describe some basic lemmas. It is convenient to represent the common quantities appearing in both formulae (3.2) and (3.3) by using the projection matrix $P(x) = D(x)^{-1}A^t B(x)^{-1}A\ D(x)^{-1}$, the projection onto the subspace spanned by the column vector of the matrix $D^{-1}(x)A^t$, and y^*, one of the optimal solutions of $<\mathrm{P}>$:

$$\eta(x)^t B(x)^{-1}\eta(x) = \mathbf{1}^t D(x)^{-1}A^t B(x)^{-1}AD(x)^{-1}\mathbf{1} = \mathbf{1}^t P(x)\mathbf{1}, \tag{4.1}$$

$$c^t B(x)^{-1}\eta(x) = y^{*t}D(x)D(x)^{-1}A^t B(x)^{-1}AD(x)^{-1}\mathbf{1} = y^{*t}D(x)P(x)\mathbf{1}, \tag{4.2}$$

$$\begin{aligned} c^t B(x)^{-1}c &= y^{*t}D(x)D(x)^{-1}A^t B(x)^{-1}AD(x)^{-1}D(x)y^*\\ &= y^{*t}D(x)P(x)D(x)y^*, \end{aligned} \tag{4.3}$$

$$c^t x = y^{*t}D(x)\mathbf{1}. \tag{4.4}$$

Here we used the equality $b^t y^* = c^t x^* = 0$ to derive (4.4). In the following we always adopt y^* such that $y^*_A > 0$ and $y^*_I = 0$. The existence of such y^* follows from the strong complimentarity. The two directions $B(x)^{-1}c$ and $B(x)^{-1}\eta(x)$ are represented as $(B(x)^{-1}A\ D(x)^{-1})D(x)y^*$ and $(B(x)^{-1}AD(x)^{-1})1$, respectively, in terms of the matrix $B(x)^{-1}AD(x)^{-1}$. This matrix is the Moore-Penrose inverse of $D(x)^{-1}A^t$ which gives the solution $B(x)^{-1}AD(x)^{-1}f$ for the following weighted least square problem:

$$\min_{\Delta x} \|D(x)^{-1}A^t\Delta x - f\|, \ \Delta x \in R^n, f \in R^m. \tag{4.5}$$

The projection matrix $P(x)$ has also a close relation to the problem. The residual of (4.5) is given by $(I - P(x))f$. In the following we investigate the properties of these matrices in a sufficiently small neighborhood of the optimal solution.

Lemma 4.1. The inequality

$$\|B_A(x)^{-1}\| \le \|A_{A1}^{-1}\|^2\|D_A(x)\|^2 \tag{4.6}$$

holds for all x in the interior of the feasible region of < D >.

Proof: We represent B_A as

$$B_A = A_{A1}D_{A1}^{-2}A_{A1}^t + A_{A2}D_{A2}^{-2}A_{A2}^t = A_{A1}D_{A1}^{-1}[I + D_{A1}UD_{A2}^{-2}U^tD_{A1}]D_{A1}^{-1}A_{A1}^t. \tag{4.7}$$

Since $D_{A1}UD_{A2}^{-2}U^tD_{A1}$ is a positive semi-definite matrix in the interior of the feasible region, $I + D_{A1}UD_{A2}^{-2}U^tD_{A1}$ is an invertible matrix. Then we have

$$B_A^{-1} = A_{A1}^{-t}D_{A1}[I + D_{A1}UD_{A2}^{-2}U^tD_{A1}]^{-1}D_{A1}A_{A1}^{-1}. \tag{4.8}$$

Since $\|[I + D_{A1}UD_{A2}^{-2}U^tD_{A1}]^{-1}\| \le 1$, we obtain $\|B_A^{-1}\| \le \|A_{A1}^{-1}\|^2\|D_{A1}\|^2 \le \|A_{A1}^{-1}\|^2\|D_A\|^2$, which is the desired estimate.

Lemma 4.2. The inequality

$$\|B_A(x)^{-1}A_AD_A(x)^{-1}\| \le \|A_{A1}^{-1}\|\|D_A(x)\| \tag{4.9}$$

holds for all x in the interior of the feasible region of < D >.

Proof: By its definition $\|B_A^{-1}A_AD_A^{-1}\|$ can be written as follows:

$$\|B_A^{-1}A_AD_A^{-1}\| = \|B_A^{-1}A_AD_A^{-2}A_A^tB_A^{-t}\|^{1/2} = \|B_A^{-1}\|^{1/2}. \tag{4.10}$$

As was shown in Lemma 4.1, $\|B_A^{-1}\| \le \|A_{A1}^{-1}\|^2\|D_A\|^2$ in the interior of the feasible region of < D >. Hence we see that $\|B_A^{-1}A_AD_A^{-1}\| \le \|A_{A1}^{-1}\|\|D_A\|$ in the interior of the feasible region of < D >.

Lemma 4.3. Let x be an interior feasible solution of < D >. Given a vector z in R^m, let

$$u = B_A(x)^{-1}A_AD_A(x)^{-1}z_A \tag{4.11}$$

and

$$w = B(x)^{-1}AD(x)^{-1}z, \tag{4.12}$$

which are the unique solutions of the least square problems

$$\min_{\tilde{u}\in R^n} \|D_A(x)^{-1}A_A^t\tilde{u} - z_A\| \text{ and } \min_{\tilde{w}\in R^n} \|D(x)^{-1}A^t\tilde{w} - z\|, \tag{4.13}$$

respectively. Then $\|u - w\| = O(\|D_A(x)\|^2)$ in a sufficiently small neighborhood of x^*.

Proof: The matrix B and $AD^{-1}z$ are represented as

$$B = (A_A\ A_I)\begin{pmatrix} D_A & 0 \\ 0 & D_I \end{pmatrix}^{-2}\begin{pmatrix} A_A^t \\ A_I^t \end{pmatrix} = B_A + A_I D_I^{-2} A_I^t \tag{4.14}$$

and

$$AD^{-1}z = A_A D_A^{-1} z_A + A_I D_I^{-1} z_I. \tag{4.15}$$

We have the following expression of B^{-1} by applying Sherman-Morrison-Woodbury formula to (4.14):

$$\begin{aligned} B^{-1} &= B_A^{-1} - B_A^{-1}A_I D_I^{-1}[I + D_I^{-1}A_I^t B_A^{-1} A_I D_I^{-1}]^{-1} D_I^{-1} A_I^t B_A^{-1} \\ &= B_A^{-1} - B_A^{-1} C B_A^{-1}, \end{aligned} \tag{4.16}$$

where

$$C(x) = A_I D_I(x)^{-1}[I + D_I(x)^{-1}A_I^t B_A(x)^{-1} A_I D_I(x)^{-1}]^{-1} D_I(x)^{-1} A_I^t. \tag{4.17}$$

Since we have $\|D_I^{-1}\| = O(1)$ near the optimal solution, we see that $\|C\| = O(1)$ in a sufficiently small neighborhood of x^*. Substituting (4.15) and (4.16) into (4.12), we obtain the following expression for w:

$$\begin{aligned} w &= B^{-1}AD^{-1}z = (B_A^{-1} - B_A^{-1}CB_A^{-1})(A_A D_A^{-1} z_A + A_I D_I^{-1} z_I) \\ &= B_A^{-1} A_A D_A^{-1} z_A + r = u + r, \end{aligned} \tag{4.18}$$

where

$$r = -B_A^{-1}CB_A^{-1}A_A D_A^{-1} z_A - B_A^{-1}CB_A^{-1}A_I D_I^{-1} z_I + B_A^{-1} A_I D_I^{-1} z_I. \tag{4.19}$$

As was shown in Lemma 4.1 and Lemma 4.2, $\|B_A^{-1}\| = O(\|D_A\|^2)$ and $\|B_A^{-1}A_A D_A^{-1}\| = O(\|D_A\|)$ in the interior of the feasible region. On the other hand, $\|D_I\|$ converges to an invertible matrix as $x \to x^*$. Hence the norm of the vector $r = u - w$ is bounded by $M\|D_A\|^2$ in a sufficiently small neighborhood of the optimal solution, where M is an appropriate constant.

Lemma 4.4. Let x be an interior feasible solution of $< \mathrm{D} >$, and let the projection matrix $\widehat{P}$ on R^m be

$$\widehat{P}(x) = \begin{pmatrix} P_A(x) & 0 \\ 0 & 0 \end{pmatrix}, \tag{4.20}$$

where

$$P_A(x) = D_A(x)^{-1} A_A^t B_A(x)^{-1} A_A D_A(x)^{-1}. \tag{4.21}$$

If x is sufficiently close to the optimal solution x^*, then

$$\|\widehat{P}(x) - P(x)\| = O(\|D_A(x)\|). \tag{4.22}$$

Proof: We write $P(x) = D(x)^{-1}A^tB(x)^{-1}AD(x)^{-1}$ by using (4.16) as follows:

$$P = \begin{pmatrix} D_A^{-1}A_A^t \\ D_I^{-1}A_I^t \end{pmatrix} (B_A^{-1} - B_A^{-1}CB_A^{-1}) \left(A_AD_A^{-1}, A_ID_I^{-1} \right)$$

$$= \begin{pmatrix} P_A + C_{11} & C_{12} \\ C_{12}^t & C_{22} \end{pmatrix} = \widehat{P} + \begin{pmatrix} C_{11} & C_{12} \\ C_{12}^t & C_{22} \end{pmatrix}, \qquad (4.23)$$

where

$$C_{11}(x) = -D_A(x)^{-1}A_A^tB_A(x)^{-1}C(x)B_A(x)^{-1}A_AD_A(x)^{-1}, \qquad (4.24)$$

$$C_{12}(x) = D_A(x)^{-1}A_A^tB_A(x)^{-1}A_ID_I(x)^{-1} - D_A(x)^{-1}A_A^tB_A(x)^{-1}C(x)B_A(x)^{-1}A_ID_I(x)^{-1}, \qquad (4.25)$$

$$C_{22}(x) = D_I(x)^{-1}A_I^tB_A(x)^{-1}A_ID_I(x)^{-1} - D_I(x)^{-1}A_I^tB_A(x)^{-1}C(x)B_A(x)^{-1}A_ID_I(x)^{-1}. \qquad (4.26)$$

From (4.17) and the order of $\|B_A^{-1}\|$ and $\|B_A^{-1}A_AD_A^{-1}\|$ given in Lemma 4.1 and Lemma 4.2, we easily see $\|C_{11}\| = O(\|D_A\|^2)$, $\|C_{12}\| = O(\|D_A\|)$ and $\|C_{22}\| = O(\|D_A\|^2)$ in a sufficiently small neighborhood of x^*. This immediately leads to the result.

