International Journal of Mathematics and Statistics Invention (IJMSI)
E-ISSN: 2321 – 4767 P-ISSN: 2321 - 4759
www.ijmsi.org Volume 2 Issue 4 || April. 2014 || PP-12-22

On interior-point methods, related dynamical systems results and cores of targets for linear programming

Ukwu Chukwunenye
Department of Mathematics
University of Jos P.M.B 2084 Jos, Nigeria

ABSTRACT: *This article examined interior-point methods for solving linear programming problems. Then it collected key relevant results, formulated analogous continuous dynamical systems, pioneered the introduction of cores of targets concept and established a sequence of lemmas which would be invoked to prove that the system's trajectories converge to the optimal solution of a linear programming problem in standard form, under appropriate conditions.These results were made possible by the exploitation of norm properties, their derivatives and the theory of ordinary differential equations, paving the way for the pursuits of solutions of some linear optimization problems that may not require exact optimal solutions but particular solutions at specified tolerance levels, within feasible guidelines or constraints.*

KEYWORDS: *Core, Dynamics, Interior, Karmarkar, Programming.*

I. INTRODUCTION AND MOTIVATION

The application of linear programming to business, management, engineering and structured decision processes has been quite remarkable. Since the development of the simplex method in 1947 by G.B. Dantzing, there has been a flurry of research activities in the designing of solution methods for linear programming, mostly aimed at realizing more effective and efficient algorithmic computer implementations and computing complexity reductions.

In the Fall of 1984, Karmarkar [1] of AT & Bell Laboratories proposed a new polynomial-time algorithm for solving linear optimization problems. The new algorithm not only possesses better complexity than the Simplex method in the worst-case scenario, but also shows the potential to rival the Simplex algorithm for large-scale, real-world applications. This development quickly captured the attention of Operations Research community. Radically different from the Simplex method, Karmarkar's original algorithm considers a linear programming problem over a simplex structure and moves through the interior of the polytope of feasible domain by transforming the space at each step to place the current solution at the center of the polytope in the transformed space. Then the solution is moved in the direction of projected steepest descent far enough to avoid the boundary of the feasible region in order to remain interior. Then the inverse transformation of the improved solution is taken to map it back to the original space to obtain a new interior solution. The process is repeated until an optimum is obtained with a desired level of accuracy.

Karmarkar's standard form for linear programming can be described as follows:

$$(LP_1)\begin{cases} \min\ c^T x \\ s.t.\ Ax = 0 \\ \quad e^T x = 1 \\ \quad x \geq 0 \end{cases} \tag{1}$$

where A is an $m \times n$ matrix of full row rank, $e = (1,1,...,1)^T$ is the column vector of n ones, c is an *n*-dimensional column vector and T denotes transpose.

The basic assumptions of Karmarkar's algorithm include:

$$Ae = 0 \tag{2}$$

the optimal objective value of (1) is zero. (3)

Notice that if we define $x^0 = \frac{e}{n}$, then assumption (2) implies that x^0 is a feasible solution of (1) and each component of x^0 has the positive value $\frac{1}{n}$.Any feasible solution *x* of (1) is called an interior feasible solution if each component of *x* is positive. This implies that *x* is not on the boundary of the feasible region; needless to say that the constraint $e^T x = 1$ in (1) implies that $\sum_{j=1}^{n} x_j = 1$.Therefore $e = (1,1,...,1)^T$, leading to the conclusion that a consistent problem in Karmarkar's standard form has a finite optimum. In Fang and Puthenpura [2], it is shown that any linear programming problem in standard form can be expressed in Karmarkar's standard form. Karmarkar's algorithm and its specifics are well-exposed in [2]. Karmarkar's algorithm is an interior-point iterative scheme for solving linear programming problems.

Interior-point methods approach the optimal solution of the linear program from the interior of the feasible region by generating a sequence of parameterized interior solutions.

Most of the interior-point methods can be categorized into three classes: the pure affine scaling methods, the potential reduction methods and the path-following methods. The specifics of these methods are described in section 3. The primary focus of this article will be on path-following methods.

The basic idea of path-following is to incorporate a barrier function into the linear objective. By parameterizing the barrier function, corresponding minimizers form a path that leads to an optimal solution of the linear program.

The main motivation for this work comes from the work of Shen and Fang [3], in which the "generalized barrier functions" for linear programming were defined to create an ideal interior trajectory for path-following. The key components such as the moving direction and the criterion of closeness required for a path-following algorithm were introduced for designing a generic path-following algorithm with convergence and polynomiality proofs under certain conditions.

This work is aimed at exploiting the convergence results in [3] to a parameterized continuous dynamical system. This would lead to the construction of appropriate energy and Lyapunov functions which would be utilized to show that the trajectories of the dynamical system converge to the optimal solution of the linear program under appropriate assumptions.

One is not aware of any interior-point dynamic solver reported in the literature. Most dynamic solvers have been used for the neural network approach. Such investigations can be referred to in Bertsekas [4], Cohen and Grossberg [5], Hopfield and Tanks [6], Wang [7, 8, 9, 10],and Zah [11]. In section 4, we formulate a continuous dynamical system and our main results captured in a sequence of lemmas.Section 5 presents our conclusions and direction of a follow-up research.

