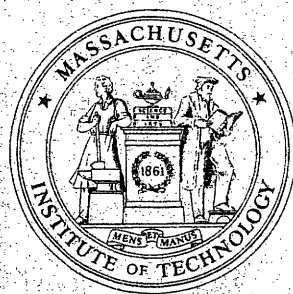# OPERATIONS RESEARCH CENTER

working paper

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
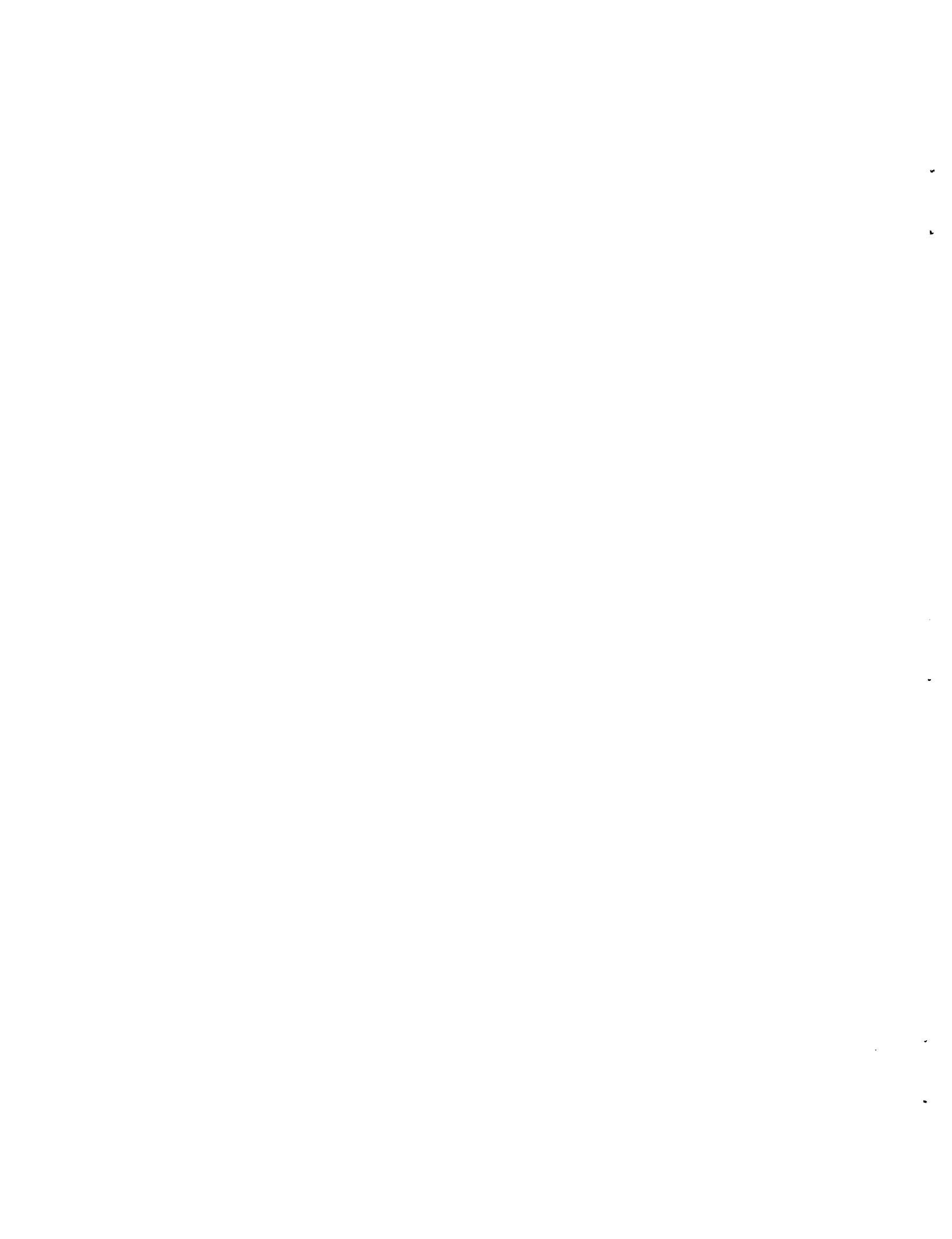
Projective Transformations for Interior Point
Methods, Part I: Basic Theory and
Linear Programming

by

Robert M. Freund

OR 179-88   June 1988

Projective Transformations for Interior Point Methods, Part I:

Basic Theory and Linear Programming

Abstract

The purpose of this two-paper study is to broaden the scope of projective methods in mathematical programming, both in terms of theory and algorithms. We start by generalizing the concept of the analytic center of a polyhedral system to the w-center of a polyhedral system, which stands for weighted center, where there are positive weights on the logarithmic barrier function for each inequality constraint defining a polyhedron X . We prove basic results regarding contained and containing ellipsoids centered at the w-center of the system X . We next shift our attention to projective transformations for transforming the polyhedron X to another polyhedron Z that turns the current point $\bar{x}$ into the w-center of the new polyhedron Z . We work throughout with a polyhedron X that contains both inequality and equality constraints of arbitrary format. We exhibit an elementary projective transformation that transforms the current point $\bar{x}$ to the w-center of Z.

This theory is then applied to two different problems: solving a linear program (of arbitrary form) and finding the w-center of a polyhedral system. Both problems are instances of a canonical optimization problem involving a weighted logarithmic barrier function. To solve a linear program, we minimize a weighted potential function, and proceed as in other projective transformation algorithms for linear programming. The advantages of our method are twofold. First, the algorithm is completely general regarding the format of the linear program, and so naturally accommodates equality and inequality constraints, upper and lower bounds, etc. Second, it works with a weighted potential function, that intrinsically rescales the problem in favor of the constraints with the largest weights. Thus, if the user has any prior judgements regarding the likelihood of particular constraints being active in the optimal solution, this judgement can be easily and systematically incorporated into the formulation of the potential function, and hence into the algorithm itself. The algorithm and analysis for the problem of solving the w-center of a polyhedral system is presented in Part II of this study.

Keywords: analytic center, w-center, projective transformation, linear program, ellipsoid, barrier penalty method.

I.   Introduction

The purpose of this two-paper study is to broaden the scope of projective methods in mathematical programming, both in terms of theory and algorithms. We start by generalizing the concept of the analytic center of a polyhedral system to the w-center of a polyhedral system, which stands for weighted center, where there are positive weights on the logarithmic barrier function for each inequality constraint defining a polyhedron X . We prove basic results regarding contained and containing ellipsoids centered at the w-center of the system X . We next shift our attention to projective transformations for transforming the polyhedron X to another polyhedron Z that turns the current point $\bar{x}$ into the w-center of the new polyhedron Z . We work throughout with a polyhedron X that contains both inequality and equality constraints of arbitrary format. We exhibit an elementary projective transformation that transforms the current point $\bar{x}$ to the w-center of Z .

This theory is then applied to two different problems: solving a linear program (of arbitrary form) and finding the w-center of a polyhedral system. Both problems are instances of a canonical optimization problem involving a weighted logarithmic barrier function. To solve a linear program, we minimize a weighted potential function, and proceed as in other projective transformation algorithms for linear programming. The advantages of our method are twofold. First, the algorithm is completely general regarding the format of the linear program, and so naturally accommodates equality and inequality constraints, upper and lower bounds, etc. Second, it works with a weighted potential function, that intrinsically rescales the problem in favor of the constraints with the largest weights. Thus, if the user has any prior judgements regarding the likelihood of particular constraints being active in the

optimal solution, this judgement can be easily and systematically incorporated into the formulation of the potential function, and hence into the algorithm itself.

In Part II of this study, the basic theory of Part I is applied to the problem of finding the w-center of a polyhedral system X . We present a projective transformation algorithm, analagous but more general than Karmarkar's algorithm, for finding the w-center of X . The algorithm exhibits superlinear convergence. At each iteration, the algorithm either improves the objective function (the weighted logarithmic barrier function) by a fixed amount or at a linear rate of improvement. This linear rate of improvement increases to unity, and so the algorithm is superlinearly convergent. At each iteration, the algorithm either detects unboundedness, or updates an upper bound on the optimal objective value of the weighted logarithmic barrier function. The direction chosen at each iteration is shown to be positively proportional to the projected Newton direction. This has two consequences. On the theoretical side, this broadens a result of Lagarias regarding the connection between projective transformation methods and Newton's method. In terms of algorithms it means that our algorithm specializes to Vaidya's algorithm if it is used with a line search, and so we see that Vaidya's algorithm is superlinearly convergent as well. Finally, we show how to use the algorithm to construct well-scaled containing and contained ellipsoids centered at near-optimal solutions to the w-center problem. After a fixed number of iterations, the current iterate of the algorithm can be used as an approximate w-center, and one can easily construct well-scaled containing and contained ellipsoids centered at the current iterate, whose scale factor is of the same order as for the w-center itself.

The W-Center of a Polyhedral System

In [13] , Karmarkar simultaneously introduced ideas regarding the center of a polyhedral system, a projective transformation that centers a given point, and a linear programming algorithm that uses this methodology to decrease a potential function involving an objective function component and a centering component. Karmarkar's ideas have since been generalized along a number of lines, both theoretical and computational. Herein, we expand on Karmarkar's methodology in at least two ways. First, we analyze the w-center of a polyhedral system

$\mathcal{X} = \left\{ x \in R^n \mid Ax \leq b, Mx = g \right\}$, defined as the the solution $\bar{x}$ to the following optimization problem:

$$P_w: \quad \text{maximize} \quad \sum_{i=1}^{m} w_i \ln s_i$$

$$\text{s.t.} \quad Ax + s = b$$

$$s > 0$$

$$Mx = g.$$

Note that P_w is a generalization of the analytic center problem first analyzed by Sonnevend [20], [21]. Also note the P_w is defined for the most general polyhedral representation, namely inequality as well as equality constraints of arbitrary form. In P_w , the weights w_i can be arbitrary positive scalars, and for convenience they are normalized so that $\sum_{i=1}^{m} w_i = 1.$

Let $\bar{w} = \min_i \left\{ w_i \right\}$. The main result for the w-center problem is that if $\bar{x}$ is the w-center, then there exist well-scaled containing and contained ellipsoids at $\bar{x}$ as follows. Let $\mathcal{X} = \left\{ x \in R^n \mid Ax \leq b, Mx = g \right\}$. Then there exist ellipsoids E_{IN}

and E_{OUT} centered at $\bar{x}$, for which $E_{IN} \subset \mathcal{X} \subset E_{OUT}$, and $(E_{IN} - \bar{x}) =$ $(\bar{w}/(1 - \bar{w}))(E_{OUT} - \bar{x})$, i.e., the inner ellipse is a scaled copy of the outer ellipse, with scaling factor $\bar{w}/(1 - \bar{w})$. When all weights are identical, $w = (1/m)\,e$ and $\bar{w} = 1/m$, and the scaling factor is $\bar{w}/(1 - \bar{w}) = 1/(m-1)$. Essentially, the scaling factor $\bar{w}/(1 - \bar{w})$ is (almost exactly) proportional to the smallest (normalized) weight w_i .

Projective W-Centering for Problems in Arbitrary Form

Numerous researchers have extended Karmarkar's projective transformation methodology, and and this study broadens this methodology as well. Gay [9] has shown how to apply Karmarkar's algorithm to problems in standard form, and how to process inequality constraints by implicitly converting them to standard form. Later, Gay [10] shows how to process problems in standard form with upper and lower bounds, as does Rinaldi [18]. Bayer and Lagarias [4] have added to the theorectical foundations for linear programming by showing that for inequality constrained problems, there exists a class of projective transformations for centering a polyhedron about a given interior point $\bar{x}$. However, their result is not constructive. Anstreicher [2] has shown a different methodology for processing problems in standard form, and in [7] the author gives a simple projective transformation that constructively demonstrates the result of Bayer and Lagarias. Even though linear programs in any one form (e.g., standard primal form) can be either linearly or projectively transformed into another form, such transformations can be computationally bothersome and awkward, and lack aesthetic appeal. Herein, we work thoughout with the most general polyhedral system, namely $\mathcal{X} = \{x \in R^n \mid Ax \le b, Mx = g\}$. It obviously contains all of the above as special cases, without transformations, addition or elimination of variables, etc. In sections

III and IV of this paper, we present an elementary projective transformation that projectively transforms a general polyhedral system

$X = \{x \in R^n \mid Ax \leq b, Mx = g\}$ to an equivalent system

$Z = \{z \in R^n \mid \tilde{A}x \leq \tilde{b}, Mx = g\}$, and that results in a given point $\bar{x}$ (in the relative interior of X) being the w-center of the polyhedral system Z . The approach taken is based on classical polarity theory for convex sets, see Rockafellar [19] and Grünbaum [12].

A Canonical Optimization Problem

The results on the w-center problem are applied to the following canonical optimization problem:

$$P_{q,p}: \underset{x,\,s}{\text{minimize}} \quad F_{q,p}(x) = \ln(q - p^T x) - \sum_{i=1}^{m} w_i \ln (b_i - A_i x)$$

$$\text{s.t.} \qquad Ax + s = b$$

$$s > 0$$

$$Mx = g$$

$$p^T x < q$$

where $X = \{x \in R^n \mid Ax \leq b, Mx = g\}$ is given. Note that problem $P_{q,p}$ has two important special cases: linear programming and the w-center problem itself. If $p = c$ is the objective function of a linear program maximization problem defined on $X = \{x \in R^n \mid Ax \leq b, Mx = g\}$, and if q is an appropriate upper bound on the optimal objective function value, then $P_{q,p}$ is just the problem of minimizing Karmarkar's potential function (generalized to nonuniform weights w_i on the constraints). If $p = 0$ and $q = 1$, then $P_{q,p}$ is just the w-center problem P_w . In

section V of this paper, we present a local improvement algorithm, completely analagous and a generalization of Karmarkar's algorithm, for solving $P_{q,p}$.

A General Linear Programming Algorithm

In Section VI of this paper, the previous results are specialized to linear programming. Herein, we solve the problem

LP : maximize $c^T x$

s.t. $Ax \leq b$

$Mx = g$

by means of weighted potential functions of the form

$$\ln(U - c^T x) - \sum_{i=1}^{m} w_i \ln(b_i - A_i x)$$

where U is a current upper bound on the optimal objective value, and the weights w_i are prespecified. The methodology for updating U is an extension of Todd and Burrell [22]. It is shown that the number of iterations of the algorithm is $O((1/\overline{w})L)$, where L is the size of the instance of the linear program, and $\overline{w} = \min_i \{w_i\}$. Thus if $w = (1/m) e$, the worst-case bound is $O(mL)$, completely analagous to other primal projective transformation algorithms for linear programming.