Before concluding this section, we give a remark on the relation between the order of $\|D_A(x)\|$, c^tx $(= c^t(x - x^*))$ and $\|x - x^*\|$, which easily follows from the facts that $D_A\mathbf{1}_A = A_A^t(x - x^*)$, $x - x^* = (A_AA_A^t)^{-1}A_AD_A\mathbf{1}_A$ and $c^tx = y_A^{*t}D_A\mathbf{1}_A$ $(y_A^* > 0)$.

Remark 4.5. We have the following inequalities for all x in the feasible region of $< \mathrm{D} >$:

$$\|A_A^t\|^{-1}\|D_A\| \leq \|x - x^*\| \leq \sqrt{k}\|(A_AA_A^t)^{-1}A_A\|\|D_A\|, \qquad (4.27)$$

$$(\min_{\kappa=1}^{k} y_{A\kappa}^*)\|D_A\| \leq c^tx \leq \sqrt{k}\|y_A^*\|\|D_A\|. \qquad (4.28)$$

This remark means that $\|D_A\|$, c^tx and $\|x - x^*\|$ are quantities of the same order in the feasible region.

5. Properties of the Iteration by the Negative Centering Displacement Vector

In this section we analyze the negative centering displacement vector $-d_C = -B^{-1}\eta$ which plays an important role in guaranteeing the quadratic convergence of Iri and Imai's method. Specifically we prove that the iteration

$$x^{(i+1)} = x^{(i)} - \mu^{(i)}B(x^{(i)})^{-1}\eta(x^{(i)}) \quad (i = 0, 1, \dots) \qquad (5.1)$$

converges quadratically to the optimal solution without violating the feasibility, if the initial interior feasible solution $x^{(0)}$ is sufficiently close to the optimal solution x^* and if we adopt appropriate step-sizes in the iteration. We denote $B(x^{(i)}), B_A(x^{(i)})$, $D(x^{(i)}), D_A(x^{(i)})$ and $\mu^{(i)}$ by B, B_A, D, D_A and μ, respectively.

Lemma 5.1. Let $\mu^{(i)} = 1$ in the iteration (5.1). Then, if $x^{(i)}$ is an interior feasible solution of $< \mathrm{D} >$ sufficiently close to the optimal solution, we have

$$\|x^{(i+1)} - x^*\| = \|x^{(i)} - B^{-1}\eta - x^*\| = O(\|x^{(i)} - x^*\|^2).$$

Proof: The vector $x^{(i)} - x^*$ satisfies the equation $A_A^t(x^{(i)} - x^*) = D_A\mathbf{1}_A$, or, equivalently,

$$D_A^{-1}A_A^t(x^{(i)} - x^*) = \mathbf{1}_A. \tag{5.2}$$

When $k > n$, this equation is overdetermined but consistent. Multiplying both sides of (5.2) by $B_A^{-1}A_A^tD_A^{-1}$, we obtain the following expression of $x^{(i)} - x^*$:

$$x^{(i)} - x^* = B_A^{-1}A_AD_A^{-1}\mathbf{1}_A. \tag{5.3}$$

Then we have

$$\begin{aligned} x^{(i+1)} - x^* &= x^{(i)} - x^* - B^{-1}AD^{-1}\mathbf{1} \\ &= B_A^{-1}A_AD_A^{-1}\mathbf{1}_A - B^{-1}AD^{-1}\mathbf{1}. \end{aligned} \tag{5.4}$$

Putting $z = \mathbf{1}$ in Lemma 4.3 and using Remark 4.5, we have

$$\|x^{(i+1)} - x^*\| \le M_1\|D_A(x^{(i)})\|^2 \le M_2\|x^{(i)} - x^*\|^2 \tag{5.5}$$

if $x^{(i)}$ is in a sufficiently small neighborhood of x^*, where M_1 and M_2 are some appropriate constants. This completes the proof.

Lemma 5.1 implies that the negative centering direction can be interpreted as the Newton direction to find the optimal solution. It suggests also a close relation between the quadratic convergence of the interior point methods and the negative centering direction. This property is quite surprising, because there seems no reason for us to expect that the direction has such a remarkable property. In the following we give the maximum step-size $\mu_{\max}^{(i)}$ which does not violate feasibility. Note that the new iterate $x^{(i+1)}$ of (5.1) remains as an "interior point" of the feasible region so long as $0 \le \mu^{(i)} < \mu_{\max}^{(i)}$. The step-size $\mu_{\max}^{(i)}$ satisfies the condition

$$A^t(x^{(i)} - \mu_{\max}^{(i)}B^{-1}\eta) - b \ge 0, \tag{5.6}$$

where at least one of the constraints in (5.6) is satisfied with equality. Then it is easy to see that $\mu_{\max}^{(i)}$ is bounded from below as follows:

$$\begin{aligned} \mu_{\max}^{(i)} &= \min_{\kappa, a_\kappa^tB^{-1}\eta>0} \frac{a_\kappa^tx^{(i)} - b_\kappa}{a_\kappa^tB^{-1}\eta} = \Big(\max_{\kappa, a_\kappa^tB^{-1}\eta>0} \frac{a_\kappa^tB^{-1}\eta}{a_\kappa^tx^{(i)} - b_\kappa}\Big)^{-1} \\ &\ge \Big(\max_{\kappa=1}^{m} \Big|\frac{a_\kappa^tB^{-1}\eta}{a_\kappa^tx^{(i)} - b_\kappa}\Big|\Big)^{-1} = \|D^{-1}A^tB^{-1}\eta\|_\infty^{-1} = \|P\mathbf{1}\|_\infty^{-1}. \end{aligned} \tag{5.7}$$

Since $D_A\mathbf{1}_A = A_A^t(x^{(i)} - x^*) = A_A^tB_A^{-1}A_AD_A^{-1}\mathbf{1}_A$ (see (5.3)), we have

$$P_A\mathbf{1}_A = D_A^{-1}A_A^tB_A^{-1}A_AD_A^{-1}\mathbf{1}_A = \mathbf{1}_A. \tag{5.8}$$

Hence it follows from the proof of Lemma 4.4 that

$$P\mathbf{1} = \begin{pmatrix} \mathbf{1}_A \\ 0 \end{pmatrix} + \begin{pmatrix} C_{11}\mathbf{1}_A + C_{12}\mathbf{1}_I \\ C_{12}^t\mathbf{1}_A + C_{22}\mathbf{1}_I \end{pmatrix}, \tag{5.9}$$

where $\|C_{11}\| = O(\|D_A\|^2)$, $\|C_{12}\| = O(\|D_A\|)$ and $\|C_{22}\| = O(\|D_A\|^2)$. Then, $\|P\mathbf{1}\|_\infty^{-1}$ is bounded from below by a quantity of $1 - O(\|D_A\|)$ near the optimal solution. This means $\mu_{\max}^{(i)}$ also is bounded from below by a quantity of $1 - O(\|D_A\|)$, say, $1 - M_3\|D_A\|$, where M_3

is an appropriate constant. Now, let $x^{(0)}$ be an interior feasible solution which is sufficiently close to the optimal solution. In the following we show that the sequence $\{x^{(i)}\}$ generated by (5.1) with $\mu^{(i)} = 1 - M_3\|D_A(x^{(i)})\|$ $(i = 0, 1, ...)$ is a sequence of interior points of the feasible region which converges quadratically to x^*.

Let $x^{(i)}$ be the i-th iterate in the iteration (5.1), which is assumed to be an interior feasible solution. As mentioned above, we have $\mu^{(i)}_{\max} > \mu^{(i)} = 1 - M_3\|D_A(x^{(i)})\|$, if $x^{(i)}$ is sufficiently close to x^*. This implies that $x^{(i+1)}$ is an interior point of the feasible region. From Lemma 5.1, we see $\|B(x^{(i)})^{-1}\eta(x^{(i)})\| = O(\|x^{(i)} - x^*\|)$. Then, by using Lemma 5.1 and Remark 4.5, $\|x^{(i+1)} - x^*\|$ is bounded from above as follows, when $\|x^{(i)} - x^*\|$ is sufficiently small:

$$\begin{aligned}\|x^{(i+1)} - x^*\| &= \|x^{(i)} - \mu^{(i)}B(x^{(i)})^{-1}\eta(x^{(i)}) - x^*\| \\ &= \|x^{(i)} - (1 - M_3\|D_A(x^{(i)})\|)B(x^{(i)})^{-1}\eta(x^{(i)}) - x^*\| \\ &\le \|x^{(i)} - B(x^{(i)})^{-1}\eta(x^{(i)}) - x^*\| + M_3\|D_A(x^{(i)})\| \cdot \|B(x^{(i)})^{-1}\eta(x^{(i)})\| \\ &\le M_2\|x^{(i)} - x^*\|^2 + M_4\|D_A(x^{(i)})\| \cdot \|x^{(i)} - x^*\| \\ &\le M_5\|x^{(i)} - x^*\|^2, \end{aligned} \tag{5.10}$$

where M_4 and M_5 are appropriate constants. Obviously, this means that, if $x^{(i)}$ is sufficiently close to the optimal solution, $x^{(i+1)}$ also remains as "an interior point of the feasible region sufficiently close to the optimal solution", so that the same argument can be applied to the new iterate $x^{(i+1)}$ recursively. Thus, we have the following theorem.

Theorem 5.2. Let $x^{(0)}$ be an initial iterate for the iteration (5.1), which is an interior feasible solution of $< \mathrm{D} >$ in a sufficiently small neighborhood of the optimal solution. If we adopt appropriate step-sizes, say, $\mu^{(i)} = 1 - M_3\|D_A(x^{(i)})\|$, then the iteration (5.1) produces a sequence of interior feasible solutions of $< \mathrm{D} >$ which converges quadratically to the optimal solution.

This is an interesting feature of the negative centering direction $-d_{\mathrm{C}}$. Note that the objective function c^tx did not appear in the analysis throughout this section. Hence the results in this section are valid near any vertex of the feasible region.

Iri and Imai demonstrated in [3] that their method has the property of quadratic convergence if it is applied to nondegenerate problems, and that the step-size converges to $m - n$ in the final stage of the iteration. Our analysis suggests that $(m - n)d_{\mathrm{N}}$ asymptotically converges to $-d_{\mathrm{C}}$ in the case where both primal and dual are nondegenerate at the optimal solution. A detailed discussion from this viewpoint will be given in §7 .

6. Properties of the Iteration of the Dual Affine Scaling Method

In this section we investigate the asymptotic behavior of the iteration of the dual affine scaling method

$$x^{(i+1)} = x^{(i)} + \mu^{(i)}\frac{d_{\mathrm{AF}}}{c^tx^{(i)}} = x^{(i)} - \mu^{(i)}\frac{B(x^{(i)})^{-1}c}{c^tx^{(i)}}. \tag{6.1}$$

Here we fix $\mu^{(i)}$ to be a constant $\mu \in (0, 1)$. As we will show later, if $x^{(i)}$ is an "interior feasible solution sufficiently close to the optimal solution", the next iterate $x^{(i+1)}$ also remains as an "interior feasible solution sufficiently close to the optimal solution", so that (6.1) is well-defined.