II. INTERIOR-POINT METHODS

2.1 Affine Scaling Methods

Let *A*, *c* and *x* be as defined in (1) and let *b* be an *m*-dimensional column vector. Consider the following linear programming problem in standard form:

$$(LP_2)\begin{cases} \min\ c^T x \\ s.t.\ Ax = b, \\ \quad x \ge 0 \end{cases} \tag{4}$$

whose linear dual problem is:

$$(DLP_2)\begin{cases} \min\ b^T y \\ s.t.\ A^T y + s = c, \\ \quad s \ge 0 \end{cases} \tag{5}$$

where $y \in R^m$ and *s* is an *n*-dimensional column vector. A feasible solution (y, s) of DLP_2 is called an interior feasible solution if $s > 0$.

Following Noble [12], the basic strategy of the affine scaling method is as follows: given an interior feasible solution, $\bar{x}$ of LP_2, construct a simple ellipsoidal approximation of its feasible region, that is centered at $\bar{x}$. Then optimize the objective function $c^T x$ over this ellipsoid and use the resulting direction with a suitable step-length to define a new algorithm iterate.

Let $\bar{x} = \{\bar{x}_1, \bar{x}_2, ..., \bar{x}_n\}$ and $\bar{X} = \mathrm{Diag}(\bar{x})$ be the diagonal matrix with $c^T x$ as the*j*-th diagonal element, for $j \in \{1, 2, ..., n\}$.

The Dikin ellipsoid at $\bar{x}$ is defined as $E_{\bar{x}} = \left\{x \in R^n : Ax = b, (x - \bar{x})^T \bar{X}^{-2}(x - \bar{x}) \le 1\right\}$. Then the affine direction at $\bar{x}$ is the solution of the following direction-finding problem:

$$\left(ADFP_{\bar{x}}\right) \begin{cases} \min\ c^T d \\ s.t.\ \ Ad = 0, \\ \qquad d^T \bar{X}^{-2} d \le 1. \end{cases} \tag{6}$$

It is noted in [6] that $E_{\bar{x}}$ is always contained in the feasible region of LP_1 for any interior feasible solution $\hat{x}$ of LP_1.

For a pure affine scaling algorithm, there is no polynomial-time proof, but it is known to converge when A has full row rank (Saigal [13] and [3]).

2.2 Potential Reduction Methods

These methods are typically designed to find successive iterates by solving the following optimization problem:

$$\left(PRM_1\right) \begin{cases} \min\ f(x, y, s) = q \text{ In}\left(c^T x - b^T y\right) - \sum_{j=1}^{n} \text{In}\left(x_j\right) \\ s.t.\ \ Ax = b, x > 0, \\ \qquad A^T y + s = c,\ s \ge 0 \end{cases} \tag{7}$$

where q is a positive parameter of the function f and $f(x, y, s)$ is called a potential function.

Remarks

The difference $c^T x - b^T y$ is the duality gap between the solution pair (x, y) and it is common knowledge that the gap is always nonnegative. The optimum of (LP_1) is attained when the duality gap is reduced to zero. Note that $c^T x - b^T y \to 0$, if $q\text{In}\left(c^T x - b^T y\right) \to -\infty$

The second part, $\sum_{j=1}^{n} \text{In}\left(x_j\right)$ of f, approaches $-\infty$ as $x_j \to 0^+$ for some j, where x_j is the j^{th} component of *x*. Therefore an optimal solution of (PRP_1) must be repelled from the boundary of the feasible domain. The potential function is a surrogate for the goal of reducing the duality gap to zero from the interior.

[1] used a specific form of the potential reduction method.

2.3 Path-following Methods

Consider a linear program:

$$(P) \begin{cases} \min\ c^T x \\ s.t.\ Ax = b \\ \qquad x \ge 0 \end{cases} \tag{8}$$

Let $W = \left\{x \in R^n : Ax = b,\ x \ge 0\right\}$ and $W_0 = \left(x \in R^n : Ax = b,\ x > 0\right)$ be the interior feasible domain.

Let the following assumptions hold:

A has full row rank, (9)

$W \equiv \overline{W}_0 \ne \Phi$, where $\overline{W}_0$ is the closure of W_0, and Φ is the empty set (10)

W is compact. (11)

2.3.1 Definition

A function $\phi : W \to \bar{R}$ is called a generalized barrier function for linear programming (GBLP), if $(P1)\ \phi : W \to \bar{R}$ is proper, strictly convex and differentiable, where $\bar{R}$ is the extended real line.

Remark

The properness property of ϕ is equivalent to the requirement that $\phi(x)$ be strictly bounded below by $-\infty$ for all $x \in W$ and be strictly bounded above by ∞ for some $x \in W$.

(P2) if the sequence $\{x^k\} \subset W_0$ converges to *x* with the i^{th} component, $x_i = 0$ then $\lim_{k\to\infty} \left(\nabla\phi\left(x^k\right)\right)_i = -\infty$

(P3) the effective domain of ϕ contains W_0. Equivalently, $W_0 \subset \{\ x \in W\ s.t\ \phi(x) \text{is finite}\ \}$.