The main contributions of this algorithm are twofold. First, it is completely general and can readily be implemented for problems in arbitrary form, without the addition or elimination of variables, constraints, etc. Second, it is general in terms of the weights w_i . These weights can be chosen beforehand to reflect a user's prior

judgement regarding the relative likelihood that a given constraint will be active (or lie near) the optimal solution set. The ability to choose values of weights beforehand has a negative and a positive consequence. On the negative side, the worst case bound on the number of iterations is increased by choosing widely varying weights, which will force $1/\overline{w} >> m$. On the positive side, the weights are a means to systematically and permanently rescale the problem so that constraints of relative prior-judged importance are accorded that importance. This permanent rescaling can be seen best as follows: in the absence of equality constraints, the direction at a give iteration of the algorithm is computed using the matrix $Q = A^T \overline{S}^{-1} W \overline{S}^{-1} A$, where $\overline{S}^{-1}$ is the diagonal matrix of inverses of the slack variables $\overline{s} = b - A \overline{x}$ for the current point $\overline{x}$, and W is the diagnonal matrix with diagonal entries $w_1, ..., w_m$. The weights therefore scale the contribution of the matrix $A_i^T A_i / \overline{s}_i^2$ to the matrix Q, wehre A_i is the ith row of A. This is directly analagous to rescaling the problem in an affine scaling algorithm, see [6], [3], [25]. Karmarkar's algorithm and other algorithms based on centers implicitly assign equal weight and likelihood to any constraint being active in the optimal solution. The methodology developed herein is designed to add flexibility in this assignment.

An Algorithm for the W-Center Problem

Part II of this study [8] applies the methodology and theory regarding the w-center, projecting to the w-center, and the local improvement algorithm for the canonical optimization problem $P_{q,p}$, to an algorithm to solve the w-center problem P_w. Other algorithms for this problem have been developed by Censor and Lent [5] and by Vaidya [23]. We present a projective transformation algorithm for finding the w-center that uses the exact methodology used for problem $P_{q,p}$ and

for linear programming. The algorithm can be run with either a fixed steplength (in the projective space) or with a line search.

This algorithm produces upper bounds on the optimal objective value at each iteration, and these bounds are used to prove that the algorithm is superlinearly convergent. We also show that the direction chosen at each iteration is proportional to the projected Newton direction. Thus, if the algorithm is run with a line search, it specializes to Vaidya's algorithm. Although Vaidya has shown that his algorithm exhibits linear convergence, our approach and analysis demonstrate that his algorithm is actually superlinearly convergent, verifying a conjecture of Vaidya [24] that his algorithm might exhibit stronger convergence properties.

We also show that after a fixed number of iterations of the algorithm, that one can construct "well-scaled" containing and contained ellipsoids at the current iterate of the algorithm. If $\bar{x} \in \mathcal{X} = \{x \in R^n \mid Ax \leq b, Mx = g\}$ is the current iterate, one can easily construct ellipsoids F_{IN} and F_{OUT} centered at $\bar{x}$, with the property that $F_{IN} \subset \mathcal{X} \subset F_{OUT}$, and $(F_{OUT} - \bar{x}) = (1.75/\bar{w} + 5) (F_{IN} - \bar{x})$. When all weights are identical, then this scale factor is $(1.75m + 5)$ which is $O(m)$. In general, the order of this scale factor is $O(1/\bar{w})$, which is the same as for the ellipses E_{IN} and E_{OUT} centered at the optimal solution to P_w, whose scale factor is $(1 - \bar{w})/\bar{w} = 1/\bar{w} - 1$.

Notation

Throughout this paper, A is an $m \times n$ matrix with $m \geq n$ and M is a $k \times n$ matrix with $k \leq n$. If s or w are m–vectors, then S or W refer to diagonal $m \times m$ matrices whose diagonal entries are the corresponding components of s or w.

Let e   denote the vector of ones of appropriate dimension, and let e_i denote the ith

unit vector.

II.   The W-Center of a Polyhedral System

For given data (A, b, M, g), let $X = \{x \in R^n \mid Ax \le b, Mx = g\}$. X then is a

polyhedron, bounded or not, lying in R^n. Let

$(X; S) = \{(x,s) \in R^n \times R^m \mid Ax + s = b, s \ge 0, Mx = g\}$ and

$S = \{s \in R^m \mid s \ge 0, s = b - Ax \text{ for some } x \text{ satisfying } Mx = g\}$. We will often refer to S

as the slack space of X or the slack space corresponding to X.

We define int X as the interior of X relative to the affine space

$\{x \in R^n \mid Mx = g\}$, and we define int $(X; S) =$

$\{(x,s) \in R^n \times R^m \mid Ax + s = b, s > 0, Mx = g\}$ and int S analogously.

Note that if X is bounded, so is S; and if S is bounded, either X is bounded or

contains a line. In either case, we say that X has bounded slack. If A has full rank, X

is bounded if and only if S is bounded.

Let $w \in R^m$ be a vector such that $w > 0$ and w has been normalized so that

$e^T w = 1$.

Consider the problem

P_w: maximize $\sum_i w_i \ln s_i$

s.t. $Ax + s = b$

$Mx = g$

$s > 0$.

This problem is a (weighted) generalization of the analytic center problem, posed by Sonnevend [20], [21], and used extensively in interior point algorithms for solving linear programming problems, see Renegar [17], Gonzaga [11], and Monteiro and Adler [15], [16], among others.

Under the assumption that X is bounded and int $\mathsf{X} \neq \emptyset$, then P_w will have a unique solution, $\overline{x}$, which we denote as the *w-center* of the polyhedral system X. To be more precise, we should say that $\overline{x}$ is the w-center of the linear system defined by (A, b, M, g), since the solution $\overline{x}$ to P_w is a function of the particular polyhedral representation of X as the intersection of halfspace and hyperplanes, and not just of the set X. However, it will be convenient in terms of notation to refer to $\overline{x}$ as the w-center of X, as long as it is understood that X represents a specific intersection of halfspaces and hyperplanes. The Karush-Kuhn-Tucker (K-K-T) conditions are necessary and sufficient for optimality in P_w, and thus $\overline{x}$ is the w-center of X if and only if $\overline{x}$ satisfies

(2.1a) $A \overline{x} + \overline{s} = b$

(2.1b) $M \overline{x} = g$ (2.1)

(2.1c) $\overline{s} > 0$

(2.1d) $\overline{w}^T \overline{S}^{-1} A = \overline{\pi}^T M$ for some $\overline{\pi} \in R^k$.

Let $\overline{w} = \min_i \{w_i\}$ be the smallest component of w. Note $\overline{w} \leq 1/m$ because of the normalization condition $e^T w = 1$. Generalizing Sonnevend [20], [21], we have the following properties of the w-center of X, that characterize inner and outer ellipsoids centered at $\overline{x}$.

Theorem 2.1. Let $X = \{x \in R^n \mid Ax \le b, Mx = g\}$, and let $\bar{x}$ be the w-center of X. Let $E_{IN} = \{x \in R^n \mid Mx = g, (x - \bar{x})^T A^T \bar{S}^{-1} W \bar{S}^{-1} A(x - \bar{x}) \le \bar{w}/(1 - \bar{w})\}$

and $E_{OUT} = \{x \in R^n \mid Mx = g, (x - \bar{x})^T A^T \bar{S}^{-1} W \bar{S}^{-1} A(x - \bar{x}) \le (1 - \bar{w})/\bar{w}\}$.

Then $E_{IN} \subset X \subset E_{OUT}$.

Before proving this theorem, we make the following remark:

Remark 2.1. $(E_{IN} - \bar{x}) = (\bar{w}/(1 - \bar{w}))(E_{OUT} - \bar{x})$, i.e., the inner ellipse is a scaled copy of the outer ellipse, with scaling factor $\bar{w}/(1 - \bar{w})$. If $w = (1/m)e$, then $\bar{w} = 1/m$, and so the scaling factor is $\bar{w}/(1 - \bar{w}) = 1/(m-1)$.

The proof of Theorem 2.1 is aided by the following three propositions:

Proposition 2.1. If $\bar{x}$ is the w-center of X, then $\bar{S}$ lies in the simplex $\Delta_w = \{s \in R^m \mid s \ge 0, w^T \bar{S}^{-1} s = 1\}$

Proof: If $s \in S$, then $w^T \bar{S}^{-1} s = w^T \bar{S}^{-1}(b - Ax)$ for some $x \in X$, and so $w^T \bar{S}^{-1} s = w \bar{S}^{-1}(\bar{s} + A\bar{x} - Ax) = w^T \bar{S}^{-1}\bar{s} + w^T \bar{S}^{-1}A(x - \bar{x})$. From (2.1d), this latter expression equals $w^T \bar{S}^{-1}\bar{s} + \bar{\pi}^T M(x - \bar{x}) = w^T \bar{S}^{-1}\bar{s} = w^T e = 1$, since $M(x - \bar{x}) = g - g = 0$. ∎

Proposition 2.2. Suppose $r \in R^m$ and r satisfies $w^T r = 0$ and

$$r^T W r \le \bar{w}/(1 - \bar{w}).$$ Then $|r_i| \le 1$ for each i.

Proof. It suffices to show that $r_i \le 1$, $i = 1, \ldots, m$. For each i, consider the program

$$\max \ r_i$$
$$\text{s.t.} \quad r^T W r \le w_i/(1 - w_i) \qquad (\alpha)$$
$$w^T r = 0. \qquad (\beta)$$

The optimal solution to this program is

$r^* = (1/(1-w_i))(-w_i e + e_i)$, with K-K-T multipliers

$\alpha = (1-w_i)/(2w_i)$ and $\beta = 1$, which satisfy the K-K-T conditions

$e_i = 2\alpha Wr + \beta w$. Notice that $r_i^* = 1$. Thus if $r^T Wr \leq \bar{w}/(1-\bar{w}) \leq$

$w_i/(1-w_i)$ and $w^T r = 0$, then $r_i \leq 1$. ∎

Proposition 2.3. Let $\bar{x}$ be the w-center of $\mathcal{X}$. If $s \in R^m$ satisfies $w^T \bar{S}^{-1}s = 1$ and $(s-\bar{s})^T \bar{S}^{-1}W \bar{S}^{-1} (s-\bar{s}) \leq \bar{w}/(1-\bar{w})$, then $0 \leq s_i \leq 2\bar{s}_i$.

Proof. Let s be as given in the proposition. Let $r = \bar{S}^{-1}(s-\bar{s})$. Then r satisfies the hypotheses of Proposition 2.2, and hence $|r_i| \leq 1$, $i = 1, \ldots, m$. Thus $0 \leq s_i \leq 2\bar{s}_i$, $i = 1, \ldots, m$. ∎

Proof of Theorem 2.1. We first prove that $\mathcal{X} \subset E_{OUT}$. By Proposition 2.1, $S \subset \Delta_w$. The extreme points of Δ_w are $(\bar{s}_i/w_i) e_i$, $i = 1, \ldots, m$. Note that each extreme point satisfies $((\bar{s}_i/w_i) e_i - \bar{s})^T \bar{S}^{-1}W \bar{S}^{-1} ((\bar{s}_i/w_i) e_i - \bar{s}) = (1-w_i)/w_i \leq (1-\bar{w})/\bar{w}$.

Thus, by the convexity of Δ_w , every $s \in S$ satisfies $(s-\bar{s})^T \bar{S}^{-1}W \bar{S}^{-1} (s-\bar{s}) \leq \bar{w}/(1-\bar{w})$. But $(s-\bar{s}) = -A(x-\bar{x})$, so $(x-\bar{x})A^T \bar{S}^{-1}W \bar{S}^{-1} A(x-\bar{x}) \leq \bar{w}/(1-\bar{w})$. This shows that $\mathcal{X} \subset E_{OUT}$.

We next show that $E_{IN} \subset \mathcal{X}$. Let $x \in E_{IN}$, and let s be the slack corresponding to x, i.e., $s = b - Ax$. Then $(s-\bar{s})^T \bar{S}^{-1}W \bar{S}^{-1} (s-\bar{s}) = (x-\bar{x})A^T \bar{S}^{-1}W \bar{S}^{-1} A(x-\bar{x}) \leq \bar{w}/(1-\bar{w})$.

Also, similar to Proposition 2.1, it is straightforward to show that

$w^T \bar{S}^{-1} s = 1$. Thus by Proposition 2.3, $s \geq 0$. Thus $Ax \leq b$, and since $x \in E_{IN}$, $Mx = g$. Thus $x \in X$. ∎

Proposition 2.4. Let $\bar{x}$ be the w–center of X. For each $i = 1, \ldots, m$, for any $x \in X$, $(b_i - A_i x) \leq \bar{s}_i / w_i$.

Proof: For any $x \in X$, let $s = b - Ax$. By Proposition 2.1, $w^T \bar{S}^{-1} s = 1$, $s \geq 0$, so $s_i \leq (\bar{s}_i / w_i)$, i.e., $b_i - A_i x \leq (\bar{s}_i / w_i)$ ∎

Remark 2.2. Assume that X is bounded and int $X \neq \emptyset$. Given $w > 0$, the w–center of X is unique. However, for a given vector $\bar{x}$, there will be many weight vectors w such that $\bar{x}$ is the w–center. Let $\bar{s}$ be the slack at $\bar{x}$. If $\bar{s} > 0$, then

$$\left\{ w \in R^m \mid w > 0, \text{ and } \bar{x} \text{ is the w–center of } X \right\} =$$
$$\left\{ w \in R^m \mid w > 0, \text{ and } w^T \bar{S}^{-1} A = \pi^T M \text{ for some } \pi \in R^k \right\}.$$

The above remark serves as a basis for the following intriguing composition theorem.

Theorem 2.2. Let $X = \left\{ x \mid Ax \leq b, Mx = g \right\}$ and assume X is bounded and int $X \neq \emptyset$, and let $\mathcal{X} = \left\{ \lambda \in R^m \mid A^T \lambda = M^T \pi \text{ for some } \pi \in R^k, \lambda > 0 \right\}$.

Then for any $w \in R^m$, $w > 0$, there is a unique slack vector $\bar{s} \in S$ and a unique $\bar{\lambda} \in \mathcal{X}$ such that $w_i = \bar{\lambda}_i \bar{s}_i$, $i = 1, \ldots, m$.

Proof. For a given $w \in R^m$, $w > 0$, let $\bar{x}$ be the (unique) w–center of X and $\bar{s}$ its corresponding slack. Then upon setting $\bar{\lambda} = \bar{S}^{-1} w$, we have by (2.1) that $\bar{\lambda}^T A = \pi^T M$ for some $\pi \in R^k$. ∎

The last result of this section characterizes the behavior of the weighted–logarithmic barrier function $\sum_{i=1}^{m} w_i \ln (b_i - A_i x)$ near the w–center of $\mathcal{X}$. This lemma parallels similar results for the uniformly weighted center in Karmarkar [13] and Vaidya [23] .

