JOURNAL OF COMPLEXITY 3, 258–269 (1987)

On the Asymptotic Behavior of the Projective Rescaling Algorithm for Linear Programming

MICHAEL SHUB*

Mathematical Sciences Department, IBM Thomas Watson Research Center, Yorktown Heights, New York 10598

Received April 20, 1987

INTRODUCTION

This paper extends some of the results of Megiddo and Shub (1986) to include the case of the projective rescaling vector field and a discrete version which is one of Karmarkar's (1984) algorithms. First it is shown that under nondegeneracy conditions every interior orbit of the projective rescaling vector field is tangent to the inverse of the reduced cost vector at the optimal vertex. This is accomplished by showing that for a nondegenerate problem in Karmarkar standard form, the linear and projective rescaling vector fields agree through quadratic terms; then the results of Megiddo and Shub (1986) apply. Using the quadratic expression for a nondegenerate problem in Karmarkar standard form, the asymptotic rate of approach of the discrete algorithm to the optimum is shown to be $1 - \alpha\gamma$ for all starting points in a cone around the central trajectory and near the optimum. Here,

$$0 < \alpha \leq \frac{1}{\sqrt{m((m-n)/n) + \gamma(n-m)/n)}},$$

where

$$\begin{pmatrix} A \\ e^{\mathrm{T}} \end{pmatrix} x = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \in R^{m-1} \times R = R^{m},$$

* Partially supported by a National Science Foundation Research Grant.

0885-064X/87 $3.00

$\binom{A}{e^{\mathrm{T}}}$ an $m \times n$ matrix and $x \geq 0$ define the polytope, and γ is a fixed constant. Thus the domains where the asymptotic rates of approach have been determined for Karmarkar (1984), Barnes (1985), and Renegar (1986) are essentially the same and the best possible rate for Karmarker is also essentially the same as the proven rates for the other two. See also Megiddo and Shub (1986). Finally, we end the paper with some open problems concerning Karmarkar's algorithm.

Jeff Lagarias, in work in progress which was reported on at Columbia University, has a result which also implies that all orbits of the projective rescaling vector field are tangent at the optimum.

1. The Asymptotics of the Projective Rescaling Vector Field

For $x \in \mathbb{R}^n$, let $D = D_x$ be the diagonal matrix with entries $x_1 \ldots x_n$. For a matrix M, let P_M denote the orthogonal projection on the null space of M. If M is $m \times n$, M: $\mathbb{R}^n \to R^m$ and the null space of $M = \{x \in \mathbb{R}^n \mid Mx = 0\}$.

Karmarkar's standard form for a linear programming problem is

$$\begin{aligned} \text{Minimize} \quad & c^{\mathrm{T}}x \\ \text{subject to} \quad & Ax = 0 \\ & e^{\mathrm{T}}x = 1 \\ & x \geq 0, \end{aligned}$$

where A is an $(m\text{-}1) \times n$ matrix $(1 \leq m \leq n)$; $x, c \in \mathbb{R}^n$; and $e = (1 \ldots 1)^{\mathrm{T}} \in \mathbb{R}^n$. Moreover, it is assumed that the minimum of $c^{\mathrm{T}}x$ over the polytope $S = \{x \in \mathbb{R}^n \mid Ax = 0,\ e^{\mathrm{T}}x = 1,\ x \geq 0\}$ is 0. We will not use this last hypothesis before Theorem 1.

Let $\dot{S}$ denote the interior of the polytope, $\dot{S} = \{x \in \mathbb{R}^n \mid Ax = 0, e^{\mathrm{T}}x = 1, x > 0\}$, and let $\tilde{A} = \binom{A}{e^{\mathrm{T}}}$. The linear, projective, μ-barrier method vector fields for μ a real parameter and τ-vector field are

$$\begin{aligned} \xi_l &= DP_{\tilde{A}D}Dc \\ \xi_p &= (D - xx^{\mathrm{T}})P_{AD}Dc \\ V_\mu &= DP_{\tilde{A}D}(Dc - \mu e) \\ \tau &= P_{AD}Dc. \end{aligned}$$

They are defined in $\dot{S}$. Megiddo and Shub (1986) prove that the vector fields ξ_l, ξ_p, V_μ, and $D\tau$ extend differentially to all of S, the extensions of

ξ_l, ξ_p, and V_μ are tangent to any face in which the point of the polytope lies. Moreover, with the nondegeneracy condition that for any m columns $\tilde{A}_{i_1} \ldots \tilde{A}_{i_m}$ of $\tilde{A}$, the $m \times m$ matrix $(\tilde{A}_{i_1}\tilde{A}_{i_2} \ldots \tilde{A}_{i_m})$ is invertible, then the vector fields ξ_l, ξ_p, V_μ, and $\tau(x)$ extend real analytically to all of S.