Let $\mu > 0$ and define an augmented primal problem (P_u) associated with a *GBLP* function follows:

$$
(P_u)\begin{cases} \min\ c^T x + \mu\phi(x) \\ s.t.\ Ax = b \\ \qquad x \ge 0 \end{cases} \tag{12}
$$

Then we have the following results from [3]:

i. (P_u) has a unique optimal solution, denoted by $x(\mu)$, in W_0,

ii. $\{c^T z(\mu)\}$ is a monotone decreasing function in μ,

iii. The set $\{x(\mu) : \mu > 0\}$ characterizes an interior, continuous curve in W_0.

iv. Given a decreasing positive sequence μ_k such that $\lim_{k\to\infty} \mu_k = 0$, if $x^* = \lim_{k\to\infty} x(\mu_k)$ then x^* is the optimum of (*P*).

v. Suppose that $\bar{x}$ is a given interior feasible solution to (*P*) and Δx solves the problem:

$$
\left(P_\mu^1\right)\begin{cases} \min\left[c + \mu\nabla\phi(\bar{x})^T\ \Delta x\right] \\ s.t.\ \ A\Delta x = 0 \\ \left\|\ \tilde{X}^{-1}\Delta x\ \right\|^2 \le\ \beta^2\ <\ 1, \end{cases} \tag{13}
$$

where $\tilde{X}$ is any positive definite symmetric matrix and $0 < \beta < 1$. Then, Δx defines a moving direction at $\bar{x}$ below:

$$
\Delta x = -\tilde{X}\left[I - \tilde{X}A^T\left(A\tilde{X}^2A^T\right)^{-1}A\tilde{X}\right]\tilde{X}\left(c + \mu\Delta\phi(\bar{x})\right).
$$

Under appropriate condition on $\tilde{X}$, *c* and ϕ, it is proved in [12], that any convergent feasible sequence of solutions to (P_u) must converge to the optimal solution to (*P*) as $\mu \to 0^+$.

From Bazaraa and Shetty[14], we also have the following result:

2.3.2 Lemma

Let *X* be a nonempty closed set in R^n and *f*, *g*, $g_1, g_2, \cdots, g_m$ be continuous functions on R^n, where $f, g,\ g_1, g_2, \cdots, g_m$ are scalar functions and g is an *m*-dimensional vector function whose components are $g_1, g_2, \cdots, g_m$ Suppose that the set $\{x \in X : g(x) < 0\}$ is not empty and that *B* is a barrier function that is continuous on $\{x : g(x) < 0\}$, where $B(x) = \sum_{i=1}^{m}\hat{\phi}[g_i(x)]$ and $\hat{\phi}$ is a function of one variable that is continuous over $\{y : y < 0\}$ and satisfies $\hat{\phi}(y) \ge 0$, if $y < 0$, and $\lim_{y\to 0} -\hat{\phi}(y) = \infty$. Furthermore, suppose that for any given $\mu > 0$, if $\{x_k\}$ in *X* satisfies $g(x_k) < 0$ and $f(x_k) + \mu B(x_k) \to \theta(\mu)$ where $\theta(\mu)$ is defined by:

$$\theta(\mu) = \inf\{ f(x) + \mu B(x) : g(x) < 0,\ x \in X \}, \tag{14}$$

then $\{x_k\}$ has a convergent subsequence. Moreover,

(i) For each $\mu > 0$, there exists an $x_\mu \in X$ with $g(x_\mu) < 0$ such that
$\theta(\mu) = f(x_\mu) + \mu B(x_\mu) = \inf\{ f(x) + \mu\, B(x) : g(x) < 0,\ x \in X\}$,

(ii) $\inf\{ f(x) : g(x) \le 0,\ x \in X \} \le \inf\{\theta(\mu) : \mu > 0\}$,

(iii) For $\mu > 0$, $f(x_\mu)$ and $\theta(\mu)$ are nondecreasing functions of μ and $B(x_\mu)$ is a nonincreasing function of μ.

2.3.3 Theorem

Let $f : R^n \to R$ and $g : R^n \to R^m$ be continuous functions, and let *X* be a nonempty closed set R^n. Suppose that the set $\{x \in X : g(x) < 0\}$ is not empty. Furthermore, suppose that the primal problem of minimizing $f(x)$ subject to $g(x) \le 0,\ x \in X$ has an optimal solution $\bar{x}$ with the following property $(\bar{P})$: Given any neighborhood *N* around $\bar{x}$, there exists an $x \in X \cap N$ such that $g(x) < 0$. Then

minimum$\{ f(x)\ g(x) \le 0, x \in X\} = \lim_{\mu \to 0} \theta(\mu) = \inf_{\mu > 0} \theta(\mu)$, where $\theta(\mu)$ is defined by (14).