Lemma 2.1. Let $\bar{x}$ be the w–center of $\mathcal{X}$, let $\bar{s} = b - A \bar{x}$, and let $d \in R^n$ be a direction that satisfies $Md = 0$, and $d^T A^T \bar{S}^{-1} W \bar{S}^{-1} Ad \le \bar{w}/(1 - \bar{w})$. Then for all α satisfying $0 \le \alpha < 1$,

$$\sum_{i=1}^{m} w_i \ln (b_i - A_i(\bar{x} + \alpha d)) \ge \sum_{i=1}^{m} w_i \ln(\bar{s}_i) - \frac{\alpha^2}{2(1 - \alpha)} \left(\frac{\bar{w}}{(1 - \bar{w})} \right)$$

The proof of Lemma 2.1 makes use of the following inequality, repeated here as

Proposition 2.5. If $|\varepsilon| \le \alpha < 1$, then $\ln (1 + \varepsilon) \ge \varepsilon - \frac{\varepsilon^2}{2(1 - \alpha)}$.

Proof: The Taylor series for $\ln (1 + \varepsilon)$ yields

$$\ln (1 + \varepsilon) = \varepsilon - \sum_{j = 2}^{\infty} \frac{(-\varepsilon)^j}{j} \ge \varepsilon - \sum_{j = 2}^{\infty} \frac{|\varepsilon|^j}{j} \ge \varepsilon - \sum_{j = 2}^{\infty} \frac{|\varepsilon|^j}{2}$$

$$= \varepsilon - \frac{|\varepsilon|^2}{2(1 - |\varepsilon|)} \ge \varepsilon - \frac{\varepsilon^2}{2(1 - \alpha)} \quad . \blacksquare$$

Proof of Lemma 2.1: Let $r = \bar{S}^{-1} Ad$. Then $w^T r = w^T \bar{S}^{-1} Ad = \pi^T Md = 0$ for some $\pi \in R^k$ by (2.1d). Furthermore $r^T W r \le \bar{w}/(1 - \bar{w})$. Thus by Proposition 2.2, $|r_i| \le 1$, $i = 1, \ldots , m$.

Now, $\sum_{i=1}^{m} w_i \ln (b_i - A_i(\bar{x} + \alpha d)) = \sum_{i=1}^{m} w_i \ln (\bar{s}_i (1 - \alpha r_i))$

$$= \sum_{i=1}^{m} w_i \ln (\bar{s}_i) \quad + \sum_{i=1}^{m} w_i \ln (1 - \alpha r_i)$$

$$\geq \sum_{i=1}^{m} w_i \ln (\bar{s}_i) + \sum_{i=1}^{m} w_i (-\alpha r_i) - \sum_{i=1}^{m} w_i \frac{(\alpha r_i)^2}{2(1 - \alpha)} \quad \text{(by Proposition 2.5)}$$

$$= \sum_{i=1}^{m} w_i \ln \bar{s}_i \quad - \alpha w^T r \quad - \frac{\alpha^2 r^T W r}{2(1 - \alpha)}$$

$$\geq \sum_{i=1}^{m} w_i \ln \bar{s}_i \quad - \frac{\alpha^2}{2(1 - \alpha)} \frac{\bar{w}}{(1 - \bar{w})} . \quad \blacksquare$$

III. Projective Transformations

The approach to projective transformations developed herein is based on polars of convex sets, see e.g. Rockafellar [19] and Grünbaum [12]. In order to motivate and clarify the exposition, we will (for the moment) assume that $X = \{x \in R^n \mid Ax \leq b\}$ has an interior and is bounded (and hence, in the notation of the paper, that (M, g) is nil).

Let $\bar{x}$ be a given element of int X, and $\bar{s} = b - A\bar{x}$. Then the polar of $(X - \bar{x})$ is given by $Y = \{y \in R^n \mid y^T(x - \bar{x}) \leq 1 \text{ for all } x \in X\}$ and by a theorem of the alternative, this is equivalent to $Y = \{y \in R^n \mid y = A^T \lambda \text{ for some } \lambda \geq 0 \text{ satisfying } \lambda^T \bar{s} = 1\}$. Thus Y consists of nonnegative weighted combinations of rows of A, where the weights λ must satisfy $\lambda^T \bar{s} = 1$. Note also that the set Y should formally be subscripted by $\bar{x}$, since it depends parametrically on $\bar{x}$ through the slack $\bar{s}$. However, for notational convenience, we will not use this subscript, except where the choice of $\bar{x} \in$ int X is not clear from the context.

The following properties of Y can readily be established, see [19] or [12].

(3.1a) Y is a combinatorial dual of X.

(3.1b) $\mathcal{Y}$ is bounded.

(3.1c) $\mathcal{Y}$ contains the origin in its interior.   (3.1)

(3.1d) For any $y \in \text{int}\, \mathcal{Y}$, $y^T(x - \bar{x}) < 1$ for all $x \in \mathcal{X}$.

(3.1e) $\text{int}\,\mathcal{Y} = \left\{y \mid y = A^T\lambda \text{ for some } \lambda > 0 \text{ satisfying } \lambda^T \bar{s} = 1\right\}$.

Property (3.1e) gives a convenient way to generate points in $\text{int}\, \mathcal{Y}$, namely by taking strictly positive combinations λ of rows of A, for which $\lambda > 0$ and $\lambda^T \bar{s} = 1$.

We now consider a projective transformation of $\mathcal{X}$ by using a vector y in the interior of $\mathcal{Y}$, as follows. Let $\bar{x} \in \text{int}\, \mathcal{X}$ be given, and $y \in \text{int}\, \mathcal{Y}$ be given. Let the projective transformation $g(\cdot) : \mathcal{X} \to R^n$ be defined by

$$z = g(x) = \bar{x} + \frac{x - \bar{x}}{1 - y^T(x - \bar{x})} , \qquad \text{for all } x \in \mathcal{X}.$$

Note that because $y^T(x - \bar{x}) < 1$ for all $x \in \mathcal{X}$ (because $y \in \text{int}\, \mathcal{Y}$), then $g(\cdot)$ is well–defined.

To see that $g(\cdot)$ is a projective transformation, rewrite $g(\cdot)$ as

$$g(x) = \frac{(I - \bar{x}y^T)x + (\bar{x}y^T \bar{x})}{(-y^T)x + (1 + y^T \bar{x})}$$

Also note that $g(\bar{x}) = \bar{x}$, i.e. $g(\cdot)$ leaves $\bar{x}$ fixed, and that $g(\cdot)$ preserves directions from $\bar{x}$, i.e., if $\bar{x} + \alpha d \in \mathcal{X}$, then

$$g(\bar{x} + \alpha d) - g(\bar{x}) = \gamma d \text{ for some } \gamma > 0.$$

Let $\mathcal{Z}$ be the polyhedron

$$\mathcal{Z} = \left\{z \in R^n \mid (A - \bar{s}y^T)z \leq b - \bar{s}y^T \bar{x}\right\} \text{ and}$$

define $(\mathcal{Z}\,;\mathcal{T}) = \left\{(z, t) \in R^n \times R^m \mid (A - \bar{s}y^T \bar{x})z + t = b - \bar{s}y^T \bar{x} , t \geq 0\right\}$

and $\mathcal{T}$ analogously. The following lemma shows the equivalence of $(\mathcal{X};\mathcal{S})$ and $(\mathcal{Z};\mathcal{T})$ under the projective transformation $g(\cdot)$.

Lemma 3.1. Let $\mathcal{X} = \{x \in R^n \mid Ax \le b\}$ be bounded, and let $\bar{x} \in int\,\mathcal{X}$ be given, and $y \in int\,\mathcal{Y}$ be given. Let $Z = \{z \in R^k \mid (A - \bar{s}y^T)z \le b - \bar{s}y^T\bar{x}\}$. Then

i) The projective transformation $g(x) = \bar{x} + \dfrac{x - \bar{x}}{1 - y^T(x - \bar{x})}$ maps $\mathcal{X}$ onto $\mathcal{Z}$, $g(\cdot)$ maps the faces of $\mathcal{X}$ onto faces of $\mathcal{Z}$, and $\mathcal{Z}$ is of the same combinatorial type as $\mathcal{X}$.

ii) The inverse of $g(\cdot)$, given by $x = h(z) = \bar{x} + \dfrac{z - \bar{x}}{1 + y^T(z - \bar{x})}$, maps $\mathcal{Z}$ onto $\mathcal{X}$.

Proof. Because $y \in int\,\mathcal{Y}$, $y^T(x - \bar{x}) < 1$ for all $x \in \mathcal{X}$, so $g(\cdot)$ is well–defined. If $z = g(x)$, then

$$(A - \bar{s}y^T)z = A\bar{x} - \bar{s}y^T\bar{x} + \frac{Ax - \bar{s}y^Tx - A\bar{x} + \bar{s}y^T\bar{x}}{1 - y^T(x - \bar{x})}$$

$$\le A\bar{x} - \bar{s}y^T\bar{x} + \frac{b - \bar{s}y^Tx - A\bar{x} + \bar{s}y^T\bar{x}}{1 - y^T(x - \bar{x})}$$

$$= A\bar{x} - \bar{s}y^T\bar{x} + \bar{s} = b - \bar{s}y^T\bar{x},$$

so $z \in \mathcal{Z}$. Furthermore, the inequality in constraint i of $Ax \le b$ is satisfied strictly or not if and only if the same inequality is satisfied in $(A - \bar{s}y^T)z \le b - \bar{s}y^T\bar{x}$. It is straightforward to show that $h(\cdot)$ is the inverse of $g(\cdot)$ and both maps then are onto. ∎

Remark 3.1. If we use the symbol P° to denote the polar of a set P, then $\mathsf{y} = (\mathsf{X} - \bar{x})^\circ$. Then Lemma 3.1 is a constructive representation of the fact that the translation of the polar X° of X is equivalent to a projective transformation of the set X, see Grünbaum [12]. One can easily verify in Lemma 3.1 that $\mathsf{y} = (\mathsf{X} - \bar{x})^\circ$, and $\mathsf{Z} - \bar{x}$ is the polar of the translated polar $\left((\mathsf{X} - \bar{x})^\circ - y \right)$, i.e.,

$$\mathsf{Z} = \bar{x} + \left[(\mathsf{X} - \bar{x})^\circ - y \right]^\circ .$$

Lemma 3.1 easily extends to the more general case of X, i.e.,

$$\mathsf{X} = \left\{ x \in R^n \mid Ax \leq b, \ Mx = g \right\}, \text{ where } \mathsf{X} \text{ is not necessarily bounded.}$$

Define

$$(\mathsf{Z}; \mathsf{T}) = \left\{ (z, t) \in R^n \times R^m \mid (A - \bar{s}y^T)z + t = b - \bar{s}y^T \bar{x}, \ t \geq 0, \ Mz = g \right\} \tag{3.2}$$

with Z and T defined analogously.

The maps $g(\cdot)$ and $h(\cdot)$ can be extended to map between $(\mathsf{X}; \mathsf{S})$ and $(\mathsf{Z}; \mathsf{T})$ as follows. For $(x; s) \in (\mathsf{X}; \mathsf{S})$ define

$$(z, t) = g(x; s) = \left(\bar{x} + \frac{x - \bar{x}}{1 - y^T(x - \bar{x})} ; \ \frac{s}{1 - y^T(x - \bar{x})} \right) \in (\mathsf{Z}; \mathsf{T}). \tag{3.3}$$

For $(z; t) \in (\mathsf{Z}; \mathsf{T})$, define

$$(x, s) = h(z; t) = \left(\bar{x} + \frac{z - \bar{x}}{1 + y^T(z - \bar{x})} ; \ \frac{t}{1 + y^T(z - \bar{x})} \right) \in (\mathsf{X}; \mathsf{S}). \tag{3.4}$$

If $g(\cdot)$ and $h(\cdot)$ are given by (3.3) and (3.4), we will refer to y as the **projection parameter** for $g(\cdot)$ and $h(\cdot)$, and will refer to $g(\cdot)$ and $h(\cdot)$ as the projective transformation **induced** by y. It is straightforward to verify that $g(x; s) \in (\mathsf{Z}; \mathsf{T})$ and $h(z, t) \in (\mathsf{X}; \mathsf{S})$. We shall refer to $g(x), g(s), h(z), h(t)$ as specific vector components of the above maps.

In the case when X is bounded, Lemma 3.1 directly extends to:

Theorem 3.1. Let $\mathsf{X} \in \{x \in R^n \mid Ax \leq b, Mx = g\}$ be a bounded polyhedron. Let $\bar{x} \in \text{int}\,\mathsf{X}$ be given, and let $\bar{s} = b - A\,\bar{x}$. Let $y \in R^n$ be chosen so that

$y = A^T\lambda$ for some $\lambda > 0$ satisfying $\lambda^T \bar{s} = 1$. Let $(Z;\mathcal{T})$, $g(\cdot)$, and $h(\cdot)$ be defined as in (3.2), (3.3), (3.4). Then:

    i)   The transformation $g(x;s)$ is well-defined for all $(x\,;s) \in (\mathsf{X};\mathsf{S})$.

    ii) $g(\cdot;\cdot)$ maps $(\mathsf{X};\mathsf{S})$ onto the set $(Z;\mathcal{T})$, and

        Z and X are of the same combinatorial type.

    iii)   the inverse of $g(\cdot;\cdot)$ is given by $h(\cdot;\cdot)$ and $h(\cdot;\cdot)$ maps $(Z;\mathcal{T})$ onto $(\mathsf{X};\mathsf{S})$. $\blacksquare$

The proof of Theorem 3.1 is the same as that of Lemma 3.1. One need only check that $Mz = g$ if and only if $Mx = g$.

Theorem 3.1 has assumed that the set $\mathsf{X} \in \{x \in R^n \mid Ax \leq b, Mx = g\}$ is bounded. In applying the theorem to polyhedra encountered in practice, this may not be a valid assumption, and one may not even know whether the set X is bounded or not.

Theorem 3.2. Let $\mathsf{X} \in \{x \in R^n \mid Ax \leq b, Mx = g\}$, let $\bar{x} \in \text{int}\,\mathsf{X}$ be given, and let $y \in R^n$ be given such that $y = A^T\lambda$, $\lambda \geq 0$, $\lambda^T \bar{s} = 1$. Let $(Z;\mathcal{T}), g(\cdot)$ and $h(\cdot)$ be defined as in (3.2), (3.3), and (3.4). Then:

i)   The projective transformation $g(\cdot)$ induced by y is well-defined for all $(x; s) \in$ int $(X; S)$.

ii) $g(\cdot)$ maps int X onto int $Z \cap \{z \in R^n \mid -y^T z < 1 - y^T \bar{x}\}$.

iii) The inverse of $g(\cdot)$ is given by $h(\cdot)$, and maps int $Z \cap \{z \in R^n \mid -y^T z < 1 - y^T \bar{x}\}$ onto int X.