DEFINITION. For $x \in S$, let $\mu(x) = (x^{\mathrm{T}}\tau(x)) = e^{\mathrm{T}}(D_x\tau(x))$.

PROPOSITION 1. *Let*

$$\begin{aligned} \textit{minimize} \quad & c^{\mathrm{T}}x \\ \textit{subject to} \quad & Ax = 0 \\ & e^{\mathrm{T}}x = 1 \\ & x \geq 0 \end{aligned}$$

be a linear programming problem. Let v be a vertex of P, and suppose that $c^{\mathrm{T}}v = 0$. Then the derivatives of the vector field $D_x\tau(x)$ and the function $\mu(x)$ are both zero at v,

$$\mu'(v) = 0, \qquad (D_x\tau)'(v) = 0.$$

Proof. Since $\mu(x) = e^{\mathrm{T}}(D_x\tau(x))$ it is sufficient to prove that $(D_x\tau)'(v) = 0$. Now we establish some notation.

(i) Let I_1 be the set of indices i such that $v_i > 0$ and I_2 the set of indices i such that $v_i = 0$.

(ii) Let R_i, $i = 1,2$, be defined by

$$R_i = \{x \in \mathbb{R}^n \mid x_j = 0 \text{ for } j \notin I_i\}.$$

(iii) Let $E \equiv$ null space A. Let $E_x \equiv D_x^{-1}E =$ null space AD_x.

(iv) Let $E_1 \equiv E \cap R_1$. Since v is a vertex of P, this implies that $E_1 \cap \{y \in R_1 \mid y > 0 \text{ and } \Sigma y_i = 1\} = \{v\}$ and thus E_1 has dimension one and consists of multiples of v.

(v) Let F_x be the orthogonal complement of $D_x^{-1}E_1$ in E_x.

(vi) Let η_{F_x} be the orthogonal projection of Dc onto F_x. Let $\eta_{E_{1x}}$ be the orthogonal projection of Dc on $D_x^{-1}E_1$. Thus $\eta_{E_{1x}}$ can also be expressed as the orthogonal projection of Dc on $D^{-1}v$. That is,

$$\eta_{E_{1x}} = \frac{(Dc)^{\mathrm{T}}D^{-1}v}{(D^{-1}v)^{\mathrm{T}}D^{-1}v} = \frac{(c^{\mathrm{T}} \cdot v)}{(D^{-1}v)^{\mathrm{T}}D^{-1}v} D^{-1}v = 0.$$

Now $\tau(x) = \eta_{F_x} + \eta_{E_{1x}}$. By Megiddo and Shub (1986, Appendix F) the derivative $(D_x\eta_{F_x})'(v) = 0$. Thus $(D_x\tau)'(v) = (D_x\eta_{E_{1x}})'v = 0$ by (vi).

Q.E.D.

As in Megiddo and Shub (1986), we use the comparison between V_μ and ξ_p first observed by Gill *et al.* (1985).

PROPOSITION 2 (Gill *et al.*, 1985). *The vector fields* V_o *and* ξ_l *are equal. The vector field* ξ_p *may be expressed as* $\xi_p(x) = V_{\mu(x)}(x)$.

Proof. The first statement is immediate from the definitions. The second follows quickly by a short computation which we repeat here:

$$V_{\mu(x)}(x) = D_x\Big(P_{AD}(Dc) - \frac{(Dc)^{\mathrm{T}}P_{AD}x}{(P_{AD}x)^{\mathrm{T}}P_{AD}x} P_{AD}x - \mu(x)P_{AD}e + \mu(x)\frac{e^{\mathrm{T}}P_{AD}x}{(P_{AD}x)^{\mathrm{T}}P_{AD}x} P_{AD}x\Big).$$

Now P_{AD} is symmetric and $e \in$ null space AD; thus $(Dc)^{\mathrm{T}}P_{AD}x = (P_{AD}Dc)^{\mathrm{T}}x = (\tau(x)^{\mathrm{T}}x) = \mu(x)$ and $e^{\mathrm{T}}P_{AD}x = P_{AD}e^{\mathrm{T}}x = e^{\mathrm{T}}x = 1$. Hence,

$$V_{\mu(x)}(x) = D_xP_{AD}(Dc) - \mu(x)e) = D_xP_{AD}DC - (x^{\mathrm{T}}P_{AD}Dc)x = \xi_p(x).$$

Q.E.D.