Letting $\theta(\mu) = f(x_\mu) + \mu B(x_\mu)$, where $x_\mu \in X$ and $g(x_\mu) < 0$, then the limit of any convergent subsequence of $\{x_\mu\}$ is an optimal solution to the primal problem, and furthermore $\mu B(x_\mu) \to 0$ as $\mu \to 0^+$.
The next section establishes and collects a sequence of lemmas needed in the proof of the asymptotic behavior of the system's trajectories.

III. CONTINUOUS DYNAMICAL SYSTEMS RESULTS

In this section, we formulate an analogous continuous dynamical system and establish a sequence of lemmas which would be invoked to prove that the system's trajectories converge to the optimal solution of (*P*) under some appropriate conditions.

Let *A* be an m x n matrix of full row rank. Let R_+^n denote the set $\{x \in R^n : x \ge 0\}$. For $x \in W$, let $z = x + r$ for some $r \in R^n$. Then $Ax = b \Leftrightarrow Az - Ar = b.$ Taking $r = -A^T (AA^T)^{-1} b$, we see that $Az = 0$. Therefore, the system $Az = 0$ is consistent if and only if the system $Ax = b$ is consistent. In the sequel we let $r \in R^n$ be such that $Ar = -b$ and $z = x + r$.

Let β be a constant such that $0 < \beta < 1$. Let $\tilde{X}$ be any positive definite symmetric matrix of order *n* and let *p* and *q* be norm conjugates of each other such that $1 \le p, q \le \infty$ and $\frac{1}{p} + \frac{1}{q} = 1.$

Let a_j denote the j^{th} row of *A* for $j \in \{1, 2, ..., m\}$. For fixed *A*, define the map $\|A(\cdot)\| : R^n \to R_+^1$ by

$$\| Az \|_p = \left(\sum_{j=1}^{m} | a_j z |^p \right)^{\frac{1}{p}}.$$

<u>Remark</u>

$\|Az\|_p$ is called the *p*-norm of the function Az. Unless explicitly stated we use the 2-norm in this article.

Define $\|A\|_2 = \max_{z \ne 0} \left\{ \dfrac{\|Az\|_2}{\|z\|_2} \right\}$. For all $x \in W$ and for any given $r \in R^n$ such that $z = x + r$ consider the function:

$$E_{p,p,r,\mu,\lambda}(z) \equiv E(z)$$
$$= c^T (z - r) + p\|Az\|_p + r\sum_{j=1}^{n} (z - r)_j^- + \mu\phi(z - r) + \lambda\left(\|\tilde{X}^{-1} z\|_2^2 - \beta^2 \right) \tag{15}$$

where

$$w^{-} = \max\{-w, 0\} = \begin{cases} -w \text{ if } w < 0 \\ 0 \text{ if } w \geq 0 \end{cases} \text{ and } w^{+} = \max\{w, 0\} = \begin{cases} w \text{ if } w > 0 \\ 0 \text{ if } w \leq 0 \end{cases},$$

for $w \in R$; $\mu > 0$, $r \geq 0$, $p > 0, \lambda \geq 0$.

Note the following:

(i) Property (P1) of definition 2.3.1 implies that, $\phi(z-r) > -\infty$ for all z in the null-space of A and $\phi(z-r) < \infty$ for at least one z in the null-space of A.

(ii) Properties (P1) and (P2) of definition 2.3.1 imply that $\| \phi(z-r) \|_2 < \infty$, for.

(iii) Property (P2) of definition 2.3.1 implies that if $\{z^k : k = 1, 2, ...\}$ is any positive convergent sequence in the null-space of A such that $\lim_{k\to\infty}(z^k - r) = \hat{z} - r$, with $(\hat{z} - r)_j = 0$, for some $j \in \{1, 2, ..., n\}$, then $\lim_{k\to\infty}(\nabla_z \phi(z^k - r))_j = -\infty$. The latter ensures that the minimization of $E(z)$ is never achieved at the boundary of the set $\{z : z - r \geq 0\}$, using a gradient projection method in the minimization program.

(iv) $p \| Az \|_p$ is penalty for the violation of $z \in N(A)$.

(v) $r\sum_{j=1}^{n}(z-r)_j^{-}$ penalizes violations of $z - r \geq 0$.

(vi) $\lambda\left(\| \tilde{X}^{-1}z \|_2^2 - \beta^2\right)$ is the Lagrangian term associated with the constraint $\| \tilde{X}^{-1}z \|_2^2 - \beta^2 \leq 0$.

The stage is now set for our dynamical system formulation and the establishment of key lemmas needed for the proof of our main result.

Let $0 < \beta < 1$ and

$S_1 = \{z \in R^n : Az - 0\} \equiv N(A)$, $S_2 = \left\{z \in R^n : \left\|\tilde{X}^{-1}z\right\|_2^2 - \beta^2 \leq 0\right\}$, and $T = \{z \in R^n : z - r \geq 0\}$,

where r is given and defined as on the previous page. Let t and t_0 be any pair of time variables such that $t \geq t_0 \geq 0$, and let z_0 be an n-dimensional column vector.

For a differentiable function $D : R^n \to R^1$, let $\nabla_z D(z)$ denote the gradient of $D(z)$ with respect to z. Observe that $\nabla_z D(z) \in R^n$ for each $z \in R^n$.