Proof. Suppose $y = A^T \lambda$ where $\lambda \geq 0$ and $\lambda^T \bar{s} = 1$. To prove (i) it suffices to show that $y^T(x - \bar{x}) < 1$ for all $x \in$ int X. For any $x \in$ int X, let $s = b - Ax$. Then $s > 0$. $y^T(x - \bar{x}) = \lambda^T A(x - \bar{x}) < \lambda^T b - \lambda^T A \bar{x} = \lambda^T \bar{s} = 1$. To show (ii), note that for any $x \in X$, if $z = g(x)$, then $-y^T z =$

$$-y^T \bar{x} + \frac{-y^T(x - \bar{x})}{1 - y^T(x - \bar{x})} < -y^T \bar{x} + 1, \text{ so that}$$

$g(x) \in \{z \in R^n \mid -y^T z < 1 - y^T \bar{x}\}$. Also if $x \in$ int X, then $t = g(s) > 0$, so that $g(x) \in$ int Z. The proof of (iii) follows by direct substitution. ∎

Remark 3.2. Points in Z that satisfy $-y^T z = 1 - y^T \bar{x}$ correspond to rays of X. To see this, suppose $z \in Z$ and $-y^T z = 1 - y^T \bar{x}$. Then $z \neq \bar{x}$, so that the vector $r = z - \bar{x} \neq 0$. Because $z \in Z$, $(A - \bar{s}y^T)z \leq b - \bar{s}y^T \bar{x}$ and hence $Az \leq b - \bar{s}y^T \bar{x} + \bar{s}y^T z$. Then $Ar = Az - A\bar{x} \leq b - A\bar{x} - \bar{s}y^T \bar{x} + \bar{s}y^T z = \bar{s} - \bar{s}y^T \bar{x} + \bar{s}(-1 + y^T \bar{x}) = 0$. Thus r is a ray of X. ∎

Suppose we are given data $\left(\hat{A}, \hat{b}, \hat{M}, \hat{g}\right)$ and consider the set

$$\hat{X} = \left\{\hat{x} \in R^n \mid \hat{A}\hat{x} \leq \hat{b}, \hat{M}\hat{x} = \hat{g}\right\}.$$

Define A to be the $m \times n$ matrix consisting of $\hat{A}$ followed by a row of zeroes, and define $b = \left(\hat{b}, 1\right)^T$, where m is chosen appropriately. Then

$$X = \left\{x \in R^n \mid Ax \leq b, Mx = g\right\}$$

is equal to $\hat{X}$, and the last constraint of $Ax \le b$ is $0^T x \le 1$ which is satisfied at strict equality for all x. Now define $(X;S)$ and $(Z;T)$ as usual. Then Theorem 3.2 can be strengthened.

Theorem 3.3. Let $X = \{x \in R^n \mid Ax \le b, Mx = g\}$ where the last (i.e., m^{th}) row of (A, b) is $(0; 1)^T$. Let $\bar{x} \in int\, X$, let $y = A^T\lambda$ for some $\lambda > 0$ satisfying $\lambda^T \bar{s} = 1$, and let $(Z;T)$, $g(\cdot)$ and $h(\cdot)$ be defined as in (3.2), (3.3), (3.4). Then

i)   The projective transformation $g(\cdot)$ given by (3.3) is well-defined for all $(x, s) \in (X, S)$.

ii) $g(\cdot)$ maps $int\, X$ onto $int\, Z$ and maps faces of X onto those faces G of Z that do not meet $\{z \in R^n \mid (A - \bar{s}y^T)_m z = (b - \bar{s}y^T \bar{x})_m\}$.

iii) The inverse of $g(\cdot)$ is given by $h(\cdot)$ in (3.4), and $h(\cdot)$ maps $int\, Z$ onto $int\, X$. $h(\cdot)$ maps faces of Z that do not meet $\{z \in R^n \mid (A - \bar{s}y^T)_m z = (b - \bar{s}y^T \bar{x})_m\}$ onto bounded faces of X.

iv)   If $z \in Z$ and $(A - \bar{s}y^T)_m z = (b - \bar{s}y^T \bar{x})_m$, then $r = z - \bar{x}$ is a ray of X.

Proof: (i). If the last row of (A, b) is $(0; 1)^T$, then the last row of the equivalent constraint in $(A - \bar{s}y^T)z \le b - \bar{s}y^T \bar{x}$ is $-y^T z \le 1 - y^T \bar{x}$. To prove (i), note that if $x \in X$, $s = b - Ax$, then $s_m = 1$. Thus $y^T(x - \bar{x}) = \lambda^T A(x - \bar{x}) < \lambda^T b - \lambda A \bar{x} = \lambda^T \bar{s} = 1$, the strict inequality following from the fact that $\lambda > 0$ and $Ax \le b$, $Ax \ne b$ due to the mth constraint. Thus $g(\cdot)$ is well-defined.

(ii)   Let F be a face of X. Let $x \in F$ and $z = g(x)$.

Then $-y^T z = -y^T \bar{x} + \dfrac{-y^T(x - \bar{x})}{1 - y^T(x - \bar{x})} < -y^T \bar{x} + 1$ for all $x \in F$. The mth

constraint of $(A - \bar{s}y^T)z \le (b - y^T \bar{x})$ is just $-y^T z \le 1 - y^T \bar{x}$. Thus no element z of $g(F)$ satisfies $(A - \bar{s}y^T)_m z \le (b - y^T \bar{x})_m$ at equality.

(iii).   Suppose z lies on a face G that does not meet

$$\left\{ z \in R^n \mid (A - \bar{s}y^T)_m z = (b - \bar{s}y^T)_m \right\}.$$

Then for all $z \in G$, $-y^T z < 1 - y^T \bar{x}$, so that $1 + y^T(z - \bar{x}) > 0$, and $h(\cdot)$ is well–defined on G. In order to show that $h(G)$ is a bounded face of X, it suffices to show that Z is bounded.

Let $(z, t) \in (Z; T)$. Then since $y = A^T \lambda$, $\lambda > 0$, and $\lambda^T \bar{s} = 1$, then for any $t \in T$,

$$\lambda^T t = \lambda^T(b - \bar{s}y^T \bar{x}) - \lambda^T(A - \bar{s}y^T)z = \lambda^T(b - \bar{s}y^T \bar{x})$$

$$= \lambda^T(A\bar{x} + \bar{s} - \bar{s}y^T \bar{x}) = \lambda^T \bar{s} = 1, \text{ since } \lambda^T(A - \bar{s}y^T) = 0 \text{ and } \lambda^T \bar{s} = 1.$$

Then the slack corresponding to z is bounded, lying on the simplex $\{t \in R^m \mid t \ge 0, \lambda^T t = 1)$. Hence any ray r of Z would have to satisfy $(A - \bar{s}y^T)r = 0$. The last row of this system says $-y^T r = 0$. Thus $Ar = 0$, and since A is assumed to have full column rank, $r = 0$, and Z contains no rays, so is bounded.

(iv).   This follows from the comments in Remark 3.2. ∎

We now comment on the generality of the projective transformation $g(\cdot)$ presented in (3.3). Although it would appear that $g(\cdot)$ takes on a very restrictive form, this is not really true. The next theorem shows that if $g(\cdot)$ is any projective

transformation of $\mathbb{X}$ into R^n that leaves $\bar{x}$ fixed and preserves directions from $\bar{x}$, then $g(\cdot)$ can be expressed as in (3.3) for a suitable choice of $y = A^T\lambda$ that satisfies $\lambda \geq 0$ and $\lambda^T \bar{s} < 1$.

Theorem 3.4. Let $\mathbb{X} = \{x \in R^n \mid Ax \leq b, Mx = g\}$ be given, and let $\bar{x} \in \text{int } \mathbb{X}$ be given. Let $g(x)$ be a projective transformation from $\mathbb{X}$ into R^n that satisfies

1) $g(\bar{x}) = \bar{x}$, i.e., $g(\cdot)$ leaves $\bar{x}$ fixed,

2) $\displaystyle\lim_{\alpha \to 0} \frac{g(\bar{x} + \alpha d) - g(\bar{x})}{\alpha} = d$, i.e., $g(\cdot)$ preserves directions from $\bar{x}$.

Then there exists $y \in R^n$ such that $y = A^T\lambda$, $\lambda \geq 0$, $\lambda^T \bar{s} < 1$, for which

$$g(x) = \bar{x} + \frac{(x - \bar{x})}{1 - y^T(x - \bar{x})} \quad \text{for all } x \in \mathbb{X}.$$

Proof: If $g(\cdot) : \mathbb{X} \to R^n$, then

$$g(x) = \frac{Gx + h}{-k^T x + l}$$

for some $n \times n$ matrix G, n–vectors h and k, and scalar l. By rescaling, if necessary, we can assume that $-k^T \bar{x} + l = 1$, so that the denominator is of the form $1 - k^T(x - \bar{x})$. Then since $g(\bar{x}) = \bar{x}$, $h = \bar{x} - G\bar{x}$, so that

$$g(x) = \frac{G(x - \bar{x}) + \bar{x}}{-k^T(x - \bar{x}) + 1}.$$

Then $\displaystyle\frac{g(\bar{x} + \alpha d) - g(\bar{x})}{\alpha} = \frac{(G + \bar{x}k^T)d}{1 - \alpha k^T d}$ and

$$\lim_{\alpha \to 0} \frac{(G + \bar{x}k^T)d}{1 - \alpha k^T d} = (G + \bar{x}k^T)d = d$$

implies G can be chosen so that

$$G = (I - \bar{x}k^T)$$

This yields $g(x) = \dfrac{(I - \bar{x}k^T)(x - \bar{x}) + \bar{x}}{1 - k^T(x - \bar{x})} = \bar{x} + \dfrac{x - \bar{x}}{1 - k^T(x - \bar{x})}$

Now, since $g(\cdot)$ must be well-defined, $k^T(x - \bar{x}) < 1$ for all $x \in \mathcal{X}$. By a theorem of the alternative, this implies that $k = A^T\lambda + M^T\pi$ for some $\lambda \geq 0$, with $\lambda^T \bar{s} < 1$, and π unrestricted in sign. Let $y = A^T\lambda$. Then for all $x \in \mathcal{X}$, $k^T(x - \bar{x}) = k^T y$, since $M(x - \bar{x}) = 0$, so that

$$g(x) = \bar{x} + \frac{x - \bar{x}}{1 - y^T(x - \bar{x})} \quad . \blacksquare$$

In Section V, we will consider an optimization problem of the form

$$\text{minimize} \quad F(x) = \ln(q - p^T x) - \sum_{i=1}^{m} w_i \ln s_i$$

$$\tag{3.5}$$

$$\begin{aligned}
\text{s.t.} \quad & Ax + s = b \\
& s > 0 \\
& Mx = g \\
& p^T x < q .
\end{aligned}$$

Let $\bar{x}$ be a given point in int $\mathcal{X}$ and let $\bar{s} = b - A\bar{x}$,

Then if y is chosen so that $y = A^T\lambda$, $\lambda^T \bar{s} = 1$, and $g(\cdot) : (\mathcal{X}; \mathcal{S}) \rightarrow (\mathcal{Z}; \mathcal{T})$ is the projective transformation induced by y and given by (3.3), program (3.5) is transformed to

$$\text{minimize} \quad G(z) = \ln\left([q-(q-p^T\,\bar{x})\,y^T\,\bar{x}] \;-\; [p-(q-p^T\,\bar{x})y]^T z\right) - \sum_{i=1}^{m} w_i \ln t_i$$

$$\text{s.t.} \qquad (A - \bar{s}y^T)z \;+\; t \;=\; (b - \bar{s}y^T\,\bar{x})$$

$$t \;>\; 0 \qquad\qquad (3.6)$$

$$Mz \;=\; g$$

$$[p-(q-p^T\,\bar{x})y]^T z \;<\; [q-(q-p^T\,\bar{x})y^T\,\bar{x}\,]$$

Lemma 3.2.

i) Under the conditions of Theorem 3.1 or Theorem 3.3, programs 3.5 and 3.6 are equivalent. For any $x \in$ int $\mathbf{X}$, $F(x) = G(g(x))$. For any $z \in$ int $\mathbf{Z}$, $G(z) = F(h(z))$.

ii) Under the conditions of Theorem 3.2, for any $x \in$ int $\mathbf{X}$, $F(x) = G(g(x))$. For any $z \in$ int $\mathbf{Z}$ satisfying $-y^T z < 1 - y^T\,\bar{x}$, $G(z) = F(h(z))$.

Proof: Follows from direct substitution. ∎

IV. Projective Transformations to w–center a given Interior Point

Suppose we are given the set $\mathbf{X} = \left\{ x \in R^n \mid Ax \le b, \, Mx = g \right\}$, and a point $\bar{x} \in$ int $\mathbf{X}$, and we wish to find a projection parameter y so that $\bar{x}$ is the w–center of the projectively transformed polyhedron $\mathbf{Z}$.

Theorem 4.1. Let $w > 0$ be an m–vector such that $e^T w = 1$. Let $\mathbf{X} = \left\{ x \in R^n \mid Ax \le b, \, Mx = g \right\}$, let $\bar{x} \in$ int $\mathbf{X}$, let $\bar{s} = b - A\,\bar{x}$, and let $y = A^T\,\bar{S}^{-1}w$. Then $\bar{x}$ is the w–center of $\mathbf{Z} = \left\{ z \in R^n \mid (A - \bar{s}y^T)z \le b - \bar{s}y^T\,\bar{x}, \, Mz = g \right\}$.

Proof: By setting $\lambda = \bar{S}^{-1}w$, the vector y satisfies $y = A^T\lambda$, $\lambda > 0$, and $\lambda^T\,\bar{s} = w^T\,\bar{S}^{-1}\,\bar{s} = w^T e = 1$. Note that $(\bar{x},\,\bar{s}) \in (\mathbf{Z};\mathbf{T})$. We must verify that

conditions (2.1a) – (2.1d) are satisfied in order to assert that $\bar{x}$ is the w–center of Z. (2.1a), (2.1b) and (2.1c) are obviously true. Condition (2.1d) states that $w^T \bar{S}^{-1}(A - \bar{s}y^T) = \pi^T M$ for some $\pi \in R^k$. Let $\pi = 0$. Then $w^T \bar{S}^{-1}(A - \bar{s}y^T) = y - y = 0 = \pi^T M$. Thus $\bar{x}$ is the w–center of Z. ∎

Remark 4.1. In the above theorem, $y = A^T \bar{S}^{-1}w$. Thus the system $(A - \bar{s}y^T) z \le b - \bar{s}y^T \bar{x}$ can be written as

$$(A - \bar{s}w^T \bar{S}^{-1}A)z \le b - \bar{s}y^T \bar{x} \ .$$

The constraint matrix is thus a rank–one modification of A, and need not be explicitly computed. The original inequality system for X is $Ax \le b$ and can be written as $A(x - \bar{x}) \le \bar{s}$; the inequality system for Z can be written as $(I - \bar{s}w^T \bar{S}^{-1}) A(z - \bar{x}) \le \bar{s}$. Thus the latter system in this form is a rank-one modification of the original inequality system.

Next, note that Theorem 4.1 is a constructive form and a generalization of an existence theorem of Lagarias [14], which in our notation asserts the existence of a projective transformation that will result in $\bar{x}$ begin the $(1/m)e$ – center of the transformed polyhedron. Theorem 4.1 is constructive, and covers the more general case of non–equal positive weights w_i , $i = 1 , \ldots , m$, and both inequality and equality constraints. Theorem 4.1 is also a generalization of the projective transformation construction in [7].