It is interesting to note here that the function $\mu(x)$ may take negative values. Nguyen Hoan (personal communication) has constructed such examples. While the μ-barrier method is generally defined only for positive μ the vector field V_μ makes sense for all μ.

The vector fields ξ_p, ξ_l, V_μ, and $D_x\tau$ are generically real analytic, but they are only proven to be differentiable in all cases. It is an open problem as to whether they are twice differentiable under all possible degeneracies. In the next proposition we assume that they are twice differentiable.

PROPOSITION 3. *Suppose that the vector fields* ξ_p, V_μ, $\mu \in \mathbb{R}$, *and* $D_x\tau$ *are twice differentiable. Let* v *be a vertex of the polytope* S, *and suppose* $e^{\mathrm{T}}v = 0$. *Then*

(a) $\xi_p(v) = V_o(v) = \xi_l(v) = 0$.
(b) $\xi_p'(v) = V_o'(v) = \xi_l'(v) = 0$.
(c) $\xi_p''(v) = V_o''(v) = \xi_l''(v)$.

Proof. (a) By Megiddo and Shub (1986) the vector fields are tangent to any face in which a point lies, so $\xi_p(v) = V_\mu(v) = \xi_l(v) = 0$ for any vertex. Thus, also $(\partial V\mu/\partial\mu)(v) = 0$.

(b) Differentiating $\xi_p(x) = V_{\mu(x)}(x)$ gives, as in Megiddo and Shub (1986, Proposition 74) that

$$\frac{d\xi_p(x)}{dx} = \frac{\partial V_\mu}{\partial \mu}\frac{d\mu}{dx} + \frac{\partial V_\mu}{\partial x}, \qquad \frac{\partial V}{\partial \mu}(v) = 0 \text{ and } \frac{\partial V_\mu}{\partial x}(v) = -\mu(v)I$$

by Megiddo and Shub (1986, Proposition 6.5).

Using the notation of Proposition 1, $\tau(x) = \eta_{F_x} + \eta_{E_{1x}}$. Now $\eta_{F_v} = 0$ by Megiddo and Shub (1986, Appendix F) and $\eta_{E_{1v}} = 0$ since $c \cdot v = 0$. Thus $\tau(v) = 0$ and $\mu(v) = 0$.

(c) Differentiating one more time gives

$$\frac{d^2\xi_p(v)}{dx^2} = \frac{\partial^2 V_\mu}{\partial\mu^2}(v)\left[\frac{d\mu(v)}{dx}\right]^2 + \frac{\partial V\mu}{\partial\mu}(v)\frac{d^2\mu(v)}{dx^2} + \frac{\partial^2 V\mu}{\partial\mu\partial x}\frac{d\mu}{dx} + \frac{\partial^2 V\mu}{\partial x^2}\quad(v).$$

The derivative $(d\mu/dx)(v) = 0$ by Proposition 1, and $(\partial V\mu/\partial\mu)(v) = 0$ by part (a) so we have $\xi_p''(v) = V_0''(v)$ and Proposition 2 finishes the proof.

Q.E.D.