Consider the following dynamical system:

$$\dot{z}(t) = -\nabla_z E_{p,p,r,\mu,\lambda}(z(t)); t \geq t_0 \geq 0 \tag{16}$$

$$z(t_0) = z_0 \tag{17}$$

$$z_0 \in S_2 \cap \text{int}(T). \tag{18}$$

Then $\sum_{j=0}^{n}(z_0 - r)_j^{-} = 0$, using the definition, $w^{-} = \max\{-w, 0\}$ and the fact that $z_0 - r > 0$, by virtue of z_0 being in int(T).

System (16) can be treated as a control system of the form:

$$\dot{z} = -c + A^T \upsilon^{(1)} + \upsilon^{(3} \tag{19}$$

where $\upsilon^{(1)} \equiv \upsilon^{(1)}(z) = -p\nabla_{Az}\|Az\|_p$, $\upsilon^{(2)} = \upsilon^{(2)}(z) = -\mu\nabla\phi(z-r)$ and

$\upsilon^{(3)} \equiv \upsilon^{(3)}(z) = -\lambda\nabla_z\left(\left\|\tilde{X}^{-1}z\right\|_2^2 - \beta^2\right)$.

$\upsilon^{(1)}$, $\upsilon^{(2)}$, and $\upsilon^{(3)}$ can be regarded as controls. These controls will be implemented such that the trajectories of (16), (17) and (18) will be forced into the feasible region $S_1 \cap S_2 \cap \text{int}(t)$ and maintained there while moving in a direction that decreases $c^T\left(z-r\right)$. The following sequence of lemmas will be found useful in the sequel.

Let I_n be the identity matrix of order *n* and let $P = I_n - A^T\left(AA^T\right)^{-1}A$ be the projection matrix onto the null-space of *A*.

3.1 Lemma

The dynamics of system (16) when restricted to $S_1 \cap S_2 \cap \operatorname{int}\left(t\right)$ are described by:

$$\dot{z} = -P\nabla_z E\left(z\right) \tag{20}$$

$$Az = 0 \tag{21}$$

<u>Proof</u>

When sliding along S_1, motion is described by $Az = 0$ and $A\dot{z} = 0$.

From (20) $A\dot{z} = -Ac + AA^T\upsilon^{(1)} + A\upsilon^{(2)} + A\upsilon^{(3)} = 0.$

Since $\left(AA^T\right)^{-1}$ exists, (rank A = *m* by (9)) we deduce that:

$$\upsilon^{(1)} = -\left(AA^T\right)^{-1}A\left[-c+\upsilon^{(2)}+\upsilon^{(3)}\right] \tag{22}$$

Substitute (22) into (19) to get $\dot{z} = \left[I_n - A^T\left(AA^T\right)^{-1}A\right]\left(-c+\upsilon^{(2)}+\upsilon^{(3)}\right)$.

Now, $\nabla_{Az}\left\|Az\right\|_p = 0$ since $z \in S_1 \Rightarrow Az = 0$. Also $\nabla_z \sum_{j=0}^{n}\left(z_0 - r\right)_j^- = 0$ since $z \in \operatorname{int}\left(T\right) \Rightarrow \sum_{j=0}^{n}\left(z_0 - r\right)_j^- = 0$,

showing that $\nabla_z E\left(z\right) = c - \upsilon^{(2)} - \upsilon^{(3)}$.

Therefore:

$$\dot{z} = -P\nabla_z E\left(z\right) \tag{23}$$

Observe from (15) that the term $P\nabla_z\left(\left\|\tilde{X}^{-1}z\right\|_2^2 - \beta^2\right)$ appears on the right side of (26).

The next lemma gives a result on $P\nabla_z\left(\left\|\tilde{X}^{-1}z\right\|_2^2 - \beta^2\right)$.

3.2 Lemma

If $0 \neq z \in S_1 \cap S_2$, then $P\nabla_z\left(\left\|\tilde{X}^{-1}z\right\|_2^2 - \beta^2\right) \neq 0$

<u>Proof</u>

Observe that $P\nabla_z\left(\left\|\tilde{X}^{-1}z\right\|_2^2 - \beta^2\right) = P\left(2\tilde{X}^{-2}z\right)$. Suppose that $P\nabla_z\left(\left\|\tilde{X}^{-1}z\right\|_2^2 - \beta^2\right) = 0$ for some $0 \neq z \in S_1 \cap S_2$. Then $P\left(2\tilde{X}^{-2}z\right) = 0$. Thus $P\tilde{X}^{-2}z = 0$ and:

$$\left(I_n - P\right)\nabla_z\left(\left\|\tilde{X}^{-1}z\right\|_2^2 - \beta^2\right) = \nabla_z\left(\left\|\left(\tilde{X}^{-1}z\right)\right\|_2^2 - \beta^2\right) = 2\tilde{X}^{-2}z \tag{24}$$

However, for any $z \in S_1$:

$$\left[\left(\left(_n - P\right)\nabla_z\left(\left\|\tilde{X}^{-1}z\right\|_2^2 - \beta^2\right)\right]^T z = \left[\left(AA^T\right)^{-1}A\nabla_z\left(\left\|\tilde{X}^{-1}z\right\|_2^2 - \beta^2\right)\right]^T Az = 0 \tag{25}$$

since $Az = 0$. It then follows from (24) and (25) that $z^T\tilde{X}^{-2}z = 0$. We conclude from the positive definiteness of $\tilde{X}$ that $z = 0$. This contradicts the nontriviality assumption on *z* and the lemma is established.