Relation of Theorem 4.1 to Karmarkar's Standard Form

The projective transformation development of Section III and this section distinguishes between the polyhedral space X and the polyhedron's slack space S. A given projective transformation $g(\cdot)$ given by (3.3) will map points in X to points in Z, and points in the slack space S of X to points in the slack space T of Z.

The distinction between X and S, and Z and T, becomes important in relating the Theorem 4.1 to Karmarkar's form for solving a linear program.

Consider Karmarkar's original form

$$X = \left\{ x \in R^n \mid \tilde{A}x = 0, \; e^T x = 1, \; x \geq 0 \right\} .$$

Then for every $x \in X$, the corresponding slack S on the inequality contraints is $s = x$, so that $S = X$. If $\bar{x} \in X$ is a point in $\text{int}\, X$, then $\bar{s} = \bar{x} > 0$, and directly adopting Theorem 4.1 with $w = (1/n)\, e$,

$$M = \begin{bmatrix} \tilde{A} \\ e^T \end{bmatrix} , \quad g = \begin{bmatrix} 0 \\ 1 \end{bmatrix} , \quad A = -I, \quad \text{and} \quad b = 0, \quad \text{yields}$$

$$Z = \left\{ z \in R^n \mid \tilde{A}z = 0, \; e^T z = 1, \; \left[-I + (1/n)\, \bar{x} e^T \bar{X}^{-1} \right] z \leq \bar{x} \right\} ,$$

which does not look at all like Karmarkar's transformed space, which we denote by $\mathcal{R}$:

$$\mathcal{R} = \left\{ r \in R^n \mid A\bar{X}r = 0, \; e^T r = 1, \; r \geq 0 \right\} .$$

It is straightforward to verify that the space T of slacks on Z is characterized as:

$$T = \left\{ t \in R^n \mid At = 0, \; e^T \bar{X}^{-1} t = 1, \; t \geq 0 \right\} .$$

Thus $\mathcal{R}$ is just an (affine) rescaling of T, i.e., $r = \bar{X}^{-1} t \in \mathcal{R}$ for any $t = T$ and $t = \bar{X} r \in T$ for any $r \in \mathcal{R}$. In this way, we see that Theorem 4.1 specializes to Karmarkar's projective transformation when viewed in terms of the slack spaces involved.

The next two sections and Part II of this study are devoted to applications of the results in sections II, III, and IV. In section V, we define a canonical optimization problem and present a local improvement algorithm. In section VI, this material is

applied to the linear programming problem. In Part II of this study, this material is applied to the problem of finding the w–center of X , i.e., the point that solves the program:

$$P_w: \quad \text{maximize} \quad \sum_{i=1}^{m} w_i \ln (b_i - A_i x)$$

$$\text{subject to} \quad Ax + s = b$$

$$s > 0$$

$$Mx = g.$$

Section V develops the basic algorithmic step used in the applications in Section VI and in Part II.

V. Local Improvement of a Cononical Optimization Problem

In this section, we consider the one–step improvement algorithm for the problem:

$$P_{q,p}: \quad \underset{x,\,s}{\text{minimize}} \quad F_{q,p}(x) = \ln(q - p^T x) - \sum_{i=1}^{m} w_i \ln (b_i - A_i x)$$

$$\text{s.t.} \quad Ax + s = b \qquad\qquad (5.1)$$

$$s > 0$$

$$Mx = g$$

$$p^T x < q$$

where $X = \{x \in R^n \mid Ax \le b, \, Mx = g\}$ is given, as well as the data q, p , and the weight vector $w > 0$ which satisfies $e^T w = 1$. We make $\underline{no}$ assumptions regarding the data (q, p).

In the spirit of Karmarkar's algorithm, suppose $\bar{x}$ is the w–center of X and that $p^T \bar{x} < q$. Then we can improve the value of $F_{q,p}(x)$ be taking a step in the

direction $\bar{d}$ that maximizes q^Tx over the inner ellipse E_{IN} defined in Section II. The direction $\bar{d}$ is the solution to the following optimization problem:

$$\text{maximize} \quad p^Td$$

$$\text{s.t.} \quad d^TA^T \ \bar{S}^{-1}\overline{W} \ \bar{S}^{-1}Ad \ \leq \ \overline{w}/(1-\overline{w}) \tag{5.2}$$

$$Md = 0 \qquad .$$

If program (5.2) has a unique solution, that solution is given by

$$\bar{d} = \frac{Gp \ \sqrt{\overline{w}/(1-\overline{w})}}{\sqrt{p^TGp}} \tag{5.3}$$

where G is the matrix defined by:

$$Q = A^T \ \bar{S}^{-1}\overline{W} \ \bar{S}^{-1}A \tag{5.4}$$

$$G = \left[Q^{-1} - Q^{-1}M^T(MQ^{-1}M^T)^{-1}MQ^{-1}\right] \tag{5.5}$$

If A has full column rank and M has full row rank, then G is well-defined. It is straightforward to check that G is positive semi-definite and that $p^TGp = 0$ if and only if p lies in the row space of M, i.e., if $p = M^T\pi$ for some $\pi \in R^k$. In this latter case, due to the presumption that $\bar{x}$ is the w-center of X, then $\bar{x}$ is the unique optimal solution to program $P_{q,p}$ given in (5.1). Furthermore, <u>unless</u> $\bar{x}$ is the optimal solution $P_{q,p}$, then the denominator of (5.3) is well-defined, and $\bar{d}$ as given in (5.3) – (5.5) is the unique solution to program (5.2).

The extent of improvement in $F_{q,p}(x)$ by moving from $\bar{x}$ in the direction $\bar{d}$ will depend on the optimal objective value of the program (5.2). This value is given by $p^T \bar{d} = \sqrt{p^TGp} \ \sqrt{\overline{w}/(1-\overline{w})}$, and is proportional to the scalar γ defined below:

$$\gamma = \frac{p^T \bar{d}(1 - \bar{w})/\bar{w}}{(q - p^T \bar{x})} \qquad (5.6)$$

The value of γ will depend on how much of a guaranteed improvement we can expect as we move from $\bar{x}$ in the direction $\bar{d}$. The numerator in (5.6) is $p^T \bar{d}$ times $(1 - \bar{w})/\bar{w}$, a normalization factor. The denominator in (5.6) represents the maximum improvement in $p^T x$ that we could possibly hope for.

Theorem 5.1. Let $\bar{x}$ be the w–center of $\mathbf{X}$, let $\bar{d}$ be the direction that solves (5.2) which is given by (5.3), (5.4), (5.5) when A and M have full rank. Let γ be determined by (5.6), and suppose that $p^T \bar{x} < q$. Then

i)   if $\gamma \geq (1 - \bar{w})/\bar{w}$, then $P_{q,p}$ is unbounded from below.

ii)   if $\gamma < (1 - \bar{w})/\bar{w}$, then

$$F_{q,p}(\bar{x} + \alpha \bar{d}) \leq F_{q,p}(\bar{x}) + \left(\frac{\bar{w}}{1 - \bar{w}}\right)\left[-\gamma\alpha + \frac{\alpha^2}{2(1 - \alpha)}\right] \quad \text{for all } 0 \leq \alpha < 1.$$

Before proving this theorem, we state two immediate consequences.

Corollary 5.1. If $\alpha = 1 - \dfrac{1}{\sqrt{1 + 2\gamma}}$, then

$$F_{q,p}(\bar{x} + \alpha \bar{d}) \leq F_{q,p}(\bar{x}) - \left(\frac{\bar{w}}{1 - \bar{w}}\right)\left(1 + \gamma - \sqrt{1 + 2\gamma}\right)$$

Proof: The value of $\alpha = 1 - \dfrac{1}{\sqrt{1 + 2\gamma}}$ is that value of α which maximizes $\gamma\alpha - \dfrac{\alpha^2}{2(1 - \alpha)}$ over the interval $0 \leq \alpha < 1$. The result follows by direct substitution . ∎

Corollary 5.2. If $p^T \bar{x} < q \leq \max \left\{ p^T x \mid x \in E_{OUT} \right\}$, where E_{OUT} is the ellipse

containing X centered at $\bar{x}$ defined in Theorem 2.1, then $\gamma \geq 1$, and so

$$F_{q,p}(\bar{x} + .42 \, \bar{d}) \leq F_{q,p}(\bar{x}) - (.267) (\bar{w}/(1 - \bar{w})) .$$

Proof: We first prove that $\gamma \geq 1$. Note that by design of program (5.2), its

solution $\bar{d}$ satisfies

$$p^T(\bar{x} + \bar{d}(1 - \bar{w})/ \bar{w}) = \max_{x \in E_{OUT}} p^T x \geq q . \text{ Thus}$$

$$p^T \bar{x} + p^T \bar{d}(1 - \bar{w})/ \bar{w} \geq q , \text{ whereby } \gamma = \frac{p^T \bar{d}}{(q - p^T \bar{x})} \left(\frac{1 - \bar{w}}{\bar{w}} \right) \geq 1 .$$

Next, substituting $\gamma = 1$ and $\alpha = .42$ in Theorem 5.1 (ii), we obtain the desired

conclusion. ∎

(Note that Corollary 5.2 is a slight sharpening of the bound of $1/5$ improvement

if $\alpha = 1/3$ in Todd and Burrell [22] .)

Proof of Theorem 5.1 : Suppose first that $\gamma \geq (1 - \bar{w})/ \bar{w}$. Let $x_\alpha = \bar{x} + \alpha \, \bar{d}$. Because

$\gamma \geq (1 - \bar{w})/ \bar{w}$, then $p^T \bar{d} \geq q - p^T \bar{x}$. Thus $q - p^T x_\alpha = q - p^T \bar{x} - \alpha p^T \bar{d} \leq$

$(q - p^T \bar{x})(1 - \alpha)$. Thus, as $\alpha \to 1$, then $\ln(q - p^T x_\alpha) \to -\infty$. As a consequence of

(5.2) and Theorem 2.1, $x_\alpha \in X$ for all $0 \leq \alpha \leq 1$. In order that $F_{q,p}(\bar{x} + \alpha \, \bar{d})$ be

bounded from below as $\alpha \to 1$, we need $-w_i \ln (b_i - A_i x_\alpha) \to +\infty$ as $\alpha \to 1$ for

each $i = 1, \ldots, m$, i.e., $A \bar{d} = \bar{s}$. We now show that this cannot occur. Note that

since $\bar{x}$ is the w–center of X, then $w^T S^{-1} A = \bar{\pi}^T M$ for some $\bar{\pi} \in R^k$, by (2.1d).

Then, if $A \bar{d} = \bar{s}$, we would have $1 = w^T \bar{S}^{-1} \bar{s} = w^T \bar{S}^{-1} A \bar{d} = \bar{\pi}^T M \bar{d} = 0$, since $M \bar{d} = 0$ (from 5.2), a contradiction. Thus $P_{q,p}$ is unbounded from below.

Next, suppose $\gamma < \dfrac{1 - \bar{w}}{\bar{w}}$, and again let $x_\alpha = \bar{x} + \alpha \bar{d}$ for $0 \le \alpha < 1$. Then

$$q - p^T x_\alpha = (q - p^T \bar{x})\left(1 - \alpha\gamma\left(\frac{\bar{w}}{1 - \bar{w}}\right)\right) \text{ and hence}$$

$$\ln(q - p^T \bar{x}_\alpha) \le \ln(q - p^T \bar{x}) + \ln\left(1 - \alpha\gamma\left(\frac{\bar{w}}{1 - \bar{w}}\right)\right) \le \ln(q - p^T \bar{x}) - \alpha\gamma\left(\frac{\bar{w}}{1 - \bar{w}}\right),$$

because $\ln(1 + \varepsilon) \le \varepsilon$ for any $\varepsilon > -1$.

Furthermore, from Lemma 2.1,

$$-\sum_{i=1}^{m} w_i \ln(b_i - A_i x_\alpha) \le -\sum_{i=1}^{m} w_i \ln(\bar{s}) + \frac{\alpha^2}{2(1 - \alpha)}\left(\frac{\bar{w}}{1 - \bar{w}}\right).$$

Combining these two inequalities yields

$$F_{q,p}(\bar{x} + \alpha \bar{d}) \le F_{q,p}(\bar{x}) + \left(\frac{\bar{w}}{1 - \bar{w}}\right)\left[-\gamma\alpha + \frac{\alpha^2}{2(1 - \alpha)}\right]. \blacksquare$$

The last two results of this section give two bounds on the value of $p^T x$ for $x \in X$, if the w–center of X is known.

Lemma 5.1. Let $\bar{x}$ be the w–center of X. Let $p \in R^n$ be given. Then for all $x \in X$,

$$p^T x \le p^T \bar{x} + p^T \bar{d}(1 - \bar{w})/\bar{w} ,$$

where $\bar{d}$ is given by (5.3).

Proof: By design, $\bar{x} + \bar{d}$ is the solution to problem of maximizing x over $x \in E_{IN}$. By Remark 2.1, $\bar{x} + \bar{d}(1 - \bar{w})/\bar{w}$ maximizes $p^T x$ over $x \in E_{OUT}$. Because $X \subset E_{OUT}$, $p^T x \le p^T \bar{x} + p^T \bar{d}(1 - \bar{w})/\bar{w}$ for all $x \in X$. $\blacksquare$

Note that Lemma 5.1 parallels the bounding methodology in Anstreicher [1].

Lemma 5.2. Let $\bar{x}$ be the w–center of $\mathcal{X}$. Let $p \in R^n$ be given. Then for all $x \in \mathcal{X}$

$$p^T x \leq p^T \bar{x} + \max_i \left(- \bar{S}^{-1} AGp \right)_i$$

where G is given by (5.4), (5.5).