In the following we show that the sequence produced by (6.1) converges linearly to the optimal solution always from the unique direction. This direction is the limiting tangential direction of the central trajectory when the trajectory approaches the optimal solution. It will be also demonstrated that the "dual variable" converges to an optimal solution of < P >. This interesting property of centering was first observed by Megiddo and Shub for the trajectory of the continuous version of the primal affine scaling method in the paper [5], where the global and local behavior of the primal affine scaling method is studied extensively under nondegeneracy assumption. Our result is an extension of their result to the discrete case, but does not require the assumption of nondegeneracy of constraints.

Firstly, we observe that the displacement vector $B^{-1}c/c^tx$ is written, by using (4.4), as follows:

$$\frac{B^{-1}c}{c^tx}=\frac{B^{-1}AD^{-1}Dy^*}{\mathbf{1}^tDy^*}. \tag{6.2}$$

As noted in §4, the vector y^* is an optimal solution of < P > such that $y^*_A>0$ and $y^*_I=0$. The optimal solution y^* may not be determined uniquely when the constraints are degenerate at the optimal solution, but this redundancy is irrelevant in the following analysis. From (4.16) and $A_Ay^*_A=c$, we see that (6.2) is represented as follows:

$$\frac{B^{-1}c}{c^tx}=B_A^{-1}A_AD_A^{-1}\frac{D_Ay^*_A}{\mathbf{1}_A^tD_Ay^*_A}+B_A^{-1}r_0. \tag{6.3}$$

Here $r_0(x)=-CB_A^{-1}A_AD_A^{-1}D_Ay^*_A/\mathbf{1}^t_AD_Ay^*_A\in R^n$ and $\|r_0\|=O(\|D_A\|)$ in a sufficiently small neighborhood of x^*, which follows from $\|D_Ay^*_A/\mathbf{1}^t_AD_Ay^*_A\|_1=1$, Lemma 4.2, and the order, $\|C\|=O(1)$, given in the proof of Lemma 4.3. Substituting (6.3) into (6.1) and then multiplying the both sides by A^t_A and subtracting b_A, we obtain the following recursive formula for the vector $D_A^{(i)}\mathbf{1}_A\equiv D_A(x^{(i)})\mathbf{1}_A$:

$$D_A^{(i+1)}\mathbf{1}_A=D_A^{(i)}\mathbf{1}_A-\mu D_A^{(i)}P_A^{(i)}\frac{D_A^{(i)}y^*_A}{\mathbf{1}^t_AD_A^{(i)}y^*_A}-\mu D_A^{(i)}r_1^{(i)}, \tag{6.4}$$

where $P_A^{(i)}=P_A(x^{(i)})$, $P_A(x)$ is as defined in §4, $r_1^{(i)}=r_1(x^{(i)})$, $r_1(x)=D_A(x)^{-1}A^t_AB_A(x)^{-1}r_0(x)\in R^k$ and $\|r_1\|=O(\|D_A\|^2)$. Here we define the weighted slack variables $\alpha(x)$ by $D_A(x)y^*_A$, and denote $P_A(x)\alpha(x)$ by $\hat{\alpha}(x)$. Multiplying both sides of (6.4) by the matrix diag$[y^*_A]$, we have the following equivalent recursive formula with respect to $\alpha^{(i)}=\alpha(x^{(i)})$ and $\hat{\alpha}^{(i)}=\hat{\alpha}(x^{(i)})$:

$$\alpha_\kappa^{(i+1)}=(1-\mu r_{1\kappa}^{(i)})\alpha_\kappa^{(i)}-\mu\frac{\alpha_\kappa^{(i)}\hat{\alpha}_\kappa^{(i)}}{\sum_{\lambda=1}^k\alpha_\lambda^{(i)}}\quad(\kappa=1,...,k). \tag{6.5}$$

Note that $\|\alpha^{(i)}\|_1=\mathbf{1}^t_A\alpha^{(i)}=\mathbf{1}^t_AD_A^{(i)}y^*_A=c^tx^{(i)}>0$. Let $\beta(x)$ be the vector $\alpha(x)/\alpha(x)^t\mathbf{1}_A=\alpha(x)/c^tx=D_A(x)y^*_A/\mathbf{1}^t_AD_A(x)y^*_A$, and $\hat{\beta}(x)$ be the vector $P_A(x)\beta(x)$. Dividing (6.5) by $\alpha^{(i)t}\mathbf{1}_A>0$, we have

$$\frac{\alpha_\kappa^{(i+1)}}{\sum_{\lambda=1}^k\alpha_\lambda^{(i)}}=(1-\mu r_{1\kappa}^{(i)})\beta_\kappa^{(i)}-\mu\beta_\kappa^{(i)}\hat{\beta}_\kappa^{(i)}, \tag{6.6}$$

where $\beta^{(i)}=\beta(x^{(i)})$ and $\hat{\beta}^{(i)}=\hat{\beta}(x^{(i)})$. From (6.6) we have the following recursive formula for $\beta^{(i)}$ and $\hat{\beta}^{(i)}$:

$$\beta_\kappa^{(i+1)}=\frac{\alpha_\kappa^{(i+1)}}{\sum_{\sigma=1}^k\alpha_\sigma^{(i+1)}}=\frac{\alpha_\kappa^{(i+1)}/\sum_{\lambda=1}^k\alpha_\lambda^{(i)}}{\sum_{\sigma=1}^k\{\alpha_\sigma^{(i+1)}/\sum_{\rho=1}^k\alpha_\rho^{(i)}\}}=\frac{(1-\mu r_{1\kappa}^{(i)})\beta_\kappa^{(i)}-\mu\beta_\kappa^{(i)}\hat{\beta}_\kappa^{(i)}}{\sum_{\lambda=1}^k\{(1-\mu r_{1\lambda}^{(i)})\beta_\lambda^{(i)}-\mu\beta_\lambda^{(i)}\hat{\beta}_\lambda^{(i)}\}}$$

$$=\frac{(1-\mu r_{1\kappa}^{(i)})\beta_\kappa^{(i)}-\mu\beta_\kappa^{(i)}\hat{\beta}_\kappa^{(i)}}{1-\mu\sum_{\lambda=1}^k\beta_\lambda^{(i)}\hat{\beta}_\lambda^{(i)}-\mu\sum_{\tau=1}^k r_{1\tau}^{(i)}\beta_\tau^{(i)}}\quad(\kappa=1,...,k). \tag{6.7}$$

We use the relations (6.8) ~(6.11) which follow from $P_A\mathbf{1}_A = \mathbf{1}_A$. Note that the first and the third inequalities in (6.10) follow from (6.8).

$$\sum_{\kappa=1}^{k}\beta_\kappa = \beta^t\mathbf{1}_A = \beta^t P_A\mathbf{1}_A = \sum_{\kappa=1}^{k}\hat\beta_\kappa = 1, \tag{6.8}$$

$$\sum_{\kappa=1}^{k}\alpha_\kappa = \alpha^t\mathbf{1}_A = \alpha^t P_A\mathbf{1}_A = \sum_{\kappa=1}^{k}\hat\alpha_\kappa = c^t x, \tag{6.9}$$

$$1 \geq \sum_{\kappa=1}^{k}\beta_\kappa^2 = \|\beta\|^2 \geq \beta^t P_A\beta = \sum_{\kappa=1}^{k}\beta_\kappa\hat\beta_\kappa = \beta^t P_A^2\beta = \|\hat\beta\|^2 = \sum_{\kappa=1}^{k}\hat\beta_\kappa\hat\beta_\kappa \geq \frac{1}{k}, \tag{6.10}$$

$$\sum_{\kappa=1}^{k}\alpha_\kappa^2 = \|\alpha\|^2 \geq \alpha^t P_A\alpha = \sum_{\kappa=1}^{k}\alpha_\kappa\hat\alpha_\kappa = \alpha^t P_A^2\alpha = \|\hat\alpha\|^2 = \sum_{\kappa=1}^{k}\hat\alpha_\kappa\hat\alpha_\kappa. \tag{6.11}$$

In the following we show that the iteration (6.1) produces a sequence of interior feasible solutions which converges linearly to the optimal solution if the initial iterate is an interior feasible solution sufficiently close to the optimal solution. We prove this in the following two steps; (i) $x^{(i+1)}$ remains as an interior feasible solution if $x^{(i)}$ is an interior feasible solution such that $c^t x^{(i)}$ is sufficiently small, and (ii) If $c^t x^{(0)}$ is sufficiently small, the value of the objective function converges linearly to 0, where the asymptotic reduction rate is at least $1-\mu/k$.

In order to see (i), we show that $D^{(i+1)}\mathbf{1} = (D_A^{(i+1)}\mathbf{1}_A\ D_I^{(i+1)}\mathbf{1}_I) > 0$ if $x^{(i)}$ is an interior feasible solution where $c^t x^{(i)}$ is sufficiently small. The value of $D_A^{(i)}\mathbf{1}_A$ changes according to (6.4). We just recall the expression (6.6) to see $D_A^{(i+1)}\mathbf{1}_A > 0$. Since $\alpha = D_A y_A$ and $y_A > 0$, sign of each component of (6.4) coincides with that of (6.6). Due to (6.10), the formula (6.6) is bounded from below as follows:

$$(\text{the formula (6.6)}) \geq (1-\mu\|r_1^{(i)}\| - \mu\|\hat\beta^{(i)}\|)\beta_\kappa^{(i)} \geq (1-\mu(1+\|r_1^{(i)}\|))\beta_\kappa^{(i)}, \tag{6.12}$$

which obviously is strictly positive if $\|D_A^{(i)}\|$ is sufficiently small, i.e., if $c^t x^{(i)}$ is sufficiently small. On the other hand, the value of $D_I^{(i+1)}\mathbf{1}_I$ changes according to the formula

$$D_I^{(i+1)}\mathbf{1}_I = D_I^{(i)}\mathbf{1}_I - \mu A_I^t\frac{B(x^{(i)})^{-1}c}{c^t x^{(i)}}. \tag{6.13}$$

From (6.3), Lemma 4.2, Remark 4.5 and the assumption $0<\mu<1$, we see that $\|\mu B^{-1}c/c^t x\| = O(c^t x)$ in a sufficiently small neighborhood of the optimal solution. Since the value of each component of $D_I^{(i)}\mathbf{1}_I$ is bounded from below by a positive number near the optimal solution, we see that each component in (6.13) is strictly positive if $x^{(i)}$ is in a sufficiently small neighborhood of the optimal solution. Thus $x^{(i+1)}$ remains also as an interior feasible solution if $x^{(i)}$ is an interior feasible solution such that $c^t x^{(i)}$ is sufficiently small.