Recall that a_j is the j^{th} row of the matrix *A*, for $j = 1, 2, ..., m$. For $p \geq 1$, the next lemma gives the form of the gradient of $\left\|Az\right\|_p$ for any *z* that is not in the null-space of *A*.

3.3 Lemma

For any $p \geq 1$

and $\forall z \notin S_1$,

$$\nabla_z\left(\|Az\|_p\right) = \frac{A^T}{\|Az\|_p^{p-1}}\left(|a_1 z|^{p-1}\operatorname{sgn}(a_1 z), |a_2 z|^{p-1}\operatorname{sgn}(a_2 z), \cdots, |a_m z|^{p-1}\operatorname{sgn}(a_m z)\right)^T \qquad (26)$$

Proof

Write $z = (z_1, z_2, ..., z_m)^T$.

$$\|Az\|_p = \left(\sum_{j=1}^{m}|a_j z|^p\right)^{\frac{1}{p}} \text{ and } \nabla_z\left(\|Az\|_p\right) = \left(\frac{\partial}{\partial z_1}\|Az\|_p, \frac{\partial}{\partial z_2}\|Az\|_p, \cdots, \frac{\partial}{\partial z_n}\|Az\|_p\right)^t \qquad (27)$$

If $z \in S_1$, then

$$\frac{\partial}{\partial z_k}\|Az\|_p = \frac{\partial}{\partial \sum_{j=1}^{m}|a_j z|^p}\left(\sum_{j=1}^{m}|a_j z|^p\right)^{\frac{1}{p}} \sum_{j=1}^{m}\frac{\partial}{\partial |a_j z|}\left(|a_j z|^p\right)\frac{\partial}{\partial (a_j z)}|a_j z|\frac{\partial}{\partial z_k}(a_j z)$$

$$= \frac{1}{p}\left(\sum_{j=1}^{m}|a_j z|^p\right)^{\frac{1}{p}-1}\sum_{j=1}^{m} p\,|a_j z|^{p-1}\left(\operatorname{sgn}(a_j z)\right)a_{jk}$$

$$= \frac{\left(\sum_{j=1}^{m}|a_j z|^p\right)^{\frac{1}{p}}}{\sum_{j=1}^{m}|a_j z|^p}\sum_{j=1}^{m}|a_j z|^{p-1} a_{jk}\operatorname{sgn}(a_j z) = \frac{1}{\|Az\|_p^{p-1}}\sum_{j=1}^{m}|a_j z|^{p-1} a_{jk}\operatorname{sgn}(a_j z)$$

Hence:

$$\nabla_z\left(\|Az\|_p\right)$$

$$= \frac{1}{\|Az\|_p^{p-1}}\left(\sum_{j=1}^{m}|a_j z|^{p-1} a_{j1}\operatorname{sgn}(a_j z), \sum_{j=1}^{m}|a_j z|^{p-1} a_{j2}\operatorname{sgn}(a_j z), \cdots, \sum_{j=1}^{m}|a_j z|^{p-1} a_{jn}\operatorname{sgn}(a_j z)\right)^t$$

$$= \frac{1}{\|Az\|_p^{p-1}}\begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & & & \\ a_{1n} & a_{2n} & \dots & a_{mn}\end{pmatrix}\begin{pmatrix}|a_1 z|^{p-1}\operatorname{sgn}(a_1 z) \\ |a_2 z|^{p-1}\operatorname{sgn}(a_2 z) \\ \vdots \\ |a_m z|^{p-1}\operatorname{sgn}(a_m z)\end{pmatrix} = \frac{A^T}{\|Az\|_p^{p-1}}\begin{pmatrix}|a_1 z|^{p-1}\operatorname{sgn}(a_1 z) \\ |a_2 z|^{p-1}\operatorname{sgn}(a_2 z) \\ \vdots \\ |a_m z|^{p-1}\operatorname{sgn}(a_m z)\end{pmatrix} \qquad (29)$$

3.4 Lemma ([12], p429)

For any square matrix *M*,

$$\|M\|_2 = \left\{\text{maximum eigenvalue of } M^T M\right\}^{1/2}$$

The following definition is a generalization of the notion of “cores of targets” used in Ukwu [15].

3.5 Definition

The core of the target S_1 for the dynamical system:

$$\dot{z} = -\nabla_z E(z) \qquad (30)$$

is the set $\{z_0 \in R^n\}$ of all initial points that can be driven to the target S_1 in finite time and maintained there, thereafter by an appropriate implementation of some feasible control procedure.