Proof: Let $x \in \mathcal{X}$ be given, and let $s = b - Ax$. Because $\bar{x}$ is the w–center of $\mathcal{X}$, $M(x - \bar{x}) = 0$ and $w^T \bar{S}^{-1} A = \bar{\pi}^T M$ for some $\bar{\pi} \in R^k$. We first show that

$$p^T(x - \bar{x}) = -s^T \bar{S}^{-1} W \bar{S}^{-1} AGp \tag{5.7}$$

To see this, note that $p^T(x - \bar{x}) = (x - \bar{x})^T QGp = (x - \bar{x})^T (A^T \bar{S}^{-1} W \bar{S}^{-1} A)Gp =$

$(\bar{s} - s)^T \bar{S}^{-1} W \bar{S}^{-1} AGp = w^T \bar{S}^{-1} AGp - s^T \bar{S}^{-1} W \bar{S}^{-1} AGp$

$= \bar{\pi}^T MGp - s^T \bar{S}^{-1} W \bar{S}^{-1} AGp = -s^T \bar{S}^{-1} W \bar{S}^{-1} AGp$, since $MG = O$. Thus

$$\max \; p^T(x - \bar{x}) = \max \; -p^T GA^T \bar{S}^{-1} W \bar{S}^{-1} s \; \leq \; \max \; - p^T GA^T \bar{S}^{-1} W \bar{S}^{-1} s$$

s.t.	$Ax + s = b$	s.t.	$Ax + s = b$	s.t.	$w^T \bar{S}^{-1} s = 1$
	$Mx = g$		$Mx = g$		$s \geq 0$
	$s \geq 0$		$w^T \bar{S}^{-1} s = 1$		
			$s \geq 0$		

$$= \max \; -p^T GA^T \bar{S}^{-1} \lambda \qquad = \max_i \left(- \bar{S}^{-1} AGp \right)_i$$

$$e^T \lambda = 1$$

$$\lambda \geq 0$$

The first equality in the above string follows from (5.7) and because $w^T \bar{S}^{-1}s = 1$ is a redundant constraint. The inequality results due to a relaxation of the constraints. The next equality follows by substituting $\lambda = W \bar{S}^{-1}s$. The final equality is the solution to the linear program in λ. Thus for any $x \in X$,

$$p^T(x - \bar{x}) \leq \max_i \left(- \bar{S}^{-1}AGp \right)_i \quad \blacksquare$$

Note that Lemma 5.2 parallels the bounding methodology in Todd and Burrell [ToBu].

Lemma 5.3. The bound in Lemma 5.2 is sharper than the bound in Lemma 5.1.

Proof: Note that $Gp = p^T \bar{d}((1 - \bar{w})/ \bar{w}) \bar{d}$. Thus the bound in Lemma 5.2 can be expressed as

$$p^T x \leq p^T \bar{x} + p^T \bar{d}((1 - \bar{w})/ \bar{w}) \max_i \left(- \bar{S}^{-1}A \bar{d} \right)_i .$$

It thus suffices to show that $\max_i \left(- \bar{S}^{-1}A \bar{d} \right)_i \leq 1$. Let $r = - \bar{S}^{-1}A \bar{d}$. Then because $\bar{x}$ is the w–center of X, $w^T r = 0$. Also, by (5.2), $r^T Wr \leq \bar{w}/(1 - \bar{w})$. Thus by Proposition 2.2, $r_i \leq 1$ for every $i = 1, \ldots , m$. $\blacksquare$

VI.   A Linear Programming Algorithm

In this section, we apply the material of sections II through V to solve the linear programming problem:

LP : maximize c^Tx

s.t. $Ax \leq b$

$Mx = g$

Note that this form of LP has as special cases the Dantzig "primal standard form" (by setting $A = -I$, and $b = 0$) , the standard "dual form" (by setting (M, g) to be nil), and problems with upper and/or lower bounds (by imbedding submatrices of I or $-I$ in A). The algorithm we employ is thus directly applicable (with no transformations, either explicit or implicit) to a linear program of arbitrary general format.

We do, however, make the following assumptions regarding the LP:

6.1a) There is a known upper bound $\overline{U}$ on the optimal objective value of LP.

6.1b) There is a known point $\overline{x}$ that satisfies $\overline{s} = b - A\overline{x} > 0$ and $M\overline{x} = g$.

6.1c) The last row of (A, b) is $(0, \dots, 0 ; 1)$.

(6.1)

6.1d) A is $m \times n$ and has rank n.

6.1e) M is $k \times n$ and has rank k.

6.1f) The set of optimal solutions to LP is a bounded set.

Assumptions (6.1a) appears to be somewhat restrictive. However, in applications, a practictioner usually knows a bound on the optimal objective function for any linear program that he or she happens to be working with. If no bound is available, one can modify the method of Anstreicher [1] for finding such a bound. Assumption (6.1b) states that we are given a feasible point in the interior of all of the

inequality constraints. This further presumes that the set of always–active inequality constraints is either null or has been identified previously and has been moved into the "Mx = g" constraints.

Assumption (6.1c) is trivial. One can always append a row of zeroes to A and a component of 1 to b. This assumption is essentially an engineering construction to get around certain problems caused by an unbounded feasible region, and allows us to apply Theorem 3.3.

Assumptions (6.1d) and (6.1e) assert that our constraint matrices have full rank. Although this is not necessary, it avoids multiple solutions to problems such as 5.2, and it avoids the need to state results in terms of pseudoinverses. For the more general setting, where full rank is not assumed, see Gay [9], [10]. The most troublesome assumption is (6.1f). We have no convenient way to avoid this assumption, which may be violated in practical problems. However, this assumption is endemic to all treatments of projective transformation algorithms that have appeared (for example, Karmarkar [13], Anstreicher [1], Gay [9], and [10]).

Before presenting the full algorithm with unknown optimal objective value, we first present a version of the algorithm that works when the optimal objective value of the LP is known. The data for the problem then is [A, b, M, g, c, U, w, $\bar{x}, \varepsilon$] where $\bar{x}$ lies in int $X = \{x \in R^n \mid Ax < b, Mx = g\}$, U is the known optimal objective value, w is a vector of weights satisfying $e^T w = 1$, and w > 0, and $\varepsilon > 0$ is a tolerance on the optimality gap. Recall that by assumption (6.1c), that the last row of (A, b) is presumed to be $(0, \ldots, 0; 1)$.

The linear programming algorithm is:

Algorithm 1. (Known optimal objective function value)

Step 0. Choose $w \in R^m$ so that $w > 0$ and $e^T w = 1$. Set $\bar{w} = \min_i \{w_i\}$.

Step 1. Set $\bar{s} = b - A\bar{x}$, $y = A^T \bar{S}^{-1} w$.

Step 2. (Projective Transformation) Set $\tilde{A} = A - \bar{s}y^T$

$$\tilde{c} = c - (U - c^T \bar{x})y$$

Step 3. (Compute direction in Z space) Set $\tilde{Q} = \tilde{A}^T \bar{S}^{-1} W \bar{S}^{-1} \tilde{A}$

$$\text{Set } \tilde{G} = \tilde{Q}^{-1} - \tilde{Q}^{-1} M^T (M \tilde{Q}^{-1} M^T)^{-1} M \tilde{Q}^{-1}$$

$$\text{Set } d = \frac{\tilde{G}\tilde{c}}{\sqrt{\tilde{c}^T \tilde{G}\tilde{c}}} \sqrt{\frac{\bar{w}}{1 - \bar{w}}}$$

Step 4. (Take step in transformed space Z)

$$\text{Set } \alpha = .42$$

$$\text{Set } z_{NEW} = \bar{x} + \alpha d$$

Step 5. (Transform back to original space X)

$$x_{NEW} = \bar{x} + \frac{z_{NEW} - \bar{x}}{1 + y^T(z_{NEW} - \bar{x})}$$

Step 6. (Stopping Criterion) Set $\bar{x} = x_{NEW}$. If $U - c^T \bar{x} \leq \varepsilon$, stop. Otherwise go to Step 1.

At Step 0, a weight vector in w is chosen. Implicit in Karmarkar's original algorithm and its variants is the choice of $w = (1/m) e$, so that equal weight is given to every inequality constraint. However, the user may have some prior belief that

certain constraints are more likely or less likely to be active at an optimal solution. These priors could then be chosen to attach more or less weight to a particular constraint. See the introduction for a further discussion regarding the choosing of weights.

The performance measure of the algorithm will be the following potential function:

$$F_{U,c}(x) = \ln(U - c^T x) - \sum_{i=1}^{m} w_i \ln(b_i - A_i x) .$$ At Step 2 of the algorithm the

polyhedron $\mathbf{X}$ is projectively transformed to the polyhedron $\mathbf{Z}$ defined in (3.2). By Theorem 3.3, the projective transformation $g(\cdot)$ given by (3.3) maps int $\mathbf{X}$ onto int $\mathbf{Z}$. The objective function constraint $c^T x \leq U$ is projectively transformed to

$$\left[c - (U - c^T \bar{x})y \right]^T z \leq \left[U - (U - c^T \bar{x})y^T \bar{x} \right] , \text{ i.e., } \tilde{c}^T z \leq \tilde{U}, \text{ where } \tilde{c} \text{ is defined in}$$

Step 2 and $\tilde{U} = U - (U - c^T \bar{x})y^T \bar{x}$, see (3.5) and (3.6). Thus, by Lemma 3.2, the problem

$$\text{minimize} \quad F_{U,c}(x)$$

$$\text{subject to} \quad x \in \mathbf{X}$$

is equivalent (under the projective transformation $g(\cdot)$) to

$$\text{minimize} \quad G_{\tilde{U},\tilde{c}}(z) = \ln(\tilde{U} - \tilde{c}^T z) - \sum_{i=1}^{m} w_i \ln(\tilde{b}_i - \tilde{A}_i z)$$

$$\text{subject to} \quad z \in \mathbf{Z}$$

where $\tilde{A}$ is defined in Step 2, and $\tilde{b} = b - \bar{s}y^T \bar{x}$. According to Theorem 4.1, $\bar{x}$ is the w–center of $\mathbf{Z}$.

In Step 3, we define the direction d by optimizing over the inner ellipse E_{IN} defined about the w–center $\bar{x}$ of the set $\mathbf{Z}$. This direction is given by (5.3), (5.4),

and (5.5), where $\tilde{A}$ is substituted for A, $\tilde{Q}$ is substituted for Q, $\tilde{G}$ for G, and $\tilde{c}$ for p. In Remark 6.2, we will show that the operations outlined in this step can be performed, unless $\bar{x}$ solves LP. In Step 4, the new point in Z is defined by moving a step length α in the direction d, where $\alpha = .42$.

Because $\tilde{U}$ is the maximum value of $\tilde{c}^T z$ for $z \in Z$, and $Z \subset E_{OUT}$, $\tilde{U} \leq \max\{\tilde{c}^T z \mid z \in E_{OUT}\}$. From Corollary 5.2 we have

$$G_{\tilde{U}, \tilde{c}}(z_{NEW}) \leq G_{\tilde{U}, \tilde{c}}(\bar{x}) - (.267)\left(\frac{\bar{w}}{1 - \bar{w}}\right).$$

In Step 5, z_{NEW} is projectively transformed back to X by $x_{NEW} = h(z_{NEW})$. According to Lemma 3.2,

$$F_{U, c}(x_{NEW}) \leq F_{U, c}(\bar{x}) - (.267)\left(\frac{\bar{w}}{1 - \bar{w}}\right).$$

Note that if $w = (1/m)e$, then $\bar{w} = 1/m$ and the guaranteed decrease at each step is the $.267/(m - 1)$. Furthermore, if $1/\bar{w} = O(m)$, then the guaranteed decrease at each step is $.267/O(m)$. In general, if $1/\bar{w}$ is $O(m^k)$, then the guaranteed decrease at each step is $O(m^{-k})$. Thus if L is the size of the problem instance, the algorithm can be terminated after $O(Lm^k)$ iterations, so long as the set of optimal solutions is bounded. The number of operations needed to perform each iteration is $O(m^3)$ (from Step 3), so that the overall complexity of the algorithm is $O(Lm^{3+k})$ operations. However, using Karmarkar's methodology (or the modification due to Anstreicher [2]) for solving for an inexact solution to the least–squares problem (5.2), the number of operations should be able to be reduced to $O(Lm^{2.5+k})$. With $w = (1/m)e$, then $k = 1$, and the number of operations is $O(Lm^{3.5})$.

Remark 6.1. Use of line search. Steps 4 and 5 can be replaced by a line search, as was suggested by Todd and Burrell [22]. Because the projective transformation g(·) preserves directions from $\bar{x}$ one can perform the line search in the space $\mathcal{X}$ directly. Specifically, Steps 4 and 5 can be replaced by finding a value of δ for which $F_{u,c}(\bar{x} + \delta d)$ is minimized. As shown in Todd and Burrell [22], there will be only one local minimizer of $F_{u,c}(\bar{x} + \delta d)$ for $\delta \geq 0$. The search could be started at the value $\delta = \dfrac{.42}{1 + .42\, y^T d}$, which corresponds to $\alpha = .42$ in the projectively transformed space.

Remark 6.2. Computing d in Step 3. We first will show that the matrices needed in Step 3 have appropriate rank. First, we show that $\tilde{A}$ has full (column) rank. If $\tilde{A}z = 0$, then $Az = \bar{s}y^T z$. However, the last row of A is $(0, \ldots, 0)$, so that $\bar{s}_m y^T z = 0$, and so $y^T z = 0$, because $\bar{s}_m = 1$. Thus $Az = 0$. But by (6.1d), $z = 0$. Therefore, $\tilde{A}$ has rank n, and thus $\tilde{Q} = \tilde{A}^T S^{-1} W S^{-1} \tilde{A}$ has rank n and its inverse exists. Finally, because M has full row rank, $(M \tilde{Q}^{-1} M^T)$ has rank k, and its inverse exists.

Next, we show that the denominator in the computation of Step 3 is positive, unless $\bar{x}$ solves LP. As mentioned in Section V, the matrix $\tilde{G}$ is positive semi-definite, and $\tilde{c}^T \tilde{G} \tilde{c} = 0$ if and only $\tilde{G}\tilde{c} = 0$ if and only if $\tilde{c}$ lies in the row space of M, i.e., if $\tilde{c} = M^T \bar{\pi}$ for some $\bar{\pi} \in R^k$. Thus $\tilde{c}^T z$ is constant (and equals $\bar{\pi}^T g$) for all $z \in Z$, and so $\bar{x}$ is optimal in the transformed linear program:

$\tilde{LP}$:   maximize $\tilde{c}^T z$

     s.t. $z \in Z$

However, because U is the optimal value of LP, then $U - (U - c^T \bar{x}) y^T \bar{x}$ is the optimal value of $\tilde{LP}$, i.e., $\tilde{c}^T \bar{x} = U$, whereby $\bar{x}$ solves LP.

Remark 6.3. Computing d in Step 3 efficiently.

At first glance, the computation of d appears to require working with the matrix $\tilde{Q} = \tilde{A}\,\overline{S}^{-1}W\,\overline{S}^{-1}\tilde{A}$, and $\tilde{A} = A - \overline{s}y^T$ can be **completely** dense if y has all nonzero components (which it will have in all likelihood), whereby $\tilde{Q}$ can be completely dense, even if $Q = A^T\,\overline{S}^{-1}W\,\overline{S}^{-1}A$ is not dense. This could result in a formidable amount of time to compute d, if n is large. Below, we show how to compute d working only with the matrix Q, and thus avoiding the density problems imposed by $\tilde{A}$ and $\tilde{Q}$.