Now we proceed to prove (ii). Summing up the elements of (6.6) and noting the relation (6.10), we see

$$\frac{c^t x^{(i+1)}}{c^t x^{(i)}} = \frac{\alpha^{(i+1)t}\mathbf{1}_A}{\alpha^{(i)t}\mathbf{1}_A} \leq (1+\mu\|r_1^{(i)}\|) - \mu\|\hat\beta^{(i)}\|^2 \leq (1+\mu\|r_1^{(i)}\|) - \frac{\mu}{k}, \tag{6.14}$$

which shows that the asymptotic reduction rate of the objective function per iteration is at least $1 - \mu/k$.

Thus, if the value of the objective function is sufficiently small at the initial iterate, the iteration (6.1) produces a sequence of interior feasible solutions which converges linearly to the optimal solution.

Now, we deal with the centering property of the iteration (6.1). We show that every sequence generated by (6.1) which converges to x^* shares the unique asymptotic direction from which the sequence approaches x^*. In order to characterize this asymptotic direction, let us consider the following strictly convex programming problem $< \mathrm{Q} >$:

$$\text{minimize} \quad -\sum_{\kappa=1}^{k} \log(a_\kappa^t z_\kappa), \qquad \text{subject to} \quad c^t z = 1, \quad A_A^t z \geq 0, \tag{6.15}$$

where $z \in R^k$. Since the problem $< \mathrm{D} >$ has an interior feasible solution, $< \mathrm{Q} >$ has a relative interior feasible solution. (To see this, let x be an interior feasible solution of $< \mathrm{D} >$, and put $z = (x - x^*)/c^t(x - x^*) = (x - x^*)/c^t x$. Noting that $A_A^t x - b_A = A_A^t(x - x^*) > 0$, it is easy to check z is actually a relative interior feasible solution of $< \mathrm{Q} >$.) From the uniqueness of the optimal solution of $< \mathrm{D} >$, we see that $< \mathrm{Q} >$ has a bounded feasible region. Then $< \mathrm{Q} >$ has a unique optimal solution in the relative interior of its feasible region. Let us denote by z^* the optimal solution of $< \mathrm{Q} >$. The point z^* is the unique solution of the following equation which is equivalent to the Karush-Kuhn-Tucker condition for $< \mathrm{Q} >$:

$$A_A \tilde{D}_A(z)^{-1} \mathbf{1}_A = kc, \tag{6.16}$$

where $\tilde{D}_A(z) = \mathrm{diag}[A_A^t z]$. Note that $-z^*$ is the limiting tangential direction of the central trajectory of $< \mathrm{D} >$ at the optimal solution x^*.

We want to show that the convergent sequence $\{x^{(i)}\}$ generated by (6.1) has the following asymptotic property:

$$\frac{(x^{(i)} - x^*)^t}{\|x^{(i)} - x^*\|} \frac{z^*}{\|z^*\|} \to 1 \quad \text{as } i \text{ goes to } \infty. \tag{6.17}$$

This implies that the sequence is always forced to approach the optimal solution from the unique direction, the direction of z^*, i. e., the sequence is pulled toward the central trajectory as it approaches the optimal solution. We begin with the following lemma.

Lemma 6.1. Consider the iteration (6.1) with the step-size $\mu^{(i)}$ chosen to be a constant $\mu \in (0,1)$. If $x^{(0)}$ is an interior feasible solution sufficiently close to the optimal solution, $\hat{\beta}^{(i)} = P_A^{(i)} \beta^{(i)}$ converges to $\mathbf{1}_A/k$ as $x^{(i)}$ does to the optimal solution x^*.

Proof: We analyze the behavior of $\prod_{\kappa=1}^{k} \beta_\kappa^{(i)}$. Let us consider the ratio

$$\frac{\prod_{\kappa=1}^{k} \beta_\kappa^{(i+1)}}{\prod_{\lambda=1}^{k} \beta_\lambda^{(i)}} = H(\hat{\beta}^{(i)}, \mu, r_1^{(i)}) = \frac{\prod_{\kappa=1}^{k}(1 - \mu\hat{\beta}_\kappa^{(i)} - \mu r_{1\kappa}^{(i)})}{(1 - \mu\sum_{\lambda=1}^{k} \hat{\beta}_\lambda^{(i)} \hat{\beta}_\lambda^{(i)} - \mu\sum_{\tau=1}^{k} r_{1\tau}^{(i)} \beta_\tau^{(i)})^k}, \tag{6.18}$$

where

$$H(\tilde{\beta}, \mu, \tilde{r}_1) \equiv \frac{\prod_{\kappa=1}^{k}(1 - \mu\tilde{\beta}_\kappa - \mu\tilde{r}_{1\kappa})}{(1 - \mu\sum_{\lambda=1}^{k} \tilde{\beta}_\lambda \tilde{\beta}_\lambda - \mu\sum_{\tau=1}^{k} \tilde{r}_{1\tau}\tilde{\beta}_\tau)^k} \quad (\tilde{\beta}, \tilde{r}_1 \in R^k). \tag{6.19}$$

The expression (6.18) follows from (6.7) and (6.10). The function $H(\tilde{\beta}, \mu, \tilde{r}_1)$ is approximated by the function

$$G(\tilde{\beta}, \mu) \equiv \frac{\prod_{\kappa=1}^{k}(1 - \mu\tilde{\beta}_\kappa)}{(1 - \mu\sum_{\kappa=1}^{k} \tilde{\beta}_\kappa \tilde{\beta}_\kappa)^k} \tag{6.20}$$

on the set

$$T \equiv \{\tilde{\beta} \in R^k | \tilde{\beta}^t \mathbf{1}_A = 1, \|\tilde{\beta}\| \leq 1\} \tag{6.21}$$

if $\|\tilde{r}_1\|$ is small enough. This applies to our case. Since $\hat{\beta}^{(i)} \in T$ for all i and $\|r_1^{(i)}\| = O(\|D_A^{(i)}\|^2) = O((c^t x^{(i)})^2)$, the function $H(\hat{\beta}^{(i)}, \mu, r_1^{(i)})$ is approximated by $G(\hat{\beta}^{(i)}, \mu)$, if $x^{(i)}$ is in a sufficiently small neighborhood of x^*. Specifically, we have the following inequality

$$G(\hat{\beta}^{(i)}, \mu)(1 - M_1 c^t x^{(i)}) < H(\hat{\beta}^{(i)}, \mu, r_1^{(i)}) < G(\hat{\beta}^{(i)}, \mu)(1 + M_1 c^t x^{(i)}) \tag{6.22}$$

if $c^t x^{(i)}$ is sufficiently small, where M_1 is an appropriate constant.

As we show in Appendix, the function $G(\tilde{\beta}, \mu)$ has the unique minimum point $\mathbf{1}_A/k$ on the set T. Since $G(\mathbf{1}_A/k, \mu) = 1$ and G is continuous on T, it is easy to see that "for any given $\varepsilon \in (0, 1]$, there exists a positive number $\zeta > 0$ such that

$$G(\tilde{\beta}, \mu) \geq 1 + \zeta \tag{6.23}$$

holds for all $\tilde{\beta} \in T$ such that $\varepsilon \leq \|\tilde{\beta} - \mathbf{1}_A/k\|$."

Now we are ready to show the lemma. If we take $x^{(0)}$ to be an interior feasible solution sufficiently close to the optimal solution, the inequality

$$c^t x^{(i)} \leq c^t x^{(0)} \sigma^i \tag{6.24}$$

and the inequality (6.22) hold for all $i = 0, 1, ...,$ where $\sigma \in (0, 1)$ is a constant. Then $\prod \beta_\kappa^{(i)}$ is bounded from below as follows:

$$\begin{aligned}\prod_{\kappa=1}^{k} \beta_\kappa^{(i)} &= \prod_{\kappa=1}^{k} \beta_\kappa^{(0)} \prod_{l=0}^{i-1} H(\hat{\beta}^{(l)}, \mu, r_1^{(l)}) \geq \prod_{\kappa=1}^{k} \beta_\kappa^{(0)} \prod_{l=0}^{i-1} G(\hat{\beta}^{(l)}, \mu) \prod_{l=0}^{i-1} (1 - M_1 c^t x^{(0)} \sigma^l) \\ &\geq (1 - M_1 c^t x^{(0)})^{\frac{1}{1-\sigma}} \prod_{\kappa=1}^{k} \beta_\kappa^{(0)} \prod_{l=0}^{i-1} G(\hat{\beta}^{(l)}, \mu) \\ &\geq (1 - M_1 c^t x^{(0)})^{\frac{1}{1-\sigma}} \prod_{\kappa=1}^{k} \beta_\kappa^{(0)}. \end{aligned} \tag{6.25}$$

Here we used the inequality

$$\prod_{l=0}^{i-1} (1 - M_1 c^t x^{(0)} \sigma^l) = \exp(\sum_{l=0}^{i-1} \log(1 - M_1 c^t x^{(0)} \sigma^l)) \geq (1 - M_1 c^t x^{(0)})^{\frac{1}{1-\sigma}}, \tag{6.26}$$

which is easily observed from

$$\log(1 - t) \geq t \frac{\log(1 - M_1 c^t x^{(0)})}{M_1 c^t x^{(0)}}, \quad t \in [0,\ M_1 c^t x^{(0)}]. \tag{6.27}$$

Since $\mathbf{1}_A^t \beta^{(i)} = 1$ and $\beta^{(i)} > 0$, we have $\prod_{\kappa=1}^{k} \beta_\kappa^{(i)} \leq 1$. Hence we see from (6.25) that

$$\prod_{l=0}^{i-1} G(\hat{\beta}^{(l)}, \mu) \leq M_2 \tag{6.28}$$

where

$$M_2 = \frac{1}{(1 - M_1 c^t x^{(0)})^{\frac{1}{1-\sigma}} \prod_{\kappa=1}^{k} \beta_\kappa^{(0)}}. \tag{6.29}$$

If the sequence $\{\hat{\beta}^{(i)}\}$ $(i = 1, \ldots)$ does not converge to $\mathbf{1}_A/k$, there exists an $\varepsilon > 0$ for which we have an infinite subsequence $\{\hat{\beta}^{(i_l)}\}$ $(l = 1, \ldots)$ of $\{\hat{\beta}^{(i)}\}$ such that $\|\hat{\beta}^{(i_l)} - \mathbf{1}_A/k\| \geq \varepsilon$. From the property of the function $G(\hat{\beta}, \mu)$ described before, we see that $G(\hat{\beta}^{(i)}, \mu) \geq 1$ for all i and $G(\hat{\beta}^{(i_l)}, \mu) \geq 1 + \zeta > 1$ for $\{i_l\}$, where ζ is a positive constant determined from ε. Obviously this implies that $\prod_{p=0}^{i-1} G(\hat{\beta}^{(p)}, \mu)$ diverges to infinity when $i \to \infty$. This contradicts to (6.28). Hence for any $\varepsilon > 0$, there exist only a finite number of $\hat{\beta}^{(i)}$ such that $\|\hat{\beta}^{(i)} - \mathbf{1}_A/k\| \geq \varepsilon$, which implies that the sequence converges to $\mathbf{1}_A/k$.