Denote this set by $\mathrm{core}\left(S_1\right)$. The following lemma demonstrates that under certain conditions $\mathrm{core}\left(S_1\right)$ is nonempty.

3.6 Lemma

If:

$$z^T A^T \left(AA^T\right)^{-1} A\nabla_z E\left(z\right) \geq k\left\|Az\right\|_p \tag{31}$$

for some $p \geq 1$ and $\forall z \notin S_1$, with some $k > 0$, then the trajectories of:

$$\dot{z} = -\nabla_z E\left(z\right) \tag{32}$$

$$z\left(0\right) = z_0 \tag{33}$$

hit S_1 in finite time and remain there thereafter.

<u>**Proof**</u>

Note that for any $z \in R^n$, $z = Pz + \left(I_n - p\right)z$ with $Pz \in N\left(A\right)$ and $\left(I_n - P\right)z \in$ column space of *A*. $\left\|\left(I_n - P\right)z\right\|_2$ is the Euclidean distance of *x* from $N\left(A\right)$. Now,

$$\frac{d}{dt}\left\|A^T\left(AA^T\right)^{-1}Az\right\|_2^2 = 2\left\|A^T\left(AA^T\right)^{-1}Az\right\|_2 \frac{d}{dt}\left\|A^T\left(AA^T\right)^{-1}Az\right\|_2$$

$$= 2\left[A^T\left(AA^T\right)^{-1}Az\right]^{T\dot{z}} = 2z^T A^T\left(AA^T\right)^{-1}A\dot{z} = -2z^T A^T\left(AA^T\right)^{-1}A\nabla_z E\left(z\right) \leq -2k\left\|Az\right\|_p$$

$\forall z \notin S_1$ with some $k > 0$, by the hypothesis of the lemma. Therefore:

$$\frac{d}{dt}\left\|A^T\left(AA^T\right)^{-1}Az\right\|_2 \leq \frac{-k}{\left\|A^T\left(AA^T\right)^{-1}Az\right\|_2}\left\|Az\right\|_p \tag{34}$$

We now use the fact that all norms in R^n are equivalent to establish the existence of constant $l > 0$ such that $\left\|Az\right\|_p \geq l\left\|Az\right\|_2$, $\forall z \notin S_1$ and $\forall_p \geq 1$ (see Stoer and Bulirsch [16], p. 185). Consequently:

$$-\left\|Az\right\|_p \leq -l\left\|Az\right\|_2 \tag{35}$$

It now follows from (34) and (35) that:

$$\frac{d}{dt}\left\|A^T\left(AA^T\right)^{-1}Az\right\|_2 \leq \frac{-kl}{\left\|A^T\left(AA^T\right)^{-1}Az\right\|_2}\left\|Az\right\|_2 \leq \frac{-kl}{\left\|A^T\left(AA^T\right)^{-1}\right\|_2\left\|Az\right\|_2}\left\|Az\right\|_2$$

$$= \frac{-kl}{\left\|A^T\left(AA^T\right)^{-1}\right\|_2} \tag{36}$$

Showing that $\left\|A^T\left(AA^T\right)^{-1}Az\right\|_2$ is a time decreasing function. Integrating the differential inequality from time 0 to time t yields:

$$\left\|A^T\left(AA^T\right)^{-1}Az\left(t\right)\right\|_2 \leq \left\|A^t\left(AA^T\right)^{-1}Az_0\right\|_2 - \frac{kl}{\left\|A^T\left(AA^T\right)^{-1}\right\|_2}t \tag{37}$$

The inequality (37) holds for all:

$$t \leq \frac{\left\|A^T\left(AA^T\right)^{-1}\right\|_2^2}{kl}\left\|A^T\left(AA^T\right)^{-1}Az_0\right\|_2 \tag{38}$$

Hence for each $z_0 \notin S_1$, there exists:

$$\bar{t} \in \left(0, \frac{\left\| A^T \left(AA^T \right)^{-1} \right\|_2^2}{kl} \left\| Az_0 \right\|_2 \right) \tag{39}$$

Such that:

$$\left\| A^T \left(AA^T \right)^{-1} Az\left(\bar{t} \right) \right\|_2 = 0 \tag{40}$$

showing that $A^T \left(AA^T \right)^{-1} Az\left(\bar{t} \right) = 0$,which in turn proves that $z\left(\bar{t} \right) \in S_1$. Therefore, if for any initial point z_0, we denote the solution through $\left(0, z_0 \right)$ by $z\left(t \right) \equiv z\left(t, z_0 \right)$ and define:

$$t\left(z_0 \right) = \inf \left\{ t \geq 0 : z\left(t, z_0 \right) \in S_1 \right\} \tag{41}$$

$$z\left(t \right) = Py\left(t \right),\, t \geq t\left(z_0 \right) \tag{42}$$

for some function $y(.)$, then $z(t) \in S_1, \forall t \geq t(z_0)$, and hence $z_0 \in \text{core}(S_1)$.