It is elementary to see that $\tilde{Q} = Q - yy^T$, so that $\tilde{Q}$ is a symmetric rank–one update of Q. Furthermore, d is the solution to the following optimization problem:

$$\underset{\tilde{d}}{\text{maximize}}\quad \tilde{c}^T\tilde{d}$$

$$\text{subject to}\quad \tilde{d}^T\tilde{Q}\tilde{d} \le \overline{w}/(1 - \overline{w})$$

$$M\tilde{d} = 0 \ .$$

The Karush–Kuhn–Tucker conditions lead to the following determination of d:

First, solve the system

$$\begin{bmatrix} Q - yy^T & -M^T \\ M & 0 \end{bmatrix}\begin{pmatrix} d^1 \\ \pi^1 \end{pmatrix} = \begin{bmatrix} \tilde{c} \\ 0 \end{bmatrix} \tag{6.2}$$

and then rescale d^1 to $d = \dfrac{d^1\left(\sqrt{\overline{w}/(1 - \overline{w})}\right)}{\sqrt{\tilde{c}^T d^1}}$ \tag{6.3}

Because $Q - yy^T$ is a rank–one modification of Q, we can solve the above by solving the two problems

$$\begin{bmatrix} Q & -M^T \\ M & 0 \end{bmatrix} \begin{pmatrix} d^2 \\ \pi^2 \end{pmatrix} = \begin{bmatrix} \tilde{c} \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} Q & -M^T \\ M & 0 \end{bmatrix} \begin{pmatrix} d^3 \\ \pi^3 \end{pmatrix} = \begin{bmatrix} y \\ 0 \end{bmatrix} . \tag{6.4}$$

By invoking the Sherman–Morrison formula for inverting a rank–one update of a matrix, we set $\quad d^1 = d^2 + \dfrac{y^T d^2}{1 - y^T d^3} \, d^3 \, , \quad \pi^1 = \pi^2 + \dfrac{y^T d^2}{1 - y^T d^3} \, \pi^3$. $\qquad$ (6.5)

It is straightforward to show that $-y^T d^3 < 1$ in the above calculations, and that (d^1, π^1) in (6.5) solves (6.2). Then d^1 can be rescaled by (6.3). Note that solving systems (6.4) necessitates computations with the matrix $\quad Q = A^T \, \overline{S}^{-1} W \, \overline{S}^{-1} A \quad$ and not $\quad \tilde{Q}$.

Remark 6.4. The composition of the direction d.

From (6.4) and (6.5), we see that d^1 (and hence d) is composed of the weighted sum of two directions d^2 and d^3 . The direction d^2 is the projection (in the Q-norm) of $\quad \tilde{c} = c - (U - c^T \overline{x}) \, y \quad$ onto the null space of M, and d^3 is the Q-norm projection of y onto the null space of M. Let G be the Q-norm projection matrix, i.e., $G = Q^{-1} - Q^{-1} M^T (M Q^{-1} M^T)^{-1} M Q^{-1}$. Then

$$d^1 = G \left(c + \left(\frac{y^T d^2}{1 - y^T d^3} - (U - c^T \overline{x}) \right) y \right) ,$$

and so d^1 is composed of a weighted sum of the Q-norm projection of c onto the null space of M and the Q-norm projection of $-y$. The direction Gc corresponds to an affine scaling direction (see Remark 6.6), and the direction $-Gy$ corresponds to a Newton direction for finding the w-center of X (see Part II of this study) . Thus, we see the relation that d^1 (and hence d) is composed of an affine scaling direction plus a Newton direction for finding the w-center of X .

Remark 6.5. Specializing the Algorithm to standard form LP with Lower and Upper Bounds.

In the case when the problem LP has the form

$$\text{maximize } c^T x$$

$$\text{s.t.} \quad Mx = g$$

$$l \leq x \leq u,$$

with some components of l or u equal to $-\infty$ or $+\infty$, respectively, the matrix A of inequality constraints will be composed of two submatrices, one a submatrix of the $n \times n$ identity matrix, I, the other a submatrix of $-I$. Thus $Q = A^T \overline{S^{-1}} W \overline{S^{-1}} A$ will be a diagonal matrix $\overline{D}$ and the expression for d in Step 3 as given in (6.4) simplifies to

$$d^2 = \left[\overline{D}^{-1} - \overline{D}^{-1} M^T (M \overline{D}^{-1} M^T)^{-1} M \overline{D}^{-1} \right] \overline{c}$$

$$d^3 = \left[\overline{D}^{-1} - \overline{D}^{-1} M^T (M \overline{D}^{-1} M^T)^{-1} M \overline{D}^{-1} \right] y$$

(6.6)

with d^1 and then d computed from (6.5) and (6.3). Gay [10] has previously shown that the addition of upper bounds to a standard form linear program only adds to the expression of the diagonal matrix in the computation of the new direction for a standard form linear program. This methodology parallels his own.

Remark 6.6. Relation of Algorithm to Other Variants of Karmarkar's Algorithm

Notice that if y is set equal to zero at Step 1, instead of setting $y = A^T \overline{S}^{-1} w$, that no projective transformation is involved, and the steps of the algorithm then

correspond exactly to an affine scaling algorithm (see Dikin [6], Barnes [3], Vanderbei et al. [25]), with $w = (1/m)\,e$.

Also, Algorithm 1 corresponds to the inequality constrained algorithm of [7] for the case $w = (1/m)\,e$ and (M, g) are nil. From the discussion at the end of Section IV, the projective transformations inherent in Algorithm 1 corresponds to an (affine) scaling of Karmarkar's transformations, as applied to the slack spaces S and T. It thus follows from Theorem 2.4 of Bayer and Lagarias [4] that the steps of Algorithm 1 specialize to that of Karmarkar [13] for the case $w = (1/m)\,e$. Direct (and laborious) computation can alternatively be used to show this fact. In a similar manner, one can show that Algorithm 1 specializes to the projective variants of Gay [9] and Rinaldi [18].

We now present the linear programming algorithm that assumes only an upper bound $\overline{U}$ on the (unknown) optimal objective value. The data for the problem is $[A, b, M, g, c, \overline{U}, w, \overline{x}, \varepsilon]$ where $\overline{x}$ lies in $\mathrm{int}\,\mathcal{X} = \left\{ x \in R^n \mid Ax \le b,\ Mx = g \right\}$, $\overline{U}$ is bound on the optimal objective value, w is a vector of weights satisfying $w > 0$ and $e^T w = 1$, and $\varepsilon > 0$ is a tolerance on the optimality gap. Recall that by assumption (6.1c), that the last row of (A, b) is presumed to be $(0, \ldots, 0\,;\,1)$. The algorithm presented below is a modification of Algorithm 1, for the case of an unknown optimal objective value.

Algorithm 2. (Unknown optimal objective function value)

Step 0. Choose $w \in R^m$ so that $w > 0$ and $e^T w = 1$. Set $\overline{w} = \min_i \left\{ w_i \right\}$.

Step 1. Set $\overline{s} = b - A\,\overline{x},\ y = A^T\,\overline{S}^{-1} w$.

Step 2. (Projective Transformation) Set $\tilde{A} = A - \overline{s}y^T$

Step 3a. (Update Objective Value Bound $\overline{U}$)

Define $\tilde{Q} = \tilde{A}^T \overline{S}^{-1} W \overline{S}^{-1} \tilde{A}$

Define $\tilde{G} = \tilde{Q}^{-1} - \tilde{Q}^{-1} M^T (M\tilde{Q}^{-1}M^T)^{-1} M\tilde{Q}^{-1}$

Define the function $\theta^2(\beta) = \max_i \left(- \overline{S}^{-1} \tilde{A}\tilde{G} \, [c - (\beta - c^T \overline{x}] \, y \,) \right)_i - \beta + c^T \overline{x}$.

If $\theta^2(\overline{U}) \geq 0$, let $\overline{U}$ remain as is.

If $\theta^2(\overline{U}) < 0$, then find a value $\overline{\beta}$ of β in the interval $[\, c^T \overline{x}, \, \overline{U}]$

$$\text{for which } \theta^2(\overline{\beta}) = 0 .$$

Set $\overline{U} \leftarrow \overline{\beta}$.

Set $\tilde{c} = c - (\overline{U} - c^T \overline{x}) \, y$.

Step 3b. (Compute direction in Z space)

$\tilde{c} = c - (\overline{U} - c^T \overline{x}) \, y$

$$d = \frac{\tilde{G}\tilde{c}}{\sqrt{\tilde{c}^T \tilde{G}\tilde{c}}} \sqrt{\frac{\overline{w}}{1 - \overline{w}}}$$

Step 4. (Take step in transformed space Z)

$\alpha = .42$

$z_{NEW} = \overline{x} + \alpha d$

Step 5. (Transform back to original space X)

$$x_{NEW} = \overline{x} + \frac{z_{NEW} - \overline{x}}{1 + y^T(z_{NEW} - \overline{x})}$$

Step 6. (Stopping Criterion) Set $\overline{x} \leftarrow x_{NEW}$. If $\overline{U} - c^T \overline{x} \leq \varepsilon$, stop.

Otherwise go to step 1.

This algorithm is identical to Algorithm 1 except for Step 3a, where the objective function value bound is possibly updated. Notice in Step 3a that $\overline{U}$ decreases (though not strictly) at each iteration. Also notice that solving $\theta^2(\overline{\beta}) = 0$ can always be accomplished if $\theta^2(\overline{U}) < 0$. To see this, note first that because $w^T \overline{S}^{-1} \widetilde{A} = 0$ in Step 3a, that $\theta^2(c^T \overline{x}) = \max_i \left\{ - \overline{S}^{-1} \widetilde{A} \widetilde{G} c \right\}_i \geq 0$. Thus, because $\theta^2(\overline{U}) < 0$ and $\theta^2(\cdot)$ is a continuous (piecewise–linear) function, there exists $\overline{\beta} \in [c^T \overline{x}, \overline{U}]$ for which $\theta^2(\overline{\beta}) = 0$. The determination of $\overline{\beta}$ can be accomplished in $O(m^2)$ operations. Note that Step 3a is an exact analog of the method of Todd and Burrell [22] for updating the objective function bound.

Complexity of Algorithm 2.

Algorithm 2 is essentially a generalization of the algorithm in Todd and Burrell [22], and hence our analysis parallels their own. Our first job is to show that the updating procedure defined in Step 3a produces a valid bound for $c^T x$ over $x \in \mathsf{X}$. By presumption, at the start of step 3a, $\overline{U}$ is a valid obund for $c^T x$ over $x \in \mathsf{X}$. If $\theta^2(\overline{U}) \geq 0$, then $\overline{U}$ remains unchanged, and is still a valid bound. Thus we need only be concerned of the case when $\theta^2(\overline{U}) < 0$. In this case, we solve for $\overline{\beta} \in [c^T \overline{x}, \overline{U}]$ such that $\theta^2(\overline{\beta}) = 0$. (Our previous discussion has established that such a value of β exists and can be computed in $O(m^2)$ operations.)

Proposition 6.1. Each time the upper bound $\overline{U}$ in Algorithm 2 is strictly decreased, the new value of $\overline{U}$ is a valid upper bound for $c^T x$ over $x \in \mathsf{X}$. In particular, the vectors

$$\overline{\pi} = \overline{S}^{-1} W \overline{S}^{-1} \widetilde{A} \widetilde{G} [c - (\overline{U} - c^T \overline{x})] + (\overline{U} - c^T \overline{x}) \overline{S}^{-1} w$$

$$\overline{\lambda} = (M \widetilde{Q}^{-1} M^T)^{-1} M \widetilde{Q}^{-1} [c - (\overline{U} - c^T \overline{x}) y]$$

(6.7)

are a feasible solution to the dual of LP, with dual objective value equal to $\overline{U}$.

Proof: Whenever the value of $\overline{U}$ is strictly decreased, then $\theta^2(\overline{U}) = 0$. Let us define $r = \tilde{G}[c - (\overline{U} - c^T \overline{x})y]$. Then $\theta^2(\overline{U}) = 0$ is equivalent to

$$\max_i \left\{ \overline{S}^{-1} \tilde{A}r \right\}_i - \overline{U} + c^T \overline{x} = 0 \tag{6.8}$$

In order to show that $(\overline{\pi}, \overline{\lambda})$ are feasible for the dual, we need to show that $A^T \overline{\pi} + M^T \overline{\lambda} = c$, and $\overline{\pi} \geq 0$. First, note that since $w^T \overline{S}^{-1} \tilde{A} = 0$, then

$$\tilde{A}^T \overline{S}^{-1} W \overline{S}^{-1} \tilde{A}r = A^T \overline{S}^{-1} W \overline{S}^{-1} \tilde{A}r - y \, s^T \overline{S}^{-1} W \overline{S}^{-1} \tilde{A}r$$

$$= A^T \overline{S}^{-1} W \overline{S}^{-1} \tilde{A}r - y \, \overline{w}^T \overline{S}^{-1} \tilde{A}r = A^T \overline{S}^{-1} W \overline{S}^{-1} \tilde{A}r.$$

Thus $A^T \overline{\pi} = A^T \overline{S}^{-1} W \overline{S}^{-1} \tilde{A}r + (\overline{U} - c^T \overline{x}) A^T \overline{S}^{-1}w$

$$= \tilde{A}^T \overline{S}^{-1} W \overline{S}^{-1} \tilde{A}r + (\overline{U} - c^T \overline{x})y$$

$$= \tilde{Q}r + (\overline{U} - c^T \overline{x})y$$

$$= \tilde{Q}\tilde{G}[c - (\overline{U} - c^T \overline{x})y] + (\overline{U} - c^T \overline{x})y$$

$$= \left[I - M^T(M\tilde{Q}^{-1}M^T)^{-1}M\tilde{Q}^{-1} \right] [c - (\overline{U} - c^T \overline{x})y] + (\overline{U} - c^T \overline{x})y$$

$$= [c - (\overline{U} - c^T \overline{x})y] - M^T \overline{\lambda} + (\overline{U} - c^T \overline{x})y = c - M^T \overline{\lambda}.$$

Thus $A^T \overline{\pi} + M^T \overline{\lambda} = c$.

Next, we show that $\overline{\pi} \geq 0$. Note that

$$\overline{\pi} = \overline{S}^{-1}W\left(\overline{S}^{-1}\tilde{A}r + (\overline{U} - c^T \overline{x})e \right)$$

However, from (6.8) we have $\overline{U} - c^T \overline{x} \geq \left(\overline{S}^{-1}\tilde{A}r \right)_i$ for each $i = 1, \ldots, m$, whereby $\overline{\pi} \geq 0$. Thus $(\overline{\pi}, \overline{\lambda})$ is feasible for the dual.