Given a vertex v of the polytope S, I_1 the set of indices such that $v_i > 0$, and I_2 the set of indices i such that $v_i = 0$, the space $R_1 = \{x \in \mathbb{R}^n \mid x_j = 0$ for $j \notin I_1\}$ is the space of the basic variable and $R_2 = \{x \in \mathbb{R}_n \mid x_j = 0$ for $j \notin I_2\}$ is the space of nonbasic variables. The nondegeneracy hypotheses on the matrix $\tilde{A}$ imply that any vertex the basic variables may be solved for in terms of the nonbasic. That is, if we let $R^n = R_1 \times R_2$ and N: $R^n \to R_2$ the projection, then $N(P)$ is a polytope P_2 contained in the positive orthant of R_2 with a vertex at 0 and there is a linear map L: $R_2 \to R_1$ such that P is the graph of $L + v$ over P_2. Symbolically, $P = \{(\begin{smallmatrix} L(p) \\ p \end{smallmatrix}) + v$ for $p \in P_2\}$. Then all vector fields and discrete iterations we consider may be projected into P_2. The orbits in P are simply the graphs over the orbits in P_2. Given the vector field or discrete iteration W we consider the vector field or iteration NW on P_2 and say we have expressed W in the nonbasic variables. The reduced cost vector $\tau = (L^t, Id)c$, where c has been written in the $R_1 \times R_2$ coordinates, expresses the cost vector in terms of the R_2 coordinates.

THEOREM 1. *Given a nondegenerate linear programming problem in Karmarker standard form:*

(i) The projective rescaling vector field $\xi_P(x)$ may be expressed in terms of the (n-m) nonbasic variables at an optimal vertex as

$$N\xi_P(x) = +\ \tilde{c}x^2 + o(\|x\|^2),$$

where $\tilde{c}$ is the (n-m)-dimensional reduced cost vector and the j^{th} component of $\tilde{c}x^2$ is $\tilde{c}_j x_j^2$.

(ii) If, moreover, the optimum is unique every interior solution curve of the differential equation $\dot{x} = N\xi_P(x)$ is tangent at the optimum to the vector $1/\tilde{c}$, that is, the vector whose j^{th} component is $1/\tilde{c}_j$.

Proof. This is now immediate from Sections 4 and 5 of Megiddo and Shub (1986). The vectorfield $N\xi_P(x) = \tilde{c}x^2 + o(\|x\|^2)$ so Proposition 3 proves (i). The uniqueness of the optimum guarantees that no $\tilde{c}_j = 0$, so $1/\tilde{c}$ makes sense.

2. The Asymptotics of Discrete Algorithm

With minor modification the discrete version of Karmarkar's algorithm used to prove his polynomial convergence theorem (Karmarker, 1984) is given as a transformation, Y, of the polytope to itself which takes a step from the point x in the direction $-\xi_p(x)$ with step size $\phi(x)$ a nonnegative continuous function

$$Y(x) = x - \phi(x)\xi_P(x).$$

The function $\phi(x)$ is defined as follows:

Let

$$\bar{A} = \begin{pmatrix} AD \\ e^{\mathrm{T}} \end{pmatrix}.$$

Let

$$\eta_P(x) = P_{\bar{A}}Dc.$$

Then

$$\phi(x) = \frac{\gamma}{\|\eta_P(x)\| - yx^{\mathrm{T}}\eta_P(x)},$$

where $0 < \gamma$ is a fixed real constant, originally chosen as $\frac{1}{4}$ (see Megiddo and Shub (1986) for this derivation). We will need some information about $\phi(x)$.

Lemma 1. (i) $\eta_P(x) = \tau(x) - (c^{\mathrm{T}}x/n)e$,
(ii) $x^{\mathrm{T}}\eta_P(x) = \mu(x) - c^{\mathrm{T}}x/n$,
(iii) $\xi_P(x) = D_x\tau(x) - \mu(x)x = D_x\eta_P(x) - (x^{\mathrm{T}}\eta_P(x))x$,
(iv) $\eta_P(x) = D_x^{-1}\xi_P(x) - (c^{\mathrm{T}}x/n)e + \mu(x)e$.

Proof. Part (i) follows from the fact that e is in null space AD. Thus the orthogonal projection of Dc into null space AD $\cap$ null space e^{T} is achieved by projecting first into null space AD and then subtracting the projection on e. That is,

$$\eta_P(x) = \tau(x) - \frac{\tau(x)^{\mathrm{T}}e}{e^{\mathrm{T}}e}\, e = \tau(x) - \frac{(D_x c)^{\mathrm{T}}e}{e^{\mathrm{T}}e}\, e = \tau(x) - \frac{c^{\mathrm{T}}x}{n}\, e.$$

The second equality is true once again since $e \in \text{null}\, AD$ and thus $\tau(x)^{\mathrm{T}}e = (D_x e)^{\mathrm{T}}e$.