Now we prove the main theorem on the centering property.

Theorem 6.2. The property (6.17) holds for the sequence $\{x^{(i)}\}$, which is produced by the iteration (6.1) initiated at an interior feasible solution of $<$ D $>$ in a sufficiently small neighborhood of the optimal solution. Furthermore, the sequence of the "dual variable" $D(x^{(i)})^{-2}A^tB(x^{(i)})^{-1}c$ converges to an optimal solution of $<$ P $>$.

Proof: We introduce the new variable $z = (x - x^*)/c^t(x - x^*)$, and denote $(x^{(i)} - x^*)/c^t(x^{(i)} - x^*) = (x^{(i)} - x^*)/\alpha^{(i)t}\mathbf{1}_A$ by $z^{(i)}$. Since $A_A^t x^{(i)} - b_A = A_A^t(x^{(i)} - x^*) > 0$, we see that $z^{(i)}$ is a relative interior feasible solution of $<$ Q $>$. In terms of $z^{(i)}$, the variable $\beta^{(i)}$ is written as $\beta_\kappa^{(i)} = \alpha_\kappa^{(i)}/\alpha^{(i)t}\mathbf{1}_A = y_{A\kappa}^* a_\kappa^t(x^{(i)} - x^*)/c^t x^{(i)} = y_{A\kappa}^* a_\kappa^t z^{(i)}$. As is seen from (6.25), $\prod_\kappa \beta_\kappa^{(i)} = \prod_\kappa y_{A\kappa}^* \prod_\lambda a_\lambda^t z^{(i)}$ is bounded from below by a positive constant throughout the iteration if the initial iterate $x^{(0)}$ is taken sufficiently close to the optimal solution. This implies that there exists a positive constant δ which is a common lower bound for each sequence $\{a_\kappa^t z^{(i)}\}$ $(\kappa = 1, \ldots, k)$. Thus $\{z^{(i)}\}$ is a sequence in a compact subset of the feasible region of $<$ Q $>$, say, $\Omega \equiv \{z \in R^n | A_A^t z \geq \delta \mathbf{1}_A (> 0), c^t z = 1\}$. Let us choose an accumulation point of $z^{(i)}$, and denote it by $\hat{z}$.

In the following we show $\hat{z} = z^*$. Since each component of $A_A^t z^{(i)} = A_A^t(x^{(i)} - x^*)/c^t x^{(i)}$ is uniformly bounded from below by $\delta > 0$, the sequence $\{\|c^t x^{(i)} A_A D_A(x^{(i)})^{-1}\|\}$ is bounded from above by a positive number, say, M_3. Noting this fact, we see that $\|c - A_A\tilde{D}_A(z^{(i)})^{-1}\mathbf{1}_A/k\|$ (see the note following (6.16) for the definition of $\tilde{D}_A(z)$) is bounded from above as follows:

$$\begin{aligned}
\|c - \frac{1}{k}A_A\tilde{D}_A(z^{(i)})^{-1}\mathbf{1}_A\| &= \|c - \frac{c^t x^{(i)}}{k}A_A D_A(x^{(i)})^{-1}\mathbf{1}_A\| \\
&= \|c^t x^{(i)} A_A D_A(x^{(i)})^{-1}(D_A(x^{(i)})^{-1}A_A^t B_A^{-1}\frac{c}{c^t x^{(i)}} - \frac{\mathbf{1}_A}{k})\| \\
&\leq \|c^t x^{(i)} A_A D_A(x^{(i)})^{-1}\| \|D_A(x^{(i)})^{-1}A_A^t B_A^{-1}\frac{c}{c^t x^{(i)}} - \frac{\mathbf{1}_A}{k}\| \\
&= \|c^t x^{(i)} A_A D_A(x^{(i)})^{-1}\| \|P_A(x^{(i)})\beta^{(i)} - \frac{\mathbf{1}_A}{k}\| \\
&\leq M_3\|P_A(x^{(i)})\beta^{(i)} - \frac{\mathbf{1}_A}{k}\|.
\end{aligned} \tag{6.30}$$

Due to Lemma 6.1, the second factor of the rightmost hand of (6.30) goes to 0 as $i \to \infty$, and this implies that $\|c - A_A\tilde{D}_A(z^{(i)})^{-1}\mathbf{1}_A/k\|$ converges to 0 asymptotically. Since $A_A\tilde{D}_A(z)^{-1}\mathbf{1}_A$ is continuous on the set Ω, it is easy to see that the accumulation point $\hat{z}$ is the unique solution z^* of the equation (6.16). Thus the whole sequence $\{z^{(i)}\}$ converges to z^*.

From the preceeding discussion, it is easy to see that

$$\frac{(x^{(i)}-x^*)^t}{\|x^{(i)}-x^*\|}\frac{z^*}{\|z^*\|}=\frac{z^{(i)t}}{\|z^{(i)}\|}\frac{z^*}{\|z^*\|} \tag{6.31}$$

goes to 1 as $x^{(i)}$ converges to x^*, which proves (6.17).

We show that the sequence of the "dual variables" $\{D(x^{(i)})^{-2}A^t\ B(x^{(i)})^{-1}c\}$ converges to the optimal solution, $\tilde{y}=(\tilde{y}_A,\tilde{y}_I)\equiv(\tilde{D}_A(z^*)^{-1}\mathbf{1}_A/k,0)$, of < P >. (See (6.16) to check $\tilde{y}$ is actually an optimal solution.) In terms of P and y^*, the dual variable is written as $D(x^{(i)})^{-2}A^tB(x^{(i)})^{-1}c=D(x^{(i)})^{-1}P(x^{(i)})D(x^{(i)})y^*$. In the following we omit the argument $x^{(i)}$ in the matrix $D_A,B_A,P,P_A,\ldots$ etc. We see, from (4.23), that the dual variable is written as follows:

$$\begin{aligned}D^{-1}PDy^*&=\begin{pmatrix}D_A^{-1}&0\\0&D_I^{-1}\end{pmatrix}\begin{pmatrix}P_A+C_{11}&C_{12}\\C_{12}^t&C_{22}\end{pmatrix}\begin{pmatrix}D_Ay_A^*\\0\end{pmatrix}\\&=\begin{pmatrix}D_A^{-1}P_AD_Ay_A^*\\0\end{pmatrix}+\begin{pmatrix}-D_A^{-2}A_A^tB_A^{-1}CB_A^{-1}c\\D_I^{-1}C_{12}^tD_Ay_A^*\end{pmatrix},\end{aligned} \tag{6.32}$$

where $C(x),C_{11}(x),C_{12}(x)$ and $C_{22}(x)$ are defined in (4.17), (4.24), (4.25) and (4.26), respectively. Since

$$D_A^{-1}P_AD_Ay_A^*=c^tx^{(i)}D_A^{-1}P_A\frac{D_Ay_A^*}{\mathbf{1}_A^tD_Ay_A^*}=c^tx^{(i)}D_A^{-1}P_A\beta=\tilde{D}_A(z^{(i)})^{-1}P_A\beta \tag{6.33}$$

and $P_A\beta$ converges to $\mathbf{1}_A/k$ as i goes to infinity, we see that $D_A^{-1}P_AD_Ay_A^*\to\tilde{D}_A(z^*)^{-1}\mathbf{1}_A/k$ as $i\to\infty$, which implies that the first term in the rightmost hand of (6.32) converges to $\tilde{y}$. Now, we want to show that the second term converges to zero when $i\to\infty$. Since $\|C_{12}^t\|=O(\|D_A\|)$, we see the lower echelon $D_I^{-1}C_{12}^tD_Ay_A^*$ of the second term converges to zero as i tends to infinity. On the other hand, the norm of the upper echelon is bounded from above as follows:

$$\begin{aligned}\|D_A^{-2}A_A^tB_A^{-1}CB_A^{-1}c\|&=\|\frac{\tilde{D}_A(z^{(i)})^{-1}}{c^tx^{(i)}}D_A^{-1}A_A^tB_A^{-1}CB_A^{-1}c\|\\&\le\|\tilde{D}_A(z^{(i)})^{-1}\|\|D_A^{-1}A_A^tB_A^{-1}\|\|C\|\|\frac{B_A^{-1}c}{c^tx^{(i)}}\|\\&=\|\tilde{D}_A(z^{(i)})^{-1}\|\|D_A^{-1}A_A^tB_A^{-1}\|\|C\|\|B_A^{-1}A_AD_A^{-1}\beta\|.\end{aligned} \tag{6.34}$$

Since $\tilde{D}_A(z^{(i)})\mathbf{1}_A\to\tilde{D}_A(z^*)\mathbf{1}_A>0$, $\|\tilde{D}_A(z^{(i)})^{-1}\|$ is bounded from above when $i\to\infty$. On the other hand, we have $\|C(x)\|=O(1)$ as is shown in the proof of Lemma 4.3. From Lemma 4.2 and $\|\beta\|_1=1$, we see $\|B_A^{-1}A_AD_A^{-1}\beta\|=O(\|D_A\|)$ when $i\to\infty$. Hence the rightmost hand of (6.34) converges to zero as $i\to\infty$, which completes the proof.

7. Local Convergence Properties of Iri and Imai's Method and Yamashita's Method

In this section we analyze the asymptotic behavior of Iri and Imai's method and Yamashita's method taking note of the coefficients for d_{N} and d_{P} in the expressions (3.2) and (3.3). The followings are the two key lemmas in the discussions of this section.

Lemma 7.1. Let x be an interior point of the feasible region. Then,

$$\frac{c^t x - c^t B(x)^{-1}\eta(x)}{c^t x} \tag{7.1}$$

is a quantity of $O(\|D_A\|)$ in a sufficiently small neighborhood of the optimal solution.

Proof: Using (4.2) and (4.4), we write (7.1) as follows:

$$\frac{c^t x - c^t B^{-1}\eta}{c^t x} = \frac{y^{*t} D(I-P)\mathbf{1}}{\mathbf{1}^t D y^*} = (\beta^t\ 0)\{(I-\hat{P}) + (\hat{P} - P)\}\mathbf{1}, \tag{7.2}$$

where $\beta = D_A y_A^* / (y_A^{*t} D_A \mathbf{1}_A)$. Then, by using (5.8), Lemma 4.4 and the fact $\|\beta\|_1 = 1$, we have the desired results.

Lemma 7.2. Let x be an interior feasible solution of < D >. Then we have

$$\|B(x)^{-1}\eta(x)\| = \Theta(\|\frac{B(x)^{-1}c}{c^t x}\|) \tag{7.3}$$

in a sufficiently small neighborhood of the optimal solution x^*.