Furthermore, the dual objective value at $(\overline{\pi}, \overline{\lambda})$ is

$$b^T \bar{\pi} + g^T \bar{\lambda} = \bar{s}^T \bar{\pi} + \bar{x}^T A^T \bar{\pi} + \bar{x}^T M^T \bar{\lambda}$$

$$= \bar{s}^T \left(\bar{S}^{-1} W \bar{S}^{-1} \tilde{A} r \right) + (\bar{U} - c^T \bar{x}) \bar{s}^T \bar{S}^{-1} w$$

$$+ \bar{x}^T (c - M^T \bar{\lambda}) + \bar{x}^T M^T \bar{\lambda}$$

$$= w^T \bar{S}^{-1} \tilde{A} r + (\bar{U} - c^T \bar{x}) + c^T \bar{x}$$

$$= 0 + (\bar{U} - c^T \bar{x}) + c^T \bar{x} = \bar{U} . \blacksquare$$

The performance measures of the algorithm are potential functions of the form

$$F_{\bar{U}, c}(x) = \ln(\bar{U} - c^T x) - \sum_{i=1}^{m} w_i \ln (b_i - A_i x) .$$ At Step 2 of the algorithm the

polyhedron X is projectively transformed to the polyhedron Z defined in (3.2). By

Theorem 3.3, the projective transformation $g(\cdot)$ given in (3.3) maps int X onto

int Z. The objective function bound constraint $c^T x \leq \bar{U}$ is projectively transformed

to $\tilde{c}^T z \leq \tilde{U}$, where $\tilde{c} = [c - (\bar{U} - c^T \bar{x}) y]$ is defined in Step 3b, and $\tilde{U} =$

$\bar{U} - (\bar{U} - c^T \bar{x}) y^T \bar{x}$, see (3.5) and (3.6). Thus, by Lemma 3.2, the problem

$$\text{minimize} \quad F_{\bar{U}, c}(x)$$

$$\text{subject to} \quad x \in X ,$$

is equivalent (under the projective transformation $g(\cdot)$) to

$$\text{minimize} \quad G_{\tilde{U}, \tilde{c}}(z) = \ln(\tilde{U} - \tilde{c}^T z) - \sum_{i=1}^{m} w_i \ln (\tilde{b}_i - \tilde{A}_i z)$$

$$\text{subject to} \quad z \in Z ,$$

where $\tilde{A}$ is defined in Step 2, and $\tilde{b} = b - \bar{s} y^T \bar{x}$. According to Theorem 4.1, $\bar{x}$ is

now the w-center of Z.

In Step 3b, we define the direction d by optimizing over the inner ellipse E_{IN} defined about the w-center $\bar{x}$ of the set Z . This direction is given by (5.3), (5.4), (5.5), where $\tilde{A}$ is substituted for A, $\tilde{Q}$ is substituted for Q, $\tilde{G}$ for G, and $\tilde{c}$ for p. In step 4, the new point z_{NEW} is defined by moving the steplength α in the direction d, where $\alpha = .42$. We next aim to show that

$$G_{\tilde{U},\tilde{c}}(z_{NEW}) \leq G_{\tilde{U},\tilde{c}}(\bar{x}) - (.267)(\bar{w}/(1-\bar{w})) \qquad (6.9)$$

This will follow directly from Corollary 5.2 if we can show

Proposition 6.2. If $\bar{U}$ is defined as in the Step 3a of Algorithm 2, then

$$\tilde{U} \leq \max\left\{ \tilde{c}^T z \mid z \in E_{OUT} \right\} , \text{ where } \tilde{U} = \bar{U} - (\bar{U} - c^T\bar{x})y^T\bar{x} , \text{ and }$$

$$\tilde{c} = c - (\bar{U} - c^T\bar{x})y .$$

Proof: At the end of Step 3a of Algorithm 2, we have a value $\bar{U}$ for which $\theta^2(\bar{U}) \geq 0$, either by modifying $\bar{U}$ or leaving $\bar{U}$ as is. Thus

$$\bar{U} \leq c^T\bar{x} + \max_i \left(-\bar{S}^{-1}\tilde{A}\tilde{G}\tilde{c}\right)_i .$$

This in turn implies that

$$\tilde{U} = \bar{U} - (\bar{U} - c^T\bar{x})y^T\bar{x} \leq [c - (\bar{U} - c^T\bar{x})y]^T\bar{x} + \max_i \left(-\bar{S}^{-1}\tilde{A}\tilde{G}\tilde{c}\right)_i ,$$

i.e.

$$\tilde{U} \leq \tilde{c}^T\bar{x} + \max_i \left(-\bar{S}^{-1}\tilde{A}\tilde{G}\tilde{c}\right)_i .$$

However, by Lemma 5.3,

$$\tilde{c}^T \bar{x} + \max_i \left(-\bar{S}^{-1}\tilde{A}\tilde{G}\,\tilde{c} \right)_i \leq \tilde{c}^T \bar{x} + \tilde{c}^T d\,(1 - \bar{w})/\,\bar{w} =$$

$$\max\left\{ \tilde{c}^T z \mid z \in E_{OUT} \right\} .$$

Thus $\tilde{U} \leq \max \left\{ \tilde{c}^T z \mid z \in E_{OUT} \right\}$. ∎

Thus (6.9) is true, and by Lemma 3.2,

$$F_{\bar{U},c}(x_{NEW}) \leq F_{\bar{U},c}(\bar{x}) - (.267)(\bar{w}/(1 - \bar{w})). \tag{6.10}$$

We next use the following inequality borrowed from Todd and Burrell [22] :

Proposition 6.3. (Todd and Burrell [22]). If U is the optimal value of LP, and $\bar{U}^1 \geq \bar{U}^2 \geq U$, and $x^\circ \in \text{int} \times$ is given, then for all $x \in \text{int} \times$,

$$F_{\bar{U}^1,c}(x) \leq F_{\bar{U}^1,c}(x^\circ) - \gamma$$

implies

$$F_{\bar{U}^2,c}(x) \leq F_{\bar{U}^2,c}(x^\circ) - \gamma .$$

Proof: Parallels that in Todd and Burrell [22] . ∎

Inequality 6.10 and Proposition 6.3 are combined to yield the following:

Lemma 6.1. Let $x^\circ, x^1, \ldots,$ be the iterates of Algorithm 2, and let $U^\circ, U^1, \ldots$ be the values of the objective function upper bound at the end of Step 3a. Then

$$F_{U^k,c}(x^{k+1}) \leq F_{U^k,c}(x^\circ) - k(\bar{w}/(1 - \bar{w}))(.267). \text{ ∎}$$

Thus, precisely as in Algorithm 1, the algorithm can be terminated after $O(Lm)$ steps if $w = (1/m)\,e$ (and hence $\bar{w}/(1 - \bar{w}) = 1/(m-1)$) so long as the set of optimal

solutions is bounded. And in general, if $1/\bar{w} = O(m^k)$, the algorithm can be terminated after $O(Lm^k)$ iterations.

Exactly as in Algorithm 1, steps 4 and 5 can be replaced by a line search. See Remark 6.1 for details. Also, exactly as in Algorithm 1, the main computational effort lies in computing d efficiently. See Remarks 6.2 and 6.3 for details. Finally, we note that as a consequence of Proposition 6.1, Algorithm 2 provides dual variables for the objective function bounds at each iteration.

As remarked earlier, Algorithm 2 is simply a modification to Algorithm 1 that uses the same essential methodology of Todd and Burrell [22] for updating the objective function bound. Instead of using the function $\theta^2(\beta)$, we could have used the function $\theta^1(\beta)$ given by

$$\theta^1(\beta) = \sqrt{[c - (\beta - c^T \bar{x})y]^T \tilde{G} [c - (\beta - c^T \bar{x})y]} \cdot \sqrt{(1 - \bar{w})/\bar{w}} - \beta + c^T \bar{x}$$

This function is the analog of Anstreicher's method [1] for updating the objective function bound.
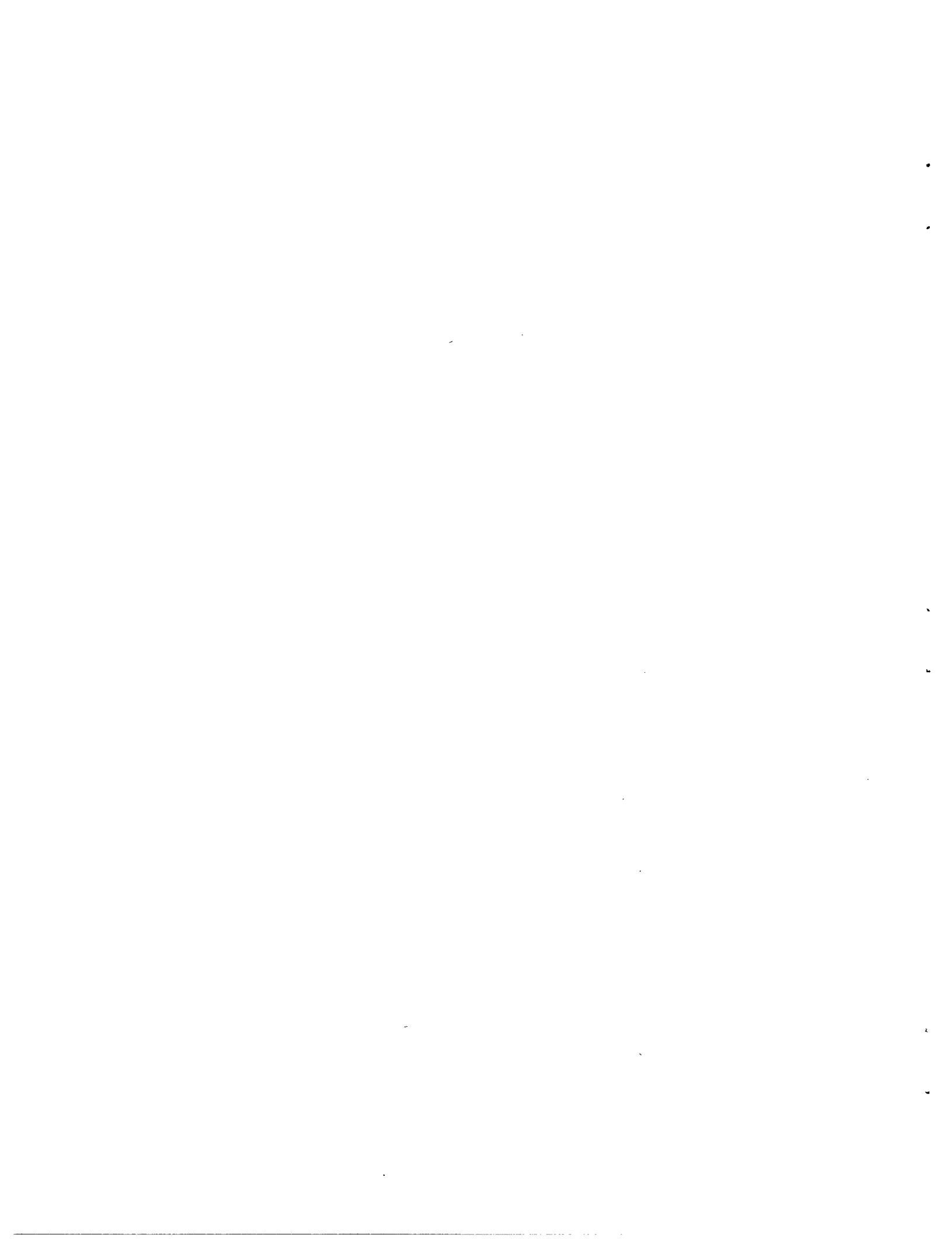
References

[1] Anstreicher, K.M. 1986. A monotonic projective algorithm for fractional linear programming. **Algorithmica** 1, 483-498.

[2] Anstreicher, K.M. 1987. A standard form variant, and safeguarded linesearch, for the modified Karmarkar algorithm, Yale School of Organization and Management, New Haven, Conn.

[3] Barnes, E.R. 1986. A variations on Karmarkar's algorithm for solving linear programming problems. **Mathematical Programming** 36, 174-182.

[4] Bayer, D. A., and J.C. Lagarias. 1987. Karmarkar's linear programming algorithm and Newton's algorithm, AT&T Bell Laboratories, Murray Hill, New Jersey.

[5] Censor, Y. , and A. Lent. 1987. Optimization of 'log x' entropy over linear equality constraints, **SIAM Journal of Control and Optimization** 25, 921-933.

[6] Dikin, I.I. 1967. Iterative solution of problems of linear and quadratic programming. **Dokl. Akadem. Nauk. SSSR** 174, 747-748 [English translation: **Sov. Math. Dokl.** 8, 674-675.

[7] Freund, R. 1988. An analog of Karmarkar's algorithm for inequality constrained linear programs, with a 'new' class of projective transformations for centering a polytope, **Operations Research Letters** 7, 9-14.

[8] Freund, R. 1988. Projective transformations for interior point methods, part II: an algorithm for finding the weighted center of a polyhedral system, M.I.T. Operations Research Center working paper OR 180-88.

[9] Gay, D. 1987. A variant of Karmarkar's linear programming algorithm for problems in standard form. **Mathematical Programming** 37, 81-90.

[10] Gay, D. 1987. Pictures of Karmarkar's linear programming algorithm. Computing Science Technical Report No. 136. AT&T Bell Laboratories, Murray Hill, N.J.

[11] Gonzaga, C.C. 1987. An algorithm for solving linear programming problems in $O(n^3L)$ operations. Memorandum UCB/ERL M87/10. Electronics Research Laboratory, University of California, Berkeley, California.

[12] Grunbaum, B. 1967. **Convex Polytopes**. Wiley, New York.

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

[14] Lagarias, J.C. 1987. The nonlinear geometry of linear programming III. Projective Legendre transform coordinates and Hilbert Geometry. AT&T Bell Laboratories, Murray Hill, N.J.

[15] Monteiro, R.C., and I. Adler. 1987. An $O(n^3 L)$ primal-dual interior point algorithm for linear programming, Dept. of Industrial Engineering and Operations Research, University of California, Berkeley.

[16] Monteiro, R.C., and I. Adler. 1987. An $O(n^3 L)$ interior point algorithm for convex quadratic programming, Dept. of Industrial Engineering and Operations Research, University of California, Berkeley.

[17] Renegar, J. 1988. A polynomial time algorithm, based on Newton's method, for linear programming. **Mathematical Programming** 40 , 59-94.

[18] Rinaldi, g. 1985. The projective method for linear programming with box-type constraints. Instituto di Analisi dei Sistemi ed Informatica del CNR, Viale Manzoni 30, 00185 Rome, Italy.

[19] Rockafellar, R.T. 1970. **Convex Analysis**. Princeton University Press, Princeton, N.J.

[20] G. Sonnevend. 1985. An 'analytical centre' for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming, preprint, Dept. of Numerical Analysis, Institute of Mathematics Eotvos University, 1088, Budapest, Muzeum Korut, 6-8.

[21] G. Sonnevend. 1985. A new method for solving a set of linear (convex) inequalitites and its applications for identification and optimization, preprint, Dept. of Numerical Analysis, Institute of Mathematics Eotvos University, 1088, Budapest, Muzeum Korut, 6-8.

[22] M.J. Todd, and B. Burrelll. 1986. An extension of Karmarkar's algorithm for linear programming using dual variables, **Algorithmica** 1 409-424.

[23] Vaidya, P. 1987. A locally well-behaved potential function and a simple Newton-type method for finding the center of a polytope, AT&T Bell Laboratories, Murray Hill, N.J.

[24] Vaidya, P. 1988. private communication.

[25] Vanderbei, R.J., M.S. Meketon, and B.A. Freedman. 1986. A modification of Karmarkar's linear programming algorithm. **Algorithmica** 1, 395-407.