Part (ii) follows from (i), the definition of $\mu(x)$, and the fact that $e^{\mathrm{T}}x = 1$.

Part (iii) follows from the definition of $\xi_P(x)$ and (i) and (ii).

Part (iv) follows from (iii) by applying D_x^{-1} and using (ii).

For a nondegenerate problem we express the iteration Y in terms of the nonbasic variables at the optimal vertex as

$$NY(x) = x - N\phi(x)N\xi_P(x),$$

where x is in the positive orthant R_2^+ of R_2 as in Section 1. We are interested in the iterates $(NY)^q(x)$.

THEOREM 2.

$$\begin{array}{lrl} \textit{Let} & \min & c^{\mathrm{T}}x \\ \textit{subject to} & Ax &= 0 \\ & e^{\mathrm{T}}x &= 1 \end{array}$$

be a nondegenerate linear programming problem in Karmarkar standard form with $\binom{A}{e}$ *an* $m \times n$ *matrix and with a unique optimal vertex. Then there is a neighborhood* U_1 *of the origin in* R_2^+ *the space of the nonbasic variables and a neighborhood* $U_2 \subset U_1$ *of the intersection of* U_1 *with the line* $\tilde{C}x = \gamma e$, $\gamma \geq 0$, *such that*

(*i*) *The set* U_2 *contains a definite angle at the origin.*

(*ii*) *For every* $x \in U_2$

$$\lim \frac{\tilde{C}_j((NY)^q(x))_j}{\tilde{C}_i((NY)^q(x))_i} = 1$$

for all i *and* j.

(*iii*) *There exist constants* $K_1, K_2 > 0$ *and an* α,

$$0 < \alpha \leq \frac{1}{\sqrt{m((n-m)/n)} + \gamma((n-m)/n)},$$

such that

$$K_1(1 - \alpha\gamma)^q \leq \|(NY)^q(x)\| \leq K_2(1 - \alpha\gamma)^q.$$

The proof of Theorem 2 occupies the remainder of this section. $\tilde{C}x$

means the vector whose j^{th} component is $\tilde{C}_j x_j$. The $\tilde{C}_i$ are strictly positive since we have a unique minimum. Thus we may find a neighborhood U_3 of the line segment with bounded angle away from the coordinate planes; that is, there exists a constant, $K_3 > 0$, such that $|x_i/x_j| > K_3$ for $x \in U_3$ and U_3 contains a definite angle about $\tilde{C}x = \gamma e$ at the origin.

LEMMA 2. *Suppose we are in the circumstances of Theorem 2. Then*

$$N\eta_P(x) = \tilde{C}x - \frac{\tilde{C}^{\mathrm{T}}x}{n} N(e) + o(\|x\|) \qquad \text{for } x \in U_3.$$

Proof. By Theorem 1,

$$N\xi_P(x) = \tilde{C}x^2 + o(\|x\|^2)$$
$$D_x^{-1}N\xi_P(x) = \tilde{C}x + o(\|x\|^2) \qquad \text{in } U_3.$$

Now use Lemma 1(iv) and Proposition 1.

In the notation of Section 1, the polytope P is the graph of $L + v$ over the polytope P_2 in R_2^+. We define the inner product

$$\langle\ ,\ \rangle_1 \text{ on } R_2^+ \text{ by } \langle x_1, x_2\rangle = \langle (x_1, L(x_1)), (x_2, L(x_2))\rangle,$$

where $\langle\ ,\ \rangle$ is the usual inner product on $R^n = R_1 \times R_2$. We let $\|\ \|_1$ be the norm associated to $\langle\ ,\ \rangle_1$, i.e., $\|x\|_1 = \langle x, x\rangle_1^{1/2}$ for $x \in R_2$.