Proof: We prove this lemma by showing that $\|B^{-1}\eta\| = \Theta(\|D_A\|)$ and $\|B^{-1}c/c^t x\| = \Theta(\|D_A\|)$ hold in a sufficiently small neighborhood of x^*. From (4.18), (4.19) and (6.3), we see that the two vectors are written as follows:

$$B^{-1}\eta = B_A^{-1} A_A D_A^{-1} \mathbf{1}_A + B_A^{-1} r_2, \tag{7.4}$$

$$\frac{B^{-1}c}{c^t x} = B^{-1} A D^{-1} \begin{pmatrix} \beta \\ 0 \end{pmatrix} = B_A^{-1} A_A D_A^{-1} \beta + B_A^{-1} r_0 = \frac{B_A^{-1} c}{c^t x} + B_A^{-1} r_0, \tag{7.5}$$

where

$$r_2(x) = -C B_A^{-1} A_A D_A^{-1} \mathbf{1}_A - C B_A^{-1} A_I D_I^{-1} \mathbf{1}_I + A_I D_I^{-1} \mathbf{1}_I, \tag{7.6}$$

and r_0 is defined in the note following (6.3). Noting that $\|r_0(x)\| = O(\|D_A\|)$ and $\|r_2(x)\| = O(1)$ in a sufficiently small neighborhood of the optimal solution, we have

$$\|B_A^{-1} r_0\| = O(\|D_A\|^3) \text{ and } \|B_A^{-1} r_2\| = O(\|D_A\|^2). \tag{7.7}$$

From Remark 4.5 and (5.3), we see that

$$\|A_A^t\|^{-1}\|D_A\| \le \|B_A^{-1} A_A D_A^{-1} \mathbf{1}_A\| \le \sqrt{k}\|(A_A A_A^t)^{-1} A_A\| \|D_A\|. \tag{7.8}$$

Due to (7.4), (7.7) and (7.8), we have $\|B^{-1}\eta\| = \Theta(\|D_A\|)$. Next we deal with $B^{-1}c/c^t x$. From (6.10) we see that

$$1 \ge \frac{c^t}{c^t x} B_A^{-1} \frac{c}{c^t x} = \frac{y_A^{*t} D_A}{\mathbf{1}_A^t D_A y_A^*} P_A \frac{D_A y_A^*}{\mathbf{1}_A^t D_A y_A^*} = \beta^t P_A \beta \ge \frac{1}{k}. \tag{7.9}$$

From (7.9) and the inequality $c^t x = \mathbf{1}_A^t D_A y_A^* \ge (\min_\kappa y_{A\kappa}^*)\|D_A\|$, we have

$$\|B_A^{-1} A_A D_A^{-1} \beta\| = \|\frac{B_A^{-1} c}{c^t x}\| \ge \frac{c^t x}{\|c\|}\frac{1}{k} \ge \frac{(\min_{\kappa=1}^k y_{A\kappa}^*)\|D_A\|}{k\|c\|} \tag{7.10}$$

in the interior of the feasible region of $< \mathrm{D} >$. On the other hand, from $\|\beta\| \le 1$ and (4.9), we see that $\|B_A^{-1}c/c^tx\| = \|B_A^{-1}A_AD_A^{-1}\beta\| \le \|B_A^{-1}A_AD_A^{-1}\|\|\beta\| \le \|A_{A1}^{-1}\|\|D_A\|$ holds in the interior of the feasible region of $< \mathrm{D} >$. Thus, we obtain the inequality,

$$\frac{(\min_\kappa y^*_{A\kappa})\|D_A\|}{k\|c\|} \le \|\frac{B_A^{-1}c}{c^tx}\| = \|B_A^{-1}A_AD_A^{-1}\beta\| \le \|A_{A1}^{-1}\|\|D_A\| \tag{7.11}$$

in the interior of the feasible region. Now $\|B^{-1}c/c^tx\| = \Theta(\|D_A\|)$ follows from (7.5), (7.7) and (7.11). This completes the proof.

Now, with the help of these lemmas we analyze the asymptotic properties of Iri and Imai's method and Yamashita's method.

Theorem 7.3. The displacement vector d_{N} of Iri and Imai's method converges to $-B^{-1}\eta$ $/(m-k)$ as x approaches the optimal solution x^*.

Proof: The displacement vector of Iri and Imai's method is written as follows:

$$\begin{aligned} d_{\mathrm{N}}(x) &= \rho_0(x)(-\rho_1(x)B(x)^{-1}c + B(x)^{-1}\eta(x)), \\ \rho_0(x) &= \frac{1}{1-g(x)^t(-\rho_1(x)B(x)^{-1}c + B(x)^{-1}\eta(x))}, \\ \rho_1(x) &= \frac{c^tx - c^tB(x)^{-1}\eta(x)}{\dfrac{(c^tx)^2}{m+1} - c^tB(x)^{-1}c}, \\ g(x) &= (m+1)\frac{c}{c^tx} - \eta(x). \end{aligned} \tag{3.2$'$}$$

We first observe that the direction of d_{N} converges to that of $B^{-1}\eta$. Since $\|B^{-1}c/c^tx\| = \Theta(\|B^{-1}\eta\|)$ near the optimal solution, it is enough to show ρ_1c^tx tends to zero as x converges to x^*. The factor ρ_1c^tx is written as follows:

$$\rho_1c^tx = c^tx\cdot\frac{c^tx - c^tB^{-1}\eta}{\dfrac{(c^tx)^2}{m+1} - c^tB^{-1}c} = \frac{c^tx - c^tB^{-1}\eta}{c^tx}\frac{1}{\dfrac{1}{m+1} - \dfrac{c^tB^{-1}c}{(c^tx)^2}}. \tag{7.12}$$

The second factor of the right hand side is $O(1)$ in a sufficiently small neighborhood of the optimal solution, because, from Lemma 4.4 and (6.10), we have

$$\begin{aligned} 1 \ge \beta^tP\beta &= \frac{c^tB^{-1}c}{(c^tx)^2} = (\beta^t\ 0)\hat{P}\begin{pmatrix}\beta\\0\end{pmatrix} + (\beta^t\ 0)(P-\hat{P})\begin{pmatrix}\beta\\0\end{pmatrix} \\ &\ge \beta^tP_A\beta - \|P-\hat{P}\| \ge \frac{1}{k} - O(\|D_A\|) > \frac{1}{m+1}. \end{aligned} \tag{7.13}$$

Then, from Lemma 7.1 and (7.13), it is easy to see that the order of (7.12) is $O(\|D_A\|)$ near the optimal solution. This implies that the displacement vector of Iri and Imai's method converges to $B^{-1}\eta$ in its direction.

To complete the proof of the theorem we show that the coefficient ρ_0 of $B^{-1}\eta$ converges to $-1/(m-k)$. We give the limit value of the quantities $g^tB^{-1}\eta$ and $g^tB^{-1}c$, which appear in the denominator of $\rho_0(x)$, as $x \to x^*$. It follows from Lemma 4.4 and the fact $P_A\mathbf{1}_A = \mathbf{1}_A$ that $\eta^tB^{-1}\eta = \mathbf{1}^tP\mathbf{1}$ converges to $\mathbf{1}_A^tP_A\mathbf{1}_A = k$. Due to Lemma

7.1, we see that $c^tB^{-1}\eta/c^tx \to 1$. Hence $g^tB^{-1}\eta$ converges to $m+1-k$. On the other hand, since $\rho_1 c^tx$ converges to 0, it follows from (7.13) and $c^tB^{-1}\eta/c^tx \to 1$ that $\rho_1 g^tB^{-1}c = (\rho_1 c^tx)g^tB^{-1}c/c^tx = (\rho_1 c^tx)\{(m+1)c^tB^{-1}c/(c^tx)^2 - \eta^tB^{-1}c/c^tx\}$ also converges to 0. Therefore, ρ_0 converges to $-1/(m-k)$ as x approaches x^*, which completes the proof.

Now, we give a proof for the quadratic convergence property of Iri and Imai's method. The next theorem was proved originally by Iri and Imai in the case where both primal and dual are nondegenerate at the optimal solution, but, our result is new in that it requires no condition on degeneracy of constraints.

Theorem 7.4. If the optimal solution is unique, i.e., the objective function is nondegenerate, Iri and Imai's method (with an exact line search) converges quadratically to the optimal solution.

Proof: Let the current point be $x^\dagger$, which is an interior feasible solution. We put $\tilde{d}_N(x^\dagger) = |\rho_0(x^\dagger)|^{-1}d_N(x^\dagger) = \text{sign}(\rho_0(x^\dagger))(-\rho_1(x^\dagger)B(x^\dagger)^{-1}c + B(x^\dagger)^{-1}\eta(x^\dagger))$, and $x(t) = x^\dagger + t\tilde{d}_N$. Since the coefficient $\rho_0(x^\dagger)$ converges to $-1/(m-k)$ as $x^\dagger$ does to x^* (cf. Theorem 7.3), we see that $\text{sign}(\rho_0(x^\dagger)) = -1$ in a sufficiently small neighborhood of the optimal solution. As was shown in the proof of Lemma 7.2, we have $\rho_1(x)c^tx = O(\|D_A(x)\|)$. We see, from Lemma 4.4,

$$\|(P(x) - \hat{P}(x))\mathbf{1}\| = O(\|D_A(x)\|). \tag{7.14}$$

Since $\text{sign}(\rho_0(x^\dagger)) = -1$ near x^* and $P_A(x)\mathbf{1}_A = \mathbf{1}_A$, the vector $D(x(t))\mathbf{1}$ is written as follows if $x^\dagger$ is in the neighborhood of the optimal solution:

$$\begin{aligned}
\begin{pmatrix} D_A(x(t))\mathbf{1}_A \\ D_I(x(t))\mathbf{1}_I \end{pmatrix} &= D(x^\dagger)\mathbf{1} + tA^t\tilde{d}_N(x^\dagger) \\
&= D(x^\dagger)\mathbf{1} - t\{-\rho_1(x^\dagger)c^tx^\dagger A^t\frac{B^{-1}(x^\dagger)c}{c^tx^\dagger} + A^tB^{-1}(x^\dagger)\eta(x^\dagger)\} \\
&= D(x^\dagger)\mathbf{1} - tD(x^\dagger)\{-\rho_1(x^\dagger)c^tx^\dagger P(x^\dagger)\begin{pmatrix}\beta(x^\dagger)\\ 0\end{pmatrix} + P(x^\dagger)\mathbf{1}\} \\
&= D(x^\dagger)(\mathbf{1} - t\hat{P}(x^\dagger)\mathbf{1}) \\
&\quad - tD(x^\dagger)\{-\rho_1(x^\dagger)c^tx^\dagger P(x^\dagger)\begin{pmatrix}\beta(x^\dagger)\\ 0\end{pmatrix} + (P(x^\dagger) - \hat{P}(x^\dagger))\mathbf{1}\} \\
&= \begin{pmatrix} D_A(x^\dagger)\mathbf{1}_A \\ D_I(x^\dagger)\mathbf{1}_I \end{pmatrix} - t\begin{pmatrix} D_A(x^\dagger)(\mathbf{1}_A + r_{3A}(x^\dagger)) \\ D_I(x^\dagger)r_{3I}(x^\dagger) \end{pmatrix},
\end{aligned} \tag{7.15}$$

where $\hat{P}$ is defined in (4.20), $r_3(x) = -\rho_1 c^txP(x)(\beta(x)^t\ 0)^t + (P(x) - \hat{P}(x))\mathbf{1} \in R^m$ and $\|r_3(x)\| = O(\|D_A(x)\|)$, which follows from Lemma 4.4 and the fact that $\rho_1(x)c^tx = O(\|D_A\|)$.