In a subsequent paper, we will impose appropriate conditions under which the trajectories of:

$$\dot{z} = -\nabla_z E\left(z \right) \tag{43}$$

$$z\left(0 \right) = z_0 \tag{44}$$

converge to $z_\mu - r$, where $z_\mu - r$ is an optimal solution of (5) for each $\mu > 0$. Then we can appeal to Theorem 2.2 of [3] to assert the convergence to the optimal solution of (P), noting that a positive decreasing sequence of parameters $\mu^{\bar{k}}$ with limit 0 may be used in place of μ.

IV. CONCLUSION

In section 1 of this work, we reviewed the interior-point approach to solving linear programming problems and in this direction briefly discussed the pioneering linear programming effort of [1], in which he proposed a polynomial-time algorithm for solving linear programming problems. In section 2, we presented three principal interior-point methods of solving linear programming problems in standard form: namely, affine scaling, potential reduction and path-following methods. Karmarkar’s work, [1] fell into the second group of methods.

This work was motivated partly by interior-point concepts and largely by the path-finding methods in [3] for solving linear programming problems. Many real-life problems which could be formulated as linear programming problems are dynamic in nature; for example, the inventory level of some item at a given time and changes in demand levels of some consumer goods due to price fluctuations and seasonal variation. Also on-line optimization may be required in many application areas, such as satellite guidance, robotics and oil outputs from oil wells and refinery operations. Some of these problems may not require exact optimal solutions but particular solution at specified tolerance levels, within feasible guidelines or constraints. In particular solutions at positive levels may be desired for all decision variables, implying that interior solutions are desired. These and many other problems of the continuous variety could be more realistically modelled by continuous dynamical systems. Unfortunately, research in this direction has been based mainly on neural network approach, none of which is interior-point oriented. In section 3, we formulated an interior-point based dynamical system for solving linear programming problems in standard form. The key ideas for this formulation came from the examination of [3]. Then, we stated that under certain conditions, the solutions of our dynamical system would converge to the solution of a corresponding linear programming problem in standard form. We proceeded to establish a sequence of lemmas which would be needed to prove that our solutions would have the right convergence property. These results were made possible by the exploitation of norm properties, their derivatives and the theory of differential equations.

Sequel to this paper, appropriate energy and Lyapunov functions will be constructed and utilized to show that the trajectories of the dynamical system converge to the optimal solution of the linear program under appropriate assumptions. This approach holds a lot of promise for an extension of our result to neural networks where dynamical systems, energy and Lyapunov functions are used extensively.

REFERENCES

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

[2] Fang, S.C., and Puthenpura, S.(1984).*"Linear optimization and extensions," Prentice Hall, Englewood Cliffs, N.J.*

[3] Shen, R.L. and Fang, S.C.(1994).*"On the generalized path-following methods for linear programming," Optimization, Vol. 30, pp. 235-249.*

[4] Bertsekas, D.P.(1975).*"Necessary and sufficient conditions for a penalty methods to be exact," Mathematical programming, Vol. 9, pp. 87-99.*

[5] Cohen, A., and Grossberg, S.(1983).*"Absolute stability of global pattern formation and parallel memory storage by competitive neural networks," IEEE Transactions on Systems, Man, and Cybernetics, Sept./Oct., Vol. SMC-13, No. 5, pp. 815-825.*

[6] Hopfield, J. and Tanks, D.(1983).*"Simple neural optimizations networks: An A/D converter signal decision circuit, and, and a linear programming circuit," IEEE Trans., CAS-33, (5), pp.533-541.*

[7] Wang, J. (1991).*"A time-varying recurrent neural system for convex programming," Proc. Int. Joint Conf. on Neural Networks, Seattle, Vol. 1, pp. 147-152.*

[8] Wang, J.(1992).*"Analogue neural network for solving the assignment problem," Electronics Letters, May , Vol. 28, No.11, pp. 1047-1050.*

[9] Wang, J.(1993).*"Analysis and design of recurrent neural network for linear programming," IEEE Trans. On Circuits and Systems-I, Sept., Vol. 40, No. 9, pp. 613-618.*

[10] Wang, J.(1991).*"On the asymptotic properties of recurrent neural networks for optimization," Int. J. Pattern Recognition and Artificial Intelligence, Vol. 5, No. 4, pp. 581-601.*

[11] Zah, H. (1995). "*Solving linear programming problems with neural networks: A comparative study," IEEE Transactions on Neural Networks, Jan. Vol. 6, No.1. pp. 94-105.*

[12] Noble, B.(1969). *"Applied linear algebra," Prentice-Hall, Inc. N.J.*

[13] Saigal, R., (1996).*"The primal power affine scaling method," Annals of Operations Research, Vol. 62, pp. 375-417.*

[14] M. Bazaraa, M. and Shetty, C.(1979).*"Nonlinear programming: theory and algorithms," John Wiley and Sons, N.Y.*

[15] Ukwu, C.(1987).*"Compactness of cores of targets for linear delay systems," J. Math.Analy.and Appl., Aug., Vol. 125, No 2, pp. 323-330.*

[16] Stoer, J., and Bulirsch, R.(1991).*"Introduction to numerical analysis," 2nd Ed., Springer-Verlg, N.Y.*