In terms of the nonbasic variables then

$$NY(x) = x - \frac{\gamma}{\|N\eta_P(x)\|_1 + \gamma(\tilde{C}^{\mathrm{T}}x/n) + \gamma\mu_1(x)} (N\xi_P(x)),$$

where $d\mu_1/dx = 0$. Thus

$$NY(x) = x - \frac{\gamma}{\|\tilde{C}x - (\tilde{C}^{\mathrm{T}}x/n)e + o\|x\|\,\|_1 + \gamma(\tilde{C}^{\mathrm{T}}x/n) + o\|x\|}$$
$$(\tilde{C}^2x^2 + o\|x\|^2)$$
$$= x - \frac{\gamma}{\|\tilde{C}x - \tilde{C}^{\mathrm{T}}x/n)e\|_1 + \gamma(\tilde{C}^{\mathrm{T}}x/n) + o\|x\|_1} (\tilde{C}x^2 + o\|x\|_1^2).$$

If we make the change of variable $y = \tilde{C}x$ then the iteration becomes

$$\text{(I) } NY(y) = y - \frac{\gamma}{\|y - (e^{\mathrm{T}}y/n)e\|_1 + \gamma(e^{\mathrm{T}}y/n) + o\|y\|} y^2 + o\|y\|^2.$$

As in Megiddo and Shub, Sect. 8) we study this equation in "polar coordinates." For $y \ni R_2^+$ let $R(y) = \|y - (e^T y/n)e\|_1 + \gamma(e^T y/n)$ and let $B \subset R_2^+$ be the set of points $\{y \mid R(y) = 1\}$. Finally, let $\alpha = 1/R(e) = n/(m\|e\|_1 + \gamma(n - m))$.

LEMMA 3.

$$\|e\|_1 \geq \sqrt{n \frac{(n-m)}{m}}$$

$$\alpha \leq \frac{1}{\sqrt{m((n-m)/n)} + \gamma((n-m)/n)}.$$

Proof. The vector $(e, L(e))$ is tangent to the simplex, thus $e^T L(e) = -(n - m)$ and

$$\|e\|_1 = \|(e, L(e)\| = (\|e\|^2 + \|L(e)\|^2)^{1/2} = ((n-m) + \|L(e)\|^2)^{1/2}$$

as

$$|e^T L(e)| = n - m \; \|L(e)\|^2 \geq \left(\frac{n-m}{m}\right)^2 m = \frac{(n-m)^2}{m}.$$

Thus

$$\|e\|_1 \geq \left((n-m) + \frac{(n-m)^2}{m}\right)^{1/2} = \sqrt{n \frac{(n-m)}{m}},$$

$$\alpha = \frac{1}{(m/n)\|e\|_1 + \gamma(n-m)/n)}$$

so substituting the inequality for $\|e\|_1$ gives the inequality for α.

LEMMA 4. *For $\gamma \leq \sqrt{2}/2$, $\alpha\gamma < \frac{1}{2}$.*

Proof.

$$\frac{1}{\sqrt{m \frac{(n-m)}{n}} + \gamma \frac{(n-m)}{n}} < \frac{1}{\sqrt{m \frac{(n-m)}{n}}} \leq \frac{1}{\sqrt{\frac{n-1}{n}}} \leq \sqrt{2}.$$

Remark. If the point $(1/n, \ldots, 1/n) \in R^n$ is in P then $(e, Le + v) = (1/n, \ldots, 1/n)$ with $\Sigma v_i = 1$. Thus it is easy to see that

$$\|e\|_1 \leq n \sqrt{\frac{n-1}{n}}$$

and consequently

$$\alpha \geq \frac{1}{m\sqrt{(n-1)/n} + \gamma((n-m)/n)}.$$

To change to "polar coordinates," let $(\sigma, \mu) \in R_+ \times B$ and let $y = \sigma u$. Then

$$W(\sigma, u) = \left(R(NY(\sigma u)), \frac{1}{R(NY(\sigma u))} NY(\sigma u)\right)$$

expresses NY in the (σ, u) coordinates. It follows from (I) that there is an $r_0 > 0$ such that W: $[0, r] \times S \to [0, r] \times S$ for $0 < r \leq r_0$. Moreover, W and Z are tangent on $0 \times B$, where

$$Z(\sigma, u) = (R(u - \gamma u^2), \frac{1}{R(u - \gamma u^2)} u - \gamma u^2).$$

The map Z takes rays to rays, and the ray through αe is fixed by Z; that is, $(0, \alpha e)$ is a fixed point for Z.