Substituting (7.15) into $c^tx = y^{*t}D(x)\mathbf{1} = (y_A^{*t}\ 0)D(x)\mathbf{1}$, we have

$$\frac{c^tx(t)}{c^tx^\dagger} = \frac{y_A^{*t}D_A(x(t))\mathbf{1}_A}{y_A^{*t}D_A(x^\dagger)\mathbf{1}_A} = 1 - (1 + c_0(x^\dagger))t \tag{7.16}$$

where $c_0(x) = \beta(x)^tr_{3A}(x) = O(\|D_A\|)$.

As was proved in [3], Multiplicative Barrier Function $F(x) = (c^t x)^{m+1}/\prod_{\kappa=1}^{m}(a_\kappa^t x - b_\kappa)$ is a strictly convex function in the interior of the feasible region. We see, from the proof of Proposition 3.1 of [3], that $F(x(t))$ diverges to infinity either if $x(t)$ approaches any point on the boundary of the feasible region except for x^*, or if $x(t)$ goes infinity along an infinite lay of the feasible region. Then, when regarded as a function of t, $F(x(t))$ has a unique minimum in the interior of the feasible region, otherwise, the optimal solution x^* is on the half line $x(t)$ $(t > 0)$. We exclude the second case, as it is trivial. Let t^* be the step-size obtained by performing an exact line search procedure. From (7.15) and (7.16) we see that t^* is a solution for the equality

$$\begin{aligned}\frac{1}{F(x(t))}\frac{dF(x(t))}{dt} &= \frac{d}{dt}\log F(x(t)) = \frac{d}{dt}\log\frac{F(x(t))}{F(x(0))} \\ &= \frac{d}{dt}\{(m+1)\log\frac{c^t x(t)}{c^t x^\dagger} - \log(\frac{a_\kappa^t x(t) - b_\kappa}{a_\kappa^t x^\dagger - b_\kappa})\} \\ &= \frac{d}{dt}\{(m+1)\log(1-(1+c_0(x^\dagger))t) \\ &\quad - \sum_{\kappa=1}^{k}\log(1-(1+r_{3\kappa}(x^\dagger))t) - \sum_{\tau=k+1}^{m}\log(1-r_{3\tau}(x^\dagger)t)\} \\ &= -(m+1)\frac{1+c_0(x^\dagger)}{1-(1+c_0(x^\dagger))t} \\ &\quad + \sum_{\kappa=1}^{k}\frac{1+r_{3\kappa}(x^\dagger)}{1-(1+r_{3\kappa}(x^\dagger))t} + \sum_{\tau=k+1}^{m}\frac{r_{3\tau}(x^\dagger)}{1-r_{3\tau}(x^\dagger)t} = 0. \end{aligned} \tag{7.17}$$

The solution t^* is strictly positive and unique.

Now, consider the largest step-size $t_{\max}$ which does not violate the feasibility. Then $x(t_{\max})$ is on the boundary of the feasible region. Obviously, we have $t^* < t_{\max}$. From (7.15) we see that $t_{\max} \le 1 + O(\|D_A(x^\dagger)\|)$ and $\|D_A(x(t_{\max}))\| = O(\|D_A(x^\dagger)\|^2)$ if $x^\dagger$ is in a sufficiently small neighborhood of the optimal solution. Due to Remark 4.5, this implies that $\|x(t_{\max}) - x^*\| = O(\|x^\dagger - x^*\|^2)$. In the following we demonstrate that $t^*(< t_{\max})$ is large enough to guarantee the quadratic convergence of $x(t^*)$ if $x^\dagger$ is an interior feasible solution of < D > sufficiently close to x^*. To this end, we find a step-size $\tilde{t} \le t^*$ such that

$$\|x^* - x(\tilde{t})\| \le M_1\|x^* - x^\dagger\|^2 \tag{7.18}$$

in a sufficiently small neighborhood of x^*, where M_1 is an appropriate positive constant. Since $\tilde{t} \le t^* < t_{\max}$, the existence of such $\tilde{t}$ immediately implies the quadratic convergence of Iri and Imai's method with an exact line search procedure.

Since $\|r_3(x)\|$ and $|c_0(x)|$ are the quantities of order, $O(\|D_A\|)(= O(c^t x))$, there exists a positive constant M_2 such that $\|r_3(x)\| \le M_2 c^t x$, $|c_0(x)| \le M_2 c^t x$. Let us denote by I the interval $[0, (1 + M_2 c^t x^\dagger)^{-1})$. If $x^\dagger$ is in a sufficiently small neighborhood of the optimal solution such that $M_2 c^t x^\dagger < 1$, $x(t)$ remains to be an interior feasible solution for all $t \in I$ and then $d\log F(x(t))/dt$ is bounded from above as follows:

$$\frac{d}{dt}\log F(x(t)) = -(m+1)\frac{1+c_0(x^\dagger)}{1-(1+c_0(x^\dagger))t} + \sum_{\kappa=1}^{k}\frac{1+r_{3\kappa}(x^\dagger)}{1-(1+r_{3\kappa}(x^\dagger))t} + \sum_{\tau=k+1}^{m}\frac{r_{3\tau}(x^\dagger)}{1-r_{3\tau}(x^\dagger)t}$$

$$
\begin{aligned}
&\leq -(m+1)\frac{1-M_2c^tx^\dagger}{1-(1-M_2c^tx^\dagger)t}+\sum_{\kappa=1}^{k}\frac{1+M_2c^tx^\dagger}{1-(1+M_2c^tx^\dagger)t}+\sum_{\tau=k+1}^{m}\frac{M_2c^tx^\dagger}{1-tM_2c^tx^\dagger}\\
&\leq -(m+1)\frac{1-M_2c^tx^\dagger}{1-(1-M_2c^tx^\dagger)t}+k\frac{1+M_2c^tx^\dagger}{1-(1+M_2c^tx^\dagger)t}+M_3c^tx^\dagger\\
&\leq -(m+1)\frac{1-M_2c^tx^\dagger}{1-(1-M_2c^tx^\dagger)t}+k\frac{1+M_2c^tx^\dagger}{1-(1+M_2c^tx^\dagger)t}\\
&\quad+\frac{M_3c^tx^\dagger}{\{1-(1-M_2c^tx^\dagger)t\}\{1-(1+M_2c^tx^\dagger)t\}}\\
&\equiv l(t),
\end{aligned}
\tag{7.19}
$$

where M_3 is an appropriate positive constant, say, $2(m-k)M_2$. Let $\hat{t}$ be a point such that $l(\hat{t})=0$. Since $d\log F(t)/dt \leq l(t)$ in the interval I, if $\hat{t}\in I$, then $x(\hat{t})$ is an interior feasible solution at which $d\log F(\hat{t})/dt \leq l(\hat{t})=0$ holds. On the other hand, since $d\log F(x(0))/dt<0$, the uniqueness of the solution t^* of (7.17) implies that $\hat{t}\leq t^*$ and $d\log F(x(t))/dt<0$ for any $0\leq t<\hat{t}$. Solving the equation $l(\hat{t})=0$, we obtain the following expression for $\hat{t}$:

$$
\hat{t}=\frac{1}{1-(M_2c^tx^\dagger)^2}\tilde{t},
\tag{7.20}
$$

where

$$
\tilde{t}=1-M_4c^tx^\dagger, M_4\equiv\frac{m+1+k}{m+1-k}M_2+\frac{M_3}{m+1-k}.
\tag{7.21}
$$

We have the inequality $0<\tilde{t}<\hat{t}<(1+M_2c^tx^\dagger)^{-1}$, and hence, $x(\tilde{t})$ is an interior point of the feasible region such that $d\log F(x(\tilde{t}))/dt<0$. Due to (7.15) and Remark 4.5, we also have

$$
\begin{aligned}
&\|x(\tilde{t})-x^*\|\leq M_5\|D_A(x(\tilde{t}))\mathbf{1}_A\|\\
&=M_5\|D_A(x^\dagger)\mathbf{1}_A-(1-M_4c^tx^\dagger)D_A(x^\dagger)\mathbf{1}_A-(1-M_4c^tx^\dagger)D_A(x^\dagger)r_{3A}\|\\
&\leq M_6\|D_A(x^\dagger)\|^2\leq M_7\|x^\dagger-x^*\|^2,
\end{aligned}
\tag{7.22}
$$

where M_5, M_6, M_7 are appropriate constants. Thus, the step-size $\tilde{t}$ guarantees the interior feasibililty and the quadratic convergence.

Now, since $\tilde{t}<t^*<t_{\max}$ and $\|x(t_{\max})-x^*\|=O(\|x^\dagger-x^*\|^2)$, we have $\|x(t^*)-x^*\|=O(\|x^\dagger-x^*\|^2)$, which implies the quadratic convergence of Iri and Imai's method with the exact line search procedure.

Before concluding this section, we show that the searh direction of Yamashita's method approaches asymptotically to that of d_{AF}, making a good contrast with the case of Iri and Imai's method (cf. Theorem 7.3). We note that the same property also holds with Karmarkar's method [4], which is a projective scaling method.

Theorem 7.5. The search direction of Yamashita's method converges to the direction of $-B^{-1}c$ as x approachesthe optimal solution x^*. Specifically, the displacement vector of

Yamashita's method converges to $-(c^t x/c^t B^{-1}c)B^{-1}c$.

Proof: The displacement vector of Yamashita's method is as follows:

$$
\begin{aligned}
d_{\mathrm{P}}(x) &= \xi_0(x)(-B(x)^{-1}c + \xi_1(x)B(x)^{-1}\eta(x)), \\
\xi_0(x) &= \frac{c^t x}{\xi_2(x)}, \ \xi_1(x) = \xi_3(x)(c^t x - c^t B(x)^{-1}\eta(x)), \\
\xi_2(x) &= (c^t x - c^t B(x)^{-1}\eta(x))^2 \xi_3(x) + c^t B(x)^{-1}c, \\
\xi_3(x) &= \frac{1}{m+1-\eta(x)^t B(x)^{-1}\eta(x)}.
\end{aligned} \tag{3.3)'}
$$

We will show that

$$\frac{\xi_1}{c^t x} = \xi_3 \frac{c^t x - c^t B^{-1}\eta}{c^t x} \tag{7.23}$$

converges to 0 as x does to x^*. Then, since $\|B^{-1}\eta\|$ and $\|B^{-1}c/c^t x\|$ are quantities of the same order, the direction of Yamashita's method is shown to approach that of $-B^{-1}c$. It is easily seen that $\xi_3 = (m+1-\eta^t B^{-1}\eta)^{-1}$ is a quantity of order 1, because $\eta^t B^{-1}\eta = \mathbf{1}^t P \mathbf{1} \leq m$. Then it follows from Lemma 7.1 that $\xi_1/c^t x$ converges to zero. Hence, in Yamashita's method, the component of $-B^{-1}\eta$ decreases to zero and $-B^{-1}c$ dominates asymptotically in the search direction, making a good contrast with the case of Iri and Imai's method.