PROPOSITION 4. *The derivative of Z at $(0, \alpha e)$ is*

$$Z'(0, \alpha e) = \begin{pmatrix} 1 - \alpha\gamma & 0 \\ 0 & (1 - \alpha\gamma(1 - \alpha\gamma))I \end{pmatrix}$$

Proof. The derivative of $u - \gamma u^2$ at αe is $(1\text{-}2\,\alpha\gamma)I$. The derivative of $(1/R(u - \gamma u^2))(u - \gamma u^2)$ applied to tangent vectors to B at αe is

$$\frac{1}{R(\alpha e - \gamma(\alpha e)^2)}(1 - 2\alpha\gamma)I = \frac{1}{\alpha(1 - \alpha\gamma)R(e)}(1 - 2\alpha\gamma)I$$
$$= \left(\frac{1 - 2\alpha\gamma}{1 - \alpha\gamma}\right) I = \left(1 - \frac{\alpha\gamma}{1 - \alpha\gamma}\right) I.$$

The derivative along the ray is just $R(\alpha e - \gamma(\alpha e)^2) = 1 - \alpha\gamma$. Q.E.D.

Now we return to the proof of Theorem 2.

Proof. Since $\alpha\gamma < \frac{1}{2}$, $1 - \alpha\gamma/(1 - \alpha\gamma)$ is positive and less than $1 - \alpha\gamma$, thus $(1 - \alpha\gamma/(1 - \alpha\gamma))\, I$ represents a stronger contraction that $1 - \alpha\gamma$ for $Z'(0, \alpha e)$. For the linearized system it is simple to see that any orbit under iteration becomes tangent to the eigenspaces of $1 - \alpha\gamma$ and has an ultimate rate of attraction $1 - \alpha\gamma$ toward 0. That the same is true for the nonlinear approximations Z and W to Z' follows from center manifold theory (see Shub, 1986). The theorem for NY results by the change of coordinates. The rate of attraction to zero remains unchanged under linear conjugation. Q.E.D.

3. Problems

Problem 1. Given a nondegenerate linear programming problem in Karmarkar standard form, is the asymptotic rate of approach to the optimum of the discrete algorithm $1 - \alpha\gamma$ for every interior point of the polytope?

A similar problem is stated by Megiddo and Shub (1986) for the discrete affine rescaling algorithm and is still open.

Problem 2. For fixed polytope P with center x_o, let C be a cost function and $\varepsilon > 0$. Define $S_d(C, \varepsilon)$ and $S_s(C, \varepsilon)$ to be the number of steps of the discrete algorithm described above, or the number of steps of line search on Karmarkar's potential function using the projective rescaling vectorfield starting at x_o necessary to be within ε of an optimal point. For nontrivial polytopes, even the unconstrained simplices of dimension 3 or 4 (or even 2?), is it true that for fixed ε sufficiently small, the functions $S_d(C, \varepsilon)$ and $S_s(C, \varepsilon)$ are unbounded on the space of problems? Their averages over the unit sphere in the space of problems should be finite. What are the averages?

The recent paper of Anstreicher (1987) might be relevant to this problem.

References

Anstreicher, K. M. (1987), "Linear Programming and the Newton Barrier Flow," Yale School of Organization and Management Preprint, Jan.

Barnes, E. R. (1985), "A Variation on Karmarkar's Algorithm for Solving Linear Programming Problems," Research Report RC-11136, IBM T. J. Watson Research Center, Yorktown Heights, NY 10598, May.

Gill, P. E., Murray, W., Saunders, M. A., Tomlin, J. A., and Wright, M. H. (1985), "On Projected Newton Barrier Methods for Linear Programming and an Equivalence to Karmarkar's Projective Method," Technical Report SOL 85-11, Systems Optimization Laboratory, Department of Operations Research, Stanford University, Stanford, CA 94305, July.

Karmarkar, N. (1984), A new polynomial-time algorithm for linear programming, *in* "Proceedings, 16th Annual ACM Symposium on Theory of Computing (1984)," pp. 302–311, ACM, New York; revised version: *Combinatorics* **4**, 373–395.

MEGIDDO, N., AND SHUB, M. (1986), "Boundary Behavior of Interior Point Algorithms in Linear Programming," Research Report RJ-5319 (54679), IBM Almaden Research Center, San Jose, CA, Sept.

RENEGAR, J. (1986), "A Polynomial-Time Algorithm, Based on Newton's Method, for Linear Programming," Report MSRI 07118-86, Mathematical Sciences Research Institute, Berkeley, CA.

SHUB, M. "Global Stability of Dynamical Systems," Springer-Verlag, New York, 1986.