Now we give the asymptotic formula for the coefficient $\xi_0 = c^t x/\xi_2$. Due to (7.13) and Lemma 7.1, it follows that $c^t B^{-1}c$ is the order of $\|D_A\|^2$ whereas $(c^t x - c^t B^{-1}\eta)^2$ is the order of $\|D_A\|^4$. Then the asymptotically dominant term in $\xi_2 = \xi_3(c^t x - c^t B^{-1}\eta)^2 + c^t B^{-1}c$ is $c^t B^{-1}c$, and hence, ξ_0 converges to $c^t x/c^t B^{-1}c$. Consequently, we can see that $-(c^t x/c^t B^{-1}c)B^{-1}c$ is the asymptotic displacement vector of Yamashita's method.

8.Conclusion

We have studied the local convergence properties of the iterative formulae related to some of the new interior point methods for linear programming. In particular, the iteration by the search direction of the dual affine scaling method d_{AF} and the iteration by the negative centering direction $-d_{\mathrm{C}}$ were analyzed in detail under the assumption of the uniqueness of the optimal solution. The quadratic convergence property of $-d_{\mathrm{C}}$ was demonstrated as well as the linear convergence property and the centering property of the dual affine scaling method. Furthermore, the local convergence properties of Iri and Imai's method and Yamashita's method were studied in detail. It was shown that the search direction of Iri and Imai's method approaches the negative centering direction, while the search direction of Yamashita's method approaches the direction of the dual affine scaling method. The quadratic convergence of Iri and Imai's method was also demonstrated.

There still remain several questions unanswered on local and global behavior of the interior point methods applied to degenerate problems. Their difficulties seem to arise from degenerate faces. In this paper we give a way to analyze the behavior of the methods near degenerate vertices. Developing the viewpoint and the techniques presented here, we can obtain further convergence results on the interior point methods for degenerate problems [9, 10].

Acknowledgment

The authors would like to thank Mr. Takashi Sasakawa of University of Tokyo (presently, Railway Technical Research Institute) for his valuable discussions. They are grateful to Mr. Fumiyasu Komaki of the Graduate University of Advanced Studies for careful reading of the manuscript. Thanks are also due to the referees for several helpful comments. In particular, the proof of Lemma A.1 was considerably simplified by a suggestion of one of the referees. This work was supported in part by a Grant-in-Aid for Encouragement of Young Scientists 63740128(1988) from the Ministry of Education, Science and Culture of Japan.

Reference

[1] I. Adler, N. Karmarkar, M. G. C. Resende, G. Veiga: An Implementation of Karmarkar's Algorithm for Linear Programming. *Mathematical Programming*, Vol. 44 (1989), pp. 297-335.

[2] D. A. Bayer and J. C. Lagarias: The Nonlinear Geometry of Linear Programming: II. Legendre Transform Coordinates and Central Trajectories. *Transactions of the American Mathematical Society*, Vol. 314, No. 2 (1989), pp. 527-581.

[3] M. Iri and H. Imai: A Multiplicative Barrier Function Method for Linear Programming, *Algorithmica* , Vol. 1 (1986), pp. 455-482.

[4] N. Karmarkar: A New Polynomial-Time Algorithm for Linear Programming, *Combinatorica* , Vol. 4, No. 4 (1984), pp. 373-395.

[5] N. Megiddo and M. Shub: Boundary Behavior of Interior Point Algorithms for Linear Programming. *Mathematics of Operations Research*, Vol. 14 (1989), pp. 97-146.

[6] J. Renegar: A Polynomial-Time Algorithm, Based on Newton's Method, for Linear Programming, *Mathematical Programming*, Vol. 40 (1988), pp. 59-93.

[7] Gy. Sonnevend: An "Analytical Centre" for Polyhedrons and New Classes of Global Algorithms for Linear (Smooth, Convex) Programming, *System Modelling and Optimization, Proceedings of the 12th IFIP Conference* (Budapest, Hungary, September 2-6, 1985) , eds. A. Prekopa, J. Szelezsan, and B. Strazicky, Springer-Verlag, Berlin, 1986, pp. 866-875.

[8] T. Tsuchiya: Dual Standard Form Linear Programming Problems and Karmarkar's Canonical Form (in Japanese). *Proceedings of the Research Institute of Mathematical Sciences*, Vol. 676 (1988), pp. 330-336.

[9] T. Tsuchiya: Local Analysis of Iri and Imai's Method for Degenerate Linear Programming Problems. (In preparation.)

[10] T. Tsuchiya: Global Convergence Properties of the Affine Scaling Methods for Primal Degenerate Linear Programming Problems. *Research Memorandom of the Institute of Statistical Mathematics,* No. 367 (Minato-ku, Tokyo, Japan, 1989). (To appear in *Mathematics of Operations Research.*)

[11] H. Yamashita: A Polynomially and Quadratically Convergent Method for Linear Programming, Mathematical Systems Institute Inc., *Technical Report* (Shinjuku-ku, Tokyo, Japan, 1986).

Appendix

We are concerned with the minimum point of the function

$$G(\beta,\mu)=\frac{\prod_{\kappa=1}^{k}(1-\mu\beta_\kappa)}{(1-\mu\sum_{\kappa=1}^{k}\beta_\kappa^2)^k} \tag{A.1}$$

on the set

$$T=\{\beta\in R^k|\beta^t\mathbf{1}_A=1,\|\beta\|\leq 1\}, \tag{A.2}$$

where μ is fixed to be a constant in the interval $(0,1)$. We prove the following lemma.

Lemma A. 1. The function $G(\beta,\mu)$ has the unique minimum at the point $\beta = \mathbf{1}_A/k$ on the set T.

Proof: We deal with the function $\log G(\beta,\mu)$ and find its critical point on T. The Lagrangian function for this problem is

$$\sum_{\kappa=1}^{k}\log(1-\mu\beta_\kappa) - k\log(1-\mu\sum_{\kappa=1}^{k}\beta_\kappa^2) - \lambda(\sum_{\kappa=1}^{k}\beta_\kappa - 1) \tag{A.3}$$

and a critical point $\beta = \beta^\dagger$ satisfies the equations,

$$-\frac{\mu}{1-\mu\beta_\kappa} + \frac{2k\mu\beta_\kappa}{1-\mu\sum_{\tau=1}^{k}\beta_\tau^2} - \lambda = 0 \quad (\kappa = 1,\dots,k). \tag{A.4}$$

Multiplying both sides of the equation by $(1-\mu\beta_\kappa)$ and taking the summation, we have

$$-k\mu + 2k\mu\frac{\sum_{\kappa=1}^{k}\beta_\kappa(1-\mu\beta_\kappa)}{1-\mu\sum_{\tau=1}^{k}\beta_\tau^2} - \lambda\sum_{\sigma=1}^{k}(1-\mu\beta_\sigma) = k\mu - (k-\mu)\lambda = 0, \tag{A.5}$$

since $\sum_{\kappa=1}^{k}\beta_\kappa = 1$. Thus we have $\lambda = k\mu/(k-\mu)$. Then the relation (A.4) implies $\beta > 0$. Substituting the expression of λ into (A.4) and dividing the κ-th equation by $\mu\beta_\kappa$, we obtain the following equations:

$$\gamma \equiv \frac{2k}{1-\mu u} = \frac{1}{\beta_\kappa(1-\mu\beta_\kappa)} + \frac{\delta}{\beta_\kappa} \quad (\kappa = 1,\dots,k), \tag{A.6}$$

where

$$u = \sum_{\kappa=1}^{k}\beta_\kappa^2, \quad \delta = \frac{k}{k-\mu}. \tag{A.7}$$

Note that $\delta > 0$. A critical point $\beta^\dagger$ is a solution of the equation (A.6). In order to show $\beta^\dagger = \mathbf{1}_A/k$, let us take a pair of indices $\{\sigma,\tau\}$ where $\sigma \neq \tau$. Subtracting the τ-th equation from the σ-th equation of (A.6), we have

$$(\beta_\sigma^\dagger - \beta_\tau^\dagger)\{(1-\mu(\beta_\sigma^\dagger + \beta_\tau^\dagger)) + \delta(1-\mu\beta_\sigma^\dagger)(1-\mu\beta_\tau^\dagger)\} = 0. \tag{A.8}$$

We already observed $\beta^\dagger > 0$ and $\delta > 0$. Since $0 < \mu < 1$ and $\beta^{\dagger t}\mathbf{1}_A = 1$, the second factor of the left hand side of (A.8) is strictly greater than zero, which implies $\beta_\sigma^\dagger = \beta_\tau^\dagger$. Thus, the point $\beta^\dagger = \mathbf{1}_A/k$ is the unique critical point of G on T, at which $G(\mathbf{1}_A/k) = 1$.

Since G is a continuous function on the compact set T, it has the minimum point on T, which is a critical point or a point on the boundary of T. In the following we give a lower bound for the value of G on the boundary of T. Since the inequality $\min_{\kappa=1}^{k}|\beta_\kappa| \le k^{-1/2}\|\beta\|$ holds, $G(\beta,\mu)$ is bounded from below as follows:

$$G(\beta,\mu) = \frac{\prod_{\kappa=1}^{k}(1-\mu\beta_\kappa)}{(1-\mu\sum_{\kappa=1}^{k}\beta_\kappa^2)^k} \ge \frac{(1-\mu\|\beta\|)^{k-1}(1-\mu k^{-1/2}\|\beta\|)}{(1-\mu\|\beta\|^2)^k}. \tag{A.9}$$

Therefore, when $\|\beta\| = 1$, we have, for $\mu \in (0,1)$,

$$G(\beta,\mu) \ge \frac{1-\mu k^{-1/2}}{1-\mu} > 1. \tag{A.10}$$

This shows that the value of G on the boundary of T is greater than 1, and $\beta = \mathbf{1}_A/k$ is the unique minimum point of G on T.

Takashi TSUCHIYA:
The Institute of Statistical Mathematics
4-6-7 Minami-Azabu, Minato-ku
Tokyo 106, Japan
(E-mail: tsuchiya@sun312.ism.ac.jp)